Weak convergence for the conditional distribution function in a Koziol-Green model under dependent censoring.

Auguste Gaddah[1] and Roel Braekers[1,2]

Abstract

In this paper we consider the conditional Koziol-Green model of Braekers and Veraverbeke (2007) in which they generalized the Koziol-Green model of Veraverbeke and Cadarso Suarez (2000) by assuming that the association between a censoring time and a time until an event is described by a known Archimedean copula function. They got in this way, an informative censoring model with two different types of informative censoring. Braekers and Veraverbeke (2007) derived in this model a nonparametric Koziol-Green estimator for the conditional distribution function of the time until an event, for which they showed the uniform consistency and the asymptotic normality. In this paper we extend their results and prove the weak convergence of the process associated to this estimator. Furthermore we show that the conditional Koziol-Green estimator is asymptotically more efficient in this model than the general copula-graphic estimator of Braekers and Veraverbeke (2005). As last result, we construct an asymptotic confidence band for the conditional Koziol-Green estimator. Through a simulation study, we investigate the small sample properties of this asymptotic confidence band. Afterwards we apply this estimator and its confidence band on a practical data set.

Keywords: censored data, dependent censoring, nonparametric statistics, informative censoring

1 Introduction

In a fixed design regression model, we have at the design points $0 \leq x_1 \leq \ldots \leq x_n \leq 1$, nonnegative responses $Y_1, \ldots, Y_n$ which describe the time until an event, a lifetime or a failure time. These responses are independent random variables and the distribution function of the response Y_i at x_i will be denoted by $F_{x_i}(t) = P(Y_i \leq t)$. In many clinical or industrial trials, the responses $Y_1, \ldots, Y_n$ are subject to random right censoring. For each response, there is a censoring variable C_i with conditional distribution function $G_{x_i}(t) = P(C_i \leq t)$. The observed random variables at a design point x_i are the minimum Z_i and the indicator δ_i $(i = 1, \ldots, n)$, with

$$Z_i = \min(Y_i, C_i) \qquad \text{and} \qquad \delta_i = I(Y_i \leq C_i).$$

For a given fixed design value $x \in [0, 1]$, we write F_x, G_x, H_x for the distribution function of respectively the response Y_x, the censoring variable C_x and the observed variable $Z_x = \min(Y_x, C_x)$ at x. Also we will write $\delta_x = I(Y_x \leq C_x)$. Note that for the design variables x_i, we write $Y_i, C_i, Z_i, F_i, \ldots$ instead of $Y_{x_i}, C_{x_i}, Z_{x_i}, F_{x_i}, \ldots$.

[1]Hasselt University, Center for Statistics, Campus Diepenbeek, Agoralaan, building D, 3590 Diepenbeek, Belgium
[2]Corresponding author, E-mail: roel.braekers@uhasselt.be

To estimate uniquely the conditional distribution function F_x from the observed data, we have to make an assumption about the dependence between the Y_i and C_i for each i (Tsiatis (1975)). It is very common in survival analysis to assume independence between these random variables (conditional on the covariate). However we see that in some practical situations this assumption clearly does not hold. For example in industrial testing, it may occur that some piece of equipment is taken away (is censored) because it shows some sign of future failure. In medicine, we are often interested in the time until dying from a disease. This time may be related to the time until dying from another disease. Therefore we have to model the association between the time until an event and the censoring time. In general, we assume that the joint conditional survival function is given by

$$S_x(t_1, t_2) = P(Y_x > t_1, C_x > t_2) = \mathcal{C}_x(\bar{F}_x(t_1), \bar{G}_x(t_2))$$

where $\mathcal{C}_x$ is a known copula function, depending on x and $\bar{F}_x(t_1) = P(Y_x > t_1)$ and $\bar{G}_x(t_2) = P(C_x > t_2)$ are the survival functions of Y_x and C_x.

Zheng and Klein (1995) introduced this assumption in the case without covariates and derived a copula-graphic estimator for the distribution function of the time until an event. Rivest and Wells (2001) later improved this estimator by considering only the class of Archimedean copulas and found a closed form for the copula-graphic estimator. Braekers and Veraverbeke (2005) generalized this estimator to a fixed design regression.

Therefore we assume, as in Braekers and Veraverbeke (2005), that at a fixed design value $x \in [0, 1]$, the joint survival function is given by

$$S_x(t_1, t_2) = \varphi_x^{[-1]}(\varphi_x(\bar{F}_x(t_1)) + \varphi_x(\bar{G}_x(t_2))) \tag{1}$$

where, for each x, $\varphi_x : [0, 1] \to [0, +\infty]$ is a known continuous, convex, strictly decreasing function with $\varphi_x(1) = 0$. $\varphi_x^{[-1]}$ is the pseudo-inverse of φ_x, as defined in Nelsen (1999) and given by

$$\varphi_x^{[-1]}(s) = \begin{cases} \varphi_x^{-1}(s) & 0 \leq s \leq \varphi_x(0) \\ 0 & \varphi_x(0) \leq s \leq +\infty \end{cases}.$$

From (1), we note that the conditional distribution function of the observed variable Z_x is given by

$$1 - H_x(t) = \bar{H}_x(t) = S_x(t, t) = \varphi_x^{[-1]}(\varphi_x(\bar{F}_x(t)) + \varphi_x(\bar{G}_x(t))). \tag{2}$$

In the design of some clinical trials, we see another type of informative censoring in which the distribution function of the time until an event and the censoring time are related. Koziol and Green (1976) considered a sub-model for the Kaplan-Meier estimator where they assumed that the survival function of the censoring variable is a power of the survival function of the time until event. This sub-model has the advantage that the estimator for the distribution function of the time until event has a simpler form. Veraverbeke and Cadarso Suárez (2000) extended this model to the fixed design regression situation.

In this paper we will further investigate the Koziol-Green type model given by Braekers and Veraverbeke (2007) in which they extended the conditional Koziol-Green model of Veraverbeke and Cadarso Suárez (2000) to the case where the time until an event Y_x depends on the censoring variable C_x. They used the fact that the classical Koziol-Green model is characterized by the conditional independence of Z_x and δ_x. Translating the latter property into model (1) leads to the following assumption: for each covariate value $x \in [0, 1]$,

$$\bar{G}_x(t) = \varphi_x^{[-1]}(\beta_x \varphi_x(\bar{F}_x(t))), \qquad \forall t > 0 \tag{3}$$

where $\beta_x > 0$ is a constant depending only on x. We show this derivation in a few lines. Considering Lemma 1 of Braekers and Veraverbeke (2005), we have in model (1) that

$$\bar{F}_x(t) = \varphi_x^{[-1]}\left(-\int_0^t \varphi_x'(\bar{H}_x(s))dH_x^u(s)\right) \quad \text{and} \quad \bar{G}_x(t) = \varphi_x^{[-1]}\left(-\int_0^t \varphi_x'(\bar{H}_x(s))dH_x^0(s)\right)$$

where $H_x^u(t) = P(Z_x \leq t, \delta_x = 1)$ and $H_x^0(t) = P(Z_x \leq t, \delta_x = 0)$. Using the conditional independence of Z_x and δ_x, we get that

$$\bar{F}_x(t) = \varphi_x^{[-1]}\left(P(\delta_x = 1)\varphi_x(\bar{H}_x(t))\right) \quad \text{and} \quad \bar{G}_x(t) = \varphi_x^{[-1]}\left(P(\delta_x = 0)\varphi_x(\bar{H}_x(t))\right).$$

After eliminating $\varphi_x(\bar{H}(t))$ from both equations and introducing $\beta_x = \frac{P(\delta_x=1)}{P(\delta_x=0)} = \frac{p_{x1}}{p_{x0}}$, we find the assumption (3).

When we consider both types of informative censoring, we rewrite (2) as

$$\begin{aligned} \bar{H}_x(t) &= \varphi_x^{[-1]}\left(\varphi_x(\bar{F}_x(t)) + \beta_x\varphi_x(\bar{F}_x(t))\right) \\ &= \varphi_x^{[-1]}\left((\beta_x + 1)\varphi_x(\bar{F}_x(t))\right). \end{aligned} \tag{4}$$

The remaining part of this paper is as follows. In Section 2, we give the conditional Koziol-Green estimator as it was developed by Braekers and Veraverbeke (2007). After some regularity conditions in Section 3, we prove the weak convergence of this estimator in Section 4. In Section 5, we derive some results for the conditional Koziol-Green estimator such as the efficiency of this estimator over the general copula-graphic estimator of Braekers and Veraverbeke (2005). Furthermore we develop in the same section an asymptotic confidence band for this estimator. We investigate the finite sample properties of this confidence band through a simulation study in Section 6. In Section 7, we apply the conditional Koziol-Green estimator and its asymptotic confidence band on a real data set on the duration of the hospital stay after acute myocardial infarction.

2 The conditional Koziol-Green model

In this section, we describe the conditional Koziol-Green estimator of Braekers and Veraverbeke (2007). For this estimator, it was assumed that the time until an event Y_x was on the one hand associated with the censoring time C_x via the joint conditional survival function, as given in (1). On the other hand the conditional distribution function of the censoring time was assumed to be related to the conditional distribution of the time until an event via the relation (3).

At a fixed design point $x \in]0, 1[$, Braekers and Veraverbeke (2007) found an estimator F_{xh} for the conditional distribution function F_x by rewriting the equation (4) as

$$\bar{F}_x(t) = \varphi_x^{[-1]}\left(\gamma_x\varphi_x(\bar{H}_x(t))\right) \tag{5}$$

with

$$\gamma_x = \frac{1}{\beta_x + 1} = P(\delta_x = 1)$$

and where the last equality results from the definition of β_x above.

