



HAL
open science

Estimating the survival functions in a censored semi-competing risks model

Stéphane Laurent

► **To cite this version:**

Stéphane Laurent. Estimating the survival functions in a censored semi-competing risks model. Sankhya A, 2013, pp.1-22. 10.1007/s13171-013-0023-2 . hal-00759809

HAL Id: hal-00759809

<https://hal.science/hal-00759809>

Submitted on 2 Dec 2012

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

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

Estimating the survival functions in a censored semi-competing risks model

Stéphane LAURENT

December 2, 2012

Abstract Rivest & Wells (2001) proposed estimators of the marginal survival functions in a right-censored model that assumes an Archimedean copula between the survival time and the censoring time. We study the extension of these estimators to the context of right-censored semi-competing risks data with an independent second level censoring time. We intensively use martingale techniques to derive their large sample properties under mild assumptions on the true distribution of the data. As compared to the simpler context of right-censored data, a primary difference is the need to enlarge the filtrations with respect to which we use the Doob-Meyer decompositions of counting processes.

Mathematics Subject Classification (2000) 62N01 · 62N02 · 60G44

Keywords Archimedean copula · Dependent censoring · Immersion of filtration · Kaplan-Meier estimator · Martingale · Semi-competing risks

1 Introduction

Rivest & Wells (2001) addressed the problem of the estimation of the survival functions of two failure times T and U when T is subject to right-censoring via U and when assuming a known Archimedean copula for the joint distribution of T and U . It is easy to extend the derivation of their estimators to the context of *right-censored semi-competing risks data*. The failure times T and U are termed as semi-competing risks when the available data are some repeated observations of $\min(T, U), U$. Right-censored semi-competing risks data is then the situation when the semi-competing risks T and U are subject to right-censoring via another failure time V and the data are some repeated observations of $(\min(T, U, V), \min(U, V), \mathbb{1}_{\{T \leq U, T \leq V\}}, \mathbb{1}_{\{U \leq V\}})$. For instance, T, U and V could respectively represent the time to first relapse of a disease, the time to death, and the time to lost to follow-up of an individual under medical treatment.

When assuming a one-dimensional parametric family of Archimedean copulas for the law of (T, U) and an independent right-censoring, methods to estimate the copula parameter

and the marginal survival function of T have been proposed by Fine & al (2001), Jiang & al (2005), Lakhali & al (2008). These methods consist in plugging-in an estimate of the copula parameter in an estimate of the marginal function of T depending on the copula parameter, and Lakhali & al (2008) used the above mentioned extension of the Rivest & Wells estimator.

We will study the large sample properties of these Rivest & Wells estimators extended to the context of independent right-censored semi-competing risks. Mathematically, this could appear to be only a slight generalization of the case treated by Rivest & Wells (2001), since we will partly follow their proofs, which intensively use the martingale techniques for the statistical analysis of counting processes. However, a substantial difference arises with the martingale approach as compared to the case treated by Rivest & Wells: we will need to enlarge the usual filtrations for the Doob-Meyer decompositions of the encountered counting processes (namely, we will use the *immersion* of a filtration in a bigger one). Moreover, the proofs given by Rivest & Wells (2001) often refer to other proofs in the book of Fleming & Harrington (1991), whereas the one we provide is self-contained, and a point in the proof of their theorem 2 is not clear: it seems to justify that a process is Gaussian by the fact that it is the sum of two Gaussian processes. We will use a bivariate martingale central limit theorem to give a correct proof of our extension of this theorem to the context of independent right-censored semi-competing risks. In addition, we also discuss about the joint asymptotic behaviour of the Rivest & Wells estimator of T with the Kaplan-Meier estimator of U .

To be more precise, we now specify the statistical model under study and we introduce our main notations. The observations are the realizations $\{x_i, y_i, \delta_{Ti}, \delta_{Ui}\}_{i=1}^n$ of n independent replicates of a random four-tuple $(X, Y, \Delta_T, \Delta_U)$ defined as follows. Consider three positive random variables T, U, V , which we respectively interpret as, for some individual, the time to some nonterminal event of interest (e.g. time until the first relapse of a certain disease), the censoring time of type 1 (e.g. the time until death), and the censoring time of type 2 (e.g. the time until lost to follow-up). Then the observable random variables are $X = \min(T, U, V)$, $Y = \min(U, V)$, and the censoring indicators

$$\begin{cases} \Delta_T = \mathbb{1}_{\{T \leq U, T \leq V\}} & (T \text{ is not censored}) \\ \Delta_U = \mathbb{1}_{\{U \leq V\}} & (\text{type 1 censoring occurs before type 2 censoring}). \end{cases}$$

We will always assume that V is independent of (T, U) and $0 < \mathbb{P}(T \leq U) < 1$. One goal of this paper is to provide consistent (in a sense to be precised later) estimators of the survival functions of T and U when assuming that the dependence between T and U is given by a survival Archimedean copula \mathcal{C}_ϕ with a strict generator denoted by ϕ . That means that $\phi: [0, 1] \rightarrow [0, \infty]$ is a continuous strictly decreasing convex function satisfying $\phi(0) = \infty$ and $\phi(1) = 0$, and the joint survival function of T and U is given by

$$\mathbb{P}(T > x, U > y) = \mathcal{C}_\phi(S(x), F(y))$$

where S and F respectively are the marginal survival functions of T and U , and the copula $\mathcal{C}_\phi: [0, 1]^2 \rightarrow [0, 1]$ is given by

$$\mathcal{C}_\phi(u, v) = \phi^{-1}(\phi(u) + \phi(v)).$$

The survival function of the independent censoring V is denoted by G . We will focus on the case when S and F are continuous, but no assumption is made about G . Additional assumptions on the generator ϕ will be given later; these assumptions will always be fulfilled for the Clayton and Frank families of Archimedean copulas. We will define and study estimators \hat{S}_ϕ and \hat{F}_ϕ of S and F that coincide with the Rivest & Wells estimators in the particular case when there is no type 2 censoring, which corresponds to the case when $G \equiv 1$

in our setting, i.e. when $V = \infty$ almost surely. Similarly to Rivest & Wells (2001), we will also study the large sample properties of \hat{S}_ϕ and \hat{F}_ϕ without assuming that the distribution of (T, U) is given by the survival Archimedean copula \mathcal{C}_ϕ . In addition we will study the joint large sample properties of \hat{S}_ϕ with the Kaplan-Meier estimator of F constructed on the observations $\{y_i, \delta_{U_i}\}_{i=1}^n$.

The estimators and the framework are presented in section 2. These estimators are written as stochastic integrals with respect to counting processes. We will fix a filtration and provide the corresponding intensity processes of these counting processes, i.e. their Doob-Meyer decomposition. The large sample properties of the estimators are the object of section 3. We assume the reader to be familiar with the basic martingale framework in the theory of counting processes; see e.g. chapters 1 and 2 in Fleming & Harrington (1991). Nevertheless we hope that an intuitive overview of these notions is sufficient for understanding our approach. We will use twice the following easy corollary to Lengart's inequality; we refer to Fleming & Harrington (1991) or Karatzas & Shreve (1988) for the statement of this inequality and we left the proof of this corollary to the reader.

Corollary to Lengart's inequality. *Let (M^n) be a sequence of local square integrable martingales. If $\langle M^n \rangle(t) = O_{\mathbb{P}}(n)$, then $\sup_{0 \leq s \leq t} |M^n(s)| = O_{\mathbb{P}}(\sqrt{n})$. The same statement with $o_{\mathbb{P}}$ instead of $O_{\mathbb{P}}$ also holds true.*

2 Estimators and martingale tools

We denote by Γ the survival function of $\min(T, U)$. Remark that V is independent of $\min(T, U)$ and X equals $\min(T, U)$ if and only if $\Delta_T = 1$ or $\Delta_U = 1$, therefore the Kaplan-Meier estimator $\hat{\Gamma}$ of Γ and the Kaplan-Meier estimator \hat{G} of G are available as statistics of the observations $\{x_i, \delta_{T_i}, \delta_{U_i}\}_{i=1}^n$. Similarly to Rivest & Wells (2001), we seek estimators \hat{S}_ϕ and \hat{F}_ϕ which are right-continuous step-functions satisfying $\hat{S}_\phi(0) = \hat{F}_\phi(0) = 1$ and

$$\phi^{-1}(\phi(\hat{S}_\phi(t)) + \phi(\hat{F}_\phi(t))) = \hat{\Gamma}(t) \quad \text{when } 0 \leq t \leq \tilde{x}_n := \max(x_i; 1 \leq i \leq n), \quad (1)$$

and we also require that \hat{S}_ϕ jumps at x_i when $\delta_{T_i} = 1$ and that \hat{F}_ϕ jumps at x_i when $\delta_{T_i} = 0$ and $\delta_{U_i} = 1$. Assuming no tied values, relation (1) implies that the sizes of the jumps of $\phi \circ \hat{S}_\phi$ and $\phi \circ \hat{F}_\phi$ must equal the sizes of the jumps of $\phi \circ \hat{\Gamma}$, therefore the unique estimators satisfying the desiderata are given by

