Universidade de Vigo

Generalized copula-graphic estimator

Jacobo de Uña Álvarez and Noël Veraverbeke

Report 12/02

Discussion Papers in Statistics and Operation Research

Departamento de Estatística e Investigación Operativa

Facultade de Ciencias Económicas e Empresariales
Lagoas-Marcosende, s/n · 36310 Vigo
Tfno.: +34 986 812440 - Fax: +34 986 812401
http://webs.uvigo.es/depc05/
E-mail: depc05@uvigo.es

Universidade de Vigo

Generalized copula-graphic estimator

Jacobo de Uña Álvarez and Noël Veraverbeke

Report 12/02

Discussion Papers in Statistics and Operation Research

Imprime: GAMESAL

Edita:

Universidade de Vigo
Facultade de CC. Económicas e Empresariales
Departamento de Estatística e Investigación Operativa
As Lagoas Marcosende, s/n 36310 Vigo
Tfno.: +34 986 812440

I.S.S.N: 1888-5756

Depósito Legal: VG 1402-2007

Generalized copula-graphic estimator

Jacobo de Uña-Álvarez (U. Vigo), Noël Veraverbeke (U. Hasselt)

February 2012

Abstract

In this paper a copula-graphic estimator is proposed for censored survival data. It is assumed that there is some dependent censoring acting on the variable of interest, which may come from an existing competing risk. Furthermore, the full process is independently censored by some administrative censoring time. The dependent censoring is modeled through an Archimedean copula function, which is supposed to be known. An asymptotic representation of the estimator as a sum of independent and identically distributed random variables is obtained and, consequently, a central limit theorem is established. We investigate the finite sample performance of the estimator through simulations. A real data illustration is included.

1 Introduction

Consider a bivariate competing risks model with variables Y and C representing the time up to event 1 and event 2 respectively on the same subject. In this situation, only one of the two events is observed, and the recorded event time is given by the minimum $Z = \min(Y, C)$. The event indicator $\delta = I(Y \leq C)$ which takes the value 1 when event 1 occurs ($\delta = 0$ otherwise) is also observed. It is known (Tsiatis, 1975) that the marginal distribution functions $F(t) = P(Y \leq t)$ and $G(t) = P(C \leq t)$ cannot be identified without knowledge of the dependence structure between Y and C. For example, the consistency of the Kaplan-Meier estimator of F (Kaplan and Meier (1958)) is not ensured since Y and C will be in general dependent. Due to this problem, attention in this setting has been mainly focused on the estimation of the so-called cause-specific hazard rate and subdistribution functions (Kalbfleisch and Prentice (1980)). However, estimation of F (and G) is possible when some information on the joint behaviour of Y and C is available.

Assume that there exists a known Archimedean copula $\mathcal{C}(u_1, u_2)$ which relates the joint survival function of (Y, C) to the marginal survival functions $\overline{F}(t) = 1 - F(t)$ and $\overline{G}(t) = 1 - G(t)$:

$$P(Y > t_1, C > t_2) = \phi^{-1}(\phi(\overline{F}(t_1)) + \phi(\overline{G}(t_2))).$$

The function $\phi :]0, 1] \to [0, \infty[$ is called the generator of the copula $\mathcal{C}$. It is a known continuous, convex, strictly decreasing function with $\phi(1) = 0$. The particular case $\phi(t) = -\ln t$ leads to the product copula $\mathcal{C}(u_1, u_2) = u_1 u_2$ and corresponds to independence between Y and C. A broad family of generators have been used to model dependent random variables. See Nelsen (2006). Under this assumption, Zheng and Klein (1995), see also Rivest and Wells (2001),

introduced a nonparametric estimator for $\overline{F}(t)$, termed copula-graphic estimator, generalizing the product-limit Kaplan-Meier estimator to the dependent scenario. Their estimator, however, requires the direct observation of the pair (Z, δ), which is not always possible. This may be due to limitations in the follow-up period for the subjects, losses unrelated to the competing risks of interest, and so on.

Therefore we introduce a potential censoring time D which is independent of (Z, δ). Rather than (Z, δ) we observe $(U, \rho, \rho\delta)$ where $U = \min(Z, D)$ and $\rho = I(Z \leq D)$; note that the value of δ (i.e. the event type) is observed only when Z is uncensored ($\rho = 1$). We put $\widetilde{G}$ for the distribution function of D. Denote $H(t) = P(Z \leq t)$, $\overline{H}(t) = 1 - H(t)$, and $H^1(t) = P(Z \leq t, \delta = 1)$. Then, if ϕ' exists and if H^1 is differentiable, we have from Tsiatis (1975)

$$\overline{F}(t) = \phi^{-1}\left(-\int_0^t \phi'(\overline{H}(s))dH^1(s)\right). \tag{1}$$

In Section 2, an estimator of $\overline{F}(t)$ will be obtained after plugging in proper estimators for H and H^1, based on the observed values $(U_i, \rho_i, \rho_i\delta_i)$, $i = 1, ..., n$, of $(U, \rho, \rho\delta)$.

The rest of the paper is organized as follows. In Section 2 we introduce the estimator and we establish an almost sure asymptotic representation. In Section 3 we investigate the finite-sample performance of the estimator through simulations. Section 4 gives an illustration of the method through the analysis of a real medical data set. Main conclusions are reported in Section 5. Some needed lemmas and their proofs are given in the Appendix.

2 The estimator. Main results

It becomes clear from equation (1) that, for the construction of an estimator of $\overline{F}(t)$, one needs suitable estimators of the distribution function of Z (H) and the subdistribution function of Z under restriction $\delta = 1$ (H^1). Since the censoring by D on Z is independent, one can estimate H by the Kaplan-Meier estimator H_n based on the (U_i, ρ_i), $i = 1, ..., n$. This estimator is defined through

$$1 - H_n(t) = \prod_{i=1}^{n}\left[\frac{n-i}{n-i+1}\right]^{\rho_{(i)}I(U_{(i)} \leq t)}$$

where $U_{(1)} \leq ... \leq U_{(n)}$ are the ordered U_i and $\rho_{(1)}, ..., \rho_{(n)}$ are the corresponding indicators. It can also be expressed as

$$H_n(t) = \sum_{i=1}^{n} W_{in} I(U_{(i)} \leq t)$$

where W_{in} is the Kaplan-Meier weight attached to $U_{(i)}$, which is given by

$$W_{in} = \frac{\rho_{(i)}}{n-i+1} \prod_{j=1}^{i-1} \left[\frac{n-j}{n-j+1} \right]^{\rho_{(j)}}.$$

To estimate H^1 we consider $\delta_{(i)}$ as a covariable for the possibly censored lifetime $U_{(i)}$; following Stute (1993), we have that

$$H_n^1(t) = \sum_{i=1}^{n} W_{in} I(U_{(i)} \leq t, \delta_{(i)} = 1)$$

is an estimator for $H^1(t) = P(Z \leq t, \delta = 1)$. Since $W_{in} = 0$ whenever $\rho_{(i)} = 0$, we may write

$$H_n^1(t) = \sum_{i=1}^{n} W_{in} I(U_{(i)} \leq t, \rho_{(i)}\delta_{(i)} = 1)$$

which demonstrates that $H_n^1(t)$ can be constructed from the observed values $(U_i, \rho_i, \rho_i\delta_i)$, $i = 1, ..., n$. We also mention that $H_n^1(t)$ is just the usual estimator for a cumulative incidence function in a censored competing risks model, cfr. Kalbfleisch and Prentice (1980), p. 169, eq. (7.10).