To find an estimator for the conditional distribution function $F_x(t)$, Braekers and Veraverbeke (2007) replaced in (5) the different quantities $H_x(t)$ and γ_x by estimators. As in other work with non-parametric regression (Veraverbeke and Cadarso Suárez (2000), Braekers and Veraverbeke (2005)), we consider estimators which involve a sequence of smoothing weights $\{w_{ni}(x, h_n)\}$, depending on a positive bandwidth sequence $\{h_n\}$, tending to zero, as $n \to +\infty$. In the present situation of fixed design points, it is customary to take the Gasser-Müller type weights, given by,

$$\begin{aligned} w_{ni}(x, h_n) &= \frac{1}{c_n(x, h_n)} \int_{x_{i-1}}^{x_i} \frac{1}{h_n} K\left(\frac{x-z}{h_n}\right) dz, \quad i = 1, \ldots, n \\ c_n(x, h_n) &= \int_0^{x_n} \frac{1}{h_n} K\left(\frac{x-z}{h_n}\right) dz. \end{aligned} \tag{6}$$

Here $x_0 = 0$ and K is a known probability density function (kernel).

For the conditional distribution function $H_x(t)$, we take a Stone type estimator (Stone (1977)) given by

$$H_{xh}(t) = \sum_{i=1}^{n} w_{ni}(x, h_n) \mathrm{I}(Z_i \leq t).$$

A similar estimator is taken for the exponent γ_x and is given by

$$\gamma_{xh} = \sum_{i=1}^{n} w_{ni}(x, h_n) I(\delta_i = 1).$$

Combining these estimators in (5), we find an estimator for the conditional distribution function $F_x(t)$ by

$$\bar{F}_{xh}(t) = \varphi_x^{[-1]} \left(\gamma_{xh} \varphi_x(\bar{H}_{xh}(t))\right).$$

Note that the estimator $\bar{F}_{xh}(t)$ has a simpler structure than the copula-graphic estimator of Braekers and Veraverbeke (2005) for the more general fixed design regression model under dependent censoring. Furthermore we see that the estimators for γ_x and $H_x(t)$ only depend on the δ_i and on the Z_i respectively. This result follows from assumption (3), which is equivalent to the assumption that Z_x and δ_x are conditionally independent.

If we take $\varphi_x(t) = -\log(t)$, we see that this estimator equals that of Veraverbeke and Cadarso Suárez (2000) as we expected.

3 Regularity conditions

For the design points $x_1, \ldots, x_n$ we write $\underline{\Delta}_n = \min_{1 \leq i \leq n} (x_i - x_{i-1})$ and $\bar{\Delta}_n = \max_{1 \leq i \leq n} (x_i - x_{i-1})$. The notations $\|K\|_\infty = \sup_{u \in \mathbb{R}} K(u)$, $\|K\|_2^2 = \int_{-\infty}^{+\infty} K^2(u) du$, $\mu_1^K = \int_{-\infty}^{+\infty} uK(u) du$, $\mu_2^K = \int_{-\infty}^{+\infty} u^2 K(u) du$ will be used for the kernel K.

We use the following assumptions on the design and on the kernel.

(C1) $x_n \to 1$, $\bar{\Delta}_n = O(n^{-1})$, $\bar{\Delta}_n - \underline{\Delta}_n = o(n^{-1})$.

(C2) K is a probability density function with finite support $[-M, M]$ for some $M > 0$, $\mu_1^K = 0$ and K is Lipschitz of order 1.

The assumption (C1) expresses that the chosen design points are asymptotically equidistant points, selected uniformly over the whole interval $[0, 1]$. This implies that, for $c_n(x, h_n)$ defined in (6), $c_n(x, h_n) = 1$ for n sufficiently large. Therefore we may take $c_n(x, h_n) = 1$ in all proofs of the asymptotic results.

If L is any distribution, then T_L denotes the right endpoint of its support ($T_L = \inf\{t : L(t) = L(+\infty)\}$). We note that $T_{H_x} = T_{F_x} = T_{G_x}$. To obtain our results, we need some smoothness conditions. For a fixed $0 < T < T_{F_x}$,

(C3) $\dot{F}_x(t) = \frac{\partial}{\partial x} F_x(t)$, $\ddot{F}_x(t) = \frac{\partial^2}{\partial x^2} F_x(t)$ exist and are continuous in $(x, t) \in [0, 1] \times [0, T]$

(C4) $\dot{\beta}_x = \frac{\partial}{\partial x} \beta_x$, $\ddot{\beta}_x = \frac{\partial^2}{\partial x^2} \beta_x$ exist and are continuous in $x \in [0, 1]$

The generator $\varphi_x(v)$ of the Archimedean copula needs to satisfy the following properties.

(C5) $\varphi'_x(v) = \frac{\partial}{\partial v} \varphi_x(v)$ and $\varphi''_x(v) = \frac{\partial^2}{\partial v^2} \varphi_x(v)$ are Lipschitz in the x-direction with a bounded Lipschitz constant, and $\varphi'''_x(v) = \frac{\partial^3}{\partial v^3} \varphi_x(v) \leq 0$ exists and is continuous in $(x, v) \in [0, 1] \times]0, 1]$.

These assumptions and the fact that φ_x is a generator for an Archimedean copula, give that $\varphi'_x(v)$ is monotone increasing with $\varphi'_x(v) < 0$ and $\varphi''_x(v)$ is monotone decreasing with $\varphi''_x(v) \geq 0$.

4 Weak convergence result

In this section, we prove the weak convergence of the process $(nh_n)^{1/2}(F_{xh}(\cdot) - F_x(\cdot))$ associated with the Koziol-Green estimator $F_{xh}(t)$ for the conditional distribution function $F_x(t)$. This extends the work of Braekers and Veraverbeke (2007) in which they showed the asymptotic normality in a fixed time point. As in the work of Veraverbeke and Cadarso Suárez (2000) and Braekers and Veraverbeke (2005, 2007), we first need to derive an almost sure representation for the conditional Koziol-Green estimator $F_{xh}(t)$. This result has already been found by Braekers and Veraverbeke (2007). For convenience we formulate their result as a lemma to clarify the following part of the proof.

Lemma 1. Assume the conditions (C1) - (C5), $h_n \to 0$, $\frac{\log n}{nh_n} \to 0$, $\frac{nh_n^5}{\log n} = O(1)$, $T < T_{F_x}$. Then, for $t < T_{F_x}$,

$$F_{xh}(t) - F_x(t) = \sum_{i=1}^{n} w_{ni}(x, h_n) g_{tx}(Z_i, \delta_i) + R_n(x, t)$$

where

$$g_{tx}(Z_i, \delta_i) = -\frac{\varphi_x(\bar{H}_x(t))}{\varphi'_x(\bar{F}_x(t))} (I(\delta_i = 1) - \gamma_x) + \frac{\gamma_x \varphi'_x(\bar{H}_x(t))}{\varphi'_x(\bar{F}_x(t))} (I(Z_i \leq t) - H_x(t))$$

and as $n \to +\infty$,

$$\sup_{0 \leq t \leq T} |R_n(x, t)| = O((nh_n)^{-1} \log n) \text{ a.s.}$$

From this asymptotic representation, we show the weak convergence of $(nh_n)^{1/2}(F_{xh}(\cdot) - F_x(\cdot))$ in the space $l^\infty[0,T]$ of all bounded functions on $[0,T]$ equipped with the supremum-norm. Therefore we prove the weak convergence of the main term in this representation which is a weighted sum of independent functions of the observed quantities. We postpone the proof of the following theorem to the Appendix.

Theorem 1. Assume the conditions (C1) -(C5), $T < T_{F_x}$,

(a) if $nh_n^5 \to 0$ and $(nh_n)^{-1/2} \log n \to 0$, then as $n \to \infty$,

$$(nh_n)^{1/2}(F_{xh}(\cdot) - F_x(\cdot)) \to W(\cdot|x) \text{ in } l^\infty[0,T]$$

(b) If $h_n = Cn^{-1/5}$ for some $C > 0$, then, as $n \to \infty$,

$$(nh_n)^{1/2}(F_{xh}(\cdot) - F_x(\cdot)) \to \tilde{W}(\cdot|x) \text{ in } l^\infty[0,T]$$

where $W(\cdot|x)$ and $\tilde{W}(\cdot|x)$ are Gaussian processes with covariance function given by

$$\begin{aligned}\Gamma_x(t,s) &= ||K||_2^2 \left\{ \frac{\varphi_x(\bar{H}_x(t))\varphi_x(\bar{H}_x(s))}{\varphi'_x(\bar{F}_x(t))\varphi'_x(\bar{F}_x(s))}\gamma_x(1-\gamma_x) \right. \\ &\quad \left. + \frac{\gamma_x^2\varphi'_x(\bar{H}_x(t))\varphi'_x(\bar{H}_x(s))}{\varphi'_x(\bar{F}_x(t))\varphi'_x(\bar{F}_x(s))}(H_x(s\wedge t) - H_x(t)H_x(s)) \right\}. \qquad (7)\end{aligned}$$

$W(\cdot|x)$ has a zero mean function while for $\tilde{W}(\cdot|x)$ this is given by

$$b_{tx} = \frac{1}{2}\mu_2^K C^{5/2} \left\{ \frac{-\varphi_x(\bar{H}_x(t))}{\varphi'_x(\bar{F}_x(t))}\ddot{\gamma}_x + \frac{\gamma_x\varphi'_x(\bar{H}_x(t))}{\varphi'_x(\bar{F}_x(t))}\ddot{H}_x(t) \right\}.$$

5 Some applications

The weak convergence result formulated in the previous section, can be used as a starting point to derive some practical applications. In this section, we first show that the conditional Koziol-Green estimator is asymptotically more efficient in the Koziol-Green model under dependent censoring than the copula-graphic estimator of Braekers and Veraverbeke (2005). A second application is an asymptotic confidence band for the conditional Koziol-Green estimator.

