Statistica Sinica **33** (2023), 983-1001
doi:https://doi.org/10.5705/ss.202020.0520

THE IDENTIFIABILITY OF COPULA MODELS FOR DEPENDENT COMPETING RISKS DATA WITH EXPONENTIALLY DISTRIBUTED MARGINS

Antai Wang

New Jersey Institute of Technology

Abstract: We prove the identifiability property of Archimedean copula models for dependent competing risks data when at least one of the failure times is exponentially distributed. With this property, it becomes possible to quantify the dependence between competing events based on exponentially distributed dependent censored data. We demonstrate our estimation procedure using simulation studies and in an application to survival data.

Key words and phrases: Archimedean copula models, copula graphic estimator, identifiability of competing risks data.

1. Introduction

The identifiability of competing risks models has been a challenging topic in statistical research. Suppose that T is a time to an event and C is a time to a competing event, so that one can only observe $(\min\{T, C\}, I(T < C))$. How to evaluate the true association between T and C is an important research issue. Tsiatis (1975) has proved the nonidentifiability of competing risks models for this type of data, without any model or covariate information. Wang (2012) has proved the nonidentifiability of Archimedean copula models based on dependent censored data. Heckman and Honoré (1989) have proposed a set of conditions to make a competing risks model identifiable with additional covariate information. Wang et al. (2015) established a set of simpler conditions to make an Archimedean copula model (a special class of competing risks model) identifiable with covariate information.

In survival data analysis, survival times are often assumed to be exponentially distributed. We are interested in finding models that are identifiable when the time to an event is exponentially distributed and subject to dependent censoring. It turns out that the Archimedean copula model assumption is good enough to

Corresponding author: Antai Wang, Ph.D., Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, NJ 07102, USA. E-mail: aw224@njit.edu.

make the corresponding model identifiable for dependent censored data when at least one of the failure times is exponentially distributed (as we prove here).

Another motivation for this research is related to the estimation of survival functions. The Kaplan–Meier estimator has become the major tool used in survival analysis since it was first introduced by Kaplan and Meier (1958). It is a consistent estimator of the survival function of a failure time T when T is subject to independent right censoring with a censoring time C. The available data we have is $(X_i, \delta_i) = (\min\{T_i, C_i\}, I(T_i < C_i))$, for $i = 1, 2, \ldots, n$, where T and C are assumed to be independent and the Kaplan-Meier estimator is defined as

$$\hat{S}_T(t) = \Pi_{X_i<t}\left(1 - \frac{d_i}{n_i}\right),$$

where d_i and n_i are defined as the number of failures at time X_i and the number "at risk" just prior to time X_i, respectively. Then, one can show

$$\hat{S}_T(t) \to S_T(t) = \Pr(T > t)$$

uniformly when $n \to \infty$ (see Fleming and Harrington (1991)). However, when T and C are dependent, the Kaplan–Meier estimator can no longer be applied, because it is a biased estimator of the survival function of T. Therefore, it is of great importance and research interest to model the dependence structure between T and C, and to quantify the true dependence between them effectively. Once the dependence level is determined, the marginal survival functions can be estimated consistently using the copula graphic estimator proposed by Zheng and Klein (1995) and Rivest and Wells (2001), or the estimator proposed by Wang (2014).

The remainder of the paper is organized as follows. We present our main results about the identifiability in Section 2. In Section 3, we propose a new approach to estimating the dependence parameter in Archimedean copula models based on the property we proved in Section 2. Simulation studies in Section 4 demonstrate that the proposed method works well. We end our paper with a discussion in Section 5.

2. Main Results

A bivariate random vector (T, C) follows an Archimedean copula if the joint survival function of (T, C) can be expressed as

$$S(t, c) = \psi_\theta\big\{\psi_\theta^{-1}[S_1(t)] + \psi_\theta^{-1}[S_2(c)]\big\},$$

where S_1 and S_2 are marginal survival functions of T and C, respectively, and ψ_θ is defined on $[0, \infty]$ so that $\psi_\theta(0) = 1$, $\psi'_\theta(s) < 0$, and $\psi''_\theta(s) > 0$. Then, ψ_θ^{-1} is the inverse function of ψ_θ, defined as a copula generator (see Nelsen (2007)), and θ is the unknown parameter.