Strong consistency and asymptotic normality of $H_n^1(t)$ and of the corresponding Kaplan-Meier integrals $\sum_{i=1}^{n} W_{in}\varphi(U_{(i)}, \rho_{(i)})$ for some general real-valued function φ can be found in Stute (1993) and Stute (1996). The needed conditions on the underlying variables are: (i) Z and D are independent and H and $\widetilde{G}$ have no jumps in common; and (ii) $P(\rho = 1|Z, \delta) = P(\rho = 1|Z)$. Assuming continuity, both conditions (i) and (ii) hold if D and (Z, δ) are independent, which is true in particular when D is independent of (Y, C).

In view of the above we propose the following generalized copula-graphic estimator for $\overline{F}(t)$:

$$\overline{F}_n(t) = \phi^{-1}\left(-\int_0^t \phi'(\overline{H}_n(s)) dH_n^1(s) \right) \tag{2}$$

where $\overline{H}_n = 1 - H_n$. In the special case of no additional censoring ($D = \infty$) we have $U = Z$, $\rho = 1$, $W_{in} = 1/n$, and $\overline{F}_n(t)$ becomes the classical copula-graphic estimator. If moreover Y and C are independent ($\phi(t) = -\log t$), $\overline{F}_n(t)$ becomes the Kaplan-Meier estimator based on observations of (Z, δ). Equation (2) also leads to a standard Kaplan-Meier estimator in absence of dependent censoring ($Z = Y$, $\delta = 1$), based on observations of (U, ρ).

We prove an almost sure asymptotic representation for (2) with a uniform rate for the remainder on each compact interval $[0, T]$ with $T < \min(T_F, T_G, T_{\widetilde{G}})$.

Here we use the notation T_F for the right endpoint of the support of any distribution F. Put $F_n = 1 - \overline{F}_n$. We will refer to the following conditions.

(C1) F, G, and $\widetilde{G}$ are continuous
(C2) D is independent of (Y, C)
(C3) H and H^1 have continuous first and second derivatives in $[0, T]$
(C4) The copula generator ϕ has three continuous derivatives in $]0, 1]$ and $\phi'''(t) \leq 0$ for $t \in]0, 1]$

Theorem 1. Under (C1)-(C4) we have for $t \leq T$

$$F_n(t) - F(t) = -\frac{1}{\phi'(F(t))n} \left\{ \sum_{i=1}^{n} \int_0^t \phi''(\overline{H}(s))\psi_i(s) dH^1(s) + \sum_{i=1}^{n} \widetilde{\psi}_i(t) \right\} + R_n(t)$$

where the ψ_i and $\widetilde{\psi}_i$ $(i = 1, ..., n)$ are i.i.d zero mean variables and

$$\sup_{0 \leq t \leq T} |R_n(t)| = O(n^{-3/4}(\log n)^{3/4}) \quad \text{a.s. as } n \to \infty.$$

Remark. (a) The ψ_i are defined as

$$\begin{aligned} \psi_i(s) &= \overline{H}(s) \left\{ \int_0^s \frac{I(U_i \leq y) - \widetilde{H}(y)}{(1 - \widetilde{H}(y))^2} d\widetilde{H}^1(y) + \frac{I(U_i \leq s, \rho_i = 1) - \widetilde{H}^1(s)}{1 - \widetilde{H}(s)} \right. \\ &\quad \left. - \int_0^s \frac{I(U_i \leq y, \rho_i = 1) - \widetilde{H}^1(y)}{(1 - \widetilde{H}(y))^2} d\widetilde{H}(y) \right\} \end{aligned}$$

where $\widetilde{H}(t) = P(U \leq t)$ and $\widetilde{H}^1(t) = P(U \leq t, \rho = 1)$.

(b) The $\widetilde{\psi}_i$ are defined as

$$\begin{aligned} \widetilde{\psi}_i(t) &= \widetilde{\varphi}(U_i)\gamma_0(U_i)\rho_i - E\left\{\widetilde{\varphi}(U)\gamma_0(U)\rho\right\} \\ &\quad + \gamma_1(U_i)(1 - \rho_i) - \gamma_2(U_i) \end{aligned}$$

where $\widetilde{\varphi}(u) = I(u \leq t)\phi'(\overline{H}(u))$. The functions γ_0, γ_1, γ_2 are defined in Stute (1996). In our case they become:

$$\gamma_0(u) = \frac{1}{1 - \widetilde{G}(u)},$$

$$\begin{aligned} \gamma_1(u) &= \frac{1}{1 - \widetilde{H}(u)} \int I(u < w)\widetilde{\varphi}(w)\gamma_0(w) d\widetilde{H}^{11}(w) \\ &= \frac{1}{1 - \widetilde{H}(u)} \int_u^\infty \widetilde{\varphi}(w) \frac{1}{1 - \widetilde{G}(w)} d\widetilde{H}^{11}(w) \\ &= \frac{1}{1 - \widetilde{H}(u)} \int_u^\infty \widetilde{\varphi}(w) dH^1(w), \end{aligned}$$

$$\begin{aligned}\gamma_2(u) &= \int\int \frac{I(v<u, v<w)\widetilde{\varphi}(w)\gamma_0(w)}{(1-\widetilde{H}(w))^2} d\widetilde{H}^0(v) d\widetilde{H}^{11}(w) \\ &= \int \widetilde{C}(u\wedge w)\widetilde{\varphi}(w)\frac{1}{1-\widetilde{G}(w)} d\widetilde{H}^{11}(w) \\ &= \int \widetilde{C}(u\wedge w)\widetilde{\varphi}(w) dH^1(w),\end{aligned}$$

where $\widetilde{H}^0(t) = P(U \le t, \rho = 0)$, $\widetilde{H}^{11}(t) = P(U \le t, \rho = 1, \delta = 1)$, and

$$\widetilde{C}(t) = \int_0^t \frac{d\widetilde{G}(v)}{(1-\widetilde{H}(v))(1-\widetilde{G}(v))}.$$

Also note that we have used that $d\widetilde{H}^{11}(t) = (1-\widetilde{G}(t))dH^1(t)$.

(c) The asymptotic representation in Theorem 1 leads to the asymptotic normality result for the estimator:

$$\sqrt{n}\,(F_n(t) - F(t)) \to^d N(0, \sigma(t))$$

where

$$\sigma^2(t) = \frac{1}{\phi'(F(t))^2} Var\left(\int_0^t \phi''(\overline{H}(s))\psi_i(s)dH^1(s) + \widetilde{\psi}_i(t)\right).$$

In practice, the estimator's variance ($\approx \sigma^2(t)/n$) may be approximated by resampling methods.

