Made available by Hasselt University Library in https://documentserver.uhasselt.be

Flexible Modeling in the Koziol-Green Model by a Copula Function

Non Peer-reviewed author version

BRAEKERS, Roel & GADDAH, Auguste (2011) Flexible Modeling in the Koziol-Green Model by a Copula Function. In: COMMUNICATIONS IN STATISTICS-THEORY AND METHODS, 40(7). p. 1218-1235.

DOI: 10.1080/03610920903564750
Handle: http://hdl.handle.net/1942/11776

Flexible modeling in the Koziol-Green model by a copula function.

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

December 10, 2009

Abstract

In survival analysis, the classical Koziol-Green random censorship model is commonly used to describe informative censoring. Hereby it is assumed that the distribution of the censoring time is a power of the distribution of the survival time. In this paper, we extend this model by assuming a general function between these distributions. We determine this function from a relationship between the observable random variables which is described by a copula family that depends on an unknown parameter θ. For this setting, we develop a semi-parametric estimator for the distribution of the survival time in which we propose a pseudo-likelihood estimator for the copula parameter θ. As results, we show first the consistency and asymptotic normality of the estimator for θ. Afterwards, we prove the weak convergence of the process associated to the semi-parametric distribution estimator. Furthermore we investigate the finite sample performance of these estimators through a simulation study and finally apply it to a practical data set on survival with malignant melanoma.

Key words and phrases: copula, pseudo-likelihood, random censorship, asymptotic representation, consistency, weak convergence, simulation.

---

1 Introduction

In many industrial or clinical studies, researchers are interested in the time until an event $Y \sim F$. However, due to practical constraints, it is often not possible to fully observe

[1]Corresponding author: roel.braekers@uhasselt.be

[2]Interuniversity Institute for Biostatistics and statistical Bioinformatics, Universiteit Hasselt, Agoralaan 1, 3590 Diepenbeek, Belgium, and Katholieke Universiteit Leuven, Belgium

this lifetime and we only have a lower bound. There exists a censoring variable $C \sim G$ which censors the time until an event. We observe the couple (Z, δ) with

$$Z = \min(Y, C) \sim H \text{ and } \delta = I(Y \leq C).$$

When Y and C are independent continuous random variables, Kaplan and Meier (1958) developed a nonparametric estimator for the survival function $\bar{F}(t) = 1 - F(t)$ of the time until an event. Koziol and Green (1976) described a sub-model of this general model in which they proposed that the censoring variable contains information about the lifetime. They assumed that the survival function of the censoring time is a power of the survival function of the time until the event and is given by

$$\bar{G}(t) = 1 - G(t) = (1 - F(t))^{\beta} = \bar{F}(t)^{\beta}, \qquad t > 0, \tag{1}$$

with $\beta > 0$ a constant. Abdushukurov (1987) and Cheng and Lin (1987), independly, suggested a nonparametric estimator for the distribution function in this sub-model. Csörgő (1988) developed a test to verify whether a practical data set satisfied the conditions of this Koziol-Green sub-model. This was done by using the characteristic of this sub-model that assumption (1) is equivalent to

$$Z \text{ and } \delta \text{ are independent.} \tag{2}$$

However, in some data sets, the independence assumption between the observable random variables is not satisfied. Therefore, we propose in this paper an extension of the classical Koziol-Green model where we assume that the survival function of the censoring variable is a general function of the survival function of the lifetime and is given by

$$\bar{G}(t) = \mu\left(\bar{F}(t), \theta\right), \qquad t > 0, \tag{3}$$

with, for every value of θ, $\mu(w, \theta)$ a non-decreasing function of $w \in [0, 1]$, $\mu(0, \theta) = 0$ and $\mu(1, \theta) = 1$. We select this function $\mu(w, \theta)$ such that the sub-distribution of the uncensored observations is given by

$$H^u(t) = P(Z \leq z, \delta = 1) = \mathcal{C}_\theta\left(P(\delta = 1), H(t)\right) \tag{4}$$

where $\{\mathcal{C}_\theta : \theta \in \Theta\}$ is a known copula family depending on some unknown parameters from a compact parameter space $\Theta \subset I\!R^d$. Since the censoring indicator δ is a discrete variable, we know from Sklar's theorem (Nelsen (2006)) that this copula function is not unique. From Genest and Nešlehová (2007), it is furthermore clear that a copula function alone is not sufficient to describe the association structure between Z and δ.

Their marginal distributions are also needed. However this non-uniqueness of the copula function does not influence the model which we will describe in the next section, as long as we take copula functions in (4) with the same vertical $P(\delta = 1)$-section, as studied in Klement *et al.* (2007).

The remaining of this paper is structured as follows. In Section 2, we first derive the function $\mu(w, \theta)$ from assumption (4). Afterwards we develop a semi-parametric estimator for the distribution of the time until an event in this flexible Koziol-Green model. Hereto we use a pseudo-likelihood method to find an estimate for the parameter θ of the copula function. After giving some regularity conditions in Section 3, we prove the asymptotic normality of the copula parameter θ and show the weak convergence of the semi-parametric distribution estimator $\hat{F}(t)$. In Section 4, we investigate the finite sample performance of our estimators through a simulation study. Finally we illustrate this extended Koziol-Green model on a practical data set on survival with malignant melanoma in Section 5.

2 The flexible Koziol-Green model

In this section, we propose an extension for the classical Koziol-Green model. Hereto we assume that the survival function of the censoring variable is a non-decreasing function $\mu(w, \theta)$ of the survival function of the time until an event as given in (3). We derive this function $\mu(w, \theta)$ from assumption (4). Since the time until an event Y and the censoring variable C are independent, we have that the sub-distribution $H^u(t)$ for the uncensored observations is given by

$$H^u(t) = P(Z \leq t, \delta = 1) = \int_0^t [1 - G(s)] dF(s).$$

Using assumptions (3) and (4), we get that,

$$\int_0^t \mu\left(\bar{F}(s), \theta\right) dF(s) = \mathcal{C}_\theta\left(p_1, 1 - \bar{F}(t)\mu\left(\bar{F}(t), \theta\right)\right), \quad t > 0$$

where we denote the probability of an uncensored observation by $p_1 = P(\delta = 1)$. To solve this integral-equation for the unknown function $\mu(w, \theta)$, we change the integration variable and set $\mu(w, \theta) = \exp(\nu(w, \theta))$. We obtain,

$$\int_w^1 \exp(\nu(s, \theta)) d(s) = \mathcal{C}_\theta\left(p_1, 1 - w \exp(\nu(w, \theta))\right), \quad w \in [0, 1].$$

Differentiating the above equation with respect to w and solving for $\nu(w,\theta)$ gives

$$\nu'(w,\theta) = \frac{1}{w}\left(\frac{1}{\mathcal{C}_\theta^2\left(p_1, 1 - w\exp(\nu(w,\theta))\right)} - 1\right) \tag{5}$$

with $\nu'(w,\theta) = \frac{\partial}{\partial w}\nu(w,\theta)$ and $\mathcal{C}_\theta^2(u,v) = \frac{\partial}{\partial v}\mathcal{C}_\theta(u,v)$. Let us define a non-decreasing function $m(w,\theta) = w\exp(\nu(w,\theta)) = w\mu(w,\theta)$ and $m'(w,\theta) = \frac{\partial}{\partial w}m(w,\theta)$. We get from (5) that

$$m'(w,\theta) = \frac{m(w,\theta)}{w\mathcal{C}_\theta^2\left(p_1, 1 - m(w,\theta)\right)}. \tag{6}$$

Using the theory of differential equations, we solve (6) under the conditions that $m(0,\theta) = 0$ and $m(1,\theta) = 1$ to obtain

$$w = k\left(m(w,\theta), \theta, p_1\right) = \exp\left(-\int_{m(w,\theta)}^{1} \frac{\mathcal{C}_\theta^2\left(p_1, 1-s\right)}{s} ds\right). \tag{7}$$

The function $\mu(w,\theta)$ in (3) which links the survival function of the censoring variable to the survival function of the time until an event is given by

$$\mu(w,\theta) = \frac{m(w,\theta)}{w} \tag{8}$$

where $m(w,\theta)$ is the solution of equation (7). Note that when we assume that Z and δ are independent, we get the classical Koziol-Green model with $\mu(w,\theta) = w^{\frac{p_0}{p_1}}$ and $p_0 = P(\delta = 0)$.