The first Archimedean copula model was proposed by Clayton (1978). For this model, $\psi_\alpha(s) = (1 + s)^{-1/\alpha}$, which leads to the bivariate survivor function

$$S(t, c) = \left\{ \frac{1}{S_1(t)^{-\alpha} + S_2(c)^{-\alpha} - 1} \right\}^{1/\alpha},$$

for $\alpha > 0$. Another important frailty model, the Hougaard model (Hougaard (1986)), has $\psi_\beta(s) = \exp(-s^\beta)$. Its bivariate survivor function is

$$S(t, c) = \exp\left(- \left[\{- \log(S_1(t))\}^{1/\beta} + \{- \log(S_2(c))\}^{1/\beta} \right]^\beta \right)$$

for $\beta \in (0, 1)$. In addition to the Clayton model and Hougaard model, other well-known models such as the Frank model (Genest (1987)) and the Log-copula model belong to this family.

For dependent censored data $(X = \min\{T, C\}, \delta = I(T < C))$ where (T, C) can be modelled by an Archimedean copula, we can prove the following result:

Theorem 1. . *For a dependent censored data $(X = \min\{T, C\}, \delta = I(T < C))$, assume that the distribution of (T, C) can be modeled by an Archimedean copula with the copula generator ψ_θ (θ is the dependence parameter):*

$$S(t, c) = \psi_\theta[\psi_\theta^{-1}(S_1(t)) + \psi_\theta^{-1}(S_2(c))].$$

Under the assumption that $\psi_{\theta_1}^{-1'}/\psi_{\theta_2}^{-1'}$ is a strictly increasing function for $\theta_1 < \theta_2$, there is a unique Archimedean copula model such that the distribution of T (or C) is exponential on $[0, \infty)$.

Proof: Suppose there are two marginal survival functions $S_1^{(1)}$ and $S_1^{(2)}$ of T, corresponding to different dependence levels θ_1 and θ_2, such that both of them are exponential distributions with parameters λ_1 and λ_2, respectively, so that

$$S_1^{(1)}(t) = \exp(-\lambda_1 t), \quad S_1^{(2)}(t) = \exp(-\lambda_2 t).$$

Then, we have two Archimedean copula models with the copula generators ψ_{θ_1} and ψ_{θ_2}, respectively such that

$$S^{(1)}(t, c) = \psi_{\theta_1}[\psi_{\theta_1}^{-1}(S_1^{(1)}(t)) + \psi_{\theta_1}^{-1}(S_2^{(1)}(c))]$$

and

$$S^{(2)}(t,c) = \psi_{\theta_2}[\psi_{\theta_2}^{-1}(S_1^{(2)}(t)) + \psi_{\theta_2}^{-1}(S_2^{(2)}(c))],$$

and the distributions of (X, δ) corresponding to both models are the same. Without loss of generality, we assume that $\theta_1 < \theta_2$. Because the distributions of (X, δ) are the same for both models, by Theorem 2 of Wang et al. (2015), the following equality must hold:

$$S_1^{(2)}(t) = \psi_{\theta_2}\left[\int_0^t \frac{\psi_{\theta_2}^{-1'}\{\pi(u)\}}{\psi_{\theta_1}^{-1'}\{\pi(u)\}} d\psi_{\theta_1}^{-1}(S_1^{(1)}(u))\right].$$

Taking the derivative with respect to t on both sides, we get:

$$\begin{aligned} S_1^{(2)'}(t) &= \psi'_{\theta_2}\left[\int_0^t \frac{\psi_{\theta_2}^{-1'}\{\pi(u)\}}{\psi_{\theta_1}^{-1'}\{\pi(u)\}} d\psi_{\theta_1}^{-1}(S_1^{(1)}(u))\right] \frac{\psi_{\theta_2}^{-1'}\{\pi(t)\}}{\psi_{\theta_1}^{-1'}\{\pi(t)\}} \psi_{\theta_1}^{-1'}(S_1^{(1)}(t)) S_1^{(1)'}(t) \\ &= \frac{\psi_{\theta_2}^{-1'}\{\pi(t)\}}{\psi_{\theta_1}^{-1'}\{\pi(t)\}} \frac{\psi_{\theta_1}^{-1'}(S_1^{(1)}(t))}{\psi_{\theta_2}^{-1'}(S_1^{(2)}(t))} S_1^{(1)'}(t). \end{aligned}$$

from which we have

$$\frac{\psi_{\theta_1}^{-1'}\{\pi(t)\}}{\psi_{\theta_2}^{-1'}\{\pi(t)\}} = \frac{\psi_{\theta_1}^{-1'}(S_1^{(1)}(t))}{\psi_{\theta_2}^{-1'}(S_1^{(2)}(t))} \times \frac{S_1^{(1)'}(t)}{S_1^{(2)'}(t)}.$$