Proof to Theorem 1. From (1) and (2) we have

$$\begin{aligned}F_n(t) - F(t) &= -\left\{\phi^{-1}\left(-\int_0^t \phi'(\overline{H}_n(s))dH_n^1(s)\right) - \phi^{-1}\left(-\int_0^t \phi'(\overline{H}(s))dH^1(s)\right)\right\} \\ &= \frac{1}{\phi'(F(t))}\left\{\int_0^t \phi'(\overline{H}_n(s))dH_n^1(s) - \int_0^t \phi'(\overline{H}(s))dH^1(s)\right\} + R_{n1}(t)\end{aligned}$$

where

$$R_{n1}(t) = \frac{1}{2}\frac{\phi''(\phi^{-1}(\varepsilon_1))}{\left[\phi'(\phi^{-1}(\varepsilon_1))\right]^3}\left\{\int_0^t \phi'(\overline{H}_n(s))dH_n^1(s) - \int_0^t \phi'(\overline{H}(s))dH^1(s)\right\}^2$$

with ε_1 between $-\int_0^t \phi'(\overline{H}_n(s))dH_n^1(s)$ and $-\int_0^t \phi'(\overline{H}(s))dH^1(s)$. Hence,

$$\begin{aligned}F_n(t) - F(t) &= \frac{1}{\phi'(F(t))}\left\{\int_0^t \left[\phi'(\overline{H}_n(s)) - \phi'(\overline{H}(s))\right] dH^1(s)\right. \\ &\quad + \int_0^t \phi'(\overline{H}(s))d\left[H_n^1(s) - H^1(s)\right] \\ &\quad \left. + \int_0^t \left[\phi'(\overline{H}_n(s)) - \phi'(\overline{H}(s))\right] d\left[H_n^1(s) - H^1(s)\right]\right\} \\ &\quad + R_{n1}(t)\end{aligned}$$

$$= \frac{1}{\phi'(F(t))}\left\{-\int_0^t \phi''(\overline{H}(s))\left[\overline{H}_n(s)-\overline{H}(s)\right]dH^1(s) \right. \quad (3)$$
$$\left. + \int_0^t \phi'(\overline{H}(s))d\left[H_n^1(s)-H^1(s)\right]\right\}$$

$$+R_{n1}(t)+R_{n2}(t)+R_{n3}(t)$$

where

$$R_{n2}(t) = -\frac{1}{2}\int_0^t \phi'''(\varepsilon_2)\left[\overline{H}_n(s)-\overline{H}(s)\right]^2 dH^1(s)$$

with ε_2 between $H_n(s)$ and $H(s)$, and where

$$R_{n3}(t) = \int_0^t \left[\phi'(\overline{H}_n(s))-\phi'(\overline{H}(s))\right] d\left[H_n^1(s)-H^1(s)\right].$$

Lemmas 1 to 3 in the Appendix guarantee that the remainders $R_{ni}(t)$ satisfy the uniform rate given for $R_n(t)$. Now, in the first term of (3) we plug in the asymptotic representation for the Kaplan-Meier estimator due to Lo and Singh (1986) and Major and Rejtő (1988). Note that, since H and $\widetilde{G}$ are continuous, we have for $t < T_{\widetilde{H}}$

$$H_n(t)-H(t) = \frac{1}{n}\sum_{i=1}^n \psi_i(t) + r_{n1}(t)$$

with $\sup_{0\le t\le T}|r_{n1}(t)| = O(n^{-1}\log n)$ a.s. For the second term in (3), we use the asymptotic representation for Kaplan-Meier integrals as in Stute (1996), but with an almost sure remainder term as in Sánchez-Sellero et al. (2005), to get

$$\int_0^t \phi'(\overline{H}(s))d\left[H_n^1(s)-H^1(s)\right] = \frac{1}{n}\sum_{i=1}^n \widetilde{\psi}_i(t) + r_{n2}(t)$$

with $\sup_{0\le t\le T}|r_{n2}(t)| = O(n^{-1}(\log n)^3)$ a.s. The integrability conditions in Stute (1996) and Sánchez-Sellero et al. (2005) are satisfied since $t \le T$, and the proof is complete.□

3 Simulation study

In this section we investigate the finite sample performance of the proposed estimator through simulations. We consider a situation with two dependent, exponential survival times $Y \sim Exp(\lambda_Y)$ and $C \sim Exp(\lambda_C)$, where $\lambda_Y = 1$ and $\lambda_C = 1$ or $\lambda_C = 3/4$. The variables Y and C follow a Clayton copula with generator $\varphi_\theta(t) = t^{-\theta} - 1$, $\theta > 0$, i.e. their joint survival function is given by

$$P(Y > x_1, C > x_2) = C(e^{-\lambda_1 x_1}, e^{-\lambda_2 x_2})$$

where

$$C(u_1, u_2) = \left[u_1^{-\theta} + u_2^{-\theta} - 1\right]^{-1/\theta}.$$

This copula implies a Kendall's Tau $\tau_\theta = \theta/(\theta+2)$. We consider the cases $\theta =$ 0.5, 2, 10, corresponding to association levels of 0.2, 0.5 and 0.83 respectively. Specifically, the simulation algorithm is as follows (cfr. Exercise 4.17 in Nelsen (2006)):

Step 1. Generate independent random variables $V_1, V_2 \sim Exp(1)$

Step 2. Independently generate $Z_0 \sim \Gamma(1/\theta, 1)$, and compute $U_i = (1 + V_i/Z_0)^{-1/\theta}$, $i = 1, 2$

Step 3. Finally, compute $Y = -\ln(U_1)/\lambda_Y$, $C = -\ln(U_2)/\lambda_C$

Then, we compute $Z = \min(Y, C)$ and $\delta = I(Y \leq C)$. The variable of interest is Y. The proportion of dependent censoring on Y when $\lambda_C = 3/4$ is smaller than when $\lambda_C = 1$. Besides, we introduce independent censoring through a potential censoring time $D \sim Exp(\lambda_D)$ independent of (Y, C), so the available information is $U = \min(Z, D)$, $\rho = I(Z \leq D)$, and $\rho\delta$. Cases $\lambda_D = 1$ and 2 are considered (the latter introducing a heavier censoring pattern). Sample sizes $n = 250$ and $n = 500$ are taken. In Table 1 we report the approximate proportion of independent and dependent censoring on Y for each combination of the parameters in the simulation.

$\theta =$	0.5	2	10
		$\lambda_D = 1$	
$\lambda_C = 1$	36.4 (50.0)	41.4 (50.0)	47.1 (50.0)
$\lambda_C = 3/4$	39.4 (41.7)	44.1 (38.1)	48.8 (24.1)
		$\lambda_D = 2$	
$\lambda_C = 1$	52.6 (50.0)	57.1 (50.0)	63.1 (50.0)
$\lambda_C = 3/4$	55.7 (42.1)	59.7 (39.6)	64.9 (28.7)

Table 1. Independent and dependent (in brackets) censoring rates (%) for the simulated models. Approximated values of $P(\rho = 0)$ and $P(\delta = 0|\rho = 1)$ (in brackets) are reported

In Tables 2-4 we report the bias, standard deviation, and mean squared error (MSE) of the proposed estimator in the case $\lambda_D = 1$, computed along 10,000 Monte Carlo trials, at the three quartiles t_i, $i = 1, 2, 3$, of the distribution of Y ($Exp(1)$). Together with the generalized copula-graphic estimator, we report the results corresponding to the naive Kaplan-Meier estimator of the survival function of Y which ignores the problem of dependent censoring. Results of the naive Kaplan-Meier estimator are expected to get better as the dependent censoring rate decreases (numbers in brackets in Table 1). It is also expected that the copula-graphic estimator will outperform the naive Kaplan-Meier more

clearly as the dependence degree grows (larger θ) and for larger quantiles. All these features are appreciated from these Tables 2-4. Besides, it is seen that the naive Kaplan-Meier has a bias which does not decrease when the sample size changes from $n = 250$ to $n = 500$. This systematic bias is a consequence of its misspecified product copula, being the main responsible for the large values of the MSE. The bias of the generalized copula-graphic estimator decreases for a larger n. Standard deviations of the proposed method are of the same order as for the Kaplan-Meier, although they are somehow smaller at the right tail.