Next we propose an extension of the classical Koziol-Green model. Since the time until an event Y and the censoring time C are independent, we get that the survival function of the observed time until an event Z is equal to

$$\bar{H}(t) = 1 - H(t) = (1 - F(t))(1 - G(t)) = \bar{F}(t)\mu\left(\bar{F}(t), \theta\right) = m\left(\bar{F}(t), \theta\right)$$

where $\mu(w,\theta)$ is given by (8). Solving this equation for $\bar{F}(t)$, we can write the distribution function of interest as

$$\bar{F}(t) = k(\bar{H}(t), \theta, p_1) = \exp\left(-\int_{\bar{H}(t)}^{1} \frac{\mathcal{C}_\theta^2\left(p_1, 1-s\right)}{s} ds\right). \tag{9}$$

To estimate the survival function $\bar{F}(t)$ we substitute estimators for $\bar{H}(t)$, p_1 and θ into (9). Let $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$ be an i.i.d. sample of the observed couple (Z, δ). We estimate $\bar{H}(t)$ and p_1 by their empirical counterparts defined as

$$\widehat{\bar{H}}(t) = \frac{1}{n}\sum_{i=1}^{n} I(Z_i > t) \text{ and } \widehat{p_1} = \frac{1}{n}\sum_{i=1}^{n} I(\delta_i = 1).$$

To estimate the unknown parameter θ, we define the likelihood function in this flexible Koziol-Green model and get

$$L(\theta) = \prod_{i=1}^{n} \mathcal{C}_\theta^2 \left(p_1, H(z_i)\right)^{\delta_i} \left(1 - \mathcal{C}_\theta^2 \left(p_1, H(z_i)\right)\right)^{1-\delta_i}.$$

In this function we note that p_1 and $H(t)$ are unknown and replace them by estimators. This leads to the following pseudo-likelihood function,

$$L(\theta) = \prod_{i=1}^{n} \mathcal{C}_\theta^2 \left(\tilde{p}_1, \tilde{H}(z_i)\right)^{\delta_i} \left(1 - \mathcal{C}_\theta^2 \left(\tilde{p}_1, \tilde{H}(z_i)\right)\right)^{1-\delta_i} \tag{10}$$

where $\tilde{p_1} = \frac{n}{n+1}\hat{p}_1$ and $\tilde{H}(t) = \frac{1}{n+1}\sum_{i=1}^{n} I(Z_i \leq t) = \frac{n}{n+1}\hat{H}(t)$ are rescaled empirical estimators to avoid complications at the end points of the support of the copula family. Note that $\tilde{H}(z_1), \ldots, \tilde{H}(z_n)$ are normalized ranks of $Z_1, \ldots, Z_n$. We estimate the parameter θ by maximizing (10) and denote the pseudo maximum likelihood estimator by $\widehat{\theta}$.

When we substitute the estimators for $\bar{H}(t)$, p_1 and θ into (9), we get an estimator for the survival function of the time until an event which is given by

$$\widehat{\bar{F}}(t) = k(\widehat{\bar{H}}(t), \widehat{\theta}, \widehat{p}_1) = \exp\left(-\int_{\widehat{\bar{H}}(t)}^{1} \frac{\mathcal{C}_{\widehat{\theta}}^2 \left(\widehat{p}_1, 1-s\right)}{s} ds\right).$$

Some remarks:

1. This estimator reduces to the classical Koziol-Green estimator when Z and δ are independent. In this case, we note that the vertical p_1-section of the copula family in (4) is given by
$$\mathcal{C}_\theta(p_1, v) = p_1 v, \qquad \forall v \in [0,1].$$

2. This model is more general than the semiparametric random censorship model of Dikta (1998). He assumes a parametric model for the conditional probability of an uncensored observation, given the observed lifetime. In this flexible Koziol-Green model, we have a semiparametric model for this conditional probability which is given by
$$P(\delta = 1 | Z = z) = \mathcal{C}_\theta^2 \left(p_1, H(t)\right).$$

3. In many generalizations or alternatives for the Koziol-Green model, like for example, Beirlant *et al.* (1992), Gupta *et al.* (1998) and Gupta and Gupta (2007), a

known non-decreasing function $\mu(w,\theta)$ with $\mu(0,\theta)=0$ and $\mu(1,\theta)=1$ is assumed which relates the survival function of the censoring time to the survival function of the lifetime. We can show that these models are special cases of our model since this function $\mu(w,\theta)$ uniquely determines the p_1-section of a copula family $\mathcal{C}_\theta$ in (4). For $v \in [0,1]$, we define this section as

$$\mathcal{C}_\theta(p_1, v) = \int_w^1 \mu(s,\theta)ds$$

with $w \in [0,1]$ the solution of

$$v = 1 - w\mu(w,\theta).$$

3 Some asymptotic results

In this section, we develop the asymptotic theory for the estimators $\widehat{\theta}$ and $\widehat{\bar{F}}(t)$ in the flexible Koziol-Green model. Hereto we use the results of Chen, Linton and Van Keilegom (2003) who proposed conditions under which a parameter estimator that is defined via an estimating equation depending on some nonparametric nuisance functions, is consistent and asymptotically normal. The pseudo maximum likelihood estimator $\widehat{\theta}$ which maximize (10) in this model is the solution of the following score equation

$$\sum_{i=1}^{n} \left\{ \frac{\delta_i \mathcal{C}_\theta^{2'}\left(\tilde{p}_1, \tilde{H}(z_i)\right)}{\mathcal{C}_\theta^2\left(\tilde{p}_1, \tilde{H}(z_i)\right)} - \frac{(1-\delta_i)\mathcal{C}_\theta^{2'}\left(\tilde{p}_1, \tilde{H}(z_i)\right)}{1-\mathcal{C}_\theta^2\left(\tilde{p}_1, \tilde{H}(z_i)\right)} \right\} = 0$$

where $\mathcal{C}_\theta^{2'}(u,v) = \left(\frac{\partial}{\partial\theta_1}\mathcal{C}_\theta^2(u,v), \ldots, \frac{\partial}{\partial\theta_d}\mathcal{C}_\theta^2(u,v)\right)^\tau$ is the vector of partial derivatives for each component of θ. Furthermore we define

$$\begin{aligned} G_n(\theta, p_1, H) &= \frac{1}{n}\sum_{i=1}^{n} g(\theta, p_1, H, Z_i, \delta_i) \\ G(\theta, p_1, H) &= E\left[g(\theta, p_1, H, Z, \delta)\right] \end{aligned}$$

where

$$g(\theta, p_1, H, z, \delta) = \frac{\delta \mathcal{C}_\theta^{2'}\left(p_1, H(z)\right)}{\mathcal{C}_\theta^2\left(p_1, H(z)\right)} - \frac{(1-\delta)\mathcal{C}_\theta^{2'}\left(p_1, H(z)\right)}{1-\mathcal{C}_\theta^2\left(p_1, H(z)\right)}.$$

We denote by θ_0, p_1^* and $H^*(t)$ the true parameter values of θ, p_1 and $H(t)$. Furthermore we see that $G(\theta_0, p_1^*, H^*) = 0$ and that $\widehat{\theta} = \text{argmin}_{\theta\in\Theta}\|G_n(\theta, \tilde{p}_1, \tilde{H})\|$ where $\| \ \|$ denotes the Euclidean norm.