Suppose $S_1^{(1)}(t) = S_1^{(2)}(t)$; then, $S_1^{(1)'}(t) = S_1^{(2)'}(t)$. From the above equality we must have

$$\frac{\psi_{\theta_1}^{-1'}\{\pi(t)\}}{\psi_{\theta_2}^{-1'}\{\pi(t)\}} = \frac{\psi_{\theta_1}^{-1'}(S_1^{(1)}(t))}{\psi_{\theta_2}^{-1'}(S_1^{(2)}(t))} \times \frac{S_1^{(1)'}(t)}{S_1^{(2)'}(t)} = \frac{\psi_{\theta_1}^{-1'}(S_1^{(1)}(t))}{\psi_{\theta_2}^{-1'}(S_1^{(1)}(t))},$$

which can't hold because we know

$$\pi(t) = P(T > t, C > t) = S(t,t) = \psi_\theta[\psi_\theta^{-1}(S_1(t)) + \psi_\theta^{-1}(S_2(t))],$$

where $\theta = \theta_1$ or θ_2, because $\pi(t) = P(T > t, C > t) = S(t,t)$ is the same for both models. It is obvious that $\pi(t) = S(t,t)$ is generally smaller than the probability of $P(T > t) = S_1(t)$. By the condition that $\psi_{\theta_1}^{-1'}/\psi_{\theta_2}^{-1'}$ is a strictly increasing function when $\theta_1 < \theta_2$, we must have

$$\frac{\psi_{\theta_1}^{-1'}\{\pi(t)\}}{\psi_{\theta_2}^{-1'}\{\pi(t)\}} < \frac{\psi_{\theta_1}^{-1'}(S_1^{(1)}(t))}{\psi_{\theta_2}^{-1'}(S_1^{(1)}(t))},$$

which leads to a contradiction. Therefore, we only need to consider the situation in which $S_1^{(1)}(t) \neq S_1^{(2)}(t)$. Now, consider the stochastic ordering of $S_1^{(1)}(t)$ and $S_1^{(2)}(t)$. Under the assumption that $\psi_{\theta_1}^{-1'}/\psi_{\theta_2}^{-1'}$ is an increasing function when $\theta_1 < \theta_2$, applying Proposition 2 in Rivest and Wells (2001), we can conclude

$$S_1^{(2)}(t) \leq S_1^{(1)}(t),$$

so that $\lambda_1 < \lambda_2$ (because we just proved that $S_1^{(1)}(t) \neq S_1^{(2)}(t)$). Hence,

$$\frac{S_1^{(1)'}(t)}{S_1^{(2)'}(t)} = \frac{\lambda_1 \exp(-\lambda_1 t)}{\lambda_2 \exp(-\lambda_2 t)}.$$

Now, consider the equality:

$$\frac{\psi_{\theta_1}^{-1'}\{\pi(t)\}}{\psi_{\theta_2}^{-1'}\{\pi(t)\}} = \frac{\psi_{\theta_1}^{-1'}(S_1^{(1)}(t))}{\psi_{\theta_2}^{-1'}(S_1^{(2)}(t))} \times \frac{S_1^{(1)'}(t)}{S_1^{(2)'}(t)} = \frac{\psi_{\theta_1}^{-1'}(S_1^{(1)}(t))}{\psi_{\theta_2}^{-1'}(S_1^{(2)}(t))} \frac{\lambda_1 \exp(-\lambda_1 t)}{\lambda_2 \exp(-\lambda_2 t)}.$$