$$\phi(\hat{S}_\phi(t)) = \sum_{i=1}^n \mathbb{1}_{\{x_i \leq t\}} \Delta(\phi \circ \hat{\Gamma})(x_i) \delta_{T_i}$$

and

$$\phi(\hat{F}_\phi(t)) = \sum_{i=1}^n \mathbb{1}_{\{x_i \leq t\}} \Delta(\phi \circ \hat{\Gamma})(x_i) (1 - \delta_{T_i}) \delta_{U_i}$$

where $\Delta(\phi \circ \hat{\Gamma})(s) = (\phi \circ \hat{\Gamma})(s) - (\phi \circ \hat{\Gamma}^-)(s)$ is the size of the jump of $\phi \circ \hat{\Gamma}$ at s . Note that \hat{S}_ϕ and \hat{F}_ϕ are Kaplan-Meier estimators when $\phi = -\log$. It is possible that $\hat{\Gamma}(\tilde{x}_n) = 0$, and in this case either $\hat{S}_\phi(\tilde{x}_n) = 0$ or $\hat{F}_\phi(\tilde{x}_n) = 0$. With counting processes notations,

$$\hat{S}_\phi(t) = \phi^{-1} \left[\int_0^t \mathbb{1}_{\{s \leq \tilde{x}_n\}} \Delta(\phi \circ \hat{\Gamma})(s) d\bar{N}_S(s) \right] \quad \text{and} \quad \hat{F}_\phi(t) = \phi^{-1} \left[\int_0^t \mathbb{1}_{\{s \leq \tilde{x}_n\}} \Delta(\phi \circ \hat{\Gamma})(s) d\bar{N}_F(s) \right]$$

where \bar{N}_S and \bar{N}_F are the counting processes defined by $\bar{N}_S(t) = \sum_{i=1}^n \delta_{T_i} \mathbb{1}_{\{x_i \leq t\}}$ and $\bar{N}_F(t) = \sum_{i=1}^n (1 - \delta_{T_i}) \delta_{U_i} \mathbb{1}_{\{x_i \leq t\}}$.

We will focus on \hat{S}_ϕ . The analogous asymptotic results for \hat{F}_ϕ will be mentioned at the end of section 3.

As a first step to the asymptotic study of \hat{S}_ϕ , we give in lemma 2.1 below another decreasing right-continuous step function \hat{S}_ϕ^* starting at 1 and which is asymptotically equivalent to \hat{S}_ϕ whatever the distribution of (T, U) . The advantage of \hat{S}_ϕ^* is that $\phi \circ \hat{S}_\phi^*$ is written as a stochastic integral with respect to \bar{N}_S of a left-continuous process. From now on, we denote by $\pi = \Gamma G$ the survival function of $X = \min(T, U, V)$ and by $\hat{\pi}$ the empirical survival function of the x_i . Remark that the equality $\hat{\pi} = \hat{\Gamma} \hat{G}$ can fail in the presence of ties in the jump times of $\hat{\Gamma}$ and \hat{G} , but it almost surely holds since the law of $\min(T, U)$ is continuous by assumption. Throughout the sequel, we will implicitly assume that there are no such tied values.

Lemma 2.1 *Let \hat{S}_ϕ^* be the right-continuous step function defined by*

$$(\phi \circ \hat{S}_\phi^*)(t) = -\frac{1}{n} \int_0^t \mathbb{1}_{\{\hat{\pi}(s) > 0\}} \frac{\phi'(\hat{\Gamma}^-(s))}{\hat{G}^-(s)} d\bar{N}_S(s),$$

and let $t_0 > 0$ be such that $\pi(t_0) > 0$. If ϕ is twice differentiable with continuous second derivative, then $\sup_{0 \leq t \leq t_0} |(\phi \circ \hat{S}_\phi)(t) - (\phi \circ \hat{S}_\phi^*)(t)| = O_P(n^{-1})$.

Proof First note that $\hat{\Gamma}^-(s) - \hat{\Gamma}(s) = \frac{1}{n\hat{G}(s)}$ for any $s \in \{t_i \mid \delta_{Ti} = 1\}$ since $\hat{G}(s)\hat{\Gamma}^-(s) = \hat{\pi}(s)$ and $\hat{G}(s)\hat{\Gamma}(s) = \hat{\pi}(s)$. Therefore, for such a $s \leq t_0$, by Taylor expansion, the inequality

$$\left| \Delta(\phi \circ \hat{\Gamma})(s) + \frac{1}{n\hat{G}(s)} \phi'(\hat{\Gamma}^-(s)) \right| \leq \sup_{\pi(t_0) \leq s \leq 1} |\phi''(s)| \frac{4}{(n\pi(t_0))^2}$$

holds on the event $\{\min(\hat{\Gamma}(s), \hat{G}(s)) > \pi(t_0)/2 \text{ for all } s \in [0, t_0]\}$, whose probability converges to 1 as the sample size increases. The result follows from the obvious inequality $\bar{N}_S(t_0) \leq n$. \square

Remark that

$$(\phi \circ \hat{S}_\phi^*)(t) = -\int_0^t \mathbb{1}_{\{\hat{\pi}(s) > 0\}} \hat{\Gamma}^-(s) \phi'(\hat{\Gamma}^-(s)) d\hat{\Lambda}(s) \quad (2)$$

where $\hat{\Lambda} = \frac{1}{n} \int \frac{d\bar{N}_S}{\hat{\pi}}$ is the Nelson-Aalen estimator of the cumulative *crude hazard rate* of T subject to be censored by $Y = \min(U, V)$, which is conceptually given by

$$\Lambda(ds) = \frac{\mathbb{P}(T \in ds, T \leq \min(U, V))}{\mathbb{P}(\min(T, U, V) \geq s)},$$

but since $\mathbb{P}(T \in ds, T \leq V \mid T \leq U) = \mathbb{P}(V \geq s) \mathbb{P}(T \in ds \mid T \leq U)$, we see that

$$\Lambda(ds) = \frac{\mathbb{P}(T \in ds, T \leq U)}{\mathbb{P}(T \geq s, U \geq s)}$$

is also the crude hazard rate of T subject to be censored by U . It is possible that $\int d\Lambda < \infty$, for example when the joint law of T and U is given by an Archimedean Clayton copula with generator $x \mapsto x^{-1} - 1$ acting on the marginal survival functions $S(t) = (t+1)^{-1}$ and $F(t) = \exp(-t)$.

From equality (2), it is expected that $\phi \circ \hat{S}_\phi^*$, and then also $\phi \circ \hat{S}_\phi$ owing to lemma 2.1, converges to $\phi \circ S_\phi^*$ given by

$$(\phi \circ S_\phi^*)(t) = \int_0^t \psi(\Gamma(s)) d\Lambda(s), \quad (3)$$

where, from now on, we denote by ψ the positive function $\psi(x) = -x\phi'(x)$, in accordance with the notations of Rivest & Wells (2001). The decreasing function S_ϕ^* is not always a proper survival function: consider for instance our preceding example of the integrable crude hazard rate Λ , and take $\phi = -\log$, so that $S_\phi^* = \exp(-\Lambda)$ and $S_\phi^*(\infty) > 0$. It is easy to check that $S_\phi^* = S$ when S is absolutely continuous and when the joint law of (T, U) is given by the Archimedean copula \mathcal{C}_ϕ acting on the marginal survival functions with a differentiable generator ϕ , by knowing that in that case, $\Lambda(ds) = \lambda(s) ds$ with $\lambda(t) = -\frac{1}{\Gamma(t)} \frac{\phi'(S(t))}{\phi'(\Gamma(t))} S'(t)$, as claimed by Rivest & Wells (2001). The interesting proposition below is a copy of proposition 2 in Rivest & Wells (2001).

Proposition 2.1 *If ϕ'_1/ϕ'_2 is increasing on $(0, 1)$ then $S_{\phi_2}^* \leq S_{\phi_1}^*$.*

The asymptotic behavior of $(\phi \circ \hat{S}_\phi) - (\phi \circ S_\phi^*)$ will be rigorously studied with the help of martingale techniques. Now, first of all, we seek a filtration with respect to which $\phi \circ \hat{S}_\phi^*$ has a simple Doob-Meyer decomposition. When considering the filtration $(\tilde{\mathcal{F}}_t)$ defined by

$$\tilde{\mathcal{F}}_t = \bigvee_{i=1}^n \bigvee_{0 \leq s \leq t} \sigma(\delta_{Ti} \mathbb{1}_{\{x_i \leq s\}}, (1 - \delta_{Ti}) \mathbb{1}_{\{x_i \leq s\}}),$$

it is known (Fleming & Harrington 1991) that the counting process \bar{N}_S has the Doob-Meyer decomposition $\bar{N}_S = \bar{M}_S + \bar{A}$ where \bar{M}_S is a square integrable martingale and the compensator \bar{A} is the (increasing and predictable) process given by

$$\bar{A}(t) = n \int_0^t \hat{\pi}^-(s) d\Lambda(s) \quad (4)$$

where Λ is the cumulative crude hazard rate introduced above. In our study, the compensator \bar{A} is a continuous process since we assume that T has a continuous distribution.