	$\theta =$	0.5		2		10	
		NKM	GCG	NKM	GCG	NKM	GCG
$n = 250$							
	t_1	.0138	-.0008	.0406	-.0018	.0875	-.0013
$\lambda_C = 1$	t_2	.0465	-.0029	.1145	-.0024	.1829	-.0013
	t_3	.0823	-.0070	.1733	-.0029	.2336	.0008
	t_1	.0103	-.0009	.0310	-.0013	.0609	-.0019
$\lambda_C = 3/4$	t_2	.0357	-.0020	.0844	-.0023	.0919	-.0033
	t_3	.0613	-.0060	.1150	-.0036	.0616	-.0070
$n = 500$							
	t_1	.0137	-.0005	.0406	-.0010	.0878	-.0007
$\lambda_C = 1$	t_2	.0472	-.0010	.1144	-.0015	.1830	-.0008
	t_3	.0839	-.0027	.1734	-.0018	.2331	-.0011
	t_1	.0100	-.0008	.0310	-.0008	.0609	-.0011
$\lambda_C = 3/4$	t_2	.0352	-.0014	.0840	-.0017	.0919	-.0019
	t_3	.0613	-.0033	.1142	-.0022	.0616	-.0039

Table 2. Bias of the naive Kaplan-Meier estimator (NKM) and of the generalized copula-graphic estimator (GCG) along 10,000 Monte Carlo trials. Case $\lambda_D = 1$.

	$\theta =$	0.5		2		10	
		NKM	GCG	NKM	GCG	NKM	GCG
$n = 250$							
	t_1	.0314	.0333	.0294	.0350	.0260	.0352
$\lambda_2 = 1$	t_2	.0464	.0508	.0430	.0511	.0396	.0439
	t_3	.0695	.0722	.0642	.0618	.0625	.0509
	t_1	.0304	.0318	.0299	.0339	.0277	.0329
$\lambda_2 = 3/4$	t_2	.0447	.0477	.0418	.0464	.0411	.0404
	t_3	.0623	.0635	.0587	.0536	.0521	.0440
$n = 500$							
	t_1	.0222	.0236	.0213	.0252	.0187	.0250
$\lambda_2 = 1$	t_2	.0331	.0362	.0306	.0360	.0279	.0307
	t_3	.0487	.0501	.0448	.0426	.0435	.0343
	t_1	.0215	.0225	.0209	.0236	.0194	.0230
$\lambda_2 = 3/4$	t_2	.0310	.0329	.0301	.0331	.0292	.0284
	t_3	.0426	.0430	.0408	.0369	.0367	.0308

Table 3. Standard deviation of the naive Kaplan-Meier estimator (NKM) and of the generalized copula-graphic estimator (GCG) along 10,000 Monte Carlo trials. Case $\lambda_D = 1$.

	$\theta =$	0.5		2		10	
		NKM	GCG	NKM	GCG	NKM	GCG
$n = 250$							
	t_1	.0012	.0011	.0025	.0012	.0083	.0012
$\lambda_2 = 1$	t_2	.0043	.0026	.0150	.0026	.0350	.0019
	t_3	.0116	.0053	.0342	.0038	.0585	.0026
	t_1	.0010	.0010	.0019	.0011	.0045	.0011
$\lambda_2 = 3/4$	t_2	.0033	.0023	.0089	.0022	.0101	.0016
	t_3	.0076	.0041	.0167	.0029	.0065	.0020
$n = 500$							
	t_1	.0007	.0006	.0021	.0006	.0081	.0006
$\lambda_2 = 1$	t_2	.0033	.0013	.0140	.0013	.0343	.0009
	t_3	.0094	.0025	.0321	.0018	.0562	.0012
	t_1	.0006	.0005	.0014	.0006	.0041	.0005
$\lambda_2 = 3/4$	t_2	.0022	.0011	.0080	.0011	.0093	.0008
	t_3	.0056	.0019	.0147	.0014	.0051	.0010

Table 4. MSE of the naive Kaplan-Meier estimator (NKM) and of the generalized copula-graphic estimator (GCG) along 10,000 Monte Carlo trials. Case $\lambda_D = 1$.

In Tables 5-7 we report the results for the case $\lambda_D = 2$. In this case, the proportion of independent censoring grows, while the dependent censoring rate remains similar to the case $\lambda_D = 1$ (see Table 1). Main features are similar to those in Tables 2-4. As expected, the error of both estimators increases with respect to the former case, particularly at the right tail of Y; however, the efficiency of the generalized copula-graphic estimator relative to the Kaplan-Meier does not increase since the proportion of informative censoring with $\lambda_D = 2$ is roughly the same as for $\lambda_D = 1$.

	$\theta =$	0.5		2		10	
		NKM	GCG	NKM	GCG	NKM	GCG
$n = 250$							
	t_1	.0136	-.0008	.0406	-.0013	.0877	-.0008
$\lambda_C = 1$	t_2	.0472	-.0017	.1144	-.0017	.1826	-.0002
	t_3	.0869	.0012	.1741	.0076	.2346	.0181
	t_1	.0105	-.0005	.0312	-.0008	.0612	-.0009
$\lambda_C = 3/4$	t_2	.0359	-.0016	.0838	-.0025	.0921	-.0034
	t_3	.0636	-.0047	.1140	-.0050	.0622	-.0137
$n = 500$							
	t_1	.0139	-.0002	.0407	-.0008	.0876	-.0004
$\lambda_C = 1$	t_2	.0470	-.0009	.1145	-.0010	.1834	.0002
	t_3	.0841	-.0031	.1736	.0004	.2344	.0061
	t_1	.0107	.0001	.0308	-.0008	.0609	-.0007
$\lambda_C = 3/4$	t_2	.0355	-.0010	.0837	-.0017	.0921	-.0018
	t_3	.0629	-.0044	.1117	-.0068	.0620	-.0095

Table 5. Bias of the naive Kaplan-Meier estimator (NKM) and of the generalized copula-graphic estimator (GCG) along 10,000 Monte Carlo trials. Case $\lambda_D = 2$.