The following notations and regularity conditions will be needed for the asymptotic results below. We define the upper end of the support of the observed lifetime as $\tau_H = \inf\{t : H(t) = 1\}$.

For the copula family $\{\mathcal{C}_\theta : \theta \in \Theta\}$, where Θ is a compact subset of $I\!R^d$, we make the following assumptions

$(A1)$ The function $(u, v, \theta) \to \frac{\dot{\mathcal{C}}^2_\theta(u,v)}{\mathcal{C}^2_\theta(u,v)}$ and $(u, v, \theta) \to \frac{\dot{\mathcal{C}}^2_\theta(u,v)}{1-\mathcal{C}^2_\theta(u,v)}$ are twice continuously differentiable with respect to u, v, and the components of θ, and all these derivatives are continuous in (u, v, θ).

$(A2)$ i. For all $\eta > 0$, there exists $\varepsilon > 0$ such that $\inf_{\|\theta-\theta_0\|>\eta} \|G(\theta, p_1, H)\| > \varepsilon$.
ii. The matrix $\Gamma_1 := \Gamma_1(\theta, p_1^*, H^*) = \frac{\partial}{\partial\theta} G(\theta, p_1^*, H^*)|_{\theta=\theta_0}$ is of full (column) rank.

In the first theorem we show that the pseudo-likelihood estimator $\widehat{\theta}$ is asymptotically normal. The second theorem gives an asymptotic representation for the flexible Koziol-Green estimator $\widehat{\bar{F}}(t)$ which is afterwards used to prove the weak convergence of this estimator to a zero mean Gaussian process in $l^\infty[0, T]$, the space of uniformly bounded real function on $[0, T]$ endowed with the sup-norm. The proofs of the theorems are postponed to the Appendix.

Theorem 1. Assume $(A1)$ and $(A2)$. Then as $n \to \infty$,

$$\sqrt{n}(\widehat{\theta} - \theta_0) \to N(0, \Omega)$$

where $\Omega = (\Gamma_1^{-1})^\tau V \Gamma_1^{-1}$ with

$$\begin{aligned} V &= \mathrm{Var}\Bigg(g(\theta_0, p_1^*, H^*, Z, \delta) + (I(\delta = 1) - p_1^*)\, E\left[\frac{\partial}{\partial u} g(\theta_0, u, H^*, Z', \delta') \Big|_{u=p_1^*} \right] \\ &\quad + E\left[\frac{\partial}{\partial v} g(\theta_0, p_1^*, v, Z', \delta') \Big|_{v=H^*} (I(Z \le Z') - H^*(Z')) \right] \Bigg). \end{aligned}$$

Hereby (Z', δ') is an i.i.d. copy of (Z, δ).

Theorem 2. Assume $(A1)$ and $T < \tau_H$. Then for $t < T$ we get that

$$\widehat{F}(t) - F(t) = \frac{1}{n} \sum_{i=1}^{n} \tilde{g}(t, Z_i, \delta_i) - k_2(\bar{H}^*(t), \theta_0, p_1^*)(\widehat{\theta} - \theta_0) + o_P(n^{-1})$$

where

$$\tilde{g}(t, Z, \delta) = k_1(\bar{H}^*(t), \theta_0, p_1^*)(I(Z \le t) - H^*(t)) - k_3(\bar{H}^*(t), \theta_0, p_1^*)(I(\delta = 1) - p_1^*)$$

with $k_1(z,\theta,p) = \frac{\partial}{\partial z}k(z,\theta,p)$, $k_3(z,\theta,p) = \frac{\partial}{\partial p}k(z,\theta,p)$ and

$$k_2(z,\theta,p) = \left(\frac{\partial}{\partial\theta_1}k(z,\theta,p),\ldots,\frac{\partial}{\partial\theta_d}k(z,\theta,p)\right).$$

We define higher order partial derivatives in a similar way.

Theorem 3. Assume $(A1)$, $(A2)$ and $T < \tau_H$. Then as $n \to \infty$,

$$n^{1/2}(\widehat{F}(\cdot) - F(\cdot)) \to W(\cdot) \quad \text{in} \quad l^\infty[0,T]$$

where $W(\cdot)$ is a zero mean Gaussian process with variance-covariance function Σ_{st} given by

$$\begin{aligned}
\Sigma_{st} &= \mathrm{Cov}(\widetilde{g}(s,Z,\delta),\widetilde{g}(t,Z,\delta)) + k_2\left(\bar{H}^*(t),\theta_0,p_1^*\right)(\Gamma_1^{-1})^\tau \\
&\times \mathrm{Cov}\Bigg(\widetilde{g}(s,Z,\delta), g(\theta_0,p_1^*,H^*,Z,\delta) + (I(\delta=1)-p_1^*)\,E\left[\frac{\partial}{\partial u}g(\theta_0,u,H^*,Z',\delta')\bigg|_{u=p_1^*}\right] \\
&\qquad + E\left[\frac{\partial}{\partial v}g(\theta_0,p_1^*,v,Z',\delta')\bigg|_{v=H^*}(I(Z\le Z')-H^*(Z'))\right]\Bigg) \\
&+ k_2\left(\bar{H}^*(s),\theta_0,p_1^*\right)(\Gamma_1^{-1})^\tau V\Gamma_1^{-1}k_2\left(\bar{H}^*(t),\theta_0,p_1^*\right)^\tau + k_2\left(\bar{H}^*(s),\theta_0,p_1^*\right)(\Gamma_1^{-1})^\tau \\
&\times \mathrm{Cov}\Bigg(\widetilde{g}(t,Z,\delta), g(\theta_0,p_1^*,H^*,Z,\delta) + (I(\delta=1)-p_1^*)\,E\left[\frac{\partial}{\partial u}g(\theta_0,u,H^*,Z',\delta')\bigg|_{u=p_1^*}\right] \\
&\qquad + E\left[\frac{\partial}{\partial v}g(\theta_0,p_1^*,v,Z',\delta')\bigg|_{v=H^*}(I(Z\le Z')-H^*(Z'))\right]\Bigg).
\end{aligned}$$

4 A simulation study

In this section, we set up a simulation study to investigate the finite sample performance of the flexible Koziol-Green model. Hereto we generate samples of observed couples (Z_i,δ_i), $i=1,\ldots,n$ such that

$$H^u(t) = P(Z\le t,\delta=1) = \mathcal{C}_\theta(p_1,H(t)).$$

We assume in this simulation study that the observed lifetimes Z_i, $i=1,2,\ldots,n$ have an exponential distribution ($Z_i \sim \mathrm{Exp}(\lambda)$) with $\lambda = 1.5$ and the indicators δ_i, $i=1,2,\ldots,n$

are Bernoulli distributed with a proportion of uncensored observations of 60% ($p_1 = 60\%$). Furthermore we use the Plackett copula family and BB10 copula family for the relationship (4) between the observed variables. These copula families are given by

$$\mathcal{C}_\theta(u,v) = \begin{cases} \frac{[1+(\theta-1)(u+v)]-\sqrt{[1+(\theta-1)(u+v)]^2-4uv\theta(\theta-1)}}{2(\theta-1)}, & \theta > 0,\ \theta \neq 1 \\ uv, & \theta = 1 \end{cases} \tag{11}$$

and

$$\mathcal{C}_{(\theta_1,\theta_2)}(u,v) = uv[1-\theta_1(1-u^{1/\theta_2})(1-v^{1/\theta_2})], \quad 0 \leq \theta_1 \leq 1,\ \theta_2 > 0. \tag{12}$$

For the Plackett copula family, we set $\theta = 0.2$ while for the BB10 copula family we take $\theta_1 = 0.9$ and $\theta_2 = 0.5$. We choose these copulas to investigate the influence of respectively a real valued or vector valued copula function on the flexible Koziol-Green model. **If we consider the odds ratio of the sub-distributions of the observed variables, we note that this odds ratio is constant over time and equal to θ in the Plackett copula. This practical interpretation of θ is shown in more detail in the next section. For the parameters θ_1 and θ_2 in the BB10 copula family we do not have such a practical interpretation. However by considering two parameters, we are able to model more flexible the association between the observed variables.** For a full description of these copula families we refer to Nelsen (2006) and Joe (1997) respectively.