However the filtration $\tilde{\mathcal{F}}$ is not adequate for our study, since, for example, the stochastic processes $\hat{\Gamma}^-$ and \hat{G}^- are not adapted to $\tilde{\mathcal{F}}$, therefore $\phi \circ \hat{S}_\phi^*$ is not written as a stochastic integral with respect to \bar{N}_S of a predictable process, while being left-continuous, because it is not adapted to $\tilde{\mathcal{F}}$. From now we will consider the bigger filtration $\tilde{\mathcal{H}} = (\tilde{\mathcal{H}}_t)$ defined by

$$\tilde{\mathcal{H}}_t = \bigvee_{i=1}^n \bigvee_{0 \leq s \leq t} \sigma(\delta_{Ti} \mathbb{1}_{\{x_i \leq s\}}, (1 - \delta_{Ti}) \mathbb{1}_{\{x_i \leq s\}}, \delta_{Ui} \mathbb{1}_{\{x_i \leq s\}}, (1 - \delta_{Ui}) \mathbb{1}_{\{x_i \leq s\}})$$

as the underlying filtration. In particular, $\hat{\Gamma}^-$ and \hat{G}^- are predictable processes with respect to $\tilde{\mathcal{H}}$. But we will keep the Doob-Meyer decomposition $\bar{N}_S = \bar{M}_S + \bar{A}$ of \bar{N}_S because, as we will now see (lemma 2.4), it is actually the same w.r.t. either $\tilde{\mathcal{F}}$ or $\tilde{\mathcal{H}}$, that is, \bar{M}_S remains a martingale with respect to $\tilde{\mathcal{H}}$. It would be sufficient to prove this fact by deriving the compensator of \bar{N}_S w.r.t. $\tilde{\mathcal{H}}$ in the same way as one derives its compensator w.r.t. $\tilde{\mathcal{F}}$, as done by Fleming & Harrington (1991). But it is more instructive to prove that the filtration $\tilde{\mathcal{F}}$ is *immersed* in the filtration $\tilde{\mathcal{H}}$, which is the statement of lemma 2.4. We say that a

filtration \mathcal{F} is immersed in a filtration \mathcal{G} and we write $\mathcal{F} \overset{m}{\subset} \mathcal{G}$ if all \mathcal{F} -martingales are \mathcal{G} -martingales. This property is also known in the literature as the (H) -hypothesis. We refer e.g. to Émery & Schachermayer (2001) or Mansuy & Yor (2006) and references given therein for more information on the notion of immersion, which is very useful in many aspects of stochastic calculus, but which seems to be rather unfamiliar to statisticians. The notion of immersion already appeared in the literature on survival analysis without being explicitly mentioned: it could be used to formulate the property (M) in Jacobsen (1989) about the concept of independent right-censoring. It is more popular in the literature on credit risk models: it explicitly appears e.g. in the papers by Gapeev & Jeanblanc (2008) and Bielecki, Jeanblanc & Rutkowski (2008).

We have claimed just above that it is more instructive to prove the immersion of $\tilde{\mathcal{F}}$ in $\tilde{\mathcal{H}}$ than just show that the compensator of \bar{N}_S in $\tilde{\mathcal{F}}$ is the same as its compensator in $\tilde{\mathcal{H}}$ by a direct calculation. Indeed, the immersion is a clean statement from which the stability of the compensator is a direct obvious consequence, hence it is a better explanation of why the compensator is stable. Moreover the immersion of $\tilde{\mathcal{F}}$ in $\tilde{\mathcal{H}}$ will be derived from lemma 2.3 which will be also used to derive another immersion property (lemma 2.5). Lemma 2.2 below will allow us to restrict our attention to the case $n = 1$.

Lemma 2.2 *Let \mathcal{F}^1 and \mathcal{G}^1 be two filtrations on some probability space and let \mathcal{F}^2 and \mathcal{G}^2 be two filtrations on a possibly other probability space. If $\mathcal{F}^1 \overset{m}{\subset} \mathcal{G}^1$ and $\mathcal{F}^2 \overset{m}{\subset} \mathcal{G}^2$ then $\mathcal{F}^1 \otimes \mathcal{F}^2 \overset{m}{\subset} \mathcal{G}^1 \otimes \mathcal{G}^2$.*

Proof See lemma 2 in Émery & Schachermayer (2001). \square

Owing to this lemma, in order to prove that $\tilde{\mathcal{F}} \overset{m}{\subset} \tilde{\mathcal{H}}$ it is sufficient to prove that $\mathcal{F} \overset{m}{\subset} \mathcal{H}$ where the filtrations \mathcal{F} and \mathcal{H} are defined by

$$\mathcal{F}_t = \bigvee_{0 \leq s \leq t} \sigma(\Delta_T \mathbb{1}_{\{X \leq s\}}, (1 - \Delta_T) \mathbb{1}_{\{X \leq s\}})$$

and

$$\mathcal{H}_t = \bigvee_{0 \leq s \leq t} \sigma(\Delta_T \mathbb{1}_{\{X \leq s\}}, (1 - \Delta_T) \mathbb{1}_{\{X \leq s\}}, \Delta_U \mathbb{1}_{\{X \leq s\}}, (1 - \Delta_U) \mathbb{1}_{\{X \leq s\}}).$$

Note that the event $\{X > t\}$ is an atom of both \mathcal{F}_t and \mathcal{H}_t and on the event $\{X \leq t\}$ one has $\mathcal{F}_t = \mathcal{F}_\infty$ and $\mathcal{H}_t = \mathcal{H}_\infty$.

Lemma 2.3 *Let \mathcal{G} be any filtration containing $\tilde{\mathcal{F}}$ and such that for each $t \geq 0$ the event $\{X > t\}$ is an atom of \mathcal{G}_t and $\mathcal{G}_t \supset \mathcal{F}_\infty$ on the event $\{X \leq t\}$. Then $\tilde{\mathcal{F}}$ is immersed in \mathcal{G} .*

Proof Under these assumptions, for every random variable $V \in L^1(\mathcal{F}_\infty)$ and every $t \geq 0$,

$$\mathbb{E}[V \mid \mathcal{F}_t] = \mathbb{E}[V \mid \mathcal{G}_t] = \mathbb{E}[V \mid X > t] \mathbb{1}_{\{X > t\}} + V \mathbb{1}_{\{X \leq t\}}.$$

The equality $\mathbb{E}[V \mid \mathcal{F}_t] = \mathbb{E}[V \mid \mathcal{G}_t]$ means that \mathcal{F}_∞ is conditionally independent of \mathcal{G}_t given \mathcal{F}_t , and this equality holds for every $t \geq 0$, which is equivalent to $\mathcal{F} \overset{m}{\subset} \mathcal{G}$ by lemma 5 in Émery & Schachermayer (2001). \square

Lemma 2.4 *The filtration $\tilde{\mathcal{F}}$ is immersed in the filtration $\tilde{\mathcal{H}}$. Consequently the Doob-Meyer decomposition of \bar{N}_S with respect to $\tilde{\mathcal{H}}$ is the same as its Doob-Meyer decomposition with respect to $\tilde{\mathcal{F}}$.*

Proof The immersion stems from lemma 2.2 and lemma 2.3, and the consequence obviously follows. \square

The quadratic variation $\langle \bar{M}_S \rangle$ of the square integrable martingale \bar{M}_S will be needed. It equals the compensator \bar{A} since this one is continuous; this fact is the well-known particular case when $N_1 = N_2 = N$ in the following theorem that we will need and that we prove because we do not find it in the literature.