5.1 Efficiency

At any fixed time point t, we note that the asymptotic variance of the copula-graphic estimator of Braekers and Veraverbeke (2005) has, after some lengthy but straightforward calculations the following expression when the Koziol-Green model is satisfied.

$$\delta_x(t,t) = \frac{||K||_2^2}{\varphi'_x(\bar{F}_x(t))^2} \left\{ \gamma_x(1-\gamma_x) \int_0^t \varphi'_x(\bar{H}_x(s))^2 dH_x(s) + \gamma_x^2\varphi'_x(\bar{H}_x(t))^2 H_x(t)(1-H_x(t)) \right\}$$

To show the efficiency of the conditional Koziol-Green estimator over the copula-graphic estimator, we compare both asymptotic variances and get that

$$\frac{\Gamma_x(t,t)}{\delta_x(t,t)} = \frac{\gamma_x(1-\gamma_x)\varphi_x(\bar{H}_x(t))^2 + \gamma_x^2\varphi'_x(\bar{H}_x(t))^2 H_x(t)(1-H_x(t))}{\gamma_x(1-\gamma_x)\int_0^t \varphi'_x(\bar{H}_x(s))^2 dH_x(s) + \gamma_x^2\varphi'_x(\bar{H}_x(t))^2 H_x(t)(1-H_x(t))}$$

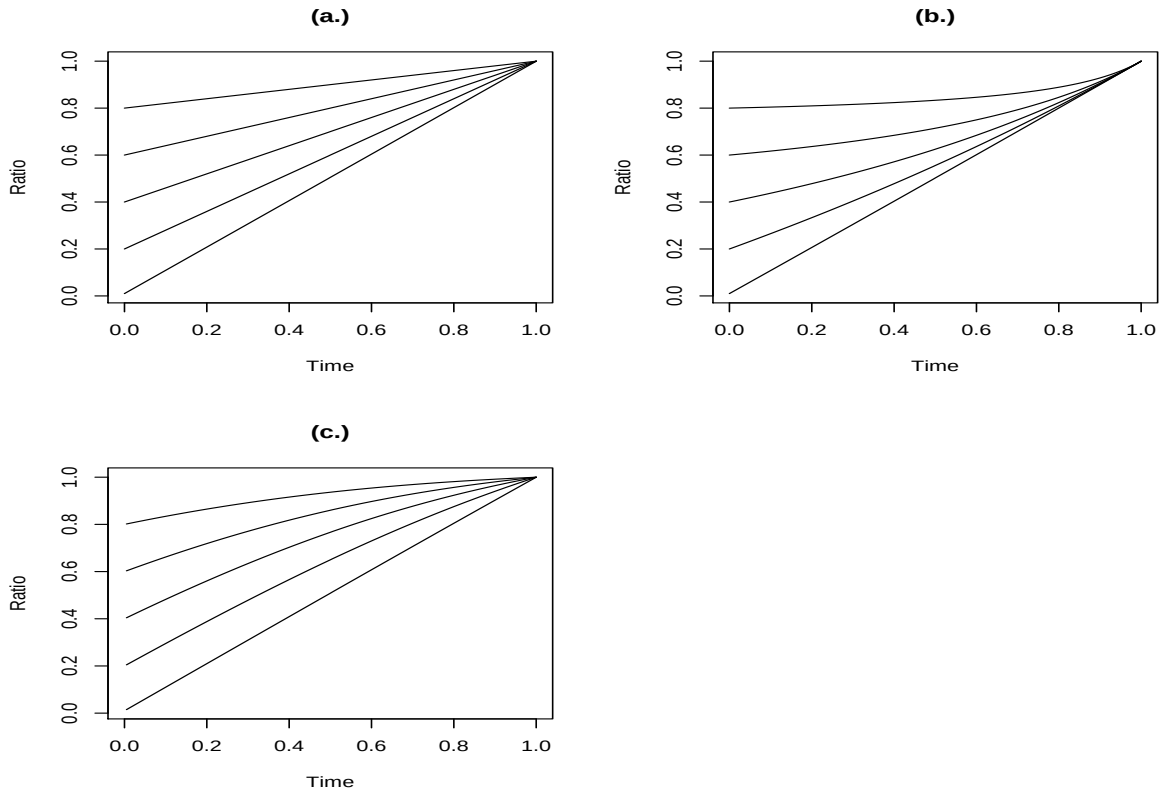

Figure 1: The upperbound for the ratio of variances, given for the independent (a.), Fréchet - Hoeffding lower bound (b.) and Clayton family copula ($\theta = 1$) (c.). Each curve presents a different percentage of uncensored observations (bottom till top: $p_{x1} = 0.01, 0.2, 0.4, 0.6, 0.8$).

$$
\begin{aligned}
&= \frac{\gamma_x(1-\gamma_x)\left(\int\limits_{\bar{H}_x(t)}^{1} |\varphi'_x(w)|dw\right)^2 + \gamma_x^2\varphi'_x(\bar{H}_x(t))^2 H_x(t)(1-H_x(t))}{\gamma_x(1-\gamma_x)\int\limits_0^t \varphi'_x(\bar{H}_x(s))^2 dH_x(s) + \gamma_x^2\varphi'_x(\bar{H}_x(t))^2 H_x(t)(1-H_x(t))} \\
&\leq \frac{\gamma_x(1-\gamma_x)H_x(t)\int\limits_0^t \varphi'_x(\bar{H}_x(s))^2 dH_x(s) + \gamma_x^2\varphi'_x(\bar{H}_x(t))^2 H_x(t)(1-H_x(t))}{\gamma_x(1-\gamma_x)\int\limits_0^t \varphi'_x(\bar{H}_x(s))^2 dH_x(s) + \gamma_x^2\varphi'_x(\bar{H}_x(t))^2 H_x(t)(1-H_x(t))} \leq 1
\end{aligned}
$$

where the inequality follows from the Cauchy-Schwartz inequality. We note that the upper bound goes to 1 if $\gamma_x \to 1$. This was expected since the estimators in both models become a conditional empirical distribution function when there is no censoring. Furthermore we see that this upper bound is 1 when $t \to +\infty$ and is $H_x(t)$ when $\gamma_x \to 0$. In Figure 1 we present the upper bound for three Archimedean copulas, the independent copula ($\varphi_x = -\log(t)$), the Fréchet-Hoeffding lower bound ($\varphi_x(t) = 1-t$) and the Clayton family copula with $\theta = 1$ ($\varphi_x(t) = \frac{1}{t} - 1$). We use in this picture the conditional distribution function $H_x(t)$ to transform the time-axis to $[0, 1]$.

For the independent copula, we see in Figure 1 straight lines for each level of censoring. The Fréchet-

Hoeffding lower bound which expresses a discordant association gives convex lines while the concordant Clayton copula shows concave lines. In each plot, we have that the lines converge to 1 at the right end and all curves are lying between the diagonal and the horizontal line at 1.

5.2 An asymptotic confidence band

From the weak convergence result in Theorem 1, we derive an asymptotic confidence band for the conditional Koziol-Green estimator $F_{xh}(t)$. Like in the work of Hollander and Peña (1989), we introduce an extra parameter λ such that we have a family of bands and which gives some flexibility in the construction of the confidence band. For example, by selecting certain values for λ we can find a more narrow asymptotic confidence band when the sample size is small or moderate, or a more conservative band when we are interested in a time t near the end of the support. The proof of this result is given in the Appendix.

Theorem 2. Assume the conditions (C1) - (C5) with $T < T_{F_x}$, $nh_n^5 \to 0$, $(nh_n)^{-1/2} \log n \to 0$ and $\lambda > 0$. For each $0 < \alpha < 1$, let $c_{\alpha xh}$ be such that, as $n \to +\infty$,

$$P\left(\sup_{0\le t\le T}\left|B_1(L_{xh}(t)) + \frac{\lambda^{1/2}\varphi_x(\bar{H}_{xh}(t))}{\gamma_{xh}\varphi_x'(\bar{H}_{xh}(t))(\bar{H}_{xh}(t)+\lambda H_{xh}(t))}B_2(\gamma_{xh})\right| \le c_{\alpha xh}\right) \to 1-\alpha, \tag{8}$$

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

$$P(F_{xh}(t) - c_{\alpha xh}D_{xh}(t) \le F_x(t) \le F_{xh}(t) + c_{\alpha xh}D_{xh}(t), \text{for all } 0 \le t \le T) \to 1-\alpha$$

where $B_1(s)$ and $B_2(s)$ are independent Brownian bridges and

$$\begin{aligned} L_{xh}(t) &= \frac{\lambda H_{xh}(t)}{\bar{H}_{xh}(t)+\lambda H_{xh}(t)} \\ D_{xh}(t) &= (nh_n\lambda)^{-1/2}\|K\|_2 \frac{\gamma_{xh}\varphi_x'(\bar{H}_{xh}(t))(\bar{H}_{xh}(t)+\lambda H_{xh}(t))}{\varphi_x'(\bar{F}_{xh}(t))}. \end{aligned}$$