Letting $t \to 0$ ($t \geq 0$) on both sides, we can see that the limit of the left-hand side equals to

$$\lim_{t\to 0} \frac{\psi_{\theta_1}^{-1'}\{\pi(t)\}}{\psi_{\theta_2}^{-1'}\{\pi(t)\}} = \frac{\psi_{\theta_1}^{-1'}\{\pi(0)\}}{\psi_{\theta_2}^{-1'}\{\pi(0)\}} = \frac{\psi_{\theta_1}^{-1'}\{1\}}{\psi_{\theta_2}^{-1'}\{1\}}$$

whereas the limit of the right-hand side is equal to

$$\frac{\psi_{\theta_1}^{-1'}(S_1^{(1)}(0))}{\psi_{\theta_2}^{-1'}(S_1^{(2)}(0))} \times \frac{\lambda_1}{\lambda_2} = \frac{\psi_{\theta_1}^{-1'}\{1\}}{\psi_{\theta_2}^{-1'}\{1\}} \frac{\lambda_1}{\lambda_2} \neq \frac{\psi_{\theta_1}^{-1'}\{1\}}{\psi_{\theta_2}^{-1'}\{1\}},$$

because $\psi_{\theta_1}^{-1'}\{1\}$ and $\psi_{\theta_2}^{-1'}\{1\}$ are well-defined finite quantities for Archimedean copula models and $\lambda_1 < \lambda_2$, where the equality cannot hold on a set with probability measure greater than zero. This yields a contradiction. Therefore, we can conclude that there is a unique Archimedean copula model such that the marginal distribution of T is exponential on $[0, \infty)$.

Remark 1. The condition $\psi_{\theta_1}^{-1'}/\psi_{\theta_2}^{-1'}$ being an increasing function for $\theta_1 < \theta_2$ is satisfied for most one parameter families of Archimedean copulas, such as the Frank model or the Clayton model. This fact has been pointed out by Rivest and Wells (2001). Note that the identifiability issue has already been dealt with using copulas in previous studies (e.g., Carrière (1995) and Escarela and Carrierè (2003)). However our work differs from theirs because in these two papers, the authors assume that both the crude survival functions and the copula function are known in order to identify the marginal survival functions. In our study, we assume only that the joint copula function is Archimedean, with the corresponding dependence parameter unspecified. Then under the weak assumption that one of the marginal survival function is exponential, the dependence can be uniquely determined based on (X, δ), and the identifiability of the marginal survival functions can be established.

3. Dependence Parameter Estimation

Using Theorem 1, we know that our model is identifiable, based on the exponential assumption of the marginal distributions of T. Now, we propose a new parameter estimator for dependent censored data $(X_i, \delta_i) = (\min\{T_i, C_i\}, I(T_i < C_i))$ under the above model/distribution assumptions; that is, the joint copula of (T, C) is an Archimedean copula, and the marginal distribution of T is an exponential distribution. Define $X_i = \min\{T_i, C_i\}$, $\delta_i = I(T_i < C_i)$, $N_i(y) = I(X_i < y, \delta_i = 1)$, $Y_i(y) = I(T_i \geq y, C_i \geq y)$, $\bar{N}(y) = \sum_{i=1}^n N_i(y)$, and $\bar{Y}(y) = \sum_{i=1}^n Y_i(y)$. According to Fleming and Harrington (1991) and Rivest and Wells (2001),

$$M_i(t) = N_i(t) - \int_0^t Y_i(s)\lambda^\sharp(s)ds$$

and

$$\bar{M}(t) = \bar{N}(t) - \int_0^t \bar{Y}(s)\lambda^\sharp(s)ds$$

are martingales with respect to the σ field $\mathcal{F}_t^i = \sigma\{I(X_i \leq t, \delta_i = 1), I(X_i \leq t, \delta_i = 0)\}$ and $\mathcal{F}_t = \bigvee_{i=1}^n \mathcal{F}_t^i$, respectively, and $\lambda^\sharp(s)$ is defined as the crude hazard function of T. We denote the corresponding crude cumulative function and the survival function by $\Lambda^\sharp(s)$ and $S^\sharp(s)$, respectively. Using Theorem 1 in Wang (2014) or formula (8) in Rivest and Wells (2001), we can express the marginal survival function of T as:

$$S_1(t) = \psi_\theta\left[\int_0^t \psi_\theta^{-1'}(\pi(u))\pi(u)d\ln(S_1^\sharp(u))\right]$$

which can be estimated by (see Wang (2014))

$$\hat{S}_1(t) = \psi_\theta\left\{-\sum_{X_i \leq t, \delta_i = 1} \psi_\theta^{-1'}\{\hat{\pi}(X_i)\}\hat{\pi}(X_i)\frac{d\bar{N}(X_i)}{\bar{Y}(X_i)}\right\},$$