Theorem 2.1 *Let N_1 and N_2 be counting processes and their Doob-Meyer decompositions $N_1 = M_1 + A_1$ and $N_2 = M_2 + A_2$ with respect to some right-continuous filtration. If the compensators A_1 and A_2 are continuous, then the covariation process $\langle M_1, M_2 \rangle$ equals the compensator of the counting process N which counts how many times N_1 and N_2 have simultaneous jumps, that is, N is defined by $N(t) = \sum_{s \leq t} \Delta N_1(s) \Delta N_2(s)$.*

Proof The proof we give is an adaptation of the proof of theorem 2.5.1 in Fleming & Harrington (1991). (the particular case when $N_1 = N_2 = N$). The local (square integrable) martingales M_1 et M_2 are right-continuous processes of bounded variation, hence the integration by parts formula for Lebesgue-Stieltjes integrals yields

$$M_1(t)M_2(t) = \int_0^t M_1^-(s) dM_2(s) + \int_0^t M_2^-(s) dM_1(s) + \sum_{s \leq t} \Delta M_1(s) \Delta M_2(s).$$

We have $\Delta M_i(s) = \Delta N_i(s)$ since A_1 and A_2 are continuous, therefore $\sum_{s \leq t} \Delta M_1(s) \Delta M_2(s) = N(t)$. Hence, denoting by $N = M + A$ the Doob-Meyer decomposition of N , one has

$$M_1(t)M_2(t) - A(t) = M(t) + \int_0^t M_1^-(s) dM_2(s) + \int_0^t M_2^-(s) dM_1(s)$$

Each process M_1 and M_2 is adapted, left-continuous and has right-hand limits, and is therefore locally bounded and predictable so that each process $\int M_1^- dM_2$ and $\int M_2^- dM_1$ is a local (square integrable) martingale. Finally the right-hand side of the equality above is a right-continuous local martingale, thereby showing that $A = \langle M_1, M_2 \rangle$. \square

Precisely, we will need to apply this theorem with the counting process \bar{N}_S which counts the jumps of \hat{S}_ϕ and the counting process \bar{N}_Γ which counts the jumps of the Kaplan-Meier estimator $\hat{\Gamma}$ of Γ , that is, $\bar{N}_\Gamma(t) = \sum_{i=1}^n \eta_i \mathbb{1}_{\{x_i \leq t\}}$ where $\eta_i = \min(1, \delta_{T_i} + \delta_{U_i})$. When considering the filtration $\bar{\mathcal{G}}$ defined by

$$\bar{\mathcal{G}}_t = \bigvee_{i=1}^n \bigvee_{0 \leq s \leq t} \sigma(\eta_i \mathbb{1}_{\{x_i \leq s\}}, (1 - \eta_i) \mathbb{1}_{\{x_i \leq s\}}),$$

it is known that \bar{N}_Γ has the Doob-Meyer decomposition $\bar{N}_\Gamma = \bar{M}_\Gamma + \bar{A}_\Gamma$ with $\bar{A}_\Gamma(t) = n \int_0^t \hat{\pi}^-(s) d\Lambda_\Gamma(s)$ where $\Lambda_\Gamma(ds) = \mathbb{P}(\min(T, U) \in ds) / \mathbb{P}(\min(T, U) \geq s)$, and in particular $\Lambda_\Gamma(ds) = \lambda_\Gamma(s) ds$ with $\lambda_\Gamma(s) = -\Gamma'(s) / \Gamma(s)$ when Γ is absolutely continuous. In our study we do not work in the filtration $\bar{\mathcal{G}}$ but in the filtration $\bar{\mathcal{H}}$ previously introduced, but that does not change the Doob-Meyer decomposition, according to the following lemma.

Lemma 2.5 *The filtration $\bar{\mathcal{G}}$ is immersed in the filtration $\bar{\mathcal{H}}$, and consequently the Doob-Meyer decomposition of \bar{N}_Γ is the same with respect to either $\bar{\mathcal{G}}$ or $\bar{\mathcal{H}}$.*

Proof Similar to the proof of lemma 2.4. \square

Note that the compensator \bar{A}_Γ is a continuous process since $\min(T, U)$ has a continuous distribution by assumption. Therefore $\langle \bar{M}_\Gamma \rangle = \bar{A}_\Gamma$. In addition to the quadratic variation processes $\langle \bar{M}_S \rangle$ and $\langle \bar{M}_\Gamma \rangle$ we will also need the covariation process $\langle \bar{M}_\Gamma, \bar{M}_S \rangle$ (with respect to the big filtration \mathcal{H}).

Lemma 2.6 *The covariation process $\langle \bar{M}_\Gamma, \bar{M}_S \rangle$ of the local square integrable \mathcal{H} -martingales \bar{M}_Γ and \bar{M}_S equals the variation process $\langle \bar{M}_S \rangle$ of \bar{M}_S .*

Proof This straightforwardly stems from theorem 2.1 since the process which counts the simultaneous jumps of \bar{N}_Γ and \bar{N}_S is nothing but \bar{N}_S . \square

Now we are ready to turn to further preliminaries for the asymptotic study of \hat{S}_ϕ . The final aim of this section is to provide the asymptotic representation of $\phi \circ \hat{S}_\phi - \phi \circ S_\phi^*$ given in proposition 2.2. It is the cornerstone of the next section. The Doob-Meyer decomposition of \bar{N}_S yields

$$(\phi \circ \hat{S}_\phi^*)(t) = -\frac{1}{n} \int_0^t \mathbb{1}_{\{\hat{\pi}(s) > 0\}} \frac{\phi'(\hat{\Gamma}^-(s))}{\hat{G}^-(s)} d\bar{M}_S(s) + \int_0^t \mathbb{1}_{\{\hat{\pi}(s) > 0\}} \psi(\hat{\Gamma}^-(s)) d\Lambda(s). \quad (5)$$

Thus, $\phi \circ \hat{S}_\phi^*$ is a local submartingale: the first term in (5) is a local square integrable martingale since $\frac{\phi'(\hat{\Gamma}^-(s))}{\hat{G}^-(s)} \mathbb{1}_{\{\hat{\pi}(s) > 0\}}$ defines a locally bounded predictable process, and the second term is an increasing predictable process. Lemma below allows us to replace the (local) martingale part of $\phi \circ \hat{S}_\phi^*$ (the first term of (5)) by a more convenient process for the asymptotic study.

Lemma 2.7 *If ϕ is twice differentiable with continuous second derivative, then, for any t_0 such that $\pi(t_0) > 0$,*

$$\sup_{0 \leq t \leq t_0} \left| \int_0^t \mathbb{1}_{\{\hat{\pi}(s) > 0\}} \frac{\phi'(\hat{\Gamma}^-(s))}{\hat{G}^-(s)} d\bar{M}_S(s) - \int_0^t \frac{\phi'(\Gamma(s))}{G(s)} d\bar{M}_S(s) \right| = o_P(\sqrt{n}).$$

Proof W.r.t. the underlying filtration \mathcal{H} , the process \bar{H} given by

$$\bar{H}(s) = \mathbb{1}_{\{\hat{\pi}(s) > 0\}} \frac{\phi'(\hat{\Gamma}^-(s))}{\hat{G}^-(s)} - \frac{\phi'(\Gamma(s))}{G(s)}$$

starts at 0, is adapted and left-continuous with finite right hand limits, hence it is a locally bounded predictable process, and therefore the process $\bar{U} = \int \bar{H} d\bar{M}_S$ is a local square integrable martingale whose variation process is given by

$$\langle \bar{U} \rangle(t) = \int_0^t \bar{H}^2(s) d\langle \bar{M}_S \rangle(s) = n \int_0^t \bar{H}^2(s) \hat{\pi}(s) d\Lambda(s).$$

Since $\Gamma(t_0) > 0$ and $G(t_0) > 0$, and since both $\hat{\Gamma}^-(s) - \Gamma(s)$ and $\hat{G}^-(s) - G(s)$ are $o_P(1)$ uniformly in $s \in [0, t_0]$, then $\bar{H}^2(s)$ is itself $o_P(1)$ uniformly in $t \in [0, t_0]$. Consequently $\langle \bar{U} \rangle(t_0) = o_P(n)$, hence the result follows from the corollary to Lenglart's inequality. \square

Now we state and prove the final result of this section, which is the main tool for the next section. Recall that $\phi \circ S_\phi^*$ is defined by (3).

Proposition 2.2 *If ϕ is twice differentiable with continuous second derivative, then, for any t_0 such that $\pi(t_0) > 0$,*

$$(\phi \circ \hat{S}_\phi)(t) - (\phi \circ S_\phi^*)(t) = -\frac{1}{n} \int_0^t \frac{\phi'(\Gamma(s))}{G(s)} d\bar{M}_S(s) + \int_0^t \psi'(\Gamma(s)) (\hat{\Gamma}^-(s) - \Gamma(s)) d\Lambda(s) + o_P(n^{-\frac{1}{2}})$$

uniformly in $t \in [0, t_0]$.

Proof By lemma 2.1 and lemma 2.7, it suffices to show that

$$\int_0^t \mathbb{1}_{\{\hat{\pi}(s) > 0\}} (\psi(\hat{\Gamma}^-(s)) - \psi(\Gamma(s))) d\Lambda(s) = \int_0^t \psi'(\Gamma(s)) (\hat{\Gamma}^-(s) - \Gamma(s)) d\Lambda(s) + o_P(n^{-\frac{1}{2}})$$

uniformly in $t \in [0, t_0]$. Since ψ' is uniformly continuous on compact subsets of $(0, 1]$, this follows from a Taylor expansion of ψ and from the fact that $\hat{\Gamma}^-(s) - \Gamma(s) = O_P(n^{-\frac{1}{2}})$ uniformly in $s \in [0, t_0]$. \square

Remark 2.1 In case when $\mathbb{P}(V = \infty) = 1$ and $\phi = -\log$ then $\hat{S}_{\cdot \log}$ is the Kaplan-Meier estimator. In that case the proposition above shows that

$$\log \hat{S}_{\cdot \log}(t) - \log \Lambda(t) = -\frac{1}{n} \int_0^t \frac{1}{\pi(s)} d\bar{M}_S(s) + o_P(n^{-\frac{1}{2}})$$

uniformly in $t \in [0, t_0]$. Also note that this still holds when $\hat{S}_{\cdot \log}$ is replaced by its left-continuous version $\hat{S}_{\cdot \log}^-$ since $\hat{S}_{\cdot \log}^-(t) = \hat{S}_{\cdot \log}(t) + O_P(n^{-1})$ uniformly for $t \in [0, t_0]$.