	$\theta =$	0.5		2		10	
		NKM	GCG	NKM	GCG	NKM	GCG
$n = 250$							
	t_1	.0341	.0363	.0319	.0381	.0284	.0383
$\lambda_C = 1$	t_2	.0582	.0642	.0532	.0647	.0493	.0560
	t_3	.1200	.1227	.1142	.1108	.1091	.0925
	t_1	.0334	.0349	.0321	.0364	.0299	.0354
$\lambda_C = 3/4$	t_2	.0556	.0510	.0522	.0586	.0506	.0494
	t_3	.1114	.1131	.1030	.0965	.0903	.0756
$n = 500$							
	t_1	.0238	.0253	.0226	.0269	.0110	.0267
$\lambda_C = 1$	t_2	.0413	.0455	.0380	.0457	.0345	.0392
	t_3	.0863	.0909	.0774	.0774	.0750	.0632
	t_1	.0235	.0246	.0229	.0260	.0210	.0250
$\lambda_C = 3/4$	t_2	.0392	.0420	.0375	.0422	.0359	.0349
	t_3	.0759	.0796	.0713	.0668	.0632	.0532

Table 6. Standard deviation of the naive Kaplan-Meier estimator (NKM) and of the generalized copula-graphic estimator (GCG) along 10,000 Monte Carlo trials. Case $\lambda_D = 2$.

	$\theta =$	0.5		2		10	
		NKM	GCG	NKM	GCG	NKM	GCG
$n = 250$							
	t_1	.0013	.0013	.0027	.0015	.0085	.0015
$\lambda_C = 1$	t_2	.0056	.0041	.0159	.0042	.0358	.0031
	t_3	.0220	.0151	.0433	.0123	.0670	.0089
	t_1	.0012	.0012	.0020	.0013	.0046	.0013
$\lambda_C = 3/4$	t_2	.0044	.0036	.0097	.0034	.0111	.0025
	t_3	.0164	.0128	.0236	.0093	.0120	.0059
$n = 500$							
	t_1	.0008	.0006	.0022	.0007	.0081	.0007
$\lambda_C = 1$	t_2	.0039	.0021	.0146	.0021	.0348	.0015
	t_3	.0145	.0083	.0361	.0060	.0605	.0040
	t_1	.0007	.0006	.0015	.0007	.0042	.0006
$\lambda_C = 3/4$	t_2	.0028	.0018	.0084	.0018	.0098	.0012
	t_3	.0097	.0064	.0176	.0045	.0078	.0029

Table 7. MSE of the naive Kaplan-Meier estimator (NKM) and of the generalized copula-graphic estimator (GCG) along 10,000 Monte Carlo trials. Case $\lambda_D = 2$.

In order to investigate the robustness of the proposed procedure, we repeat the simulation of 10,000 Monte Carlo trials of sample size $n = 250$ of the model

with $\lambda_Y = \lambda_C = 1$ and $\theta = 0.5$, but applying the GCG estimator with a wrong specification of the value of θ. Badly specified values of θ are 0.083, 0.125, 0.25, 0.75, 1, and 1.5. We also consider the naive Kaplan-Meier which can be seen as arising from a Clayton copula with $\theta = 0$. In Figure 1 we display the MSE's of the GCG estimator for the three quartiles of Y. These MSE's grow as the used value for θ departs from 0.5. Interestingly, we see that the Kaplan-Meier estimator provides the worst results for the third quartile. For the second quartile, a large overestimation of θ results in a MSE larger than that of Kaplan-Meier; this changes as the misspecification degree decreases. For the first quartile, the MSE remains roughly constant along θ; this is a consequence of the low proportion of dependent censoring at the left tail of the distribution.

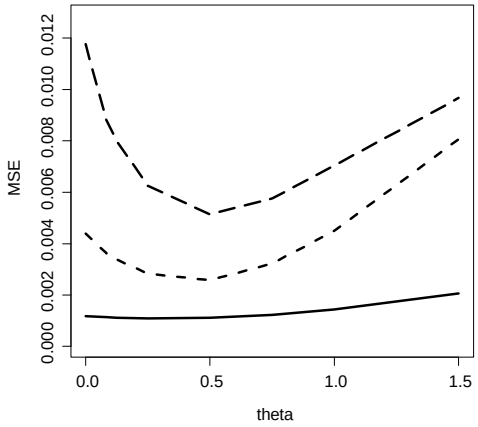

Figure 1. MSE of the GCG estimator with misspecified Clayton copula parameter (true theta is 0.5): first (solid), second (short dashed) and third (large dashed) quartiles.

In Figure 2 the bias and the standard deviation of the misspecified GCG estimator are reported. It is seen that the influence of the misspecification degree in the estimator's variance is small, while the influence on the bias is remarkable, specially for the second and the third quartiles. Indeed, this 'bias term' arising from the misspecification of the Clayton copula parameter is the main responsible for the variations in the MSE depicted in Figure 1. As expected, the absolute bias increases with $|\theta - 0.5|$, this being much more evident at the right tail of the distribution (second and third quartiles), according to the heavier effects of the dependent censoring.

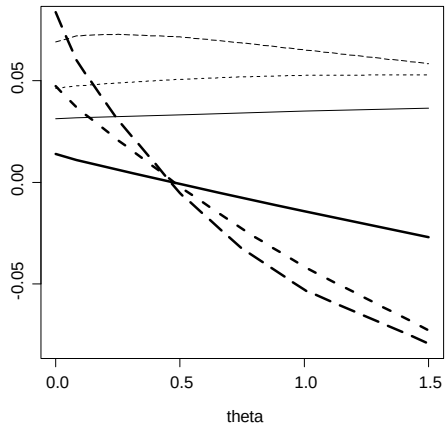

Figure 2. Bias (thick lines) and standard deviation (thin lines) of the GCG estimator with misspecified Clayton copula parameter (true theta is 0.5): first (solid), second (short dashed) and third (large dashed) quartiles.

4 Real data illustration

For illustration purposes, we consider the PBC data set reported and widely explained in Fleming and Harrington (1991), with $n = 312$ individuals. In this example, the Y variable denotes survival time (in years) for primary biliary cirrhosis (PBC) patients. Censoring from the right is provoked by the end of following-up or by liver transplantation (187 censored times, or about 60% of censoring). The number of transplants is 19, while 168 individuals were alive at the end of the study. In this setting, estimation of the survival function of Y ($\overline{F}(t)$) through the Kaplan-Meier estimator may be biased, due to the violation of the independent censoring assumption for the patients who receive a new liver (cfr. Fleming and Harrington, 1991, p. 103). Since the number of patients with a transplant is relatively small, this bias is likely to be small (same reference). However, it seems more realistic to estimate $\overline{F}(t)$ by incorporating a copula function which models the positive correlation between survival and transplantation times.

Therefore we consider our model in which C stands for time to transplant (the dependent censoring time) and D represents time to end of follow-up (independent censoring). As copula generator function we take the Clayton copula for which $\varphi_\theta(t) = t^{-\theta} - 1$, $\theta > 0$. As mentioned in Section 3, this copula implies a Kendall's Tau $\tau_\theta = \theta/(\theta+2)$. We consider the cases $\theta = 4$ and $\theta = 48$, leading to association levels of 0.67 and 0.96 respectively. In Figure 3 we display the naive Kaplan-Meier estimator (which assumes independence between Y and C) together with the proposed estimator for the two mentioned association degrees.

