



A simplified model for studying bivariate mortality under right-censoring

Svetlana Gribkova, Olivier Lopez, Philippe Saint-Pierre

► To cite this version:

Svetlana Gribkova, Olivier Lopez, Philippe Saint-Pierre. A simplified model for studying bivariate mortality under right-censoring. 2012. hal-00683483

HAL Id: hal-00683483

<https://hal.science/hal-00683483>

Preprint submitted on 28 Mar 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.

A simplified model for studying bivariate mortality under right-censoring

Svetlana GRIBKOVA¹, Olivier LOPEZ², Philippe SAINT PIERRE³

March 28, 2012

Abstract

In this paper, we provide a nonparametric estimator of the distribution of bivariate censored lifetimes, in a model where the two censoring variables differ only through an additional observed variable. This situation is motivated by a particular application to insurance, where the supplementary variable corresponds to the age difference between two individuals. Asymptotic results for our estimator are provided. The new tools that we develop are used to perform goodness-of-fit tests for survival copula models. The practical performance is illustrated through simulations and a real data analysis.

Key words: Survival analysis, nonparametric estimation, Kaplan-Meier estimator, bivariate censoring, mortality analysis, copula models, bootstrap.

Short title: Bivariate mortality under right-censoring.

¹ Laboratoire de Statistique Théorique et Appliquée, Université Pierre et Marie Curie Paris VI, 4 place Jussieu, 75005 Paris, France, E-mail: svetlana.gribkova@etu.upmc.fr

² Laboratoire de Statistique Théorique et Appliquée, Université Pierre et Marie Curie Paris VI, 4 place Jussieu, 75005 Paris, France, E-mail: olivier.lopez0@upmc.fr

³ Laboratoire de Statistique Théorique et Appliquée, Université Pierre et Marie Curie Paris VI, 4 place Jussieu, 75005 Paris, France, E-mail: philippe.saint_pierre@upmc.fr

1 Introduction

In last survivor insurance, an important issue is to infer on the joint distribution of the lifetimes of two individuals linked through an insurance contract, say (T, U) . One of the difficulties in studying such variables comes from the presence of bivariate censoring, with a proportion of censored observations which may be quite high. Therefore, most of the approaches used in this field are parametric (typically parametric survival copula models, see e.g. Shih and Louis (1995)), while nonparametric tools are rarely used, although they would be required at least to assess the validity of the proposed models. The aim of this paper is to provide a new nonparametric estimator of the joint distribution of two lifetimes under bivariate random censoring, in a framework which is adapted to the study of problems coming from the insurance field. A specificity of such problems is that an additional variable is generally present, which carries information on the model, this variable being the age difference between the two individuals under study. Using this information that is often neglected, one can define a quite simple nonparametric estimator of the distribution of the two lifetimes, which is close to Kaplan-Meier estimator (Kaplan and Meier (1958)) and to the estimator of Lin and Ying (1993).

Various approaches have been used to perform nonparametric estimation of multivariate lifetimes. Most of them focus on estimating the survival function, without focusing on the joint distribution itself. Therefore, many of them provide consistent estimators of this function, but fail to define a true distribution. For example, the estimator of the survival function proposed by Campbell and Földes (1982) is not monotonic. The nonparametric maximum likelihood (NPMLE) procedure of Hanley and Parnes (1983) leads to an estimator which is sometimes inconsistent for continuous data (Tsai et al., 1986), while the rate of convergence of a modification of this estimator suggested by Tsai et al. (1986) achieves a slow convergence rate (slower than $n^{1/2}$ where n denotes the sample size). Another NPMLE approach is proposed by van der Laan (1996), introducing some modification of the data and using an interval censoring methodology. Although this estimator is shown to be asymptotically efficient for these modified data, the convergence rate is also slower than $n^{1/2}$. On the other hand, the product-limit type estimator proposed by Dabrowska (1988), which is often used in practice (see e.g. Luciano et al. (2008), Fan et al. (2000), Wang and Wells (2000b), Gill et al. (1995), Prentice and Cai (1992)), assigns negative mass to some points on the plane (Pruitt, 1991b). Nonparametric smoothing techniques have been used by Pruitt (1991a) (but the implicit definition of the estimator leads to difficulties and a weak performance according to van der Laan (1996)), and by Akritas

and Van Keilegom (2003) (this last estimator presenting the advantage to define a true distribution, but, again, with a slower convergence rate, and the necessity of an absolutely continuous censoring variables).

Among the approaches that we mention, each of them suffers either from a too slow convergence rate, or from the fact that the corresponding estimators do not provide true probability distributions. The estimator proposed by Lopez and Saint Pierre (2011) does not present these drawbacks, but relies on an assumption on the joint distribution of the censoring variables which may not be reasonable for the particular application we have in mind. Indeed, a specificity of data-sets coming from last-survivor insurance, is that individuals usually quit the study (for some other cause than death) at the same time. This induces a specific dependence between the two censoring variables involved in the problem. The main idea of the new estimator that we propose consists of using this additional information.

The rest of the paper is organized as follows. In section 2, we present the general censoring framework that we consider. We define a non parametric estimator of the distribution of (T, U) . In section 3, we provide asymptotic results for estimating quantities of the type $E[\phi(T, U)]$ for a large class of functions ϕ (the survival function being only a particular case). A bootstrap procedure is proposed to compute the variance of the error in such estimation problems. Application of our nonparametric estimator to goodness-of-fit for copula models is considered. Section 4 illustrates our result through simulation studies and a real data analysis.

2 A simplified model for bivariate right-censoring

We first present in section 2.1 the general bivariate censoring model that will be considered in the rest of the paper. Estimation of the joint distribution of (T, U) is introduced in section 2.2.

2.1 Bivariate right-censoring

In the following, we consider two lifetimes (T, U) , and i.i.d. replications $(T_i, U_i)_{1 \leq i \leq n}$ of these random variables. In a bivariate right-censoring model, $(T_i, U_i)_{1 \leq i \leq n}$ are not directly observed. Instead, one observes

$$\begin{cases} Y_i &= \inf(T_i, C_i), \text{ and } \delta_i = \mathbf{1}_{T_i \leq C_i} \\ Z_i &= \inf(U_i, D_i), \text{ and } \gamma_i = \mathbf{1}_{U_i \leq D_i}, \end{cases}$$

where $(C_i, D_i)_{1 \leq i \leq n}$ consists of i.i.d. replications of a random bivariate censoring vector (C, D) , and $(\delta_i, \gamma_i)_{1 \leq i \leq n}$ are indicator functions allowing the distinction between censored and uncensored observations.

In many applications, such as last survivor insurance, there exists some clear relationship between the two censoring variables (C, D) . In this particular case, if T (resp. U) denotes the total lifetime of the husband (resp. his wife), C (resp. D) denotes the age at which the husband (resp. the wife) stops being under observation for any other cause than death.

Usually, censoring causes are twofold: the end of the observation period (if the person is not dead at this time, his/her lifetime is not observed), or the surrender of the contract. In both situations, one can observe that, for many cases, this event which stops observation occurs at the same time for both members of the couple. Taking the example of a pension contract with a reversion clause, one can see that surrendering the contract will automatically remove the two members of the couple from the database (unless one of them possesses additional contracts that could allow the company to keep some track on him/her, which is usually not the case due to the complexity of such a tracking process). If ε denotes the age difference between the two members of the couple, then $D = C + \varepsilon$. Moreover, the variables $(\varepsilon_i)_{1 \leq i \leq n}$ are observed for all couples.

To summarize, in such a framework, observations are made of $(Y_i, Z_i, \varepsilon_i, \delta_i, \gamma_i)_{1 \leq i \leq n}$, where the random variables ε_i represents some age difference between the two persons observed. We now state some identifiability assumptions that will hold throughout this paper.