As a consequence of remark 2.1 and lemma 2.5,

$$\log \hat{\Gamma}^-(t) - \log \Gamma(t) = -\frac{1}{n} \int_0^t \frac{1}{\pi(s)} d\bar{M}_\Gamma(s) + o_P(n^{-\frac{1}{2}}) \quad (6)$$

uniformly for $t \in [0, t_0]$, where \bar{M}_Γ is the (square integrable) martingale part in the Doob-Meyer decomposition $\bar{N}_\Gamma = \bar{M}_\Gamma + \bar{A}_\Gamma$ with respect to the filtration \mathcal{H} given above lemma 2.5.

In the next section we will study the joint asymptotic behavior of \hat{S}_ϕ and $\hat{\Gamma}^-$. Then the joint asymptotic behavior of \hat{S}_ϕ and \hat{F}_ϕ will derive from equality 1. We will also study the joint asymptotic behavior of \hat{S}_ϕ and the Kaplan-Meier estimator \hat{F} of F which is available as a statistic of the observations $\{y_i, \delta_{U_i}\}_{i=1}^n$.

3 Large sample properties

We fix some assumptions throughout this section. We consider $t_0 > 0$ such that $\pi(t_0) > 0$, we assume that ϕ is twice differentiable with continuous second derivative and that ϕ' is bounded away from 0, that is, $\phi'(1) < 0$. This last assumption will serve us to derive the asymptotic behavior of \hat{S}_ϕ from the one of $\phi \circ \hat{S}_\phi$. It is met by numerous common families of Archimedean copulas, such as Clayton, Frank, or interior power Frank families, but not by the Gumbel family (these families are defined in Nelsen, 1999).

We start by proving the consistency of \hat{S}_ϕ . Recall that $S_\phi^* = S$ and when the model is well-specified, that is, when the survival copula of (T, U) is the Archimedean copula \mathcal{C}_ϕ .

Theorem 3.1 *The process $\phi \circ \hat{S}_\phi$ is a uniformly consistent estimate of the function $\phi \circ S_\phi^*$ on $[0, t_0]$. More precisely,*

$$\sup_{0 \leq t \leq t_0} |(\phi \circ \hat{S}_\phi)(t) - (\phi \circ S_\phi^*)(t)| = O_P(n^{-\frac{1}{2}}).$$

Consequently $\sup_{0 \leq t \leq t_0} |\hat{S}_\phi(t) - S_\phi^*(t)| = O_P(n^{-\frac{1}{2}})$.

Proof The consequence easily follows from the fact that ϕ^{-1} is Lipschitz under the assumption that ϕ' is bounded away from 0. To prove the first assertion, it suffices, in view of proposition 2.2, to prove that both $\frac{1}{n} \int_0^t \frac{\phi'(\Gamma(s))}{G(s)} d\bar{M}_S(s)$ and $\int_0^t \psi'(\Gamma(s)) (\hat{\Gamma}^-(s) - \Gamma(s)) d\Lambda(s)$ are $O_P(n^{-\frac{1}{2}})$ uniformly in $t \in [0, t_0]$. This is clear for the second term because $\hat{\Gamma}^-(s) - \Gamma(s)$ is $O_P(n^{-\frac{1}{2}})$ uniformly in $s \in [0, t_0]$. For the first term, it suffices to show that $\sup_{0 \leq s \leq t_0} |\bar{M}_S(s)| = O_P(\sqrt{n})$, but this results from the corollary to Lenglart's inequality since it is clear that $\langle \bar{M}_S \rangle(t_0) = O_P(n)$. \square

Our next underlying goal is to prove that the processes $\sqrt{n}(\phi \circ \hat{S}_\phi - \phi \circ S_\phi^*)$ and $\sqrt{n}(\log \hat{\Gamma}^- - \log \Gamma)$ jointly converges weakly to an explicit continuous centered Gaussian martingale for the Skorohod topology on $D[0, t_0]$. Here and hereafter, we shortly denote by $D[0, t_0]$ the Skorohod space $D([0, t_0], \mathbb{R}^d)$ where the dimension d is clearly understood from the context. By proposition 2.2, we can equivalently replace $\sqrt{n}(\phi \circ \hat{S}_\phi - \phi \circ S_\phi^*)$ by

$$-\frac{1}{\sqrt{n}} \int_0^t \frac{\phi'(\Gamma(s))}{G(s)} d\bar{M}_S(s) + \sqrt{n} \int_0^t \psi'(\Gamma(s)) (\hat{\Gamma}^-(s) - \Gamma(s)) d\Lambda(s). \quad (7)$$

We put $\bar{W}(t) = \int_0^t \frac{\phi'(\Gamma(s))}{G(s)} d\bar{M}_S(s)$. Hence the key step is to study the joint asymptotic distribution of $n^{-\frac{1}{2}} \bar{W}$ and $\sqrt{n}(\log \hat{\Gamma}^- - \log \Gamma)$, which will be given in proposition 3.1. Note that we can replace $\log \hat{\Gamma}^- - \log \Gamma$ by its asymptotic martingale representation (6). The tool we will use is the following version of the martingale central limit theorem (see Ethier & Kurtz, 1986).

Martingale central limit theorem. *For each $n \geq 0$, let $U_1^{(n)}, \dots, U_d^{(n)}$ be d (cdlg) local square integrable martingales w.r.t. some filtration $\mathcal{F}^{(n)}$ satisfying the usual conditions (complete and right-continuous) and form the d -variate martingale $U^{(n)} = (U_1^{(n)}, \dots, U_d^{(n)})$. Assume that the variation processes $\langle U^{(n)} \rangle$ are continuous. Let $U^\infty = (U_1^\infty, \dots, U_d^\infty)$ be a continuous d -variate centered Gaussian martingale and $t_0 > 0$. If $\langle U^{(n)} \rangle(t) \xrightarrow{\mathbb{P}} \langle U^\infty \rangle(t)$ for all $t \in (0, t_0)$ and $\mathbb{E} \left[\sup_{s \leq t_0} |\Delta U^{(n)}(s)|^2 \right] \rightarrow 0$, where $|\cdot|$ denotes the Euclidean norm, then the sequence $(U^{(n)})$ converges weakly to U^∞ in $D[0, t_0]$.*

This theorem is often used in the following context. We have counting processes $N_j^{(n)}$, each having Doob-Meyer decomposition $N_j^{(n)} = M_j^{(n)} + A_j^{(n)}$ where $M_j^{(n)}$ is a local square integrable martingale and the compensator $A_j^{(n)}$ is continuous, and we consider the local square integrable martingales $U_j^{(n)}(t) = \int_0^t H_j^{(n)}(s) dM_j^{(n)}(s)$ for some given predictable locally bounded processes $H_j^{(n)}$. In this case, $\langle U_j^{(n)} \rangle(t) = \int_0^t (H_j^{(n)}(s))^2 dA_j^{(n)}(s)$ and the jumps of $U_j^{(n)}$ are the $H_j^{(n)}(\tau_i)$ where the τ_i are the jumping times of the counting process $N_j^{(n)}$. In

this context, the condition $\mathbb{E} \left[\sup_{s \leq t_0} |\Delta U^{(n)}(s)|^2 \right] \rightarrow 0$ occurs whenever the $H_j^{(n)}$ are uniformly bounded in n for each j .

By applying the martingale central limit theorem we easily get

$$\frac{1}{\sqrt{n}}(\bar{M}_S, \bar{M}_\Gamma) \rightsquigarrow \mathbb{G} = (\mathbb{G}_1, \mathbb{G}_2) \quad \text{in } D[0, t_0],$$

where \mathbb{G} is a continuous centered Gaussian martingale with $\mathbb{E}\mathbb{G}_1(t)^2 = \langle \mathbb{G}_1 \rangle(t) = \int_0^t \pi \, d\Lambda$, $\mathbb{E}\mathbb{G}_2(t)^2 = \langle \mathbb{G}_2 \rangle(t) = \int_0^t \pi \, d\Lambda_\Gamma$ and $\mathbb{E}\mathbb{G}_1(s)\mathbb{G}_2(t) = \langle \mathbb{G}_1, \mathbb{G}_2 \rangle(s \wedge t) = \int_0^{\min(s,t)} \pi \, d\Lambda$. Then we can heuristically claim that

$$\left(n^{-\frac{1}{2}}\bar{W}, n^{-\frac{1}{2}} \int \frac{1}{\pi} d\bar{M}_\Gamma \right) \rightsquigarrow (\mathbb{M}_1, \mathbb{M}_2) := \left(\int \frac{\phi' \circ \Gamma}{G} d\mathbb{G}_1, \int \frac{1}{\pi} d\mathbb{G}_2 \right) \quad \text{in } D[0, t_0], \quad (8)$$

which yield proposition 3.1 since $\log \hat{\Gamma}(t) - \log \Gamma(t) \approx -\frac{1}{n} \int_0^t \frac{1}{\pi(s)} d\bar{M}_\Gamma(s)$ in the sense of the asymptotic martingale representation (6). We will rigorously prove (8) with the martingale central limit theorem. Lemma 3.1 below is a preliminary step.