where $\psi_\theta^{-1'}(u)$ is the derivative of ψ_θ^{-1} with respect to u, and $\pi(X_i) = \Pr(X > X_i)$, where $\hat{\pi}(X_i)$ is the empirical survival function of $\pi(X)$ evaluated at X_i. The dependence parameter θ can be estimated in the following way: because $S_1(t)$ is assumed to be the survival function of an exponential distribution, $S_1(t) = \exp(-\lambda t)$, for some λ value. The cumulative hazard function of T is

$$H(t) = -\log(S_1(t)) = \lambda t,$$

and $H(t)/t = \lambda = 1/\mu$, where μ is the mean value of T. Given dependent censored data, a natural estimator of μ can be established as:

$$\hat{\mu} = \left[\frac{1}{n} \sum_{i=1}^{n} \frac{\hat{H}_1(X_i)}{X_i} \right]^{-1} = \left[\frac{1}{n} \sum_{i=1}^{n} -\frac{\log \hat{S}_1(X_i)}{X_i} \right]^{-1},$$

where $\hat{S}_1$ is the copula graphic estimator of S_1 (see Rivest and Wells (2001)) or the Wang estimator of S_1 (see Wang (2014)). Therefore, the survival function of the corresponding exponential distribution can be alternatively estimated by $\exp(-X_i/\hat{\mu})$. The dependence parameter value can then be determined as:

$$\hat{\theta}_n = \underset{\theta \in \Theta}{\operatorname{argmin}}\, Q(\theta) = \underset{\theta \in \Theta}{\operatorname{argmin}} \sum_{i=1}^{n} \left\{ \hat{S}_1(X_i) - \exp\left(-\frac{X_i}{\hat{\mu}} \right) \right\}^2.$$

Assuming the differentiability of $\hat{S}_1$ with respect to the unknown parameter θ, the corresponding estimating equation can be written as

$$\frac{dQ}{d\theta} = 2 \sum_{i=1}^{n} \left\{ \hat{S}_1(X_i) - \exp\left(-\frac{X_i}{\hat{\mu}} \right) \right\} \left\{ \frac{d\hat{S}_1(X_i)}{d\theta} - \exp\left(-\frac{X_i}{\hat{\mu}} \right) \frac{X_i}{\hat{\mu}^2} \frac{d\hat{\mu}}{d\theta} \right\} = 0.$$

The asymptotic normality of our parameter estimator follows based on the corresponding Taylor expansion of our estimating equation under necessary regularity conditions.

Theorem 2. *Under necessary regularity conditions, $\hat{\theta}_n$ is consistent and $\sqrt{n}$ $(\hat{\theta}_n - \theta_0)$ is asymptotically normal with mean zero and a finite variance σ^2.*

For the proof of Theorem 2, see the Appendix. The variance formula is quite complicated, and we recommend using a bootstrap variance estimator to estimate the variance.

4. Simulation Studies

We conducted simulation studies to demonstrate our parameter estimation procedure. We consider weak to strong dependence levels, with Kendall's τ values equal to $0.3, 0.5$, and 0.7 for the Clayton copula model when the sample sizes are $n = 500$, $n = 1500$, and $n = 3000$, respectively. We generate (T, C) from a Clayton copula model with unit exponential marginal distributions, and estimate the dependence parameter $\alpha = 2\tau/(1-\tau)$ using our proposed method described in Section 3. Based on any simulated sample, we get a bootstrap sample using the resampling approach. We estimated the parameter using our proposed method for the bootstrap samples B times to get a bootstrap variance estimate using the bootstrap estimates of the parameter. We then take the average of the bootstrap variance estimate, which should be approximately the sample variance of our proposed parameter estimates. For each simulated sample, we can get the confidence interval for our parameter using the bootstrap variance estimate. The simulation results are listed in Table 1. In Table 1, $\hat{\alpha}_1$ represents our proposed estimate of α, and $\hat{\alpha}_2$ represents the MLE estimate of α, assuming that the joint survival can be modeled by the Clayton model and the marginal distributions are all exponential distributions. Furthermore, $\hat{\alpha}_3$ represents the estimate of the parameter α by minimizing the sum of the distances between the copula-graphic estimates of S_T and the survival function estimates of T, and also the distances between the copula-graphic estimates of S_C and the survival function estimates of C, assuming that the joint survival can be modeled by the Clayton model and the marginal distributions are all exponential distributions. From Table 1, we can see that our proposed estimation approach works well for different levels of dependence under the Clayton model assumption. When the sample size increases from $n = 500$ to $n = 3000$, the proposed parameter estimates tend to be less biased in both scenarios, the bootstrap variance estimators work reasonably well, because they are close to the sample variances of the proposed parameter estimates. In general, these results demonstrate the identifiability of our models when the marginal distributions of T are exponential. In addition, we see that the MLE ($\hat{\alpha}_2$) and $\hat{\alpha}_3$ tend to be more efficient than the proposed parameter estimates by comparing the sample variances, and are also less biased (not surprisingly, the MLE is the most efficient estimate). This is understandable, because we have more information about the data (here, we assume both marginal distributions are exponential, rather than only one margin being exponentially distributed, which is the assumptions under which our proposed estimate $\hat{\alpha}_1$ is developed).