It is seen that the generalized copula-graphic estimator separates more from the Kaplan-Meier as the dependence grows. Indeed, the Kaplan-Meier curve overestimates the survival probability, because it ignores that transplant is associated with high risk of death. Crosses in Figure 3 indicate the transplantation times. As expected, the three estimators agree until the appearance of the first transplants, while they become more distinct with time, as more transplants occur. In our application, the difference between the Kaplan-Meier survival rate and that of the proposed estimator with $\theta = 48$ was above 4% from 2000 days on, reaching a maximum of 6.2% around time 3200.

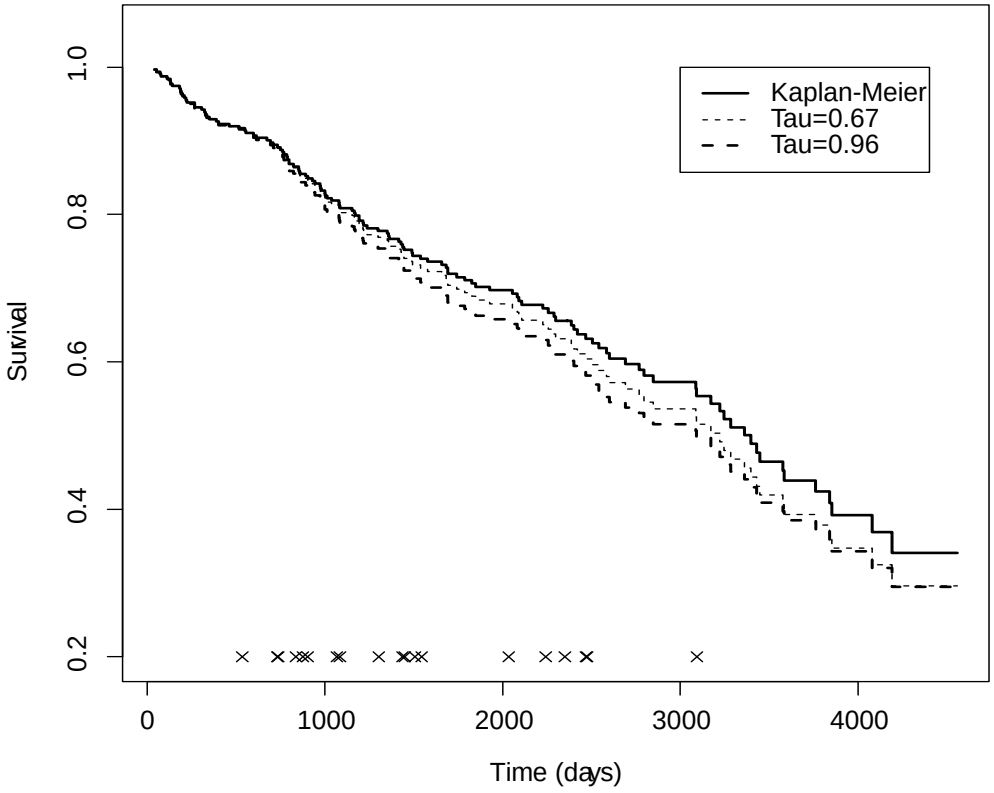

Figure 3. Kaplan-Meier estimator (solid line) and generalized copula-graphic estimators (dashed lines) for the PBC data: moderate association (thin dashed) and strong association (thick dashed line) between survival and transplant times (indicated with crosses).

5 Main conclusions

In this paper a generalized copula-graphic estimator for dependent censoring has been introduced. The estimator is suitable in situations in which, besides the dependent censoring, some independent mechanism censors the process of interest. This is the case of, e.g., the censored competing risks model. The proposed estimator reduces to the usual copula-graphic estimator when removing the independent censoring, while it equals the standard Kaplan-Meier estimator when the probability of having dependent censoring goes to zero. An asymptotic representation of the estimator as a sum of iid random variables has been obtained and, accordingly, a central limit theorem has been derived. As a by-product, an almost sure rate of convergence and the almost sure behavior of the modulus of an empirical cumulative incidence function have been established.

Simulations have shown that the proposed method performs well with finite sample sizes. As expected, the extent at which the new method may outperform the naive Kaplan-Meier (which ignores the dependent censoring) varies with the proportion of dependent censoring, as well as of the dependence degree. More benefits will be obtained from the new method under strongly dependent censoring and/or a large percentage of dependently censored data. It has been also shown that the relative efficiency of the generalized copula-graphic estimator increases at the right tail of the distribution, where the censoring effects concentrate. Robustness with respect to a misspecified copula has been explored; small departures from the true copula result in an estimator with mean squared error below that of the Kaplan-Meier. The practical performance of the estimator has been illustrated through real medical data analysis too.

A practical issue regarding the application of the estimator is that of choosing the copula function. The dependence sructure between the times to two competing risks which censor each other cannot be identified from the available data. The time-honoured Kaplan-Meier estimator is based on a specific copula (the product copula) which is not suitable for dependent competing risks. In practice, information on the association degree should be obtained from other sources such as e.g. expert knowledge or past studies. Human intervention provoking dependent or informative censoring (as in the PBC data illustration) is an example in which the expert may inform about reasonable candidates for the copula.

Acknowledgements. Work supported by the Grant MTM2008-03129 of the Spanish Ministry of Science and Innovation. The first author acknowledges support from the projects MTM2011-23204 of the Spanish Ministry of Science and Innovation (FEDER support included) and 10PXIB300068PR of the Xunta de Galicia too. The second author also acknowledges the IAP Research Network P6/03 of the Belgian State (Belgian Science Policy).

6 References

Földes A, Rejtő L (1981). A LIL type result for the product limit estimator. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 56, 75-86.

Kalbfleisch JD, Prentice RL (1980). *The Statistical Analysis of Failure Time Data*. Wiley, New York.

Kaplan EL, Meier P (1958). Nonparametric estimation from incomplete observations. Journal of the American Statistical Association., 53, 457-481.

Lo SH, Singh K (1986). The product-limit estimator and the bootstrap: some asymptotic representations. Probability Theory and Related Fields, 71, 455-456.

Major P, Rejtő L (1988). Strong embedding of the estimator of the distribution function under random censorship. Annals of Statistics, 16, 1113-1132.

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

Rivest LP, Wells MT (2001). A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. Journal of Multivariate Analysis, 79, 138-155.

Sánchez-Sellero C, González-Manteiga W, Van Keilegom I (2005). Uniform representation of product-limit integrals with applications. Scandinavian Journal of Statistics, 32, 563-581.

Schäfer H (1986). Local convergence of empirical measures in the random censorship situation with application to density and rate estimators. Annals of Statistics, 14, 1240-1245.

Stute W (1993). Consistent estimation under random censorship when covariables are present. Journal of Multivariate Analysis, 45, 89-103.

Stute W (1995). The central limit theorem under random censorship. Annals of Statistics, 23, 422-439.

Stute W (1996). Distributional convergence under random censorship when covariables are present. Scandinavian Journal of Statistics, 23, 461-471.