Lemma 3.1 *The process $n^{-\frac{1}{2}}\bar{W}$ converges weakly in $D[0, t_0]$ to the centered Gaussian martingale \mathbb{M}_1 whose variance function V_1 is given by*

$$V_1(t) = \int_0^t \frac{\psi(\Gamma(s))^2}{\pi(s)} d\Lambda(s).$$

Proof It is easy to check that $\langle \mathbb{M}_1 \rangle = V_1$. Now we apply the martingale central limit theorem. The process $n^{-\frac{1}{2}}\bar{H}$ on the interval $(0, t_0)$ where \bar{H} is given by

$$\bar{H}(s) = \frac{\phi'(\Gamma(s))}{G(s)} = -\frac{\psi(\Gamma(s))}{\pi(s)}$$

is predictable since it is deterministic, and it is bounded. Therefore the process $\bar{U} := n^{-\frac{1}{2}}\bar{W}$ is a square integrable martingale whose variation process $\langle \bar{U} \rangle$ is given by

$$\langle \bar{U} \rangle(t) = \frac{1}{n} \int_0^t \bar{H}^2(s) d\langle \bar{M}_S \rangle(s) = \int_0^t \bar{H}^2(s) \hat{\pi}(s) d\Lambda(s),$$

and $\langle \bar{U} \rangle(t)$ goes to $V_1(t)$ in probability since

$$\sup_{0 \leq s \leq t_0} \left| \bar{H}^2(s) \hat{\pi}(s) - \frac{\psi(\Gamma(s))^2}{\pi(s)} \right| \leq \frac{\max_{\Gamma(t_0) \leq x \leq 1} \{\psi(x)^2\}}{\pi(t_0)} \sup_{0 \leq s \leq t_0} \left| \frac{\hat{\pi}(s)}{\pi(s)} - 1 \right| = o_P(1)$$

The other condition $\mathbb{E} \left[\sup_{s \leq t} |\Delta \bar{U}(s)|^2 \right] \rightarrow 0$ required by the martingale central limit theorem obviously follows from the boundedness of \bar{H} . \square

Proposition 3.1 *The bivariate process*

$$\sqrt{n}(n^{-1}\bar{W}(t), \log \hat{\Gamma}(t) - \log \Gamma(t))$$

converges weakly in $D[0, t_0]$ to the centered Gaussian martingale $(\mathbb{M}_1, -\mathbb{M}_2)$ whose variance function at time t is given by:

1. $v_1(t) := \mathbb{E}\mathbb{M}_1(t)^2 = V_1(t)$ (lemma 3.1);

2. $v_2(t) := \mathbb{E}\mathbb{M}_2(t)^2 = \int_0^t \frac{1}{\pi(s)} d\Lambda_\Gamma(s)$
3. $v_{12}(t) := -\mathbb{E}\mathbb{M}_1(t)\mathbb{M}_2(t) = \int_0^t \frac{\psi(\Gamma(s))}{\pi(s)} d\Lambda(s)$.

Proof The variance function of $(\mathbb{M}_1, \mathbb{M}_2)$ is easy to derive. We have seen that $n^{-\frac{1}{2}}\bar{W} \rightsquigarrow \mathbb{M}_1$ in lemma 3.1. The weak convergence $\log \hat{\Gamma}^- - \log \Gamma \rightsquigarrow -\mathbb{M}_2$ also results from this lemma in view of remark 2.1 and the asymptotic martingale representation (6). Still due to (6), we are equivalently seeking to derive the limit law of the joint process

$$\bar{U} = (\bar{U}_1, \bar{U}_2) := \left(n^{-\frac{1}{2}}\bar{W}, -n^{-\frac{1}{2}} \int \frac{1}{\pi} d\bar{M}_\Gamma \right).$$

To do so, we apply the martingale central limit theorem. It remains to show that the covariation process $\langle \bar{U}_1, \bar{U}_2 \rangle$ of the square integrable martingales \bar{U}_1 and \bar{U}_2 goes to v_{12} and to check the condition about the jumps. The covariation process $\langle \bar{U}_1, \bar{U}_2 \rangle$ is given by $\langle \bar{U}_1, \bar{U}_2 \rangle(t) = -\int_0^t \frac{\phi'(\Gamma(s))}{\pi(s)G(s)} d\langle \bar{M}_S, \bar{M}_\Gamma \rangle(s)$, and we know that $\langle \bar{M}_S, \bar{M}_\Gamma \rangle(s) = \langle \bar{M}_S \rangle(s)$ by lemma 2.6, hence

$$\langle \bar{U}_1, \bar{U}_2 \rangle(t) = -\int_0^t \frac{\phi'(\Gamma(s))}{G(s)} \frac{\hat{\pi}(s)}{\pi(s)} d\Lambda(s) \xrightarrow{P} -\int_0^t \frac{\phi'(\Gamma(s))}{G(s)} d\Lambda(s) = v_{12}(t).$$

The condition $\mathbb{E} \left[\sup_{s \leq t} |\Delta \bar{U}(s)|^2 \right] \rightarrow 0$ required by the martingale central limit theorem obviously holds because of the boundedness of the jumps of $\sqrt{n}\bar{U}_1$ and $\sqrt{n}\bar{U}_2$. \square

From now on, we denote by \bar{D} the process defined by $\bar{D}(t) = \hat{\Gamma}^-(t) - \Gamma(t)$, so that the right member in (7) is $\sqrt{n} \int_0^t \psi'(\Gamma(s)) \bar{D}(s) d\Lambda(s)$.

Proposition 3.2 *The tridimensional process $\sqrt{n} (n^{-1}\bar{W}, \int (\psi' \circ \Gamma) \bar{D} d\Lambda, \bar{D})$ converges weakly in $D[0, t_0]$ to the centered Gaussian process $(\mathbb{M}_1, -\int (\psi' \circ \Gamma) \Gamma \mathbb{M}_2 d\Lambda, \Gamma \mathbb{M}_2)$. The variance of the second component of the limit process at t is*

$$V_2(t) = 2 \int_0^t \int_0^s \Gamma(u) \Gamma(s) \psi'(\Gamma(u)) \psi'(\Gamma(s)) v_2(u) d\Lambda(u) d\Lambda(s),$$

where v_2 is defined in the preceding lemma, and the covariance between the first two components at t is

$$V_{12}(t) = \int_0^t \int_0^s \frac{\psi(\Gamma(u))}{\pi(u)} \Gamma(s) \psi'(\Gamma(s)) d\Lambda(u) d\Lambda(s).$$

Proof First, in view of proposition 3.1 and the Delta method (proposition A.1), the process $\sqrt{n}(n^{-1}\bar{W}, \bar{D})$ converges weakly to the continuous centered Gaussian process $(\mathbb{M}_1, \Gamma \mathbb{M}_2)$. Since $(\alpha, \beta) \mapsto (\alpha, \int_0^\cdot g(\beta(s)) ds, \beta)$ is a continuous function from $D([0, t_0], \mathbb{R}^2)$ to $D([0, t_0], \mathbb{R}^3)$ whenever g is continuous (see Prigent, 2003), the announced weak convergence stems from the continuous mapping theorem. By Fubini's theorem,

$$\left(\int_0^t \Gamma(s) \psi'(\Gamma(s)) \mathbb{M}_2(s) d\Lambda(s) \right)^2 = 2 \int_0^t \int_0^s \Gamma(u) \Gamma(s) \psi'(\Gamma(u)) \psi'(\Gamma(s)) \mathbb{M}_2(u) \mathbb{M}_2(s) d\Lambda(u) d\Lambda(s),$$

and the expression of $V_2(t)$ follows from Fubini's theorem as well. The expression of V_{12} also follows from Fubini's theorem and from the expression of v_{12} given in proposition 3.1. \square

Remark 3.1 In the particular case treated by Rivest & Wells (2001), an error occurs in the expression of V_2 . This error results from the erroneous claim that the limit variance of $\sqrt{n}(\hat{\pi}(t) - \pi(t))$ is $\pi(s \wedge t) - \pi(s)\pi(t)$ instead of the correct variance $\pi(s \vee t) - \pi(s)\pi(t)$.

Theorem 3.2 below describes the asymptotic distributional behavior of our estimate of interest \hat{S}_ϕ . Recall that $S_\phi^* = S$ and $\Lambda(ds) = -\frac{1}{\Gamma(s)} \frac{\phi'(S(s))}{\phi'(\Gamma(s))} S'(s) ds$ in case when the Archimedean copula \mathcal{C}_ϕ is the survival copula of (T, U) .