5. An Illustrative Example

In this section, we illustrate our methodology by analyzing the bone marrow transplant data described in Klein and Moeschberger (1997). The data set was

Table 1. The Clayton model: performance of our parameter estimates based on 1,000 repetitions with sample sizes $N = 500$, $N = 1500$, and $N = 3000$. $\overline{\hat{\alpha}_1}$, $\overline{\hat{\alpha}_2}$, and $\overline{\hat{\alpha}_3}$ are the mean values of $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\alpha}_3$, respectively. $\widehat{var}(\hat{\alpha}_1)$, $\widehat{var}(\hat{\alpha}_2)$, and $\widehat{var}(\hat{\alpha}_3)$ represent the sample variance of $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\alpha}_3$ respectively. $\overline{\widehat{bvar}(\hat{\alpha})_1}$ represents the mean of the bootstrap variances of the proposed parameter estimate $\hat{\alpha}_1$.

α	$\overline{\hat{\alpha}_1}$	$\overline{\hat{\alpha}_2}$	$\overline{\hat{\alpha}_3}$	$\widehat{var}(\hat{\alpha}_1)$	$\overline{\widehat{bvar}(\hat{\alpha}_1)}$	$\widehat{var}(\hat{\alpha}_2)$	$\widehat{var}(\hat{\alpha}_3)$
				$N = 500$			
0.86	2.11	0.81	1.81	5.48	4.30	0.25	4.37
2.00	3.13	2.31	2.79	6.00	7.52	6.52	2.94
4.67	6.72	5.93	7.16	16.96	14.60	14.63	14.57
				$N = 1500$			
0.86	1.12	0.83	0.98	1.54	2.42	0.05	0.85
2.00	2.93	2.19	3.02	6.99	4.11	0.32	1.30
4.67	5.26	5.08	5.52	5.31	8.57	1.71	4.73
				$N = 3000$			
0.86	1.11	0.67	0.68	0.91	1.88	0.08	0.03
2.00	2.44	2.12	2.08	5.10	6.85	0.14	0.23
4.67	5.21	4.79	5.09	4.11	4.53	0.65	1.71

collected during a study in which 137 patients were followed in their recovery from leukemia after a bone marrow transplant. We are interested in the disease-free survival time T, that is, the time until a relapse of leukemia. The patients can be censored by two possible events: disease-free death or disease-free and alive at the end of the study. The censoring time C is then defined as the time until the first of these two events happen. Assuming that T is exponentially distributed and applying our estimation method, we find that the association level is $\hat{\alpha} = 3.62$ ($\hat{\tau} = 0.64$), with the bootstrap standard error 0.93 under the Clayton model assumption. The 95% confidence interval for α is CI=$[3.62-1.96*0.93, 3.62+1.96*0.93] = [1.80, 5.44]$, with $0 \notin$ CI. There is enough evidence to conclude that the dependence between T and C is significantly different from zero. This conclusion is consistent with the results given in Lakhal, Rivest and Abdous (2008) and Wang et al. (2015). In Lakhal, Rivest and Abdous (2008), the authors explore the relationship between T and the time to death (different from our censoring variable C) for the same data, and found a strong relationship using a semicompeting risks model. Based on our parameter estimate, we can estimate the survival function using the copula-graphic estimate proposed by Rivest and Wells (2001) or the Wang estimate proposed in Wang (2014) (these two estimates are equivalent asymptotically, as shown in Wang (2014)). We plot the Wang estimate and the Kaplan-Meier estimate for the same data in the same figure. From Figure 1, we can see that the Kaplan-Meier curve overestimates the