Assumption 1 *Assume that*

1. $D = C + \varepsilon$.
2. (T, U) is independent from ε , and from C , and $\mathbb{P}(T = C) = \mathbb{P}(U = C + \varepsilon) = 0$.
3. C is independent from ε .

In Assumption 1, points 2 and 3 are a direct multivariate extension of the classical identifiability assumption required to ensure consistency of Kaplan-Meier estimator in the univariate case (see Stute and Wang (1993)). In section 2.2 below, we show how one can estimate nonparametrically the distribution of (T, U) under Assumption 1.

2.2 Nonparametric estimation of the distribution of (T, U)

In this section, we show how to estimate quantities of the type $E[\phi(T, U)]$ for some function ϕ . A particular case is the joint survival function $S_F(t, u) = \mathbb{P}(T > t, U > u)$, which corresponds to $\phi(T, U) = \mathbf{1}_{T > t, U > u}$. In absence of censoring, the answer to this estimation problem consists of using empirical means. Defining $F(t, u) = \mathbb{P}(T \leq t, U \leq u)$, one can rewrite $E[\phi(T, U)] = \int \phi(t, u) dF(t, u)$, which can be consistently estimated by $\int \phi(t, u) dF_{emp}(t, u)$, where $F_{emp}(t, u) = n^{-1} \sum_{i=1}^n \mathbf{1}_{T_i \leq t, U_i \leq u}$ denotes the empirical distribution function. In our framework, the empirical distribution is unfortunately unavailable, since $(T_i, U_i)_{1 \leq i \leq n}$ are not directly observed.

We propose to rely on an estimator of the type

$$\hat{F}(t, u) = \sum_{i=1}^n \delta_i \gamma_i W_n(Y_i, Z_i, \varepsilon_i) \mathbf{1}_{Y_i \leq t, Z_i \leq u}, \quad (2.1)$$

to generalize the empirical distribution function to our framework. Using such type of estimators, one can straightforwardly define an estimator of $E[\phi(T, U)] = \int \phi(t, u) dF(t, u)$ by

$$\int \phi(t, u) d\hat{F}(t, u) = \sum_{i=1}^n \delta_i \gamma_i W_n(Y_i, Z_i, \varepsilon_i) \phi(Y_i, Z_i). \quad (2.2)$$

The idea is similar to Lopez (2012) and Lopez and Saint Pierre (2011): instead of assigning the same n^{-1} -mass to each observation (as it is the case when considering the empirical distribution function), one assigns mass at doubly uncensored observations (since only these observations are completely relevant to understand the dependence structure between T and U), while the mass $W_n(Y_i, Z_i, \varepsilon_i)$ is designed to compensate censoring.

Defining $S_G(t) = \mathbb{P}(C > t)$, (with $G(t) = \mathbb{P}(C \leq t)$) and $F^*(t, u) = \sum_{i=1}^n \delta_i \gamma_i W^*(Y_i, Z_i, \varepsilon_i) \mathbf{1}_{Y_i \leq t, Z_i \leq u}$, with $W^*(y, z, e) = n^{-1} S_G(\max(y, z - e) -)^{-1}$, one can observe that $\int \phi(t, u) dF^*(t, u)$ is an unbiased estimator of $E[\phi(T, U)]$ under Assumption 1. Indeed, for any function ψ with finite expectation, we have, under Assumption 1,

$$\begin{aligned} E[\delta \gamma \psi(Y, Z)] &= E \left[E \left[\mathbf{1}_{\max(T, U - \varepsilon) \leq C} | T, U, \varepsilon \right] \psi(T, U) \right] \\ &= E \left[S_G(\max(T, U - \varepsilon) -) \psi(T, U) \right]. \end{aligned}$$

Unfortunately, this ideal estimator F^* can not be computed in practice, since it relies on the unknown survival function S_G . Nevertheless, it is possible to estimate this function S_G . Define $\eta_i = 1 - \delta_i \gamma_i$, and $A_i = \max(T_i, U_i - \varepsilon_i)$. The variable C_i is observed as long

as $C_i < A_i$ (that is $\eta_i = 1$). Hence, C can also be considered as a right-censored variable (censored by variables A), provided that the event $\{C_i = A_i\}$ has probability 0. This is actually the case, from point 2 in Assumption 1, where we assumed that $\mathbb{P}(T = C) = \mathbb{P}(U - \varepsilon = C) = 0$. Therefore, the distribution of C can be estimated by a Kaplan-Meier estimator based on the censored sample $(B_i, \eta_i)_{1 \leq i \leq n}$, where $B_i = \inf(C_i, A_i)$. Moreover, it is important to notice that Assumption 1 ensures consistency of Kaplan-Meier estimator, that is, following Stute and Wang (1993), $\mathbb{P}(A = C) = 0$ and A independent from C .

Therefore, defining the Kaplan-Meier estimator $\hat{S}_G$ of S_G , that is,

$$\hat{S}_G(t) = \prod_{k: B_k \leq t} \left(1 - \frac{d\hat{H}_0(B_k)}{\hat{H}(B_k)} \right),$$

where $\hat{H}_0(t) = n^{-1} \sum_{i=1}^n \eta_i \mathbf{1}_{B_i \leq t}$, and $\hat{H}(t) = n^{-1} \sum_{i=1}^n \mathbf{1}_{B_i \geq t}$, a natural choice of a function W_n in (2.1) is

$$W_n(y, z, e) = \frac{1}{n\hat{S}_G(\max(y, z - e) -)}. \quad (2.3)$$

This estimator is close to the estimator proposed by Lin and Ying (1993). The difference, in our approach, is the presence of the additional random variable ε_i corresponding to the age difference.

3 Asymptotic theory

The present section is devoted to the asymptotic results on the nonparametric estimator defined in section 2.2. A Central Limit Theorem for (2.2) is provided in section 3.1. As a corollary of this result, we deduce asymptotic convergence results for estimating Kendall's τ coefficient, which is a classical dependence measure. Section 3.2 provides a bootstrap procedure in order to investigate the estimation error. In section 3.3, we derive theoretical results that may be used to perform goodness-of-fit tests for survival copula models.

3.1 An asymptotic representation Theorem for estimator (2.1)

We aim to obtain an asymptotic representation for quantities of the type (2.1). Instead of considering a single function ϕ , we focus on obtaining results that hold uniformly for functions $\phi \in \mathcal{F}$, $\mathcal{F}$ denoting a class of functions. This uniformity result is required if we wish to obtain, for example, uniform consistency results for the estimation of the distribution function. Considering this problem, the natural class of functions to be

considered is $\mathcal{F}_1 = \{(t, u) \rightarrow \mathbf{1}_{t \leq x, u \leq y} : x \in \mathcal{T}, y \in \mathcal{U}\}$, where $\mathcal{T}$ and $\mathcal{U}$ denote the support of the distribution of each marginal.

In the following, we consider a class of functions $\mathcal{F}$, with envelope Φ , satisfying Assumptions 2 and 3 below. Since our proof will rely on empirical processes theory, Assumption 2 consists of assuming that a class of functions related to $\mathcal{F}$ is Donsker, that is a class with an uniform central limit theorem property (see van der Vaart and Wellner (1996) for a precise definition of Donsker classes).

Assumption 2 *Let $\mathcal{G}$ denote the class of positive, monotonic functions bounded by 1, and $\chi(T, U, C, \varepsilon) = \delta\gamma S_G(\max(T, U - \varepsilon) -)^{-2}$. For any (t_0, u_0) in $\mathbb{R}^2$ such that $S_F(t_0, u_0) > 0$, define*

$$\mathcal{H}_{t_0, u_0} = \{(T, U, C, \varepsilon) \rightarrow \chi(T, U, C, \varepsilon) f(T, U) g(\max(T, U - \varepsilon) -) \mathbf{1}_{T \leq t_0, U \leq u_0}, f \in \mathcal{F}, g \in \mathcal{G}\},$$

and assume that $\mathcal{H}_{t_0, u_0}$ is a Donsker class of functions.