Theorem 3.2 *The process $\sqrt{n}(\phi(\hat{S}_\phi) - \phi(S_\phi^*))$ converges weakly in $D[0, t_0]$ to a centered Gaussian process with variance function*

$$\begin{aligned} V(t) &= \int_0^t \frac{\Psi(\Gamma(s))^2}{\pi(s)} d\Lambda(s) \\ &\quad - 2 \int_0^t \int_0^s \frac{\Psi(\Gamma(u))}{\pi(u)} \Gamma(s) \Psi'(\Gamma(s)) d\Lambda(u) d\Lambda(s) \\ &\quad + 2 \int_0^t \int_0^s \Gamma(u) \Gamma(s) v_2(u) \Psi'(\Gamma(u)) \Psi'(\Gamma(s)) d\Lambda(u) d\Lambda(s) \end{aligned}$$

where $v_2(t) = \int_0^t \frac{1}{\pi(s)} d\Lambda_\Gamma(s)$.

Consequently, $\sqrt{n}(\hat{S}_\phi - S_\phi^*)$ converges weakly to a centered Gaussian process with variance function $V(t)/\phi'(S_\phi^*(t))^2$.

Proof The consequence follows from the Delta method (proposition A.1). By the asymptotic representation (7) and proposition 3.2, $\sqrt{n}(\phi(\hat{S}_\phi) - \phi(S_\phi^*))$ goes to $-\mathbb{M}_1 - \int \Gamma(\Psi' \circ \Gamma) \mathbb{M}_2 d\Lambda$, hence $V(t) = V_1(t) - 2V_{12}(t) + V_2(t)$ with the notations of proposition 3.2 and lemma 3.1. \square

Remark 3.2 An error occurs in Rivest & Wells (2001) where the particular case $\Gamma = \pi$ is treated: the authors give $V_1(t) + 2V_{12}(t) + V_2(t)$ instead of $V_1(t) - 2V_{12}(t) + V_2(t)$.

Now we turn to the estimator \hat{F}_ϕ of F introduced in section 2. All our results for \hat{S}_ϕ and S_ϕ^* admit analogous statements for \hat{F}_ϕ and F_ϕ^* , where F_ϕ^* is defined similarly to S_ϕ^* with

$$\Lambda_F(ds) = \frac{\mathbb{P}(U \in ds, U \leq \min(T, V))}{\mathbb{P}(\min(T, U, V) \geq s)} = \frac{\mathbb{P}(U \in ds, U \leq T)}{\mathbb{P}(\min(T, U) \geq s)}$$

instead of Λ . Since

$$\mathbb{P}(T \in ds, T \leq U) + \mathbb{P}(U \in ds, U \leq T) = \mathbb{P}(\min(T, U) \in ds),$$

one has $\Lambda + \Lambda_F = \Lambda_\Gamma$, therefrom we find that $(\phi \circ S_\phi^*) + (\phi \circ F_\phi^*) = (\phi \circ \Gamma)$. By the similar relation (1) between \hat{S}_ϕ and \hat{F}_ϕ , we obtain the following result.

Proposition 3.3 *The process $\sqrt{n}(\phi(\hat{S}_\phi) - \phi(S_\phi^*), \phi(\hat{F}_\phi) - \phi(F_\phi^*))$ converges weakly in $D[0, t_0]$ to the centered continuous Gaussian process*

$$\left(-\mathbb{M}_1 - \int \Gamma(\Psi' \circ \Gamma) \mathbb{M}_2 d\Lambda, \mathbb{M}_1 + \int \Gamma(\Psi' \circ \Gamma) \mathbb{M}_2 d\Lambda - (\Psi \circ \Gamma) \mathbb{M}_2 \right).$$

Proof Using the asymptotic representation (7) and proposition 3.2, this results from the equality

$$\phi(\hat{F}_\phi) - \phi(F_\phi^*) = -(\phi(\hat{S}_\phi) - \phi(S_\phi^*)) + (\phi(\hat{\Gamma}) - \phi(\Gamma))$$

and from the Delta method (proposition A.1). \square

Finally, we will discuss about the Kaplan-Meier estimator \hat{F} of F constructed on the observations $\{y_i, \delta_{U_i}\}_{i=1}^n$. Even when ϕ is known, \hat{F} can be preferable to \hat{F}_ϕ since it jumps at each y_i for those i with $\delta_{U_i} = 1$, whereas \hat{F}_ϕ only jumps at the y_i for those i with $\delta_{U_i} = 1$ and $\delta_{T_i} = 0$ (hence $y_i = x_i$). We will argue that the bivariate process $\sqrt{n}(\hat{S}_\phi - S_\phi^*, \hat{F} - F)$ converges weakly to a continuous Gaussian process on $D[0, t_0]$ for any t_0 such that $\pi(t_0) > 0$. Of course we already know that this is true for the two marginal processes. With the same type of arguments used before, we see that the statement above is true if the trivariate process

$$\bar{\Theta} := \sqrt{n} \begin{pmatrix} n^{-1} \bar{W} \\ \log \hat{\Gamma} - \log \Gamma \\ \log \hat{F} - \log F \end{pmatrix}$$

converges weakly to a continuous Gaussian process. We know that $\log \hat{F} - \log F$ has an asymptotic martingale representation

$$\log \hat{F}(t) - \log F(t) = \int_0^t h(s) d\bar{M}(s) + o_P(n^{-\frac{1}{2}})$$

uniformly in $t \in [0, t_0]$, where \bar{M} is a certain square integrable martingale with respect to the filtration $\bar{\mathcal{F}}$ defined by

$$\bar{\mathcal{F}}_t = \bigvee_{i=1}^n \bigvee_{0 \leq s \leq t} \sigma(\delta_{U_i} \mathbb{1}_{\{y_i \leq s\}}, (1 - \delta_{U_i}) \mathbb{1}_{\{y_i \leq s\}})$$

and h is a measurable function bounded on $[0, t_0]$. Therefore we have an asymptotic representation

$$\bar{\Theta} \approx \sqrt{n} \begin{pmatrix} \int f d\bar{M}_S \\ \int g d\bar{M}_\Gamma \\ \int h d\bar{M} \end{pmatrix},$$

but the square integrable martingales \bar{M}_S , \bar{M}_Γ and \bar{M} are not martingales with respect to the same filtration. But since we know that each of the marginal processes of $\bar{\Theta}$ converges weakly to a centered continuous Gaussian process, then we know that the joint sequence is tight (corollary VI.3.3 in Jacod & Shiryaev, 2003), hence it suffices to check that the finite-dimensional distributions of $\bar{\Theta}$ converge to the same finite-dimensional distributions of a trivariate Gaussian process. This is achievable since each of the martingales \bar{M}_S , \bar{M}_Γ and \bar{M} is the sum of n independent identically distributed terms, each of these terms being constructed from the observations $(x_i, y_i, \delta_{T_i}, \delta_{U_i})$ for one individual i .

4 Discussion

We have derived the large sample properties of the estimator \hat{S}_ϕ defined as the extension of the Rivest & Wells estimator (2001) from the context of right-censored data to the context of right-censored semi-competing risks data. We firstly note that further work is needed to cover the case when $\phi'(1) = 0$, which occurs for the Gumbel copula.

Our approach confirm that martingales techniques are an adequate tool in the theoretical study of survival analysis (see Aalen & al, 2009, for an exposition of the use of martingale methods in survival analysis). We used the natural and helpful notion of immersion of a filtration in a bigger one and we hope that our work will contribute to the popularization of this notion.

A first possible extension of our work is to generalize it to the case when the survival function of Y is possibly improper, with a point mass at $+\infty$ representing the proportion of cured patients. In the simpler context of right-censored data, Li, Tiwari and Guha (2007) generalized the works of Rivest & Wells (2001) to this case.

A second possible extension is to allow the model to incorporate covariates. In the simpler context of right-censored data (when only $\min\{Y, C\}$ is observed and is not censored), a way to do so has been achieved by Braekers & Veraverbeke (2005), and this could be extended to the context of semi-competing risks data. They propose to estimate the survival function of $\min\{Y, C\}$ by a Beran-type estimate $\hat{F}(\cdot | z_i)$ depending on the covariate z_i , and then to define $\hat{S}_\phi(\cdot | z_i)$ similarly to \hat{S}_ϕ by using $\hat{F}(\cdot | z_i)$ instead of the Kaplan-Meier estimator \hat{F} . The large sample properties of the corresponding estimator $\hat{S}_\phi(\cdot | z_i)$ are then derived with the help of empirical processes theory.

Some methods to fit a right-censored semi-competing risks model with an unknown Archimedean copula have been proposed e.g. by Lakhal & al (2008) and Xu & al (2010). The drawback of the Lakhal & al approach is that it only allows for one-parameter Archimedean copulas. The Xu & al model approach adopts a different modeling. Heuchenne & al (2012) studied the performance of another method through simulations. Some of the results of the present paper could help to investigate the theoretical study of the large sample properties of their approach, by using the Chen & al (2003) theory.