Tsiatis A (1975). A nonidentifiability aspect of the problem of competing risks. Proceedings of the National Academy of Sciences, 72, 20-22.

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

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

7 Appendix: technical lemmas

In this Section we give the technical lemmas needed in the proof to Theorem 1.

Lemma 1. Under the conditions in Theorem 1 we have

$$\sup_{0\leq t\leq T} |R_{n1}(t)| = O(n^{-1}\log\log n) \qquad \text{a.s.}$$

Proof. It is easy to see that, with probability 1,

$$\sup_{0\le t\le T}\left|\int_0^t \phi'(\overline{H}_n(s))dH_n^1(s)-\int_0^t \phi'(\overline{H}(s))dH^1(s)\right| = O(\sup_{0\le t\le T}|H_n(t)-H(t)|) + O(\sup_{0\le t\le T}\left|H_n^1(t)-H^1(t)\right|).$$

The first term is $O(n^{-1/2}(\log\log n)^{1/2})$ a.s., cfr. Földes and Rejtő (1981). The same order bound is proved to hold for the second term in Lemma 4 below, and the proof is complete.□

Lemma 2. Under the conditions in Theorem 1 we have

$$\sup_{0\le t\le T}|R_{n2}(t)| = O(n^{-1}\log\log n) \qquad \text{a.s.}$$

Proof. With probability 1 we have

$$\sup_{0\le t\le T}|R_{n2}(t)| \le \frac{1}{2}\sup_{0\le t\le T}|H_n(t)-H(t)|^2 \max(\sup_{0\le t\le T}\left|\phi'''(\overline{H}_n(t))\right|, \sup_{0\le t\le T}\left|\phi'''(\overline{H}(t))\right|).$$

Then, the assertion of the Lemma follows from a result of Földes and Rejtő (1981).□

Lemma 2. Under the conditions in Theorem 1 we have

$$\sup_{0\le t\le T}|R_{n3}(t)| = O(n^{-3/4}(\log n)^{3/4}) \qquad \text{a.s.}$$

Proof. Divide $[0,T]$ into $k_n = O(n^{1/2}(\log n)^{1/2})$ subintervals $[t_i, t_{i+1}]$ of length $O(n^{-1/2}(\log n)^{1/2})$. Then, as in the proof of Lo and Singh (1986) we have

$$\sup_{0\le t\le T}|R_{n3}(t)| \le 2\max_{1\le i\le k_n}\sup_{t_i\le y\le t_{i+1}}\left|\phi'(\overline{H}_n(y))-\phi'(\overline{H}_n(t_i))-\phi'(\overline{H}(y))+\phi'(\overline{H}(t_i))\right|$$

$$+k_n\sup_{0\le t\le T}\left|\phi'(\overline{H}_n(t))-\phi'(\overline{H}(t))\right|\max\left|H_n^1(t_{i+1})-H_n^1(t_i)-H^1(t_{i+1})+H^1(t_i)\right|$$

$$= I + II.$$

For I we have by Taylor expansion and the fact that $\sup_{0\le t\le T}\left|\overline{H}_n(t)-\overline{H}(t)\right| = O(n^{-1/2}(\log\log n)^{1/2})$ a.s. (Földes and Rejtő, 1981):

$$I \le 2\max_{1\le i\le k_n}\sup_{t_i\le y\le t_{i+1}}\left|\phi''(\overline{H}(t_{i+1}))\right||H_n(y)-H(y)-H_n(t_i)+H(t_i)| + O(n^{-1}\log\log n).$$

Now further subdivide each interval $[t_i, t_{i+1}]$ into $a_n = O(n^{1/4}(\log n)^{-1/4})$ subintervals of length $O(n^{-3/4}(\log n)^{3/4})$. By using Bernstein's inequality we can show that this term is bounded a.s. by

$$c \max_{1\le i\le k_n} \max_{0\le j\le a_n-1} |H_n(t_{i,j+1}) - H(t_{i,j+1}) - H_n(t_i) + H(t_i)| + O(n^{-3/4}(\log n)^{3/4})$$

for some constant $c > 0$. By the modulus of continuity result for the Kaplan-Meier estimator (see Schäfer, 1986) we obtain that $I = O(n^{-3/4}(\log n)^{3/4})$ a.s. The II term is treated similarly and leads to the same order. It requires the almost sure behaviour of the modulus of continuity of the H_n^1 estimator and this follows from Lemma 5 below. In that Lemma take $a_n = n^{-1/2}(\log n)^{1/2}$.□

Lemma 4 and Lemma 5 below are needed for the proofs of Lemma 1 and Lemma 3 respectively. They have some independent interest, since they provide the almost sure rate of convergence and the almost sure behavior of the modulus of continuity for the estimator of the cumulative indicence function of Z subject to $\delta = 1$ (H_n^1).

Lemma 4. For $T < \min(T_F, T_G, T_{\widetilde{G}})$ we have

$$\sup_{0\le t\le T} \left|H_n^1(t) - H^1(t)\right| = O(n^{-1/2}(\log\log n)^{1/2}) \qquad \text{a.s.}$$

Proof. Define the following empirical estimators for the distribution function $\widetilde{H}(t) = P(U \le t)$ and for the subdistribution functions $\widetilde{H}^0(t) = P(U \le t, \rho = 0)$ and $\widetilde{H}^{11}(t) = P(U \le t, \rho = 1, \delta = 1)$:

$$\widetilde{H}_n(t) = \frac{1}{n}\sum_{i=1}^{n} I(U_i \le t), \qquad \widetilde{H}_n^0(t) = \frac{1}{n}\sum_{i=1}^{n} I(U_i \le t, \rho_i = 0),$$

$$\widetilde{H}_n^{11}(t) = \frac{1}{n}\sum_{i=1}^{n} I(U_i \le t, \rho_i = 1, \delta_i = 1).$$

Then, $H^1(t)$ can be expressed in terms of $\widetilde{H}$, $\widetilde{H}^0$ and $\widetilde{H}^{11}$ and $H_n^1(t)$ can be expressed in terms of the corresponding empiricals. Similar as in Stute (1995) we obtain

$$H_n^1(t) = \int_0^t \exp\left\{ n \int_0^u \log(1 + \frac{1}{n(1-\widetilde{H}_n(z))}) d\widetilde{H}_n^0(z) \right\} d\widetilde{H}_n^{11}(u)$$

and

$$H^1(t) = \int_0^t \exp\left\{ \int_0^u \frac{d\widetilde{H}^0(z)}{1-\widetilde{H}(z)} \right\} d\widetilde{H}^{11}(u).$$

It follows that $\sup_{0\le t\le T}\left|H_n^1(t)-H^1(t)\right|$ is smaller than

$$\sup_{0\le u\le T}\left|\exp\left\{n\int_0^u \log(1+\frac{1}{n(1-\widetilde{H}_n(z))})d\widetilde{H}_n^0(z)\right\}-\exp\left\{\int_0^u \frac{d\widetilde{H}^0(z)}{1-\widetilde{H}(z)}\right\}\right|$$

$$+2\frac{1}{1-\widetilde{G}(T)}\sup_{0\le t\le T}\left|\widetilde{H}_n^{11}(t)-\widetilde{H}^{11}(t)\right| \tag{4}$$