Assumption 3 is required only if we wish to obtain consistency on the whole support of (T, U) . It automatically holds if one considers bounded functions with compact support strictly included in the support of the distribution. It can be understood as an assumption on the tail of the distributions of T and U . Similar assumptions have been used in Gill (1983), Stute (1996), or Lopez and Saint Pierre (2011).

Assumption 3 *Assume that $E[\Phi(T, U)^2 S_G(\max(T, U - \varepsilon) -)^{-1}] < \infty$. Moreover, define*

$$\mathcal{C}(y) = \int_{-\infty}^y \frac{dG(t)}{[1 - F(t)][1 - G(t-)]^2},$$

and assume that $E[\Phi(T, U) \mathcal{C}^{1/2+v}(\max(T, U - \varepsilon) -) S_G(\max(T, U - \varepsilon) -)^{-1}] < \infty$, for some $v > 0$ (arbitrary small).

If we consider the particular case of $\mathcal{F}_1$, Assumption 2 automatically holds, provided that the moment conditions of Assumption 3 hold. Generally, this will also be the case for parametric class of functions, or sufficiently smooth classes of functions using permanence properties of Donsker classes.

We now state the main theoretical result of this section.

Theorem 3.1 *Recall that $B_i = \inf(A_i, C_i)$, where $A_i = \max(T_i, U_i - \varepsilon_i)$, and that $\eta_i = \mathbf{1}_{C_i \leq A_i}$, and let $F_A(t) = \mathbb{P}(A \leq t)$ and $H(t) = \mathbb{P}(B > t)$. Under Assumptions 1 to 3,*

$$\int \phi(t, u) d(\hat{F} - F^*)(t, u) = \frac{1}{n} \sum_{i=1}^n \psi_\phi(T_i, U_i, \varepsilon_i) + R_n(\phi), \quad (3.1)$$

where $\sup_{\phi \in \mathcal{F}} |R_n(\phi)| = o_P(n^{-1/2})$, and

$$\begin{aligned} \psi_\phi(Y_i, Z_i, \varepsilon_i) &= \int \left\{ \frac{\eta_i S_G(B_i \vee a)}{H(B_i)} - \int \frac{\mathbf{1}_{B_i \geq u} S_G(u \vee y) dF_A(u)}{H(u) F_A(u)} \right. \\ &\quad \left. + \frac{(1 - \eta_i) \mathbf{1}_{B_i > y}}{F_A(B_i)} - S_G(y) \right\} \frac{\phi(t, u) d\mathbb{P}_{(T, U, C, \varepsilon)}(t, u, c, e)}{S_G(a-)}, \end{aligned}$$

where we used $a = \max(t, u - e)$ to shorten the notation, and where $\mathbb{P}_{(T, U, C, \varepsilon)}$ denotes the true law of (T, U, C, ε) . As a consequence, since $E[\psi_\phi(Y, Z, \varepsilon)] = 0$, we have, for all $\phi \in \mathcal{F}$,

$$n^{1/2} \left(\int \phi(t, u) d\hat{F}(t, u) - E[\phi(T, U)] \right) \Rightarrow \mathcal{N}(0, \sigma_\phi^2), \quad (3.2)$$

with

$$\sigma_\phi^2 = E \left[\left\{ \frac{\delta_i \gamma_i \phi(Y_i, Z_i)}{S_G(B_i -)} - E[\phi(T, U)] + \psi_\phi(Y_i, Z_i, \varepsilon_i) \right\}^2 \right],$$

and $\Rightarrow$ denotes the weak convergence.

The proof of this result is postponed to the Appendix section. Equation (3.2) can be used to compute asymptotic confidence intervals, provided that one is able to consistently estimate the asymptotic variance σ_ϕ^2 . This can be done by replacing all the unknown distributions functions involved in the expression of σ_ϕ^2 by their empirical counterparts. However, this approximation may be too rough in practice, and bootstrap procedures seem to be more appropriate if one wishes to compute confidence interval. This bootstrap method is shown in section 3.2.

In addition to the application of Theorem 3.1 to the class $\mathcal{F}_1$ (corresponding to the estimation of the joint survival function), we show how this result may be used to provide asymptotic results for the estimation of Kendall's τ coefficient. Kendall's τ coefficient is a classical dependence measure which can be defined in the following way. For two random variables (T, U) , $\tau = \mathbb{P}((T_1 - U_1)(T_2 - U_2) > 0) - \mathbb{P}((T_1 - U_1)(T_2 - U_2) < 0)$, where (T_1, U_1) and (T_2, U_2) are independent replications of (T, U) . There exists a relationship between τ and the distribution function F , that is $\tau = 4 \int F(x, y) dF(x, y) - 1$, see e.g. Nelsen (2006). Therefore, a natural estimator of τ is

$$\hat{\tau} = 4 \int \hat{F}(x, y) d\hat{F}(x, y) - 1. \quad (3.3)$$

As it is shown in Wang and Wells (2000a), censoring may cause this estimator not to be consistent in some particular situations. Indeed, defining $S_H(y, z) = \mathbb{P}(Y > y, Z > z)$ the survival function of the observed times, $\mathcal{S}_1 = \{(t, u) : S_H(t, u) > 0\}$, and $\mathcal{S}_2 = \{(t, u) :$

$S_F(t, u) > 0\}$, we can see that some part of the distribution, namely $\mathcal{S}_2 - \mathcal{S}_1$ is never observed, since the corresponding observations are always censored. If this difference of sets is empty, this does not introduce bias in the estimation of τ . In other situations, some bias will arise, which can be evaluated according to the method of Wang and Wells (2000a). Corollary 3.2 below shows that this estimator admits an asymptotic representation.

Corollary 3.2 *Let ψ_F denote function ψ as defined in Theorem 3.1 for function $\phi = F$. To shorten the notation, we will denote $\psi_{t,u}$ the function corresponding to $\phi(Y, Z) = \mathbf{1}_{Y \leq t, Z \leq u}$. Assume that $\mathcal{S}_2 - \mathcal{S}_1 = \emptyset$. Then,*

$$\begin{aligned} \hat{\tau} - \tau &= 4 \left\{ \int F(t, u) d[F^* - F](t, u) + \int [F^*(t, u) - F(t, u)] dF(t, u) \right. \\ &\quad \left. + \frac{1}{n} \sum_{i=1}^n \left\{ \psi_F(Y_i, Z_i, \varepsilon_i) + \int \psi_{t,u}(Y_i, Z_i, \varepsilon_i) dF(t, u) \right\} \right\} + o_P(n^{-1/2}). \end{aligned} \quad (3.4)$$

In the representation of Corollary (3.2), each term is a sum of i.i.d. quantities with zero expectation and finite variance. Therefore, Corollary (3.2) shows that $\hat{\tau}$ is asymptotically Gaussian. Its asymptotic variance (which as a complex form) can be deduced from this representation. Nevertheless, we do not emphasize this variance, since we recommend using bootstrap procedures to investigate the law of $\hat{\tau}$ (see section 4.1).

Let us also mention that, if the assumption $\mathcal{S}_2 - \mathcal{S}_1 = \emptyset$ does not hold, the result is still true, but with τ replaced by $4 \int_{\mathcal{S}_1} F(x, y) dF(x, y) - 1$.