To generate a sample of observed couples (Z_i, δ_i), $i = 1, \ldots, n$ we use the inverse distribution function method as follows:

1. We generate two independent uniform $(0,1)$ samples u and t.
2. We set $v = (\mathcal{C}_\theta^2)^{-1}(t)$ where $\mathcal{C}_\theta^2 = \frac{\partial}{\partial v}\mathcal{C}_\theta(u,v)$ and $(\mathcal{C}_\theta^2)^{-1}$ is the inverse function of $\mathcal{C}_\theta^2$
3. We define the observed quantities $\delta_i = I(u > 1-p_1)$ and $z_i = -\frac{1}{\lambda}\log(1-v)$.

In this way, we obtain for each copula choice, 100 000 replicated samples of size n. For each of these samples, we compute $\widetilde{p}_1$ and $\widetilde{H}(t)$ and consequently obtain $\widehat{\theta}$ as the maximizer of the pseudo-likelihood

$$L(\theta) = \prod_{i=1}^{n} \mathcal{C}_\theta^2\left(\widetilde{p}_1, \widetilde{H}(Z_i)\right)^{\delta_i}\left(1-\mathcal{C}_\theta^2\left(\widetilde{p}_1, \widetilde{H}(Z_i)\right)\right)^{1-\delta_i}.$$

In Table 1 we report the average bias with standard deviation of the pseudo-likelihood estimate $\widehat{\theta}$ for different sample sizes n. From the table, we observe under each choice of

n	Plackett $\widehat{\theta}$	BB10 $\widehat{\theta}_1$	$\widehat{\theta}_2$
15	7.4268 (902.5344)	0.0490 (0.1835)	0.5326 (4.1292)
30	0.0366 (0.1976)	0.0472 (0.1183)	0.1141 (1.5730)
50	0.0190 (0.1300)	0.0360 (0.1010)	0.0444 (0.6070)
100	0.0085 (0.0833)	0.0235 (0.0858)	0.0226 (0.2594)
200	0.0041 (0.0565)	0.0142 (0.0700)	0.0194 (0.1831)
300	0.0025 (0.0453)	0.0106 (0.0604)	0.0165 (0.1485)
500	0.0018 (0.0351)	0.0070 (0.0483)	0.0108 (0.1143)
700	0.0013 (0.0293)	0.0049 (0.0409)	0.0076 (0.0947)
900	0.0009 (0.0257)	0.0038 (0.0362)	0.0060 (0.0830)

Table 1: Average bias with standard deviation (between brackets) of the pseudo-likelihood parameter estimates.

copula that the bias and standard deviation are decreasing with increasing sample size. This informally suggests that the pseudo likelihood estimate of θ is consistent.

To illustrate the performance of the flexible Koziol-Green estimator

$$\widehat{\widehat{F}}(t) = \exp\left(-\int_{\widehat{\widehat{H}}(t)}^{1} \frac{\mathcal{C}_{\widehat{\theta}}^{2}\left(\widehat{p}_1, 1-s\right)}{s} ds\right)$$

we further compute the bias associated with $\widehat{\widehat{F}}(t)$ at a pre-specified event time $t = 0.5$. In Table 2, we present the absolute bias based on the 100 000 replicates. From the table, we observe that the bias is non-increasing with increasing sample sizes under both copula families. This suggests the consistency of the flexible Koziol-Green estimates. Furthermore, we note that the bias is close to zero (i.e. 2 decimal places approximation). This shows that the flexible Koziol-Green model can readily be used to approximate the true distribution function.

5 Real data illustration: Survival with malignant melanoma

In this section, we apply the estimator developed in Section 2 to a real data set. The data comes from a historical prospective clinical study conducted in the period 1962-77 and is described in Andersen *et al.* (1993). The study took place at the university hospital of

n	Plackett	BB10
15	0.0094 (0.1302)	0.0121 (0.1219)
30	0.0043 (0.0923)	0.0068 (0.0880)
50	0.0030 (0.0716)	0.0040 (0.0689)
100	0.0016 (0.0506)	0.0023 (0.0494)
200	0.0007 (0.0360)	0.0009 (0.0350)
300	0.0005 (0.0293)	0.0006 (0.0286)
500	0.0002 (0.0227)	0.0003 (0.0221)
700	0.0002 (0.0191)	0.0002 (0.0187)
900	0.0002 (0.0169)	0.0002 (0.0165)

Table 2: Average bias with standard deviation (between brackets) of the flexible Koziol-Green estimate $\widehat{\widehat{F}}(0.5)$ under the Plackett and BB10 copulas.

Odense, Denmark and has information on 225 patients with malignant melanoma (skin cancer). The main objective of this study was to access risk factors on the survival rate following a radical operation on these patients. However, we will focus on the estimation of the survival distribution of these patients. Also, only the 205 patients with complete information are considered here. Of these patients, 57 (28%) had an event (i.e. malignant melanoma related death). Whereas, 14 (7%) died from causes other than malignant melanoma and 134 (65%) were alive at the end of the study.

In this study, we believe that death without malignant melanoma or alive at end of study could be an indirect manifestation of the operation. Therefore, we suspect that malignant melanoma free death time or end of study (i.e. censoring time) is informative to the time until death from malignant melanoma (i.e. survival time) through its distribution function. As such, the Koziol-Green model appears to be the outstanding candidate.

	Estimate	Std. Error	P-Value
Intercept	1.8347	0.4461	< 0.0001
Observed time	-0.0015	0.0002	< 0.0001

Table 3: The logistic regression model.