6 A simulation study

In this section we perform a simulation study to investigate the finite sample coverage probability of the asymptotic confidence band of Theorem 2. The covariance structure of the limiting process $W(.|x)$ in Theorem 1 precludes the possibility to readily find values of $c_{\alpha xh}$ to satisfy (8). As a consequence, exact confidence bands for $F_x(t)$ cannot be obtained. To circumvent this problem, we develop in this section, an asymptotically conservative confidence band. Therefore we start with the fact that the left-hand side of (8) satisfies the inequality

$$\begin{aligned} &P\left(\sup_{0\le t\le T}|B_1(L_{xh}(t))| + \sup_{0\le t\le T}\left|\frac{\lambda^{1/2}\varphi_x(\bar{H}_{xh}(t))}{\gamma_{xh}\varphi_x'(\bar{H}_{xh}(t))(\bar{H}_{xh}(t)+\lambda H_{xh}(t))}B_2(\gamma_{xh})\right| \le c_{\alpha xh}\right) \\ &\le P\left(\sup_{0\le t\le T}\left|B_1(L_{xh}(t)) + \frac{\lambda^{1/2}\varphi_x(\bar{H}_{xh}(t))}{\gamma_{xh}\varphi_x'(\bar{H}_{xh}(t))(\bar{H}_{xh}(t)+\lambda H_{xh}(t))}B_2(\gamma_{xh})\right| \le c_{\alpha xh}\right) \end{aligned} \tag{9}$$

Using the independence of $B_1(L_{xh}(t))$ and $B_2(\gamma_{xh})$, we convolve and rewrite the left-hand side of (9) as

$$\int_0^{c_{\alpha xh}} P\left(\sup_{0\le t\le T}|B_1(L_{xh}(t))| \le c_{\alpha xh}-y\right) dP\left(\sup_{0\le t\le T}\left|\frac{\lambda^{1/2}\varphi\left(\bar{H}_{xh}(t)\right)}{\gamma_{xh}\varphi'\left(\bar{H}_{xh}(t)\right)\left(\bar{H}_{xh}(t)+\lambda H_{xh}(t)\right)}B_2(\gamma_{xh})\right| \le y\right)$$

$$= \int_0^{c_{\alpha xh}} Q_{d_{xh}(T)}\left(c_{\alpha xh} - y\right) dP\left(|N| \leq \frac{y}{M_{xh}\left(\gamma_{xh}, L_{xh}(T), \lambda\right)}\right) \tag{10}$$

where $d_{xh}(T) = \frac{L_{xh}(T)}{1-L_{xh}(T)}$, $M_{xh}\left(\gamma_{xh}, L_{xh}(T), \lambda\right) = (\lambda\beta_{xh})^{(1/2)} \sup_{0\leq t\leq T}\left|\frac{\varphi\left(\bar{H}_{xh}(t)\right)(1-L_{xh}(t,\lambda))}{\varphi_x{}'(H_{xh}(t))\bar{H}_{xh}(t)}\right|$, $\beta_{xh} = \frac{1-\gamma_{xh}}{\gamma_{xh}}$ and N denotes a standard normal random variable.

Mimicking Hollander and Peña (1989), we define a distribution function

$$\begin{aligned} Q^{\star}\left(c_{\alpha x}, \gamma_x, L_x(T), \lambda\right) &= \sqrt{\frac{2}{\pi}} \frac{1}{M_x\left(c_{\alpha x}, \gamma_x, L_x(T), \lambda\right)} \times \\ &\int_0^{c_{\alpha x}} Q_{d_x(T)}\left(c_{\alpha x} - y\right) \exp\left(-\frac{1}{2}\left(\frac{y}{M_x\left(\gamma_x, L_x(T), \lambda\right)}\right)^2\right) dy \end{aligned}$$

where $d_x(T) = \frac{L_x(T)}{1-L_x(T)}$ and $Q_{d_x(T)}$ is defined as

$$\begin{aligned} Q_{d_x(T)}(c_x) &= 1 - 2\Phi\left(-c_x \frac{1+d_x(T)}{d_x(T)^{1/2}}\right) + 2\sum_{k=1}^{\infty}(-1)^k \exp\left(-2c_x^2k^2\right) \times \\ &\left\{\Phi\left(c_x \frac{d_x(T)+2k+1}{d_x(T)^{1/2}}\right) - \Phi\left(-c_x \frac{d_x(T)-2k+1}{d_x(T)^{1/2}}\right)\right\} \end{aligned}$$

with $\Phi(.)$ being the standard normal cumulative distribution function. By choosing $c_{\alpha x}$ to satisfy $Q^{\star}\left(c_{\alpha x}, \gamma_x, L_x(T), \lambda\right) = 1 - \alpha$, we obtain an asymptotically conservative confidence band

$$P\left[F_{xh}(t) - c_{\alpha xh} D_{xh}(t) \leq F_x(t) \leq F_{xh}(t) + c_{\alpha xh} D_{xh}(t)\right] \geq 1 - \alpha \tag{11}$$

To investigate the coverage probabilities of (11), we generate data by taking fixed and equidistant design points $x_i = \frac{i}{n}$ $(i = 1, 2, 3, \ldots, n)$. Also, we assume that the survival times Y_i $(i = 1, 2, 3, ..., n)$ are independent random variables with $Y_i \sim$ Weibull$(a_1 + a_2 x_i, b)$ such that for each design point the conditional survival function $\bar{F}_i(t)$ is given as

$$\bar{F}_i(t) = \exp\left(-\left(\frac{t}{b}\right)^{(a_1+a_2x_i)}\right)$$

for some constants a_1, a_2 such that $a_1 > \wedge(0, -a_2)$ and $b > 0$. Note that $a_1 + a_2 x_i$ characterizes the shape of the survival distribution of the i-th subject whereas b is the scale parameter.

Furthermore, we assume that the censoring intensity parameter $\beta_{x_i} = \exp(a_3 + a_4 x_i)$ $(i = 1, 2, 3, \ldots n)$ for some constants a_3 and a_4. Using the relation

$$\bar{G}_i(t) = \varphi_x^{[-1]}\left(\beta_{x_i}\varphi_x\left(\bar{F}_i(t)\right)\right),$$

we obtain informative censoring times C_i based on the Clayton and Frank copula generator functions $\varphi_x(.)$ at a pre-specified covariate level x with dependence parameter θ as follows:

1. we generate two independent uniform (0,1) random variables u and t.

2. we set $v = c_u^{-1}(t)$, where $c_u(v) = \frac{\partial}{\partial u}\left\{\varphi_x^{(-1)}\left(\varphi_x(u) + \varphi_x(v)\right)\right\}$ and c_u^{-1} is the inverse or quasi-inverse of c_u depending on whether φ_x is a strict or non-strict generator function.

3. we set $C_i = \bar{G}_i^{(-1)}(v)$ and $Y_i = \bar{F}_i^{(-1)}(u)$.

In particular, we use generators $\varphi_x(t) = \frac{1}{\theta}\left(t^{-\theta} - 1\right)$ and $\varphi_x(t) = -\log\left(\frac{\exp(-\theta t)-1}{\exp(-\theta)-1}\right)$ for the Clayton and Frank copulas respectively. We investigate the effect of the association structure on the coverage probabilities by considering different choices of θ. Note that each choice of θ will lead to a different dependence structure for the Clayton and Frank copulas. Therefore, we use Kendall's τ as a measure of dependence structure so as to compare results under the two copula families. This dependence measure is defined as

$$\tau(x) = 1 + 4\int_0^1 \frac{\varphi_x(t)}{\varphi_x'(t)}dt$$

in Nelsen (1999) such that $-1 \leq \tau(x) \leq 1$, where the dependence gets stronger as $\tau(x)$ goes away from zero. Also, we investigate the effect of the censoring intensity on the coverage probabilities. That is, for each value of $\tau(x)$, we study three different sets of parameters a_1, a_2, a_3 and a_4. In the first set $(a_1 = 1, a_2 = 0.5, a_3 = -2.2, a_4 = 2)$, we chose the parameters such that the percentage of censored observations is always smaller than 45% (i.e. light censoring). In the second set $(a_1 = 1, a_2 = 0.5, a_3 = -0.2, a_4 = 0.4)$, the percentage of censored observations is inclusively between 45 and 55% (i.e. medium censoring); whereas in the third set $(a_1 = 1, a_2 = 0.5, a_3 = 0.2, a_4 = 0.5)$, the parameters are such that the percentage of censored observations is always greater than 55% (i.e. heavy censoring). At each combination of parameters, we generate 2000 samples, each of a size n. For each of these samples, we estimate the conditional Koziol-Green survival distribution at a pre-specified covariate level x together with the corresponding 95% confidence band. We use the Gasser-Müller weights given in (6) with the biquadratic kernel $K(z) = (15/16)\left(1 - z^2\right) I\left(|z| \leq 1\right)$, since it is the most used type of weights in fixed design settings. Also, we use bandwidth $h_n = \left(\log n/n^{3/2}\right)^{2/11}$ so that as $n \to +\infty$, $nh_n \to 0$ and $(nh_n)^{-1/2}\log n \to 0$. Note that this bandwidth is based on the assumption made in Theorem 2 whereas the optimum bandwidth choice is still a topic of future research.