Proof of Corollary 3.2. Write

$$\begin{aligned} \hat{\tau} - \tau &= 4 \left\{ \int F(t, u) d[\hat{F} - F](t, u) + \int [\hat{F}(t, u) - F(t, u)] dF(t, u) \right. \\ &\quad \left. + \int [\hat{F}(t, u) - F(t, u)] d[\hat{F} - F](t, u) \right\}. \end{aligned} \quad (3.5)$$

Applying Theorem 3.1 to function F , the first term of (3.5) can be expanded as

$$\int F(t, u) d[F^* - F](t, u) + \frac{1}{n} \sum_{i=1}^n \psi_F(Y_i, Z_i, \varepsilon_i) + o_P(n^{-1/2}).$$

Moreover, again from Theorem 3.1,

$$\begin{aligned} \int [\hat{F}(t, u) - F(t, u)] dF(t, u) &= \int [F^*(t, u) - F(t, u)] dF(t, u) \\ &\quad + \frac{1}{n} \sum_{i=1}^n \int \psi_{t,u}(Y_i, Z_i, \varepsilon_i) dF(t, u) + o_P(n^{-1/2}). \end{aligned}$$

The third term of (3.5) can be rewritten as

$$\begin{aligned} & \int [F^*(t, u) - F(t, u)] d[F^* - F](t, u) + \int [\hat{F}(t, u) - F^*(t, u)] d[F^* - F](t, u) \\ & + \int [\hat{F}(t, u) - F^*(t, u)] d[\hat{F} - F^*](t, u) := \mathcal{T}_1 + \mathcal{T}_2 + \mathcal{T}_3. \end{aligned}$$

The term $\mathcal{T}_1$ is a second order degenerate U -statistics and is therefore of order $O_P(n^{-1})$. To study $\mathcal{T}_2$, apply Theorem 3.1 to the class of indicator functions $\mathbf{1}_{T \leq t, U \leq u}$ to obtain

$$\mathcal{T}_2 = \int \psi(T_i, U_i, \varepsilon_i) d[F^* - F](t, u) + o_P(n^{-1/2}).$$

The integral in this decomposition is zero, since $E[\psi_\phi(T, U, \varepsilon)] = 0$, and since $\int \phi d[F^* - F](t, u) = 0$. Finally, observe that $\mathcal{T}_3$ can be rewritten as

$$\mathcal{T}_3 = \frac{1}{n} \sum_{i=1}^n \frac{\delta_i \gamma_i [\hat{F}(Y_i, Z_i) - F^*(Y_i, Z_i)] [\hat{S}_G(B_i-) - S_G(B_i-)]}{S_G(B_i-) \hat{S}_G(B_i-)},$$

which is bounded by

$$|\mathcal{T}_3| \leq \sup_{t, u} |\hat{F}(t, u) - F^*(t, u)| \sup_b |\hat{S}_G(b) - S_G(b)| \sup_b \left| \frac{S_G(b)}{\hat{S}_G(b)} \right| \times \frac{1}{n} \sum_{i=1}^n \frac{\eta_i}{S_G(B_i-)^2}.$$

The first supremum is $O_P(n^{-1/2})$ from Theorem 3.1, the second one is $o_P(1)$ from the uniform consistency of Kaplan-Meier estimator (see Stute and Wang (1993)), while the third one is $O_P(1)$ (see Gill (1983)). Moreover, the empirical mean on the right-hand side is $O_P(1)$ provided that each term of the sum has finite expectation, which is the case from Assumption 3. Combining these facts leads to $\mathcal{T}_3 = o_P(n^{-1/2})$ and concludes the proof. ■

3.2 Bootstrap procedure

As we already mentioned, the asymptotic results of Theorem 3.1 may be difficult to use when it comes to approximate the law of $\int \phi(t, u) d\hat{F}(t, u)$. The problem comes from the complex form of the asymptotic variance, and therefore from the difficulty to estimate it in an accurate way. Therefore, the aim of the present section is to propose a bootstrap procedure that allows to circumvent this problem.

Under univariate censoring, two main methodologies have been proposed in the literature to perform bootstrap, see Efron (1981) and Reid (1981). The methodology of Efron (1981) consists of using the nonparametric estimators of the distribution of the lifetime

and of the censoring to resimulate samples. Akritas (1986) showed that only Efron's methodology was consistent. We therefore propose to adapt this strategy.

The basic idea consists of simulating variables (T, U) according to the estimated distribution defined by $\hat{F}$ (and renormalized in order to ensure that the total mass is equal to one). The censoring can be simulated similarly using $\hat{G}$, while ε is simulated according to its empirical distribution $\hat{F}_\varepsilon(t) = n^{-1} \sum_{i=1}^n \mathbf{1}_{\varepsilon_i \leq t}$. The procedure is summarized below.

To compute B bootstrap n -samples, repeat for $b = 1, \dots, B$ the following simulation scheme,

1. Simulate independent variables $(T_i^b, U_i^b)_{1 \leq i \leq n}$ under the probability distribution $\hat{F}/\hat{F}(\mathbb{R}^2)$.
2. Simulate independent variables $(\varepsilon_i^b)_{1 \leq i \leq n}$ under the probability distribution $\hat{F}_\varepsilon$.
3. Simulate independent variables $(C_i^b)_{1 \leq i \leq n}$ under the probability distribution $\hat{G}/\hat{G}(\mathbb{R}^2)$.
4. The b -th bootstrap sample is composed of $(Y_i^b, Z_i^b, \delta_i^b, \gamma_i^b, \varepsilon_i^b)_{1 \leq i \leq n}$, where $Y_i^b = \inf(T_i^b, C_i^b)$, $Z_i^b = \inf(U_i^b, C_i^b + \varepsilon_i^b)$, $\delta_i^b = \mathbf{1}_{T_i^b \leq C_i^b}$, $\gamma_i^b = \mathbf{1}_{U_i^b \leq C_i^b + \varepsilon_i^b}$.

3.3 Application to survival copula inference

Survival copula models are a common tool to model dependence between two lifetimes (T, U) . Indeed, the bivariate survival function $S_F(t, u) = \mathbb{P}(T > t, U > u)$ of the random vector (T, U) admits, by Sklar's Theorem (Sklar (1959)), a copula representation, that is

$$S_F(t, u) = \mathfrak{C}(S_T(t), S_U(u)),$$

where $S_T(t) = \mathbb{P}(T > t)$ and $S_U(u) = \mathbb{P}(U > u)$, and where $\mathfrak{C}$ is a survival copula function (see e.g. Nelsen (2006)). To understand the dependence between T and U , which is represented by the copula function $\mathfrak{C}$, it is natural to search for an estimator of $\mathfrak{C}$, usually based on a parametric or semiparametric model, see e.g. Shih and Louis (1995). Nonparametric inference is then required to assess the validity of the model.

Wang and Wells (2000b) proposed to extend the methodology of Genest and Rivest (1993) in presence of censoring. This approach relies on the estimation of the function $v \rightarrow K(v) = \mathbb{P}(S_F(T, U) \leq v)$. If we consider the particular case of Archimedean copula families (that is copulas defined as $\mathfrak{C}(u, v) = \phi^{-1}(\phi(u) + \phi(v))$ where the generator ϕ is a convex function satisfying the conditions of Theorem 4.3.4 in Nelsen (2006)), there exists a one-to-one correspondance between the generator ϕ and the function K , through the

relationship

$$K(v) = v - \frac{\phi(v)}{\phi'(v)}. \quad (3.6)$$

The basic idea of goodness-of-fit procedures based on K consists of comparing a parametric estimator, (based on an estimator $\phi_{\hat{\theta}}$ depending on the parametric model and on $\hat{\theta}$, the association parameter estimated from the data) to a nonparametric one. Wang and Wells (2000b) used an estimator based on the nonparametric estimator of Dabrowska (1988). The nonparametric estimator that we propose to use is defined as

$$\hat{K}(v) = \int \mathbf{1}_{\hat{S}_F(t,u) \leq v} d\hat{F}(t, u). \quad (3.7)$$

In Proposition 3.3, we show that the process $n^{1/2}(\hat{K}(v) - K(v))$ converges towards a Gaussian process. This kind of result is essential to legitimate goodness-of-fit techniques that will be fully discussed in section 4.2. Nevertheless, we do not focus on the estimation of the asymptotic covariance process. In practice, since its computation seems rather delicate, it is preferable to rely on bootstrap procedure (see section 4.2).