To investigate the Koziol-Green assumption, we use a logistic model of the observed time on the censoring indicator. In Table 3, we observe that these two random variables are highly related since the associated p-value is far less than the nominal 5%. This implies that, the proportionality assumption of the classical Koziol-Green model is violated and

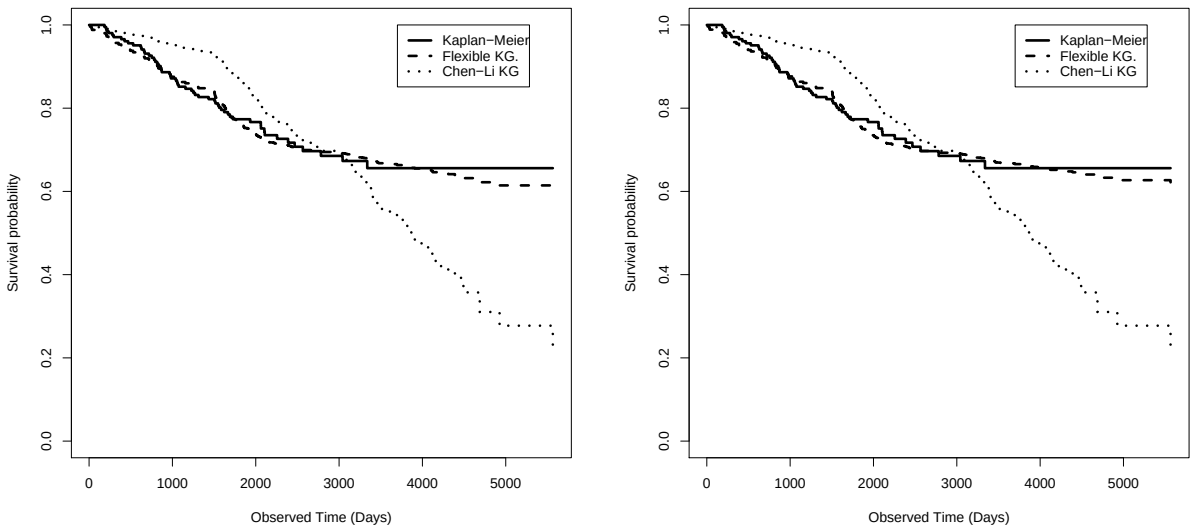

Figure 1: Extimated lifetime distribution of the malignant melanoma patients under the Plackett (left panel) and the BB10 (right panel) copulas.

that Z_i and δ_i are not independent. As a result, we assume that the sub-distribution of the time to death from malignant melanoma is related to the proportion of patients who died of malignant melanoma and the probability of patients who are alive at time t after the operation through the Plackett copula family given in (11). In this copula family we note that the parameter θ is equal to a constant odds ratio for the sub-distributions of the observed variables and is given by

$$\frac{\dfrac{P(Z \leq t, \delta = 1)}{P(Z > t, \delta = 1)}}{\dfrac{P(Z \leq t, \delta = 0)}{P(Z > t, \delta = 0)}} = \frac{P(Z \leq t, \delta = 1)P(Z > t, \delta = 0)}{P(Z > t, \delta = 1)P(Z \leq t, \delta = 0)} = \theta, \quad t > 0.$$

In this illustration, the estimated value of θ is $\widehat{\theta} = 12.62$. This means that the odds of the sub-distribution for an uncensored observation is about 13 times the odds of the sub-distribution for a censored observation. Furthermore we can show that the sub-distribution for an uncensored observation is larger than what we expect under independence. This indicates that the smaller observed survival times are mainly uncensored while the larger observed times are censored. Using this estimate (i.e. $\widehat{\theta} = 12.62$), we give in Figure 1 a graphical representation of the flexible Koziol-Green model compared to the Kaplan-Meier estimate as well as Chen and Lin version of the Koziol-Green model. In contrast to the Chen and Lin estimate, we can observe from the figure that the Kaplan-Meier and the flexible Koziol-Green estimates are very close to each other. This suggests

that the flexible Koziol-Green might be the most appropriate in this setting. Also we see that the flexible Koziol-Green estimate is at the right end higher than the Chen and Lin estimate. We expected this since $\widehat{\theta}$ is larger than 1. In addition, we redo the analysis with the BB10 ($\hat{\theta}_1 = 0.96, \hat{\theta}_2 = 0.41$) copula family, found in (12), and observe that the estimates for $\bar{F}$ based on the Plackett and BB10 copulas are very similar. Thus, this suggests that inference from the flexible Koziol-Green model is independent of the copula choice. Nevertheless, more future research is needed on this topic.

Appendix: Proofs

Proof of Theorem 1:

To prove the asymptotic normality of $\widehat{\theta}$, we verify the conditions of Theorem 2 of Chen, Linton an Van Keilegom (2003) (CLV hereafter). First we need to show that $\widehat{\theta} - \theta_0 = o_p(1)$. For this we verify the conditions of Theorem 1 of the same paper. Condition (1.1) holds by definition of $\widehat{\theta}$ while conditions (1.2) and (1.3) are satisfied by assumption $(A1)$ and $(A2)$. Finally, conditions (1.4) and (1.5) are weaker than the conditions (2.4) and (2.5) of Theorem 2 of CLV. Therefore, we will verify the conditions (2.4) and (2.5) below. Except for conditions (1.4) and (1.5), we have verified all the conditions for Theorem 1 and get that $\widehat{\theta} - \theta_0 = o_p(1)$.

Next, we verify the conditions of Theorem 2 to get the asymptotic normality. Condition (2.1) is as condition (1.1) satisfied by definition of $\widehat{\theta}$ while condition (2.2) follows from assumptions $(A1)$ and $(A2)$. For condition (2.3), we have that

$$\begin{aligned}
&\Gamma_2(\theta, p_1^*, H^*)\,[p_1 - p_1^*, H - H^*] \\
&= \lim_{\tau \to 0} \frac{1}{\tau} \left\{ G\left(\theta, p_1^* + \tau(p_1 - p_1^*), H^* + \tau(H - H^*)\right) - G(\theta, p_1^*, H^*) \right\} \\
&= E\left[\frac{\partial}{\partial u} g\left(\theta, u, H^*, Z, \delta\right)\Big|_{u=p_1^*} [p_1 - p_1^*] + \frac{\partial}{\partial v} g\left(\theta, p_1^*, v, Z, \delta\right)\Big|_{v=H^*} [H(Z) - H^*(Z)] \right]
\end{aligned}$$

Hence, we get that

$$\begin{aligned}
&G(\theta, p_1, H) - G(\theta, p_1^*, H^*) - \Gamma_2(\theta, p_1^*, H^*)\,[p_1 - p_1^*, H - H^*] \\
&= E\Bigg[g(\theta, p_1, H, Z, \delta) - g(\theta, p_1^*, H^*, Z, \delta) - \frac{\partial}{\partial u} g\left(\theta, u, H^*, Z, \delta\right)\Big|_{u=p_1^*} [p_1 - p_1^*] \\
&\quad - \frac{\partial}{\partial v} g\left(\theta, p_1^*, v, Z, \delta\right)\Big|_{v=H^*} [H(Z) - H^*(Z)] \Bigg]
\end{aligned}$$

Using a second order Taylor expansion, we find that

$$\begin{aligned}
&\|G(\theta,p_1,H)-G(\theta,p_1^*,H^*)-\Gamma_2(\theta,p_1^*,H^*)\,[p_1-p_1^*,H-H^*]\| \\
&= \left\| \frac{1}{2}\frac{\partial^2}{\partial u^2}g(\theta,u,H^{**},z,\delta)\Big|_{u=p_1^{**}}(p_1-p_1^*)^2+\frac{1}{2}\frac{\partial^2}{\partial v^2}g(\theta,p_1^{**},v,z,\delta)\Big|_{v=H^{**}}(H(z)-H^*(z))^2 \right. \\
&+ \left. \frac{\partial^2}{\partial u\partial v}g(\theta,u,v,z,\delta)\Big|_{u=p_1^{**},v=H^{**}}(p_1-p_1^*)(H(z)-H^*(z)) \right\| \\
&\leq c.\|(p_1-p_1^*,H-H^*)\|_{\mathcal{H}}^2
\end{aligned}$$

with p_1^{**} between p_1^* and p_1, H^{**} between H and H^*, and where $\|(p_1-p_1^*,H-H^*)\|_{\mathcal{H}}^2 := \max\left\{|p_1-p_1^*|^2,\|H-H^*\|_{L^2}^2\right\}$ with $\|H-H^*\|_{L^2}^2 := E|H(Z)-H^*(Z)|^2$. This shows the first part of condition (2.3). For the second part of this condition, it follows from the proof of Theorem 2 of CLV that it suffices to show that

$$\begin{aligned}
&\left\|\Gamma_2(\widehat{\theta},p_1^*,H^*)\,[p_1-p_1^*,H-H^*]-\Gamma_2(\theta_0,p_1^*,H^*)\,[p_1-p_1^*,H-H^*]\right\| \\
&= o_P(1)\left\|\widehat{\theta}-\theta_0\right\|
\end{aligned}$$