Next, we compute the coverage probability as the percentage of samples for which the confidence band at x covers its corresponding true survival distribution. In particular, we consider estimation at $x = 0.97$ and $x = 0.65$ as extreme and non-extreme covariate levels respectively in order to get some insight into the effect of x on the coverage probabilities. Also, we consider the cases $\lambda = 1$ and $\lambda = \gamma_x^2$ so as to obtain less and more conservative confidence bands respectively. In addition, we repeat the above process for different values of n (i.e. $n = 20, 30, 50, 100, 200, 300$) so as to examine also, the influence of n on the coverage probabilities. Nevertheless, we report only results corresponding to the minimum sample size (i.e. $n = 50$) for which the coverage probabilities (at extreme or non-extreme covariate level) are at least their corresponding nominal confidence level. Note that the results for $\tau = 0$ are only given in Table 1 since it represents the independent copula which is a special case for both the Clayton and Frank copula when $\theta \to 0$.

In Tables 1 and 2 we observe that use of Clayton and Frank copulas results in similar coverage probabilities at equivalent censoring intensities and dependence structures. This implies that the choice of the copula function (i.e. Clayton or Frank) does not have a significant influence on the coverage probabilities. However, assuming $\lambda = \gamma_x^2$, leads to a non-decreasing trend in the coverage probability with increasing censoring intensity. This can be explained (at least in part) by the fact that as censoring increases,

Covariate level = 0.65		Coverage (%)					
		Clayton ($\lambda = 1$)			Clayton ($\lambda = \gamma_x^2$)		
Dependence	Nominal (%)	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
$\tau = -0.99$	90.0	97.9	99.5	99.5	99.8	99.9	99.9
	95.0	99.0	99.9	99.9	99.9	99.9	99.9
	99.0	99.9	99.9	99.9	99.9	99.9	99.9
$\tau = 0.00$	90.0	98.1	98.7	99.5	99.8	99.9	99.9
	95.0	99.4	99.1	99.9	99.9	99.9	99.9
	99.0	99.9	99.9	99.9	99.9	99.9	99.9
$\tau = 0.99$	90.0	94.9	94.3	94.2	99.7	99.9	99.9
	95.0	98.2	97.5	96.4	99.9	99.9	99.9
	99.0	99.6	99.6	99.2	99.9	99.9	99.9
Covariate level = 0.97		Coverage (%)					
		Clayton ($\lambda = 1$)			Clayton ($\lambda = \gamma_x^2$)		
Dependence	Nominal (%)	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
$\tau = -0.99$	90.0	87.9	94.4	94.7	98.7	99.7	99.8
	95.0	91.8	98.5	96.8	99.5	99.9	99.9
	99.0	97.9	99.5	99.3	99.8	99.9	99.9
$\tau = 0.00$	90.0	87.5	93.5	93.5	98.4	99.5	99.9
	95.0	93.2	96.5	97.6	99.7	99.8	99.9
	99.0	96.4	99.1	99.5	99.8	99.9	99.9
$\tau = 0.99$	90.0	82.6	78.5	74.6	98.4	99.7	99.8
	95.0	86.3	85.6	82.1	98.8	99.8	99.9
	99.0	94.8	94.8	92.5	99.8	99.9	99.9

Table 1: Coverage probabilities of the asymptotic confidence band at covariate levels of 0.65 and 0.97 using the Clayton copula

Covariate level = 0.65		Coverage (%)					
		Frank ($\lambda = 1$)			Frank ($\lambda = \gamma_x^2$)		
Dependence	Nominal (%)	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
	90.0	99.1	99.6	99.7	99.9	99.9	99.9
$\tau = -0.99$	95.0	99.4	99.8	99.8	99.9	99.9	99.9
	99.0	99.9	99.9	99.9	99.9	99.9	99.9
	90.0	96.1	94.8	92.9	99.6	99.9	99.9
$\tau = 0.99$	95.0	97.8	98.1	96.5	99.9	99.9	99.9
	99.0	99.8	99.5	99.4	99.9	99.9	99.9
Covariate level = 0.97		Coverage (%)					
		Frank ($\lambda = 1$)			Frank ($\lambda = \gamma_x^2$)		
Dependence	Nominal (%)	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
	90.0	89.6	96.3	94.1	98.6	99.6	99.9
$\tau = -0.99$	95.0	93.0	98.1	98.5	99.6	99.9	99.9
	99.0	97.1	99.4	99.4	99.8	99.9	99.9
	90.0	80.6	77.2	75.2	98.3	99.4	99.9
$\tau = 0.99$	95.0	88.1	84.9	82.8	98.9	99.7	99.9
	99.0	95.8	94.5	92.7	99.9	99.9	99.9

Table 2: Coverage probabilities of the asymptotic confidence band at covariate levels of 0.65 and 0.97 using the Frank copula

the rate of deviation of the conditional Koziol-Green survival function estimate from the true survival function is negligible compared to the rate at which the bands increase with increasing censoring.

Furthermore, we observe at the extreme covariate level that, the coverage probabilities are at least their corresponding nominal only when we assume $\lambda = \gamma_x^2$. In contrast, the coverage probabilities at the non-extreme covariate level are always at least their corresponding nominal irrespective of whether we assume $\lambda = \gamma_x^2$ or $\lambda = 1$. Also for the non-extreme covariate level, assuming $\lambda = 1$ results in coverage probabilities which are at most those under the assumption that $\lambda = \gamma_x^2$. As already mentioned, assuming $\lambda = 1$ yields less (relative to $\lambda = \gamma_x^2$) conservative confidence bands. As such, the particular choice of λ depends on whether one wants a less conservative confidence band.

7 A real data example: Worcester heart attack study

In this section, we illustrate the asymptotic conditional Koziol-Green confidence band on a real data set. The data set comes from the Worcester Heart Attack Study (WHAS) which has information on more than 8000 admissions. The main objective of this study was to describe trends over time in the incident and survival rates following hospital admission of Acute Myocardial Infarction (AMI) patients. However, we will only consider the 10% random sample of the original data set presented by Hosmer and Lemeshow (1999) (described on pages 24 and 25). Only a small subset of variables as well as patients

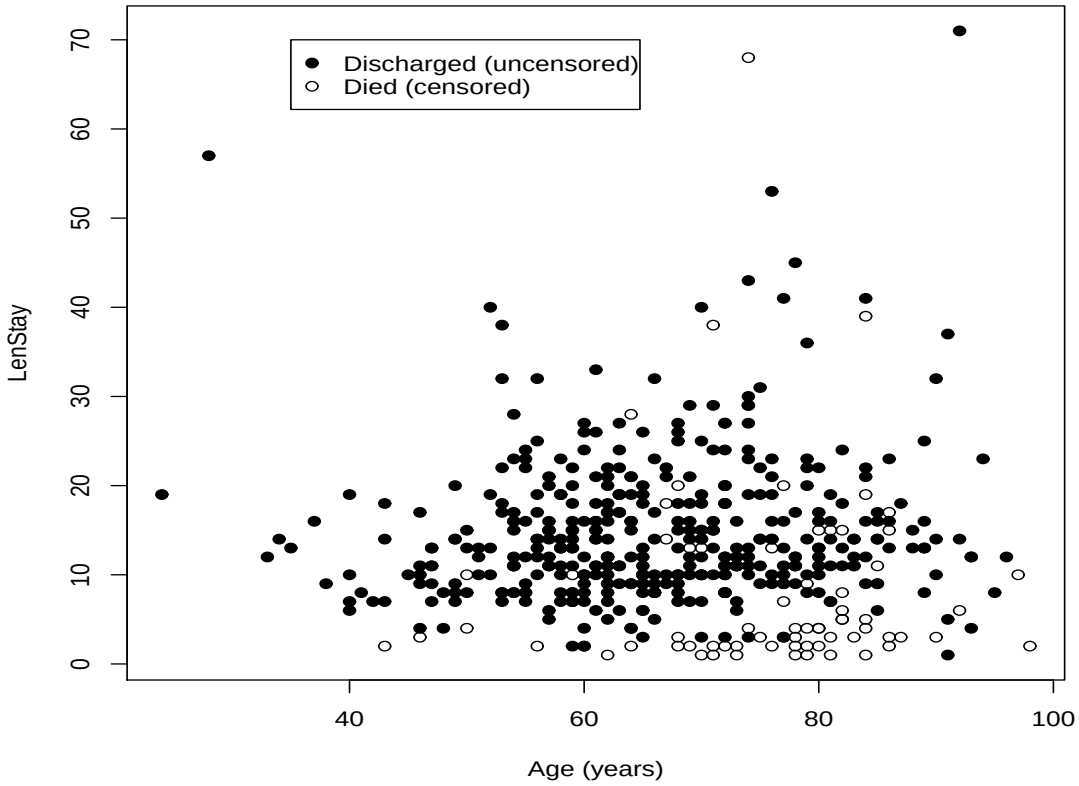

Figure 2: Scatter plot of time spent in hospital (LenStay) versus Age.

with no missing values are included in this subsample. As a result, the data set we utilize in the section has information on only 481 patients. Of these patients, 82 (17%) died while in admission (censored) whereas 399 (83%) were discharged (uncensored). Also, we will mainly be concerned about the time until discharge from hospital of such patients. Note that, the results of this section are only for illustrative purpose, and not to be compared with the analysis of the complete data set. See Hosmer and Lemeshow (1999) for more details and pointers towards the findings from the complete WHAS data set.

In this study, we observe that a patient with severe health condition is likely to die within the first few days of admission. However, if such patient does not die, then he/she is most likely to spend many days in hospital bed. Not only severe health conditions would increase the days that a patient spends in the hospital, but also, for example, an infection from the hospital can increase his/her days in the hospital bed. As such, we allege that time until discharge from hospital Y_i of a patient depends on the time until death in the hospital C_i (i.e. time until discharge has a negative influence on the time until death in the hospital).