Proposition 3.3 *Assume that:*

1. *The distribution function $K(v)$ admits a continuous bounded derivative $k(v)$.*
2. *Given $S_F(t, u) = v$, there exists a version of the conditional distribution of (Y, Z) and a countable family $\mathcal{P}$ of partitions $\mathcal{E}$ on $\mathcal{I}$ (where $\mathcal{I}$ denotes the support of (T, U)) into a finite number of Borel sets satisfying $\inf_{\mathcal{E} \in \mathcal{P}} \max_{E \in \mathcal{E}} \text{diam}(E) = 0$, such that, for all $E \in \mathcal{E}$, the mapping $v \rightarrow \mu_v(E) = k(v)\mathbb{P}((T, U) \in E | S(T, U) = v)$ is continuous.*

Then, there exists a zero-mean Gaussian process such that,

$$n^{1/2} \left(\hat{K}(\cdot) - K(\cdot) \right) \Longrightarrow - \int \int \mathbf{1}_{S_F(t,u) > \cdot} dW(t, u) - \int \int W(t, u) d\mu(\cdot, t, u).$$

Proof. From Theorem 3.1, one can deduce $n^{1/2}(\hat{S}_F(t, u) - S_F(t, u)) \Longrightarrow W(t, u)$, where W is a Gaussian process with mean zero. Consequently, Theorem 1 in Wang and Wells (2000b) applies. ■

4 Simulations and real data example

In this section, we investigate the finite sample size behaviour of our procedure. This investigation is done through simulation studies and illustrated on a real data example.

The data that we consider has been initially studied by Frees et al. (1996), and was studied by Carriere (2000), Youn and Shemyakin (1999), Youn and Shemyakin (2001) and Luciano et al. (2008). We refer to Frees et al. (1996) for a more detailed description of this dataset, containing lifetimes of two members of a couple who subscribed an insurance contract. The dataset concerns 14947 contracts from a large Canadian insurer, observed between December 29th, 1988 and December 31th, 1993¹. After elimination of same-sex contracts and of couples with more than one policies (for which we only keep one policy), 11454 contracts remain. In addition to bivariate censoring, observations are subject to left truncation. Nevertheless, we do not consider left-truncation in the approach that we develop in the present paper. Neglecting left-truncation will lead to a slight over-estimation of the lifetimes, which, from the prospective of an insurer who wishes to evaluate his liabilities in the case of a pension contract, represents a cautious approach.

In section 4.1, we discuss the problem of estimating Kendall's τ coefficient, illustrating the theoretical results of Corollary 3.2. In section 4.2, we study the practical implementation of a goodness-of-fit procedure for copula models, based on the process $\hat{K}$ defined in (3.7).

4.1 Nonparametric estimation of Kendall's τ coefficient

Real data example. Using $\hat{\tau}$ defined in equation (3.3), we find an estimated value of Kendall's τ coefficient which is $\hat{\tau} = 0.6696$, which is roughly of the same order as the values obtained by other authors on the same data-set (for example, for a specific generation, Luciano et al. (2008) obtained an estimation which is 0.6039). We used the nonparametric bootstrap procedure described in section 3.2 to approximate the law of $\hat{\tau}$. Through $B = 1000$ bootstrap replications, we obtain an estimation of the distribution of $\hat{\tau}$ which is represented in Figure 4.1 below.

¹The author wishes to thank the Society of Actuaries, through the courtesy of Edward J. Frees and Emiliano Valdez, for allowing use of the data in this paper

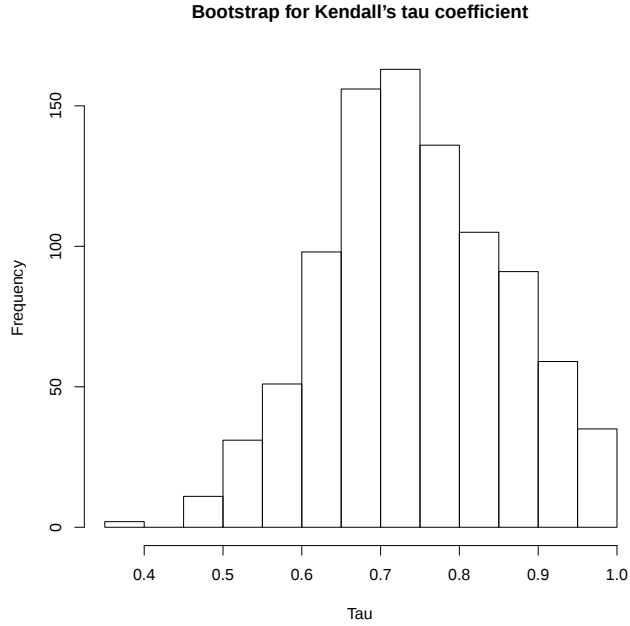


Figure 1: Histogram of the distribution of $\hat{\tau}$ using the bootstrap procedure.

We can observe on Figure 4.1 that the distribution of $\hat{\tau}$ obtained using the bootstrap procedure does not seem to be Gaussian. Therefore, it legitimates to rely rather on this bootstrap procedure than on normal approximation to investigate uncertainty in estimating τ .

Simulation study. To illustrate the convergence of $\hat{\tau}$, we present some results of a simulation study. The random lifetimes (T, U) are simulated from a Clayton copula model (see Table 2 in section 4.2 for a precise definition) with association parameter $\theta = 2$ (which corresponds to a value $\tau = 0.5$), with marginals following a Weibull distribution. Weibull distribution is parametrized through a shape parameter α and a scale parameter β , and admits a density

$$f(t) = \frac{\alpha}{\beta} \left(\frac{t}{\beta} \right)^{\alpha-1} \exp \left(-\frac{t^\alpha}{\beta^\alpha} \right),$$

for $t \geq 0$. We consider the case $\alpha = 10$ and $\beta = 1.7$. The censoring variables C are simulated according to an exponential distribution with parameter λ (with mean λ^{-1}). Different values of λ are considered in order to change the average proportion of doubly uncensored observations. Random variables ε_i are simulated according to an exponential distribution with parameter $\mu = 50$.

For each considered value of the parameter λ , we generate n -samples for different

values of n . We repeat $N = 1000$ times the simulation scheme in order to estimate the bias $E[\hat{\tau} - \tau]$, the variance $Var(\hat{\tau})$, and the mean-squared error $E[(\hat{\tau} - \tau)^2]$. Results are presented in Table 1 below.

Model	Criterion	n=1000	n=2000
$\alpha = 10$ $\beta = 1.1$ (35% of uncensored)	MSE =	0.004537	0.002548
	Bias =	0.06686	0.05020
	Variance =	6.6984e-5	2.8133e-5
$\alpha = 10$ $\beta = 1.7$ (20% of uncensored)	MSE =	0.006949	0.004482
	Bias =	0.08275	0.06650
	Variance =	1.020e-4	5.9425e-5

Table 1: Estimation of the mean-squared error and related quantities for the estimation of Kendall's τ coefficient.