Using the mean value theorem and assumption $(A1)$, we get

$$\begin{aligned}
&\left\|\Gamma_2(\widehat{\theta},p_1^*,H^*)\,[p_1-p_1^*,H-H^*]-\Gamma_2(\theta_0,p_1^*,H^*)\,[p_1-p_1^*,H-H^*]\right\| \\
&= \left\| E\left[\frac{\partial^2}{\partial\theta\partial u}g(\theta,u,H^*,Z,\delta)\Big|_{\theta=\theta^{**}}(p_1-p_1^*)(\widehat{\theta}-\theta_0)\right.\right. \\
&\quad \left.\left. + \frac{\partial^2}{\partial\theta\partial v}g(\theta,p_1^*,v,Z,\delta)\Big|_{\theta=\theta^{**}}(H(Z)-H^*(Z))(\widehat{\theta}-\theta_0)\right]\right\| \\
&\leq o_P(1)\left\|\widehat{\theta}-\theta_0\right\|
\end{aligned}$$

where θ^{**} lies between $\widehat{\theta}$ and θ_0.

For condition (2.4), we define

$$\mathcal{H}_1 = \{x \to h_1(x), h_1(\cdot) \text{ is monotone and maps onto } [0,1]\}$$

and $\mathcal{H} = [0,1]\times\mathcal{H}_1$.

Then $P(\widetilde{p}_1\in[0,1])\to 1$ and $P\left(\widetilde{H}\in\mathcal{H}_1\right)\to 1$, since $\widetilde{H}$ is monotone. Moreover, we have that

$$P\left(\left|\sqrt{n}\,(\widetilde{p}_1-p_1^*)\right|>\varepsilon\right)\leq P\left(\sqrt{n}\,|\widetilde{p}_1-\widehat{p}_1|>\frac{\varepsilon}{2}\right)+P\left(\sqrt{n}\,|\widehat{p}_1-p_1^*|>\frac{\varepsilon}{2}\right).$$

For the second term, we get that $|\widehat{p}_1-p_1^*| = o_P\left(n^{-1/2}\right)$ (Serfling (1980)), while for the first term holds

$$P\left(\sqrt{n}\,|\widetilde{p}_1-\widehat{p}_1|>\frac{\varepsilon}{2}\right)=P\left(\frac{\sqrt{n}}{n+1}>\frac{\varepsilon}{2}\right)\to 0,\quad n\to+\infty.$$

Hence, $\widetilde{p}_1 - p_1^* = o_P\left(n^{-1/2}\right)$. Analogously, we can show that

$$\sup_{0\le z\le T}\left|\widetilde{H}(z) - H^*(z)\right| = o_P\left(n^{-1/2}\right)$$

such that $\left\|\left(\widetilde{p}_1 - p_1^*\right), \left(\widetilde{H} - H^*\right)\right\|_{\mathcal{H}} = o_P\left(n^{-1/2}\right)$.

To verify condition (2.5), we use Theorem 3 of CLV, and show that

$$\begin{aligned} &\left|g_j(\theta^1, p_1^1, H^1, z, \delta) - g_j(\theta^2, p_1^2, H^2, z, \delta)\right| \\ &\quad\le\ b_j(z,\delta)\left\{\|\theta^1 - \theta^2\| + \left\|(p_1^1 - p_1^2), (H^1 - H^2)\right\|_{\mathcal{H}}\right\} \end{aligned} \tag{13}$$

where $g_j(\theta, p_1, H, z, \delta)$, $(j = 1, 2, ..., k)$ is the j-th component of the function g, $E\left[b_j(Z,\delta)^2\right] \le +\infty$ and $\int_0^\infty \sqrt{\log N(\lambda, \mathcal{H}, \|\cdot\|_{\mathcal{H}})d\lambda} \le +\infty$ with $N(\lambda, \mathcal{H}, \|\cdot\|_{\mathcal{H}})$ the covering number needed to cover $\mathcal{H}$.

By assuming $(A1)$, we see that $E\left[b_j(Z,\delta)^2\right] \le +\infty$ and (13) holds. For the covering number, we know that $\forall\ \lambda > 0$,

$$N(\lambda, \mathcal{H}, \|\cdot\|_{\mathcal{H}}) \le N_{[\ ]}(2\lambda, \mathcal{H}, \|\cdot\|_{\mathcal{H}}) \le N_{[\ ]}(2\lambda, \mathcal{H}_2, \|\cdot\|_{L_2})$$

where the last inequality follows from $[0,1]$, a compact interval in $I\!R$. $N_{[\ ]}(\lambda, \mathcal{H}, \|\cdot\|_{\mathcal{H}})$ denotes the bracketing number to cover $\mathcal{H}$. By Theorem 2.7.5 of Van der Vaart and Wellner (1996), we have that

$$\log N_{[\ ]}(2\lambda, \mathcal{H}_2, \|\cdot\|_{L_2}) \le K.\frac{1}{2\lambda}$$

with K a constant such that

$$\int_0^\infty \sqrt{\log N(\lambda, \mathcal{H}, \|\cdot\|_{\mathcal{H}})} < +\infty.$$

It remains to verify condition (2.6) of CLV. Let us consider

$$\begin{aligned} &G_n(\theta_0, p_1^*, H^*) + \Gamma_2(\theta_0, p_1^*, H^*)\left[\widetilde{p}_1 - p_1^*, \widetilde{H} - H^*\right] \\ &= \frac{1}{n}\sum_{i=1}^n g(\theta_0, p_1^*, H^*, Z_i, \delta_i) + E\left[\frac{\partial}{\partial u} g(\theta_0, u, H^*, Z, \delta)\Big|_{u=p_1^*}(\widetilde{p}_1 - p_1^*)\right] \\ &\quad + E\left[\frac{\partial}{\partial v} g(\theta_0, p_1^*, v, Z, \delta)\Big|_{v=H^*}(\widetilde{H}(Z) - H^*(Z))\right] \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n} g(\theta_0, p_1^*, H^*, Z_i, \delta_i) + E\left[\frac{\partial}{\partial u} g(\theta_0, u, H^*, Z, \delta)\Big|_{u=p_1^*} (\widehat{p}_1 - p_1^* + \widetilde{p}_1 - \widehat{p}_1)\right] \\
&\quad + E\left[\frac{\partial}{\partial v} g(\theta_0, p_1^*, v, Z, \delta)\Big|_{v=H^*} \left(\widehat{H}(Z) - H^*(Z) + \widetilde{H}(Z) - \widehat{H}(Z)\right)\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} \Bigg\{ g(\theta_0, p_1^*, H^*, Z_i, \delta_i) + (I(\delta_i = 1) - p_1^*) E\left[\frac{\partial}{\partial u} g(\theta_0, u, H^*, Z, \delta)\Big|_{u=p_1^*}\right] \\
&\quad + E\left[\frac{\partial}{\partial v} g(\theta_0, p_1^*, v, Z, \delta)\Big|_{v=H^*} (I(Z_i \le Z) - H^*(Z))\right]\Bigg\} + o_P\left(n^{-1/2}\right)
\end{aligned}
$$