Methods allowing for both an unknown copula and the incorporation of covariates also have been developed, most recently by Chen (2012).

A Delta method in the Skorohod space

The goal of this appendix is to prove a version of the Delta method in the Skorohod space $D[0, 1]$ when the limit process is continuous, because we do not find it in the literature. The proof we give is essentially the same as the proof of the Delta method for real random variables. An alternative proof consists in using the powerful functional Delta method (van der Vaart & Wellner 1996) after having said that in this context, the weak convergence in $D[0, 1]$ is equivalent to the weak convergence in the uniform sense, as in the modern theory of empirical processes; this fact results from theorem 1.7.2 in van der Vaart & Wellner (1996) and theorem 6.6 in Billingsley (1999). All our statements are proved for stochastic processes in the Skorohod space $D([0, 1], \mathbb{R}^d)$ which we shortly denote by $D[0, 1]$. Similarly we shortly denote by $C[0, 1]$ the space of continuous functions $C([0, 1], \mathbb{R}^d)$.

Lemma A.1 *Let (X_n) be a sequence of random variables in $D[0, 1]$ which converges weakly to a random variable X in $C[0, 1]$. Then the sequence $(\sup_{0 \leq s \leq 1} \|X_n(s)\|)$ converges weakly to $\sup_{0 \leq s \leq 1} \|X(s)\|$.*

Proof This follows from the continuity of $x \mapsto \sup_{0 \leq s \leq 1} \|x(s)\|$ outside a set with null measure with respect to the law of X . \square

Lemma A.2 *Let (Z_n) be a sequence of random variables in $D[0, 1]$ such that $\sup_{0 \leq t \leq 1} \|Z_n(t)\| = o_P(1)$ and the sequence $(\sqrt{n}Z_n)$ converges weakly to a random variable W in $C[0, 1]$. Let $R: \mathbb{R}^d \rightarrow \mathbb{R}$ be a function such that $R(z) = o(\|z\|)$ when $z \rightarrow 0$. Then $\sup_{0 \leq t \leq 1} |\sqrt{n}R(Z_n(t))| = o_P(1)$.*

Proof One has

$$|\sqrt{n}R(Z_n(t))| = \|\sqrt{n}Z_n(t)\| h(Z_n(t)) \leq \sup_{0 \leq s \leq 1} \|\sqrt{n}Z_n(s)\| h(Z_n(t))$$

where h is the function defined by $h(z) = |R(z)|/\|z\|$ if $z \neq 0$ and $h(0) = 0$, which is continuous at 0. The sequence $(\sup_{0 \leq s \leq 1} \|\sqrt{n}Z_n(s)\|)$ converges weakly to $\sup_{0 \leq s \leq 1} \|W(s)\|$ by lemma A.1, hence, owing to Slutsky's theorem, it suffices to show that $\sup_{0 \leq t \leq 1} h(Z_n(t)) = o_P(1)$, but this follows from the continuity of h at 0. \square

Proposition A.1 (Delta method) *Let (X_n) be a sequence of random variables in $D[0, 1]$ and θ a function belonging to $D[0, 1]$. Assume that $\sup_{0 \leq t \leq 1} \|X_n(t) - \theta(t)\| = o_P(1)$ and that the sequence $\sqrt{n}(X_n - \theta)$ converges weakly in $D[0, 1]$ to a random variable G in $C[0, 1]$. Let $g: \mathbb{R}^d \rightarrow \mathbb{R}^k$ be a differentiable function. Denote by R_μ the remainder in the first order Taylor expansion of g at μ , that is, $R_\mu(y) = g(\mu + y) - g(\mu) - g'(\mu)(y)$. If $\sup_{0 \leq t \leq 1} \|R_{\theta(t)}(y)\| = o(\|y\|)$, then $\sqrt{n}(g(X_n) - g(\theta)) = g'(\theta)\sqrt{n}(X_n - \theta) + o_P(1)$ uniformly on $[0, 1]$, therefore $\sqrt{n}(g(X_n) - g(\theta))$ converges weakly to $g'(\theta)G$ in $D[0, 1]$.*

Proof One has

$$\|\sqrt{n}(g(X_n) - g(\theta)) - g'(\theta)\sqrt{n}(X_n - \theta)\| \leq \sqrt{n}R(X_n - \theta)$$

where $R(y) = \sup_{0 \leq t \leq 1} \|R_{\theta(t)}(y)\|$, hence the result follows from the preceding lemma. \square

Denoting by H the limiting process in the above proposition, its variance function is $\mathbb{E}H(s)H(t)' = D_{\theta(s)}C(s, t)D_{\theta(t)}$ where D_z is the matrix of $g'(z)$ and $C(s, t) = \mathbb{E}G(s)G(t)'$ is the variance function of G .

References

1. Aalen, O.O, Andersen, P.K., Borgan, Ø., Gill, R.D. and Keiding, N. (2009). History of applications of martingales in survival analysis. *Electronic Journal for History of Probability and Statistics*, **5**:1 <http://arxiv.org/abs/1003.0188>
2. Andersen, P.K., Borgan, Ø., Gill, R.D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer, New York.
3. Billingsley, P. *Convergence of probability measures*. Second edition. Wiley, 1999.
4. Bielecki, T.R, Jeanblanc, M. & Rutkowski, M. (2008). *The Annals of Applied Probability*, **18**:6, 2495–2529.
5. Braekers, R. and Veraverbeke, N. (2005). A copula-graphic estimator for the conditional survival function under dependent censoring. *Canad. J. Statist.*, **33**, 429–447.
6. Chen, Y-H (2012). Maximum likelihood analysis of semicompeting risks data with semiparametric regression models. *Lifetime Data Analysis* **18**:1, 36–57.
7. Chen, X., Linton, O. and Van Keilegom, I. (2003). Estimation of semiparametric models when the criterion function is not smooth. *Econometrica*, **71**, 1591–1608.
8. Émery, M. and Schachermayer, W. (2001). On Vershik's standardness criterion and Tsirelson's notion of cosiness. *Séminaire de Probabilités XXXV* (Springer Lectures Notes in Math. 1755, Berlin 2001), 265–305
9. Ethier, S. N. and Kurtz, T. G. (1986). *Markov processes: Characterization and convergence*. John Wiley & Sons, New York, Chichester, Brisbane, Toronto, and Singapore.
10. Fine, J. P., Jiang, H., and Chappell, R. (2001). On semicompeting risks. *Biometrika*, **88**, 907–919.
11. Fleming, T.R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis*. Wiley, New York.
12. P.V. Gapeev & M. Jeanblanc (2008). On filtration immersions and credit events. CDAM Research Report Series, LSE-CDAM-2008-24.
13. Heuchenne, C., Laurent, S., Legrand, C. and I. Van Keilegom. Likelihood based inference for semi-competing risks. Discussion paper DP2011/22 of the Institute of Statistics. To appear in: Communications in Statistics - Simulations and Computations.
14. Jacobsen, M. (1982). *Statistical analysis of counting processes*. Lecture Notes in Statistics, Vol. 12, Springer-Verlag, New York, 1982
15. Jacobsen, M. (1989). Right Censoring and Martingale Methods for Failure Time Data. *Ann. Statist.*, **17**:3, 1133–1156.
16. Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*. Second edition, Springer.
17. Jiang, H., Fine, J. P., Kosorok, R. and Chappell, R. (2005). Pseudo self-consistent estimation of a copula model with informative censoring. *Scandinavian Journal of Statistics*, **33**, 1–20.
18. Kallenberg, O. (1997). *Foundations of Modern Probability*. Springer, Berlin Heidelberg New York.
19. Karatzas, I. and Shreve, S. (1988). *Brownian Motion and Stochastic Calculus*. Springer Verlag, New York.
20. Lakhal, L., Rivest, L-P. and Abdous, B. (2008). Estimating Survival and Association in a Semicompeting Risks Model. *Biometrics*, **64**:1, 180–188
21. Li, Y., Tiwari, R.C. and Guha, S. (2007). Mixture cure survival models with dependent censoring. *J. R. Statist. Soc. B*, **69**:3, 285–306.
22. Mansuy, R. and Yor, M. (2006). *Random Times and Enlargements of Filtrations in a Brownian Setting*. Lecture Notes in Mathematics **1873**, Springer.
23. Nelsen, R.B. (1999). *An introduction to copulas*. Springer-Verlag, New York.
24. Prigent, J-L (2003). *Weak convergence of financial markets*. Springer.
25. Rivest, L. and Wells, M.T. (2001). A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. *J. Multivariate Analysis*, **79**, 138–155.
26. van der Vaart, A.W. and Wellner, J.A. (1996). *Weak convergence and empirical processes*. Springer-Verlag, New York.
27. Xu, J., Kalbfleisch, J.-D. and Tai, B. (2010). Statistical analysis of illness-death processes and semicompeting risks data. *Biometrics*, **66**, 716–725.