4.2 Goodness-of-fit for semiparametric copula models

A goodness-of-fit procedure based on $\hat{K}$. Consider a parametric family of Archimedean survival copulas $\mathcal{F}_C = \{\mathfrak{C}_\theta : \theta \in \Theta\}$. We will denote ϕ_θ the Archimedean generator of copula $\mathfrak{C}_\theta$. We describe how to extend the procedure proposed by Genest and Rivest (1993) to test

$$H_0 : \mathfrak{C} \in \mathcal{F}_C,$$

against

$$H_1 : \mathfrak{C} \notin \mathcal{F}_C.$$

The principle of the test consists of computing an estimator $\hat{\theta}$ (assuming that H_0 holds) from the data, and then use $\phi_{\hat{\theta}}$ and (3.6) to compute a parametric estimator $K_{\hat{\theta}}$ of function K . Next, considering some distance d between curves, the test statistic is $\mathcal{T}_n = d(\hat{K}, \hat{K}_{\hat{\theta}})$, where $\hat{K}$ is defined in (3.7). H_0 is rejected when $\mathcal{T}_n > s_\alpha$, where s_α is a critical value that ensures that the procedure achieves level α . In the following, we will consider the particular case $d(\hat{K}, K_\theta) = [\int_0^1 (\hat{K}(v) - K_\theta(v))^2 dv]^{1/2}$.

To estimate $\hat{\theta}$, one can either rely on a semiparametric maximum likelihood procedure, as it is done in Shih and Louis (1995), or take $\hat{\theta} = \arg \min_{\theta \in \Theta} d(\hat{K}, K_\theta)$, which has been done in Luciano et al. (2008) and seems more natural in our framework. Therefore, we will use this second approach, and our test statistic may be rewritten as $\mathcal{T}_n = \min_{\theta \in \Theta} d(\hat{K}, K_\theta)$.

To compute the critical values, a bootstrap procedure is required. In our framework, Wang and Wells (2000b) proposed a bootstrap methodology, which has been shown to fail to be consistent by Genest et al. (2006). Therefore, we prefer to adapt the consistent resampling plan defined in Genest et al. (2006) to the presence of censoring. This results on using the bootstrap procedure defined in section 3.2, but replacing Step 1 by

- 1'. Simulate independent variables $(T_i^b, U_i^b)_{1 \leq i \leq n}$ under the distribution defined by $\mathfrak{C}_{\hat{\theta}}$ and with marginal distributions defined by the Kaplan-Meier estimators (univariate) of T and U ,

which corresponds to an approximation of the law of (T, U) under H_0 . Alternatively, in a full parametric modelling of the distribution of (T, U) , parametric distributions may be used instead of the nonparametric Kaplan-Meier estimators. Based on B bootstrap replications of $\mathcal{T}_n$, the critical value s_α can be determined in order to ensure a level α of the procedure.

Real data example. We consider three copula models that have been used by Luciano et al. (2008) to study the mortality of a particular generation in the data-set (in the present paper, we do not distinguish between generations). These models are Clayton, Frank copula models, and a copula called Nelsen 4.2.20 (corresponding to the copula defined in formula 4.2.20 in Nelsen (2006)). Definition of these three Archimedean families is recalled in Table 2 below. Estimators $\hat{\theta}$ are computed by minimization of the distance $d(\hat{K}, K_\theta)$.

Model	$\phi_\theta(t)$	$\mathfrak{C}_\theta(u, v)$	$\hat{\theta}$
Clayton	$\theta^{-1}(t^{-\theta} - 1)$	$(u^{-\theta} + v^{-\theta} + 1)^{-1/\theta}$	4.8991
Frank	$-\log\left(\frac{\exp(-\theta t)-1}{\exp(-\theta)-1}\right)$	$-\theta^{-1} \log\left(1 + \frac{(\exp(-\theta u)-1)(\exp(-\theta v)-1)}{(\exp(-\theta)-1)}\right)$	11.4115
4.2.20	$\exp(t^{-\theta}) - e$	$\log(\exp(u^{-\theta}) + \exp(v^{-\theta}) - e)^{-1/\theta}$	1.338

Table 2: Expression of the different copula families considered. The column $\hat{\theta}$ presents the estimated association parameter on the data-set, minimizing distance d .

The graphical comparison between $\hat{K}$ and $K_{\hat{\theta}}$ is presented in Figure 2 below. Table 3 presents the results of the test procedure described above, comparing the value of the test statistic $\mathcal{T}_n$ to the quantiles of the distribution of $\mathcal{T}_n$ under the null hypothesis (computed using our bootstrap procedure).

The model with the smallest value of the test-statistic is Nelsen's 4.2.20 copula model. Frank's copula model achieves a value of the test-statistic which is quite close from Nelsen's

Model	Test statistic	95% quantile	97.5 % quantile	99 % quantile	p-value
Clayton	0.06229	0.16638	0.17600	0.19362	0.533
Frank	0.05434	0.04330	0.04667	0.05203	0.008
Nelsen 4.2.20	0.05181	0.12243	0.12245	0.12246	0.492

Table 3: Goodness-of-fit procedure for a the three survival copula models considered (Clayton, Frank, Nelsen 4.2.20).

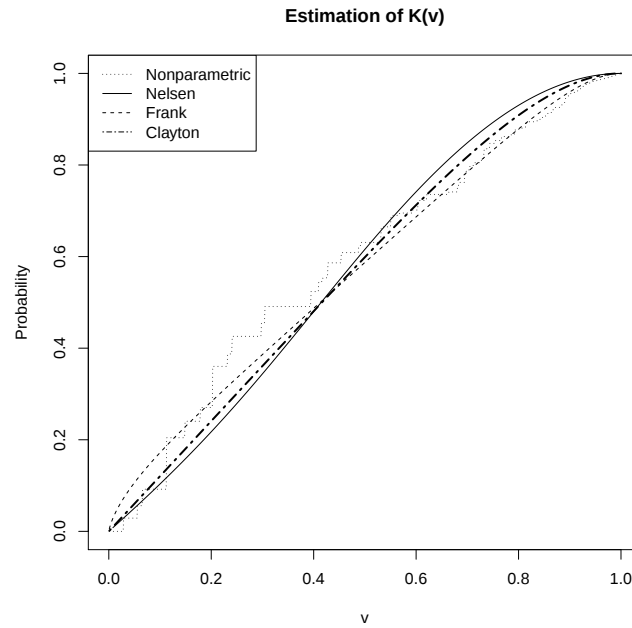


Figure 2: Graphical comparison between $\hat{K}$ and $K_{\hat{\theta}}$ for the three copula models considered.

4.2.20 case. However, in this last model, the corresponding p -value is small, while it is not the case for the two other models. Graphically, it seems that all the models that we consider have difficulties to capture the behaviour of function $\hat{K}$ for values of v between 0.2 and 0.4.

5 Conclusion

The estimator that we considered in this paper is designed for applications in which the censoring times for both individuals only differs through an observed random variable. In the example that we consider, this observed variable represents the age difference. We

only considered the case of two random lifetimes (T, U) , but the procedure can easily be generalized to more lifetimes. The main difficulty of this extension comes from the fact that the procedure requires to put mass only at fully observed observations. To obtain a sufficient number of such observations when the number of lifetimes is high, one would need to have a large value of the sample size n . In this paper, we focused on applying our results to the study of copula models. Other applications of this technique could be considered, as regression models (see Lopez and Saint Pierre (2011) for related problems), or to the evaluation of various dependence measures, see Fan et al. (2000) or Hougaard (2000).

6 Appendix

6.1 Proof of Theorem 3.1

Let (t_0, u_0) denote some point in $\mathbb{R}^2$ such that $S_F(t_0, u_0) > 0$.