From the central limit theorem, we find that

$$
\sqrt{n} G_n(\theta_0, p_1^*, H^*) + \Gamma_2(\theta_0, p_1^*, H^*)\left[\widetilde{p}_1 - p_1^*, \widetilde{H} - H^*\right] \Rightarrow \mathrm{N}(0, V)
$$

with

$$
\begin{aligned}
V &= \mathrm{Var}\Bigg(g(\theta_0, p_1^*, H^*, Z, \delta) + (I(\delta = 1) - p_1^*) E\left[\frac{\partial}{\partial u} g(\theta_0, u, H^*, Z', \delta')\Big|_{u=p_1^*}\right] \\
&\qquad + E\left[\frac{\partial}{\partial v} g(\theta_0, p_1^*, v, Z', \delta')\Big|_{v=H^*} (I(Z \le Z') - H^*(Z'))\right]\Bigg).
\end{aligned}
$$

From Theorem 2 of CLV, we get that

$$
\sqrt{n}\left(\widehat{\theta} - \theta_0\right) \to \mathrm{N}(0, \Omega)
$$

where $\Omega = (\Gamma_1^{-1})^\tau V \Gamma_1^{-1}$.

Proof of Theorem 2:

We have for $t < T$, that

$$
\widehat{F}(t) - F(t) = \bar{F}(t) - \widehat{\bar{F}}(t) = k(\bar{H}^*(t), \theta_0, p_1^*) - k(\widehat{\bar{H}}(t), \widehat{\theta}_0, \widehat{p}_1)
$$

with

$$
k(z, \theta, p) = \exp\left(-\int_z^1 \frac{\mathcal{C}_\theta^2(p, 1-s)}{s} ds\right).
$$

Using a second order Taylor expansion, we get that

$$
\begin{aligned}
\widehat{F}(t) - F(t) &= -k_1(\bar{H}^*(t), \theta_0, p_1^*)(\widehat{\bar{H}}(t) - \bar{H}^*(t)) - k_2(\bar{H}^*(t), \theta_0, p_1^*)(\widehat{\theta} - \theta_0) \\
&\quad - k_3(\bar{H}^*(t), \theta_0, p_1^*)(\widehat{p}_1 - p_1^*) - \frac{1}{2} k_{11}(\bar{H}^{**}(t), \theta^{**}, p_1^{**})(\widehat{\bar{H}}(t) - \bar{H}^*(t))^2 \\
&\quad - \frac{1}{2} k_{33}(\bar{H}^{**}(t), \theta^{**}, p_1^{**})(\widehat{p}_1 - p_1^*)^2 - \frac{1}{2}(\widehat{\theta} - \theta_0)^\tau k_{22}(\bar{H}^{**}(t), \theta^{**}, p_1^{**})(\widehat{\theta} - \theta_0) \\
&\quad - (\widehat{\bar{H}}(t) - H^*(t)) k_{12}(\bar{H}^{**}(t), \theta^{**}, p_1^{**})(\widehat{\theta} - \theta_0) - (\widehat{p}_1 - p_1^*) k_{32}(\bar{H}^{**}(t), \theta^{**}, p_1^{**})(\widehat{\theta} - \theta_0)
\end{aligned}
$$

where $\bar{H}^{**}(t), \theta^{**}, p_1^{**}$ lies respectively between $\widehat{\bar{H}}(t)$ and $\bar{H}^*(t)$, $\widehat{\theta}$ and θ_0, $\widehat{p}_1$ and p_1^*.

Since $\sup_{t\in[0,T]} \left|\widehat{H}(t) - H^*(t)\right| = o_p\left(n^{-1/2}\right)$ (Serfling (1980)), we have that $\widehat{H}(T) < 1$ for a sufficiently large n. Combining this result with assumption $(A1)$, we get that the various partial derivatives are continuous functions on the compact space $[0,T] \times \Theta \times [0,1]$ and therefore are bounded. Noting also that $\widehat{\theta} - \theta_0 = o_p\left(n^{-1/2}\right)$ and $\widehat{p}_1 - p_1^* = o_p\left(n^{-1/2}\right)$, we get that

$$\begin{aligned}
\widehat{F}(t) - F(t) &= -k_1(\bar{H}^*(t), \theta_0, p_1^*)(\widehat{\bar{H}}(t) - \bar{H}^*(t)) - k_2(\bar{H}^*(t), \theta_0, p_1^*)(\widehat{\theta} - \theta_0) \\
&\quad -k_3(\bar{H}^*(t), \theta_0, p_1^*)(\widehat{p}_1 - p_1^*) + o_P\left(n^{-1}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} \left\{k_1(\bar{H}^*(t), \theta_0, p_1^*)(I(Z_i \le t) - H^*(t)) - k_3(\bar{H}^*(t), \theta_0, p_1^*)(I(\delta_i = 1) - p_1^*)\right\} \\
&\quad -k_2(\bar{H}^*(t), \theta_0, p_1^*)(\widehat{\theta} - \theta_0) + o_P\left(n^{-1}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} \widetilde{g}(t, Z_i, \delta_i) - k_2(\bar{H}^*(t), \theta_0, p_1^*)(\widehat{\theta} - \theta_0) + o_P\left(n^{-1}\right)
\end{aligned}$$

where $\widetilde{g}(t, Z, \delta) = k_1(\bar{H}^*(t), \theta_0, p_1^*)(I(Z \le t) - H^*(t)) - k_3(\bar{H}^*(t), \theta_0, p_1^*)(I(\delta = 1) - p_1^*)$.

Proof of Theorem 3:

To show the weak convergence of the process $n^{1/2}(\widehat{F}(\cdot) - F(\cdot))$, we show the weak convergence of the first term in the asymptotic representation in Theorem 2 and use Theorem 1 which gives the asymptotic normality of $\widehat{\theta}$. Let us denote $\widetilde{W}_n(t) := \frac{1}{n}\sum_{i=1}^{n} \widetilde{g}(t, Z_i, \delta_i)$. First we show that the marginals of $\widetilde{W}_n(t)$ converge weakly to a multivariate normal distribution. Hereto we take distinct time points $0 < t_1 < t_2 < ... < t_q = T$, $q > 0$. By the central limit theorem, we know that $\left(\widetilde{W}_n(t_1), \ldots, \widetilde{W}_n(t_q)\right)$ converges to an asymptotic normal distribution with mean zero and variance-covariance matrix equal to

$$\widetilde{\Sigma}_{t_j t_k} = \mathrm{Cov}(\widetilde{g}(t_j, Z, \delta), \widetilde{g}(t_k, Z, \delta)) = E\left[\widetilde{g}(t_j, Z, \delta), \widetilde{g}(t_k, Z, \delta)\right].$$

To prove the tightness for this process, we define

$$\begin{aligned}
\mathcal{F} &= \{\widetilde{g}(t, Z, \delta) | t \in [0,T]\} \\
&= \left\{k_1(\bar{H}^*(t), \theta_0, p_1^*)(I(Z \le t) - H^*(t)) - k_3(\bar{H}^*(t), \theta_0, p_1^*)(I(\delta = 1) - p_1^*) | t \in [0,T]\right\}
\end{aligned}$$