Figure 2 is a scatter plot of the observed time spent in hospital (LenStay) verses age of the patient at admission (Age) with a distinction between censored and uncensored patients. From the figure, we observe that most of the censored observations occurred among patients whose age is in the neighborhood

Age (years)	50		75	
	Statistic	P-Value	Statistic	P-Value
Kolmogorov-Smirnov	0.5536	0.9191	1.0033	0.2664
Cramer-von Mises	0.0735	0.7213	0.2934	0.1396
Anderson-Darling	0.8386	0.4531	2.4294	0.0689

Table 3: Conditional Koziol-Green goodness-of-fit test at ages 50 and 75 years

of 80 years. This suggests possible association between censoring time and age of patients at admission. To formally investigate the applicability of the conditional Koziol-Green model, we adapt the partial Koziol-Green goodness-of-fit test of Braekers and Veraverbeke (2003) and calculate the Kolmogorov-Smirnov, the Cramer-von Mises and the Anderson-Darling types of test statistics given respectively as

$$\begin{aligned} K_{nx} &= \left(\frac{nh_n}{||K||_2^2\gamma_{xh}(1-\gamma_{xh})}\right)^{1/2} \max_{1\leq i\leq n-1} |V_{n,i}^1 - V_{n,i}| \\ W_{nx}^2 &= \frac{nh_n}{||K||_2^2\gamma_{xh}(1-\gamma_{xh})} \sum_{i=1}^{n-1} (V_{n,i}^1 - \gamma_{xh}V_{n,i})^2 w_{n(i)}(x;h_n) \\ A_{nx}^2 &= \frac{nh_n}{||K||_2^2\gamma_{xh}(1-\gamma_{xh})} \sum_{i=1}^{n-1} \frac{(V_{n,i}^1 - \gamma_{xh}V_{n,i})^2}{V_{n,i}(1-V_{n,i})} w_{n(i)}(x;h_n) \end{aligned}$$

with $||K||_2^2 = \frac{5}{7}$, $V_{n,i}^1 = \sum_{k=1}^{i} w_{n(k)}(x;h_n)I(\delta_{(k)} = 1)$ and $V_{n,i} = \sum_{k=1}^{i} w_{n(k)}(x;h_n)$, $(i = 1, 2, ...n = 481)$ where $\delta(k)$ and $w_{n(k)}(x;h_n)$ denotes respectively, the censoring indicator and Gasser-Müller weights (with the biquadratic kernel) corresponding to the ordered observed time spent in the hospital. We test at ages 50 and 75 years (i.e. $x = 50$ and 75). Hereby we take as bandwidth, $h_n = 43$. This choice is only to illustrate our method. We considered other choices $h_n = 33$ and $h_n = 53$ (not shown) but they gave similar results. A formal method to find the optimal bandwidth is a research area which we do not enter at this moment.

From Table 3, we observe that the p-values associated with the three goodness-of-fit test statistics are larger than 5% (critical level). Thus, we fail to reject the conditional independence of the Z_x and the δ_x. Therefore, we allege that the conditional Koziol-Green model may be appropriate for the data set at 50 and 75 years.

Using the Clayton and Frank copulas on this data set, we construct and compare confidence bands around the conditional Koziol-Green estimate of the survival (length of stay in hospital) function at ages 50 (middle aged patients) and 75 years (elderly patients). In the sequel, we assume $\lambda = 1$ so as to obtain less conservative (relative to $\lambda = \gamma_x^2$) confidence bands. In addition, we again use the Gasser-Müller weights with the biquadratic kernel and bandwidth $h_n = 43$. Figure 3 is a graphical representation of the conditional Koziol-Green survival distribution at ages 50 and 75 years for the AMI patients together with their corresponding 95% confidence band. In the figure, we consider two different association structures between the survival time (i.e time until discharge) and the censoring time (i.e. time until death in the hospital). Firstly, we assume that the survival time and censoring time are discordant (i.e. $\tau = -0.99$) since we expect that small death times in the hospital are related to large discharge times and vice versa; See Nelsen (1999) for a formal definition of discordance. Secondly, we assume that the

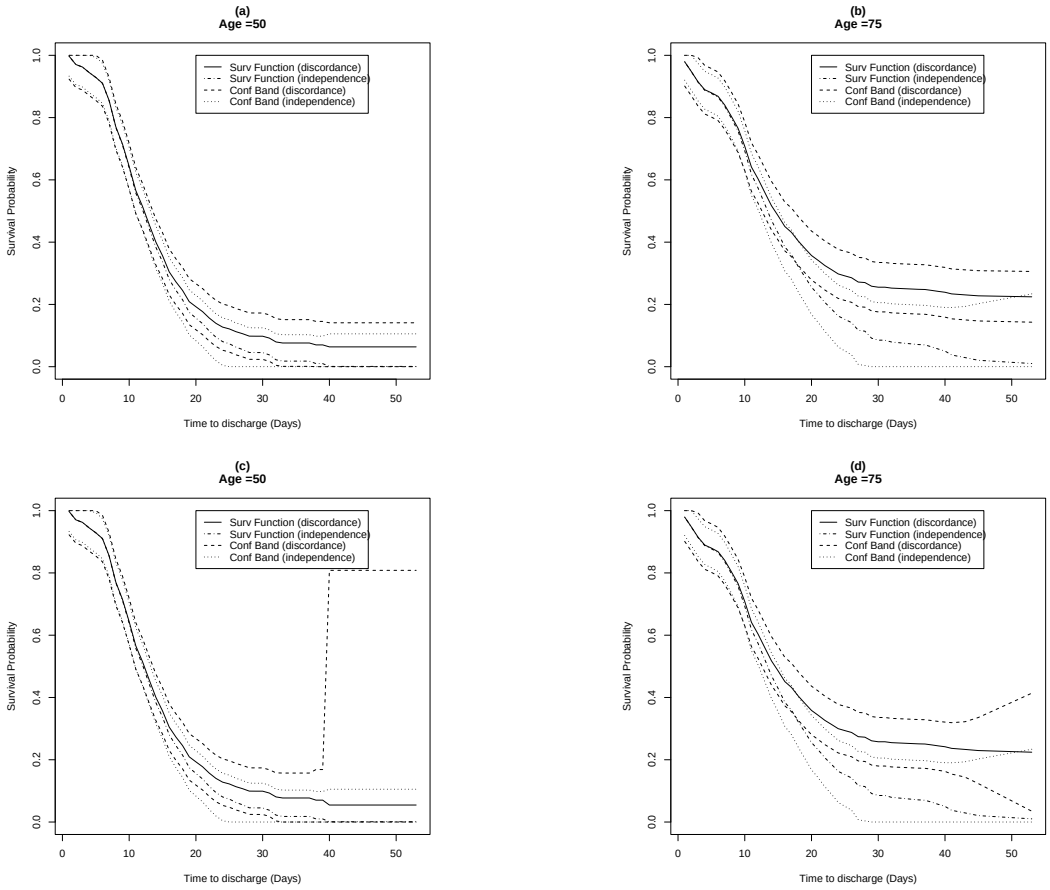

Figure 3: The conditional Koziol-Green survival function estimates (Surv Function) and associated 95% confidence bands (Conf Band) for middle aged (age = 50 years) and elderly (age = 75 years) patients under the Clayton (a and b) and Frank (c and d) copulas.

discharge time and time until death in the hospital are independent (i.e. $\tau = 0$). Note that the later assumption may be wrong for this data set. However, it is commonly used in other real data analyses. Therefore, we consider this choice only as reference for comparison with the result under the discordant association.

At 50 years, we observe under the Clayton and Frank copulas (Figure 3) that the survival distribution under the independent and discordant associations are close to each other. As a result, the confidence band constructed under the independent association clearly covers the survival distribution under the discordant association, and vice versa. This means that, ignoring the possibility of a dependence between the time until discharge from the hospital and the time until death in the hospital may not have any significant influence on the estimates based on the conditional Koziol-Green survival function and its associated 95% confidence band for middle aged patients. However, the same cannot be said about elderly patients since Figure 3 (i.e. (b) and (d)) indicate that the estimated survival distributions under independent and discordant associations at 75 years are clearly separated from each other; and that the confidence band under one form of association does not consistently cover the survival function under the other form of association.

Appendix

Before we prove the weak convergence result in Theorem 1, we give two lemmas about the asymptotic bias and variance of the conditional Koziol-Green estimator.

Lemma 2. Assume (C1), (C2), $F_x(t)$ and β_x satisfy (C3) and (C4) in $[0,T]$ with $T < T_{F_x}$ and φ_x satisfies (C5), $h_n \to 0$. Then, as $n \to +\infty$

$$\sup_{0\le t\le T}\left|\sum_{i=1}^{n} w_{ni}(x,h_n)Eg_{tx}(Z_i,\delta_i) + \frac{\mu_2^K h_n^2}{2}\left(\frac{\varphi_x(\bar{H}_x(t))}{\varphi_x'(\bar{F}_x(t))}\ddot{\gamma}_x - \frac{\gamma_x\varphi_x'(\bar{H}_x(t))}{\varphi_x'(\bar{F}_x(t))}\ddot{H}_x(t)\right)\right| = o(h_n^2) + O(n^{-1}).$$

Proof. For fixed $t \le T$,

$$\sum_{i=1}^{n} w_{ni}(x,h_n)Eg_{tx}(Z_i,\delta_i) = \frac{-\varphi_x(\bar{H}_x(t))}{\varphi_x'(\bar{F}_x(t))}(E\gamma_{xh} - \gamma_x) + \frac{\gamma_x\varphi_x'(\bar{H}_x(t))}{\varphi_x'(\bar{F}_x(t))}(EH_{xh}(t) - H_x(t))$$

By Lemma A.1.b of Van Keilegom and Veraverbeke (1997a), we get the result.