First case: $\phi(t, u) = 0$ for $t \geq t_0$ or $u \geq u_0$. One can write

$$\int \phi(t, u) d(\hat{F} - F^*)(t, u) = \frac{1}{n} \sum_{i=1}^n \frac{\delta_i \gamma_i [\hat{S}_G(B_i-) - S_G(B_i-)] \phi(Y_i, Z_i)}{S_G(B_i-)^2} + R_{1n}(\phi),$$

with

$$|R_{1n}(\phi)| \leq \sup_{a \leq a_0} \frac{|\hat{S}_G(a) - S_G(a)|^2}{S_G(a) \hat{S}_G(a)^2} \times \left(\frac{1}{n} \sum_{i=1}^n \delta_i \gamma_i \Phi(Y_i, Z_i) \right), \quad (6.1)$$

where a_0 is some point in $\mathbb{R}$ such that $S_G(a_0) > 0$. It follows from the uniform $n^{1/2}$ -consistency of $\hat{S}_G$ (see Gill (1983)) that the right-hand side in (6.1), which does not depend on ϕ , is $O_P(n^{-1/2})$. Moreover, let us observe that the functions

$$f_n(Y_i, Z_i, \varepsilon_i, \delta_i, \gamma_i) = \delta_i \gamma_i \hat{S}_G(B_i-) \phi(Y_i, Z_i) [S_G(B_i-)]^{-2},$$

$$f(Y_i, Z_i, \varepsilon_i, \delta_i, \gamma_i) = \delta_i \gamma_i S_G(B_i-) \phi(Y_i, Z_i) [S_G(B_i-)]^{-2},$$

are two elements of the Donsker class $\mathcal{H}$ defined in Assumption 3. Moreover, using the uniform convergence rate of $\hat{S}_G$, one obtains that $\|f_n - f\|_\infty \rightarrow 0$. Therefore, the asymptotic equicontinuity of Donsker classes (see Lemma 19.24 in van der Vaart (1998)) ensures that

$$\int \phi(t, u) d(\hat{F} - F^*)(t, u) = \int \frac{[\hat{S}_G(a-) - S_G(a-)] \phi(t, u) d\mathbb{P}_{(T, U, C, \varepsilon)}(t, u, c, e)}{S_G(a-)} + R_{2n}(\phi), \quad (6.2)$$

where $\sup_{\phi \in \mathcal{F}} |R_{2n}(\phi)| = o_P(n^{-1/2})$. Next, the representation (3.1) follows from Stute (1996) or Gijbels and Veraverbeke (1991), since

$$\begin{aligned} \hat{S}_G(a) - S_G(a) &= \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\eta_i S_G(B_i \vee a)}{H(B_i)} - \int \frac{\mathbf{1}_{B_i \geq u} S_G(u \vee y) dF_A(u)}{H(u) F_A(u)} \right\} \\ &\quad + \left\{ \frac{(1 - \eta_i) \mathbf{1}_{B_i > y}}{F_A(B_i)} - S_G(y) \right\} + R(y), \end{aligned}$$

where $\sup_{y \leq y_0} |R(y)| = o_P(n^{-1/2})$, where y_0 is some point such that $\mathbb{P}(B > y_0) > 0$ (by assuming that ϕ is zero for large values of t or u , we ensure that, in (6.2), only terms with b smaller than such a y_0 appear). To deduce equation (3.2), it suffices to observe that

$$\int \phi(t, u) d[\hat{F} - F](t, u) = \frac{1}{n} \sum_{i=1}^n \frac{\delta_i \gamma_i \phi(Y_i, Z_i)}{S_G(B_i -)} - E[\phi(T, U)] + \frac{1}{n} \sum_{i=1}^n \psi_\phi(Y_i, Z_i, \varepsilon_i) + o_P(n^{-1/2}).$$

Each of these two i.i.d. sums have zero mean, and the Central Limit Theorem applies.

General case: the general case follows from a combination of the first case and of Lemma 6.1, in order to make (t_0, u_0) tend to infinity. Point 1 in Lemma 6.1 is easily checked by observing that the terms in the i.i.d. sum in the right-hand side of decomposition (3.1) have finite variance. To show that points 2-4 hold, define $\mathcal{I}_{(t_0, u_0)} = \{(t, u) : t \leq t_0, u \leq u_0\}$, and observe that

$$\begin{aligned} n^{1/2} \left| \int \phi(t, u) \mathbf{1}_{(t, u) \notin \mathcal{I}_{(t_0, u_0)}} d[F - F^*](t, u) \right| &\leq \sup_a n^{1/2} \left| \frac{\hat{S}_G(a) - S_G(a)}{\mathcal{C}(a)^{1/2+v}} \right| \times \sup_a \left| \frac{S_G(a)}{\hat{S}_G(a)} \right| \\ &\quad \times \frac{1}{n} \sum_{i=1}^n \frac{\delta_i \gamma_i \mathcal{C}^{1/2+v}(B_i -) \mathbf{1}_{Y_i \geq t_0, Z_i \geq u_0} \Phi(Y_i, Z_i)}{S_G(B_i -)^2}. \end{aligned}$$

Defining $M_n = \sup_a n^{1/2} |[\hat{S}_G(a) - S_G(a)] [\mathcal{C}(a)^{-1/2-v}]| \times \sup_a |S_G(a) \hat{S}_G(a)^{-1}|$, we get $M_n = O_P(1)$ from Theorem 2.1 in Gill (1983), since $\int \mathcal{C}(a)^{-1-2v} d\mathcal{C}(a) < \infty$ (see condition (2.1) in Gill (1983)). Moreover, defining

$$\Gamma_n(t_0, u_0) = n^{-1} \sum_{i=1}^n \delta_i \gamma_i \mathcal{C}^{1/2+v}(B_i -) \mathbf{1}_{Y_i \geq t_0, Z_i \geq u_0} \Phi(Y_i, Z_i) S_G(B_i -)^{-2},$$

we see that conditions 4 and 5 in Lemma 6.1 hold, thanks to Assumption 3.

6.2 A technical Lemma

Lemma 6.1 *Let $\mathcal{I}_{(t_0, u_0)} = \{t \leq t_0, u \leq u_0\}$, and let $\mathcal{I}$ denote the support of (Y, Z) , and (t_1, u_1) the upper bound of the support. Let $\mathcal{F}$ be a class of functions. Let $P_n(t, u, \phi)$ be a*

process on $\mathcal{I}_{(t_0, u_0)} \times \mathcal{F}$. Define, for any $(t, u) \in \mathcal{I}$, $R_n(t, u, \phi) = P_n(t_1, u_1, \phi) - P_n(t, u, \phi)$. Assume that for all (t_0, u_0) such that $\mathcal{I}_{(t_0, u_0)}$ is strictly included in the interior part of $\mathcal{I}$,

$$(P_n(t, u, \phi))_{(t, u) \in \mathcal{I}_{(t_0, u_0)}, \phi \in \mathcal{F}} \Longrightarrow (W(V_\phi(t)))_{t \in \mathcal{I}_{(t_0, u_0)}, \phi \in \mathcal{F}},$$

where $W(V_\phi(t))$ is a Gaussian process with covariance function V_ϕ , and $\Longrightarrow$ denotes the weak convergence.

Assume that the following conditions hold,

1. $\lim_{(t_0, u_0) \rightarrow (t_1, u_1)} V_\phi(t_0, u_0) = V_\phi(t_1, u_1)$, with $\sup_{\phi \in \mathcal{F}} |V_\phi(t_1, u_1)| < \infty$,
2. $|R_n(t', u', \phi)| \leq M_n \times \Gamma_n(t_0, u_0)$, for all $t' \in \mathcal{I} - \mathcal{I}_{(t_0, u_0)}$,
3. $M_n = O_P(1)$,
4. $\Gamma_n(t_0, u_0) \rightarrow \Gamma(t_0, u_0)$ in probability,
5. $\lim_{(t_0, u_0) \rightarrow (t_1, u_1)} \Gamma(t_0, u_0) = 0$.

Then $P_n(t_1, u_1, \phi) \Longrightarrow \mathcal{N}(0, V_\phi(t_1, u_1))$.