We note that the function $z \to k_1(\bar{H}^*(t), \theta_0, p_1^*)(I(z \le t) - H^*(t))$ is uniformly bounded over t and is a monotone function of z. Furthermore we note that the second function $k_3(\bar{H}^*(t), \theta_0, p_1^*)(I(\delta = 1) - p_1^*)$ is uniformly bounded. Thus, by Theorem 2.7.5 of van

der Vaart and Wellner (1996), we get that the bracketing number $N_{[\,]}(\varepsilon, \mathcal{F}, L_2(P)) = O\left(\exp\left(\frac{K}{\varepsilon}\right)\right)$ with K a constant. To prove the tightness of our process, we have to verify that the class $\mathcal{F}$ is Donsker. By the result 2.5.6 of Van der Vaart and Wellner, we need to show that

$$\int_0^{+\infty} \sqrt{\log N_{[\,]}(\varepsilon, \mathcal{F}, L_2(P))}d\varepsilon < +\infty$$

Since the functions $\widetilde{g}(t, Z, \delta)$ are uniformly bounded, we only have to consider the integration interval $[0, 2M]$ where $|\widetilde{g}(t, Z, \delta)| \leq M$, for all t, z, δ. For $\varepsilon > 2M$, we take $N_{[\,]}(\varepsilon, \mathcal{F}, L_2(P)) = 1$. Hence

$$\int_0^{+\infty} \sqrt{\log N_{[\,]}(\varepsilon, \mathcal{F}, L_2(P))}d\varepsilon = \int_0^{2M} \sqrt{\log N_{[\,]}(\varepsilon, \mathcal{F}, L_2(P))}d\varepsilon = \int_0^{2M} \sqrt{\frac{K}{\varepsilon}}d\varepsilon < +\infty.$$

In this way, we have shown that the process $\widetilde{W}_n(\cdot)$ converges weakly to a Gaussian process $\widetilde{W}(\cdot)$ with zero mean and variance covariance function

$$\widetilde{\Sigma}_{st} = \mathrm{Cov}(\widetilde{g}(s, Z, \delta), \widetilde{g}(t, Z, \delta)).$$

Since we already proved that the estimator $\widehat{\theta}$ for the parameter θ_0 is asymptotically normal, we find that the process $n^{1/2}(\widehat{F}(\cdot) - F(\cdot))$ weakly converges to a Gaussian process $W(\cdot)$. To find the mean and variance function of this process, we note by Theorem 2 that

$$\begin{aligned}
\widehat{F}(t) - F(t) &= \frac{1}{n}\sum_{i=1}^{n} \widetilde{g}(t, Z_i, \delta_i) - k_2(\bar{H}^*(t), \theta_0, p_1^*)(\widehat{\theta} - \theta_0) + o_P(n^{-1}) \\
&= \frac{1}{n}\sum_{i=1}^{n} \Bigg\{ \widetilde{g}(t, Z_i, \delta_i) + k_2(\bar{H}^*(t), \theta_0, p_1^*)(\Gamma_1^{-1})^\tau \Bigg[g(\theta_0, p_1^*, H^*, Z_i, \delta_i) \\
&\qquad + (I(\delta_i = 1) - p_1^*) E\left[\frac{\partial}{\partial u} g(\theta_0, u, H^*, Z', \delta') \Big|_{u=p_1^*} \right] \\
&\qquad + E\left[\frac{\partial}{\partial v} g(\theta_0, p_1^*, v, Z', \delta') \Big|_{v=H^*} (I(Z_i \leq Z') - H^*(Z')) \right] \Bigg] \Bigg\} \\
&+ o_P(n^{-1/2})
\end{aligned}$$

where the last equation follows from the proof of Theorem 2 in CLV. From the first term on the right side, we get that the mean and variance function of the Gaussian process

$W(\cdot)$ is given by $E\left[W(t)\right] = 0$ and

$$
\begin{aligned}
\Sigma_{st} &= \mathrm{Cov}(W(s), W(t)) \\
&= \mathrm{Cov}(\widetilde{g}(s, Z, \delta), \widetilde{g}(t, Z, \delta)) + k_2\left(\bar{H}^*(t), \theta_0, p_1^*\right)(\Gamma_1^{-1})^\tau \\
&\times \mathrm{Cov}\Bigg(\widetilde{g}(s, Z, \delta), g(\theta_0, p_1^*, H^*, Z, \delta) + (I(\delta = 1) - p_1^*)\, E\left[\frac{\partial}{\partial u} g(\theta_0, u, H^*, Z', \delta')\Big|_{u=p_1^*}\right] \\
&\qquad + E\left[\frac{\partial}{\partial v} g(\theta_0, p_1^*, v, Z', \delta')\Big|_{v=H^*}(I(Z \le Z') - H^*(Z'))\right]\Bigg) \\
&+ k_2\left(\bar{H}^*(s), \theta_0, p_1^*\right)(\Gamma_1^{-1})^\tau V \Gamma_1^{-1} k_2\left(\bar{H}^*(t), \theta_0, p_1^*\right)^\tau + k_2\left(\bar{H}^*(s), \theta_0, p_1^*\right)(\Gamma_1^{-1})^\tau \\
&\times \mathrm{Cov}\Bigg(\widetilde{g}(t, Z, \delta), g(\theta_0, p_1^*, H^*, Z, \delta) + (I(\delta = 1) - p_1^*)\, E\left[\frac{\partial}{\partial u} g(\theta_0, u, H^*, Z', \delta')\Big|_{u=p_1^*}\right] \\
&\qquad + E\left[\frac{\partial}{\partial v} g(\theta_0, p_1^*, v, Z', \delta')\Big|_{v=H^*}(I(Z \le Z') - H^*(Z'))\right]\Bigg).
\end{aligned}
$$

Acknowledgement

Both authors gratefully acknowledge financial support from IAP research network P6/03 of the Belgian Government (Belgian Science Policy).

References

Abdushukurov, A.A. (1987). Nonparametric estimation in the proportional hazards model of random censorship. *Akad. Nauk. Uz Tashkent. VINITI* NO. 3448-V (in Russian).

Andersen, P.K., Borgan, Ø., Gill, R.D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer, New York.

Beirlant, J., Carbonez, A., van der Meulen, E. (1992). Long run proportional hazards models of random censorship. *Journal of Statistical Planning and Inference*, **32**, 25-44.

Chen, X., Linton, O. and Van Keilegom, I. (2003). Estimation of semiparametric models when the criterion function is not smooth. *Econometrica*, **71**, 1591-1608.

Cheng, P.E., Lin, G.D. (1987). Maximum likelihood estimation of a survival function under the Koziol-Green proportional hazards model. *Statistics and Probability Letters*, **5**, 75-80.

Csörgő, S. (1988). Testing for the proportional hazards model of random censorship. *Proceedings of the 4th Prague Symposium on Asymptotic Statistics*, Prague.

Dikta, G. (1998) On semiparametric random censorship models. *Journal of Statistical Planning and Inference*, **66**, 253-279.

Genest,C., Nešlehová, J. (2007). A primer on copulas for count data. *Austin Bulletin*, **37**, 475-515.

Gupta, R.C., Gupta, P.L., Gupta, R.D. (1998) Modeling failure time data by Lehman alternatives. *Communications in Statistics - Theory and Methods*, **27**, 887-904.

Gupta, R.C., Gupta, R.D. (2007) Proportional reversed hazard rate model and its applications. *Journal of Statistical Planning and Inference*, **137**, 3525-3536.

Joe, H. (1997). *Multivariate Models and Dependence Concepts.* Chapman & Hall.

Kaplan, E.L., Meier, P. (1958). Non-parametric estimation from incomplete observations. *Journal of American Statistical Association*, **53**, 457-481.

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

Klement, P.E., Kolesárová, A., Mesiar, R., Sempi, C. (2007). Copulas constructed from horizontal sections. *Communications in Statistics - Theory and Methods*, **36**, 2901-2911.

Nelsen, R.B. (2006). *An Introduction to Copulas.* Springer, New York.

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

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