The second term in (4) is $O(n^{-1/2}(\log\log n)^{1/2})$ a.s. For the first term in (4) we use (with obvious abreviations) that

$$\begin{aligned}\exp(a)-\exp(b) &= \exp(b)\left\{\exp(a-b)-1\right\}\\ &= \exp(b)\left\{(a-b)+\frac{1}{2}e^{\theta}(a-b)^2\right\}\end{aligned}$$

with θ between 0 and $a-b$. Note that $\exp(b)$ is uniformly bounded in $[0,T]$. Looking at $(a-b)$ we have

$$\sup_{0\le u\le T}\left|n\int_0^u \log(1+\frac{1}{n(1-\widetilde{H}_n(z))})d\widetilde{H}_n^0(z)-\int_0^u \frac{d\widetilde{H}^0(z)}{1-\widetilde{H}(z)}\right|$$

$$\le \sup_{0\le z\le T}\left|n\log(1+\frac{1}{n(1-\widetilde{H}_n(z))})-\frac{1}{1-\widetilde{H}(z)}\right|+\frac{2}{1-\widetilde{H}(T)}\sup_{0\le z\le T}\left|\widetilde{H}_n^0(z)-\widetilde{H}^0(z)\right|. \tag{5}$$

The second term in (5) is $O(n^{-1/2}(\log\log n)^{1/2})$ a.s. For the first term in (5) we use that for $x\ge 0$

$$x-\frac{1}{2}x^2\le\log(1+x)\le x.$$

It follows that the first term in (5) is bounded above by

$$\sup_{0\le z\le T}\left|\frac{1}{1-\widetilde{H}_n(z)}-\frac{1}{1-\widetilde{H}(z)}\right|+\frac{1}{2n}\sup_{0\le z\le T}\frac{1}{(1-\widetilde{H}_n(z))^2}.$$

This is $O(n^{-1/2}(\log\log n)^{1/2})$ a.s. since $\sup_{0\le z\le T}\left|\widetilde{H}_n(z)-\widetilde{H}(z)\right|$ has the same order and since $\widetilde{H}(T)<1$.□

Lemma 5. Suppose $T<\min(T_F,T_G,T_{\widetilde{G}})$. Suppose $H(t)=P(Z\le t)$ and $H^1(t)=P(Z\le t,\delta=1)$ have bounded first derivative in $[0,T]$. Let $\{a_n\}$ be a

sequence of positive constants tending to zero with $a_n n(\log n)^{-5} > \Delta > 0$ for all n sufficiently large. Then

$$\sup_{0\le t,s\le T,|t-s|\le a_n} \left|H_n^1(t) - H_n^1(s) - H^1(t) + H^1(s)\right| = O(a_n^{1/2} n^{-1/2} (\log n)^{1/2}) \qquad \text{a.s.}$$

Proof. We make the same partition of the interval $[0,T]$ as in Lemma A.5 of Van Keilegom and Veraverbeke (1997). Exploiting the monotonicity of $H^1(t)$ and $H_n^1(t)$ and also the Lipschitz continuity of $H^1(t)$, we obtain that it suffices to prove that

$$\begin{aligned} & \max_{1\le i\le m-1} \max_{-b_n<j,k<b_n} \left|H_n^1(t_{ik}) - H_n^1(t_{ij}) - H^1(t_{ik}) + H^1(t_{ij})\right| \\ = \;& O(a_n^{1/2} n^{-1/2} (\log n)^{1/2}) \qquad \text{a.s.,} \end{aligned}$$

where $\{t_{ij}\}$, $i = 1, ..., m$, $j = -b_n, ..., b_n$ is a grid of points with $m = \left[\frac{T}{a_n}\right]$ ([.] denoting the integer part) and $b_n \sim a_n^{1/2} n^{1/2} (\log n)^{-1/2}$. At this point we use the almost sure asymptotic representation for $H_n^1(t)$ as it can be derived as a special case of the more general result of Sánchez-Sellero et al. (2005):

$$H_n^1(t) - H^1(t) = \frac{1}{n} \sum_{i=1}^{n} \widetilde{\widetilde{\psi}}_i(t) + R_n(t)$$

where

$$\begin{aligned} \widetilde{\widetilde{\psi}}_i(t) \;=\; & I(U_i \le t)\gamma_0(U_i)\rho_i - E\left\{I(U \le t)\gamma_0(U)\rho\right\} \\ & + \gamma_{1t}(U_i)(1-\rho_i) - \gamma_{2t}(U_i) \end{aligned}$$

with

$$\gamma_0(u) = \frac{1}{1-\widetilde{G}(u)}, \qquad \gamma_{1t}(u) = \frac{H^1(t) - H^1(u)}{1-\widetilde{H}(u)}, \qquad \gamma_{2t}(u) = \int_0^t \widetilde{C}(u \wedge w) dH^1(w),$$

the function $\widetilde{C}(t)$ being that in the Remark of Section 2; and $\sup_{0\le t\le T} |R_n(t)| = O(n^{-1}(\log n)^3)$ a.s. It follows that it suffices to show that

$$\max_{1\le i\le m-1} \max_{-b_n<j,k<b_n} \left| \frac{1}{n} \sum_{r=1}^{n} (\widetilde{\widetilde{\psi}}_r(t_{ik}) - \widetilde{\widetilde{\psi}}_r(t_{ij})) \right| = O(a_n^{1/2} n^{-1/2} (\log n)^{1/2}).$$

To achieve this we use Bernstein's inequality as in Van Keilegom and Veraverbeke (1997). The random variables $\widetilde{\widetilde{\psi}}_r(t_{ik}) - \widetilde{\widetilde{\psi}}_r(t_{ij})$ are bounded and $Var(\widetilde{\widetilde{\psi}}_r(t_{ik}) - \widetilde{\widetilde{\psi}}_r(t_{ij}))$ is bounded by a constant times a_n. The latter fact is shown by checking six appropriate groups of terms in

$$Var(\widetilde{\widetilde{\psi}}_i(t) - \widetilde{\widetilde{\psi}}_i(s)) = Var(I(U_i \le t)\gamma_0(U_i)\rho_i + \gamma_{1t}(U_i)(1-\rho_i) - \gamma_{2t}(U_i))$$

$$+Var(I(U_i \leq s)\gamma_0(U_i)\rho_i + \gamma_{1s}(U_i)(1-\rho_i) - \gamma_{2s}(U_i))$$

$$-2Cov(\tilde{\tilde{\psi}}_i(t), \tilde{\tilde{\psi}}_i(s)).$$

For example, by direct calculation,

$$Var(I(U_i \leq t)\gamma_0(U_i)\rho_i) + Var(I(U_i \leq s)\gamma_0(U_i)\rho_i)$$

$$-2Cov(I(U_i \leq t)\gamma_0(U_i)\rho_i, I(U_i \leq s)\gamma_0(U_i)\rho_i) =$$

$$\int_{t\wedge s}^{t\vee s} \frac{dH(u)}{1-\widetilde{G}(u)} - (H(t)-H(s))^2 \leq c\,|t-s|$$

for some constant $c > 0$ by the Lipschitz continuity of H. The other groups of terms are treated similarly.□