Lemma 3. Assume (C1), (C2), (C3) and (C4) in $[0,T]$ with $T < T_{H_x}$ and φ_x satisfies (C5), $h_n \to 0$, $nh_n \to +\infty$. Then, as $n \to +\infty$

$$\sup_{0\le t\le T}\left|\sum_{i=1}^{n} w_{ni}^2(x,h_n)\mathrm{Cov}(g_{tx}(Z_i,\delta_i), g_{sx}(Z_i,\delta_i)) - \frac{1}{nh_n}\Gamma_x(t,s)\right| = o((nh_n)^{-1})$$

where $\Gamma_x(t,s)$ is given by (7).

Proof. Some straightforward calculations show that

$$\begin{aligned}&\mathrm{Cov}(g_{tx}(Z_i,\delta_i), g_{sx}(Z_i,\delta_i)) = \\ &\quad\frac{\varphi_x(\bar{H}_x(t))\varphi_x(\bar{H}_x(s))}{\varphi_x'(\bar{F}_x(t))\varphi_x'(\bar{F}_x(s))}\gamma_{x_i}(1-\gamma_{x_i}) + \frac{\gamma_x^2\varphi_x'(\bar{H}_x(t))\varphi_x'(\bar{H}_x(s))}{\varphi_x'(\bar{F}_x(t))\varphi_x'(\bar{F}_x(s))}\left(H_{x_i}(t\wedge s) - H_{x_i}(t)H_{x_i}(s)\right)\end{aligned}$$

from which the result follows via standard calculations of asymptotic variances in a fixed design regression situation.

Proof of Theorem 1. From Lemma 1 and 2, we find

$$F_{xh}(t) - F_x(t) = \sum_{i=1}^{n} w_{ni}(x,h_n)\xi_{tx}(Z_i,\delta_i) + h_n^2\bar{b}_{tx} + \bar{R}_n(t)$$

where

$$\begin{aligned}\xi_{tx}(Z_i,\delta_i) &= g_{tx}(Z_i,\delta_i) - Eg_{tx}(Z_i,\delta_i)\\ \sup_{0\le t\le T}|\bar{R}_n(t)| &= O((nh_n)^{-3/4}(\log n)^{3/4}) + o(h_n^2) \text{ a.s.}\end{aligned}$$

and

$$\bar{b}_{tx} = \frac{\mu_2^K h_n^2}{2}\left(\frac{-\varphi_x(\bar{H}_x(t))}{\varphi_x'(\bar{F}_x(t))}\ddot{\gamma}_x + \frac{\gamma_x\varphi_x'(\bar{H}_x(t))}{\varphi_x'(\bar{F}_x(t))}\ddot{H}_x(t)\right).$$

The bias $(nh_n)^{1/2}h_n^2\bar{b}_{tx}$ is $o(1)$ under conditions (a) and equals b_{tx} under conditions (b). Hence it suffices to prove the weak convergence of $W_{hx}(\cdot) = (nh_n)^{1/2}\sum_{i=1}^{n} w_{ni}(x,h_n)\xi_{\cdot x}(Z_i,\delta_i)$ to the Gaussian process $W(\cdot|x)$ with mean zero and covariance function $\Gamma_x(t,s)$.

This will be done in two steps. First we show the convergence of the finite dimensional distributions. Next we verify the asymptotic tightness by Theorem 2.11.9 (Bracketing central limit theorem) of van der Vaart and Wellner (1996).

Convergence of the finite dimensional distributions is that for any $q = 1, 2, \ldots$ and any $0 \le t_1 \le \ldots \le t_q \le T$: $(W_{hx}(t_1), W_{hx}(t_2), \ldots, W_{hx}(t_q)) \xrightarrow{D} N(0, \Gamma_x(t_i,t_j))$. Since $W_{hx}(t_i) = \sum_{k=1}^{n} W_{nki}$ where $W_{nki} = (nh_n)^{1/2} w_{nk}(x,h_n)\xi_{t_i x}(Z_k,\delta_k)$, it suffices to check that (see e.g. Araujo and Giné (1980)),

$$\lim_{n\to+\infty}\sum_{k=1}^{n} E(W_{nki}W_{nkj}) = \Gamma_x(t_i,t_j) \qquad (1 \le i,j \le q)$$

$$\lim_{n\to+\infty}\sum_{k=1}^{n}\int_{\{|W_{nk}|>\varepsilon\}} |W_{nk}|^2 dP = 0$$

for every $\varepsilon > 0$, where $|W_{nk}|^2 = \sum_{i=1}^{q} W_{nki}^2$. Now, applying Lemma 3,

$$\sum_{k=1}^{n} E(W_{nki}W_{nkj}) = (nh_n)\sum_{k=1}^{n} w_{nk}^2(x,h_n)\mathrm{Cov}(g_{t_i x}(Z_k,\delta_k), g_{t_j x}(Z_k,\delta_k)) = \Gamma_x(t_i,t_j) + o(1).$$

Since the functions $\xi_{t_i x}(Z_k,\delta_k)$ are uniformly bounded, it follows that $\max_{1\le k\le n}|W_{nk}| = O((nh_n)^{-1/2})$ a.s. and $\sum_{k=1}^{n}|W_{nk}|^2 = O(1)$ a.s., and hence,

$$\sum_{k=1}^{n}\int_{\{|W_{nk}|>\varepsilon\}} |W_{nk}|^2 dP \le O(1)P(\max_{1\le k\le n}|W_{nk}| > \varepsilon) = o(1).$$

To prove the asymptotic tightness, we denote the process $W_{hx}(t)$ as $W_{hx}(t) = \sum_{i=1}^{n} Z_{ni}(t)$ where $Z_{ni}(t) = (nh_n)^{1/2} w_{ni}(x,h_n)\xi_{tx}(Z_i,\delta_i)$.

To verify the three conditions of Theorem 2.11.9 of van der Vaart and Wellner (1996), we put on $\mathcal{F} = [0,T]$, the semimetric

$$\begin{aligned}\rho(t,t') &= \max\left\{\left|\frac{-1}{\varphi'_x(\bar{F}_x(t))} + \frac{1}{\varphi'_x(\bar{F}_x(t'))}\right|, |\varphi'_x(\bar{H}_x(t)) - \varphi'_x(\bar{H}_x(t'))|,\right.\\ &\qquad \left. |\varphi_x(\bar{H}_x(t)) - \varphi_x(\bar{H}_x(t'))|, \sup_{x'\in[0,1]}\sqrt{|H_{x'}(t) - H_{x'}(t')|}\right\}.\end{aligned}$$

In the third condition, we need the bracketing number $N_{[\,]}(\varepsilon,\mathcal{F},L_2^n)$. This number is defined as the minimal number of sets in a partition of $\mathcal{F} = [0,T] = \bigcup_j \mathcal{F}_{\varepsilon j}$ such that for every set $\mathcal{F}_{\varepsilon j}$:

$$\sum_{i=1}^{n} E\left[\sup_{t,t'\in\mathcal{F}_{\varepsilon j}} |Z_{ni}(t) - Z_{ni}(t')|^2\right] \le \varepsilon^2.$$

Let us divide $\mathcal{F} = [0,T]$ into subintervals $0 = t_0 \leq t_1 \leq \ldots \leq t_q = T$ where $\rho(t,t') \leq C\varepsilon$ for all $t,t' \in [t_{j-1},t_j], j = 1,\ldots,q$ with C some constant which we will determine further on. For the partition $\mathcal{F} = [0,t_1] \bigcup \bigcup_{j=2}^{q}]t_{j-1},t_j]$, we find after some tedious calculations that

$$
\begin{aligned}
&|Z_{ni}(t) - Z_{ni}(t')| \leq (nh_n)^{1/2} w_{ni}(x,h_n) \left(\frac{-1}{\varphi_x'(1)} |\varphi_x(\bar{H}_x(t)) - \varphi_x(\bar{H}_x(t'))| \right. \\
&+ (2\varphi_x(\bar{H}_x(T)) + 2\varphi_x'(\bar{H}_x(T))) \left| \frac{-1}{\varphi_x'(\bar{F}_x(t))} + \frac{1}{\varphi_x'(\bar{F}_x(t'))} \right| - \frac{2}{\varphi_x'(1)} |\varphi_x'(\bar{H}_x(t)) - \varphi_x'(\bar{H}_x(t'))| \\
&\left. + \frac{\varphi_x'(\bar{H}_x(T))}{\varphi_x'(1)} \left(|I(Z_i \leq t) - I(Z_i \leq t')| + |H_{x_i}(t) - H_{x_i}(t')| \right) \right) \qquad (12)
\end{aligned}
$$

So

$$
\begin{aligned}
\sup_{t,t' \in \mathcal{F}_{\varepsilon j}} |Z_{ni}(t) - Z_{ni}(t')|^2 &\leq (nh_n) w_{ni}^2(x,h_n) \{ C_1 (C\varepsilon)^2 \\
&+ \left(\frac{\varphi_x'(\bar{H}_x(T))}{\varphi_x'(1)} \right)^2 |I(Z_i \leq t_j) - I(Z_i \leq t_{j-1})|^2 \}
\end{aligned}
$$

where C_1 is a constant, uniquely determined by the right hand side of (12). For the appropriate choice of C, this leads to

$$
\sum_{i=1}^{n} E \left[\sup_{t,t' \in \mathcal{F}_{\varepsilon j}} |Z_{ni}(t) - Z_{ni}(t')|^2 \right] \leq \varepsilon^2.
$$