Proof. From Theorem 13.5 in Billingsley (1999) and from condition 1, it suffices to show that, for all $\varepsilon > 0$,

$$\lim_{(t_0, u_0) \rightarrow (t_1, u_1)} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\sup_{(t, u) \in \mathcal{I} - \mathcal{I}_{(t_0, u_0)}, \phi \in \mathcal{F}} |R_n(t, u, \phi)| > \varepsilon \right) = 0. \quad (6.3)$$

Using condition 2 in the Lemma, the probability in equation (6.3) is bounded, for all $M > 0$, by

$$\mathbb{P}(|\Gamma_n(t, u) - \Gamma(t, u)| > \varepsilon/M - \Gamma(t, u)) + \mathbb{P}(M_n > M). \quad (6.4)$$

Moreover, from condition 4,

$$\limsup_{n \rightarrow \infty} \mathbb{P}(|\Gamma_n(t, u) - \Gamma(t, u)| > \varepsilon/M - \Gamma(t, u)) = \mathbf{1}_{\varepsilon/M - \Gamma(t, u) \geq 0}.$$

Since $\Gamma(t, u) \rightarrow 0$ (condition 5), we can deduce that

$$\lim_{(t, u) \rightarrow \tau_H} \limsup_{n \rightarrow \infty} \mathbb{P}(|\Gamma_n(t, u) - \Gamma(t, u)| > \varepsilon/M - \Gamma(t, u)) = 0.$$

Hence,

$$\lim_{(t, u) \rightarrow (t_1, u_1)} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\sup_{(t', u') > (t, u), \phi \in \mathcal{F}} |R_n(t', u', \phi)| > \varepsilon \right) \leq \limsup_{n \rightarrow \infty} \mathbb{P}(M_n > M).$$

As a consequence,

$$\lim_{(t,u) \rightarrow (t_1,u_1)} \limsup_{n \rightarrow \infty} P \left(\sup_{(t',u') > (t,u), \phi \in \mathcal{F}} |R_n(t',u',\phi)| > \varepsilon \right) \leq \lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} P(M_n > M) = 0,$$

using the fact that $M_n = O_P(1)$ (condition 3). ■

Acknowledgements: This work is supported by French Agence Nationale de la Recherche (ANR) ANR Grant ANR-09-JCJC-0101-01.

References

- Akritis, M. G. (1986). Bootstrapping the Kaplan-Meier estimator. *J. Amer. Statist. Assoc.*, 81(396):1032–1038.
- Akritis, M. G. and Van Keilegom, I. (2003). Estimation of bivariate and marginal distributions with censored data. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 65(2):457–471.
- Billingsley, P. (1999). *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition. A Wiley-Interscience Publication.
- Campbell, G. and Földes, A. (1982). Large-sample properties of nonparametric bivariate estimators with censored data. In *Nonparametric statistical inference, Vol. I, II (Budapest, 1980)*, volume 32 of *Colloq. Math. Soc. János Bolyai*, pages 103–121. North-Holland, Amsterdam.
- Carriere, J. F. (2000). Bivariate survival models for coupled lives. *Scand. Actuar. J.*, (1):17–32.
- Dabrowska, D. M. (1988). Kaplan-Meier estimate on the plane. *Ann. Statist.*, 16(4):1475–1489.
- Efron, B. (1981). Censored data and the bootstrap. *J. Amer. Statist. Assoc.*, 76(374):312–319.
- Fan, J., Hsu, L., and Prentice, R. L. (2000). Dependence estimation over a finite bivariate failure time region. *Lifetime Data Anal.*, 6(4):343–355.
- Frees, E. W., Carriere, J. F., and Valdez, E. A. (1996). Annuity valuation with dependent mortality. *Journal of Risk and Insurance*, 63(2):229–261.

- Genest, C., Quessy, J.-F., and Rémillard, B. (2006). Goodness-of-fit procedures for copula models based on the probability integral transformation. *Scand. J. Statist.*, 33(2):337–366.
- Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *J. Amer. Statist. Assoc.*, 88(423):1034–1043.
- Gijbels, I. and Veraverbeke, N. (1991). Almost sure asymptotic representation for a class of functionals of the Kaplan-Meier estimator. *Ann. Statist.*, 19(3):1457–1470.
- Gill, R. (1983). Large sample behaviour of the product-limit estimator on the whole line. *Ann. Statist.*, 11(1):49–58.
- Gill, R. D., van der Laan, M. J., and Wellner, J. A. (1995). Inefficient estimators of the bivariate survival function for three models. *Ann. Inst. H. Poincaré Probab. Statist.*, 31:547–597.
- Hanley, J. A. and Parnes, M. N. (1983). Nonparametric estimation of a multivariate distribution in the presence of censoring. *Biometrics*, 39(1):129–139.
- Hougaard, P. (2000). *Analysis of multivariate survival data*. Statistics for Biology and Health. Springer-Verlag, New York.
- Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.*, 53:457–481.
- Lin, D. Y. and Ying, Z. (1993). A simple nonparametric estimator of the bivariate survival function under univariate censoring. *Biometrika*, 80(3):573–581.
- Lopez, O. (2012). A generalization of kaplan-meier estimator for analyzing bivariate mortality under right-censoring and left-truncation with applications to model-checking for survival copula models. *Preprint*.
- Lopez, O. and Saint Pierre, P. (2011). Bivariate censored regression relying on a new estimator of the joint distribution function. *Preprint*.
- Luciano, E., Spreeuw, J., and Vigna, E. (2008). Modelling stochastic mortality for dependent lives. *Insurance Math. Econom.*, 43(2):234–244.
- Nelsen, R. B. (2006). *An introduction to copulas*. Springer Series in Statistics. Springer, New York, second edition.

- Prentice, R. L. and Cai, J. (1992). Covariance and survivor function estimation using censored multivariate failure time data. *Biometrika*, 79(3):495–512.
- Pruitt, R. C. (1991a). On negative mass assigned by the bivariate Kaplan-Meier estimator. *Ann. Statist.*, 19(1):443–453.
- Pruitt, R. C. (1991b). Strong consistency of self-consistent estimators: general theory and an application to bivariate survival analysis. Technical Report 543, University of Minnesota, Minneapolis.
- Reid, N. (1981). Estimating the median survival time. *Biometrika*, 68(3):601–608.
- Shih, J. H. and Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, 51(4):1384–1399.
- Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.
- Stute, W. (1996). Distributional convergence under random censorship when covariables are present. *Scand. J. Statist.*, 23(4):461–471.
- Stute, W. and Wang, J.-L. (1993). The strong law under random censorship. *Ann. Statist.*, 21(3):1591–1607.
- Tsai, W.-Y., Leurgans, S., and Crowley, J. (1986). Nonparametric estimation of a bivariate survival function in the presence of censoring. *Ann. Statist.*, 14(4):1351–1365.
- van der Laan, M. J. (1996). Efficient estimation in the bivariate censoring model and repairing NPMLE. *Ann. Statist.*, 24(2):596–627.
- van der Vaart, A. W. (1998). *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge.
- van der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence and empirical processes with applications to statistics*. Springer Series in Statistics. Springer-Verlag, New York.
- Wang, W. and Wells, M. T. (2000a). Estimation of Kendall’s tau under censoring. *Statist. Sinica*, 10(4):1199–1215.
- Wang, W. and Wells, M. T. (2000b). Model selection and semiparametric inference for bivariate failure-time data. *J. Amer. Statist. Assoc.*, 95(449):62–76. With a comment by Edsel A. Peña and a rejoinder by the authors.

Youn, H. and Shemyakin, A. (1999). Statistical aspects of joint life insurance pricing. *Proceedings of the business and statistics section of the American Statistical Association*, pages 34–38.

Youn, H. and Shemyakin, A. (2001). Pricing practices for joint last survivor insurance. *Actuarial Research Clearing House*.