Hence the bracketing number $N_{[\,]}(\varepsilon, \mathcal{F}, L_2^n)$ is equal to $O(\varepsilon^{-1})$ and we get

$$
\int_0^{\delta_n} \sqrt{\log N_{[\,]}(\varepsilon, \mathcal{F}, L_2^n)} d\varepsilon = \int_0^{\delta_n} \sqrt{\log O(\varepsilon^{-1})} d\varepsilon \to 0
$$

when $\delta_n \to 0$. We do not need to verify the second condition of Theorem 2.11.9 in van der Vaart and Wellner (1996), since our partition of $\mathcal{F} = [0,T]$ is independent of n. As last condition we have to check whether for all $\eta > 0$,

$$
\sum_{i=1}^{n} E \left[\sup_{0 \leq t \leq T} |Z_{ni}(t)| I \left(\sup_{0 \leq t \leq T} |Z_{ni}(t)| > \eta \right) \right] \to 0 \text{ as } n \to +\infty.
$$

Since $\xi_{tx}(Z_i, \delta_i)$ is bounded uniformly and $\max_{1 \leq i \leq n} w_{ni}(x,h_n) = O((nh_n)^{-1})$ a.s., we get that $\sup_{0 \leq t \leq T} |Z_{ni}(t)| = O((nh_n)^{-1/2})$ a.s., which is always smaller than η for n sufficiently large. So the first condition is also satisfied. By Theorem 2.11.9 of van der Vaart and Wellner (1996), we have that $W_{hx}(\cdot) \to W(\cdot|x)$ in $l^\infty[0,T]$.

Proof of Theorem 2: We note that we can rewrite in Theorem 1 the Gaussian process $W(\cdot|x)$ as, for a given $\lambda > 0$,

$$
\lambda^{-1/2} \|K\|_2 \frac{\gamma_x \varphi_x'(\bar{H}_x(t))(\bar{H}_x(t) + \lambda H_x(t))}{\varphi_x'(\bar{F}_x(t))} B_1(L_x(t)) + \|K\|_2 \frac{\varphi_x(\bar{H}_x(t))}{\varphi_x'(\bar{F}_x(t))} B_2(\gamma_x)
$$

where $\{B_1(s) | 0 \leq s \leq 1\}$ and $\{B_2(s) | 0 \leq s \leq 1\}$ are independent Brownian bridges and

$$
L_x(t) = \frac{\lambda H_x(t)}{\bar{H}_x(t) + \lambda H_x(t)}. \qquad (13)
$$

Using Theorem 1 together with Theorem 1 of Braekers and Veraverbeke (2007), Lemma A.2 of Van Keilegom and Veraverbeke (1997a), lemma A.1. of Braekers and Veraverbeke (2001) and Slutsky's Theorem, we have that

$$(F_{xh}(\cdot) - F_x(\cdot))D_{xh}^{-1}(\cdot) \to B_1(L_x(\cdot)) + \frac{\lambda^{1/2}\varphi_x(\bar{H}_x(\cdot))}{\gamma_x\varphi_x'(\bar{H}_x(\cdot))(\bar{H}_x(\cdot)+\lambda H_x(\cdot))}B_2(\gamma_x) \text{ in } l^\infty[0,T].$$

Analogously, we find that

$$B_1(L_{xh}(\cdot))+\frac{\lambda^{1/2}\varphi_x(\bar{H}_{xh}(\cdot))}{\gamma_{xh}\varphi_x'(\bar{H}_{xh}(\cdot))(\bar{H}_{xh}(\cdot)+\lambda H_{xh}(\cdot))}B_2(\gamma_{xh}) \to B_1(L_x(\cdot))+\frac{\lambda^{1/2}\varphi_x(\bar{H}_x(\cdot))}{\gamma_x\varphi_x'(\bar{H}_x(\cdot))(\bar{H}_x(\cdot)+\lambda H_x(\cdot))}B_2(\gamma_x)$$

in $l^\infty[0,T]$.

Let

$$\begin{aligned}
\eta_x(c) &= P\left(\sup_{0\le t\le T}\left|B_1(L_x(t)) + \frac{\lambda^{1/2}\varphi_x(\bar{H}_x(t))}{\gamma_x\varphi_x'(\bar{H}_x(t))(\bar{H}_x(t)+\lambda H_x(t))}B_2(\gamma_x)\right| \le c\right)\\
\eta_{xh}(c) &= P\left(\sup_{0\le t\le T}\left|B_1(L_{xh}(t)) + \frac{\lambda^{1/2}\varphi_x(\bar{H}_{xh}(t))}{\gamma_{xh}\varphi_x'(\bar{H}_{xh}(t))(\bar{H}_{xh}(t)+\lambda H_{xh}(t))}B_2(\gamma_{xh})\right| \le c\right).
\end{aligned}$$

Since $\sup_{0\le t\le T}|\cdot|$ is a continuous functional, we have that as $n \to +\infty$, $\eta_{xh}(c) \to \eta_x(c)$ for all c. By Lemma 4 below, we have that $\eta_x(\cdot)$ is a continuous function, and hence $\sup_{c>0}|\eta_{xh}(c) - \eta_x(c)| \to 0$ by Pólya's Theorem (see e.g. Serfling (1980)). In particular, we see that $\eta_{xh}(c_{x\alpha h}) - \eta_x(c_{x\alpha h}) \to 0$ and by the definition of $c_{x\alpha h}$ we get that $\eta_x(c_{x\alpha h}) \to 1-\alpha$ which finishes our proof.

Lemma 4. Let $\{B_1(s)|0 \le s \le 1\}$ and $\{B_2(s)|0 \le s \le 1\}$ be independent Brownian bridges. Let $L_x(t),\ (0 \le t \le T)$ be as in (13), $\lambda > 0$. Then

$$\sup_{0\le t\le T}\left|B_1(L_x(t)) + \frac{\lambda^{1/2}\varphi_x(\bar{H}_x(t))}{\gamma_x\varphi_x'(\bar{H}_x(t))(\bar{H}_x(t)+\lambda H_x(t))}B_2(\gamma_x)\right|$$

has a continuous distribution.

The proof of this lemma will not be given since it has the same structure as that of Lemma A4 of Van Keilegom and Veraverbeke (1997b) if we take $Y_x = B_1(L_x(t)) + \frac{\lambda^{1/2}\varphi_x(\bar{H}_x(t))}{\gamma_x\varphi_x'(\bar{H}_x(t))(\bar{H}_x(t)+\lambda H_x(t))}B_2(\gamma_x)$.

Acknowledgement

The authors gratefully acknowledge the financial support from the IAP research Network P6/03 of the Belgian Government (Belgian Science Policy). Furthermore the authors wish to thank the referees for their constructive comments.

References

Araujo, A., Giné, E., 1980. *the Central Limit Theorem for Real and Banach Valued Random Variables* (Wiley, New York).

Braekers, R., Veraverbeke, N., 2003. Testing for the partial Koziol-Green model with covariates, *Journal of Statistical planning and Inference* **115**, 181-192.

Braekers, R., Veraverbeke, N., 2005. A copula-graphic estimator for the conditional survival function under dependent censoring, *The Canadian Journal of Statistics* **33**, 429-447.

Braekers, R., Veraverbeke, N., 2007. A conditional Koziol-Green model under dependent censoring, Accepted to be published by *Statistics & Probability Letters*.

Hollander, M., Peña, E., 1989. Families of confidence bands for the survival function under the general random censorship model and the Koziol - Green model, *The Canadian Journal of Statistics* **17**, 59-74.

Hosmer, W.D., Lemeshow, S., 1999. *Applied Survival Analysis: Regression Modeling of Time to Event Data* (Wiley, New York)

Koziol, J.A., Green, S.B., 1976. A Cramér-von Mises statistic for randomly censored data, *Biometrika* **63**, 465-474.

Nelsen, R.B., 1999. *An introduction to copulas* (Springer-Verlag, New York).

Stone, C.J., 1977. Consistent nonparametric regression, *Annals of Statistics* **5**, 595-645.

Tsiatis, A., 1975. A nonidentifiability aspect of the problem of competing risks, *Proceedings of the National Academy of Sciences of the United States of America* **72**, 20-22.

Rivest, L., Wells, M.T., 2001. A martingale approach to the copula-graphic estimator for the survival function under dependent censoring, *Journal of Multivariate Analysis* **79**, 138-155.

Serfling, R.J., 1980. *Approximation theorems of mathematical statistics* (Wiley, New York)

van der Vaart, A.W., Welner, J.A., 1996. *Weak Convergence and Empirical Processes* (Springer, New York).

Van Keilegom, I., Veraverbeke, N., 1997a. Estimation and bootstrap with censored data in fixed design nonparametric regression, *Annals of the Institute of Statistical Mathematics* **49**, 467-491.

Van Keilegom, I., Veraverbeke, N., 1997b. Weak convergence of the bootstrapped conditional Kaplan-Meier process and its quantile process, *Communications in Statistics: Theory and Methods* 26, 853-869

Veraverbeke, N., Cadarso Suárez, C., 2000. Estimation of the conditional distribution in a conditional Koziol-Green model, *Test* **9**, 97-122.

Zheng, M., Klein, J.P., 1995. Estimates of marginal survival for dependent competing risks based on an assumed copula, *Biometrika* **82**, 127-138.