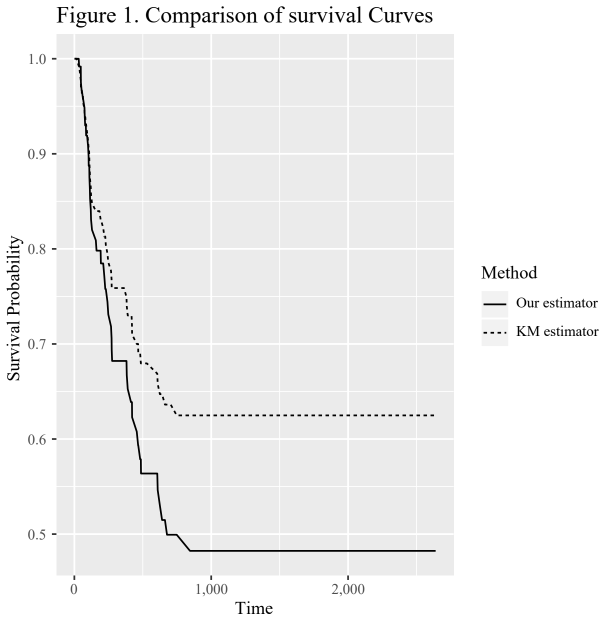

Figure 1

survival function when the association is strong. There is a significant difference between the two survival function estimates based on the same censored data. Therefore, it is important to quantify the dependence level accurately to get reasonable survival function estimates. The advantage of using our approach is that we can directly estimate the association between T and C based on a simple exponential distribution assumption of T, which is quite widely used in survival analysis. We do not need to assume a special relationship between T and C (such as $T < C$ in the semicompeting risks setting) to make the association estimable (see Lakhal, Rivest and Abdous (2008) and Fine, Jiang and Chappell (2001)).

6. Discussion

The identifiability of competing risks models has been a long-standing problem in statistical research, making the analysis of dependent competing risks data a challenging task. This research is an attempt to solve this problem by imposing a simple copula model and distribution assumption for dependent censored data. We have proved the identifiability of Archimedean copula models for dependent competing risks data when either the failure time T or the censoring time C is exponentially distributed. We have also proposed a parameter estimation approach based on our model and distribution assumptions. In survival data analysis, researchers often assume exponential distributions for time-to-event data. Therefore, our method should be quite useful for statistical analyses of dependent censored data. Once we obtain consistent estimates of the dependence parameter, the association between T and C can be tested, and significance tests can be developed based on the asymptotic theory described in Zeng, Lin and Lin (2008) for semiparametric transformation models (see Theorems 2 and 3 in their paper). The marginal survival functions can be consistently estimated using the copula-graphic estimator (Zheng and Klein (1995), Rivest and Wells (2001)), or the estimator proposed by Wang (2014).

Acknowledgments

The author would like to thank the editor and three anonymous reviewers for their encouragement and valuable comments.

Appendix

Proof of consistency of our parameter estimator: by the strong law of large numbers, it is easy to show that

$$\begin{aligned} M_n(\theta) &= \frac{1}{n}\sum_{i=1}^{n}\left[\hat{S}_1(X_i) - \exp\left(\frac{-X_i}{\hat{\mu}}\right)\right]^2 \\ &\xrightarrow{p} \mathrm{E}\left[S_1(X) - \exp\left(\frac{-X}{\mu}\right)\right]^2 = M(\theta), \end{aligned}$$

when $n \to \infty$ based on the following facts: (a) $\hat{S}_1(X_i) \xrightarrow{p} S_1(X_i)$ uniformly when $n \to \infty$ by Theorem 1 in Rivest and Wells (2001); (b) $\exp(-X_i/\hat{\mu}) \xrightarrow{p} \exp(-X_i/\mu)$ uniformly when $n \to \infty$ because of (a), the boundedness of corresponding derivatives and the strong law of large numbers. By Theorem 1 we know there is a unique $\theta = \theta_0$ value such that $S_1(X) = \exp(-X/\mu)$ and $M(\theta)$ is a continuous function of θ on a compact set Θ, therefore, we must have

$$\min_{\theta: d(\theta,\theta_0)\geq\epsilon} M(\theta) > M(\theta_0) = 0.$$

$M_n(\hat{\theta}) \leq M_n(\hat{\theta}) + o_p(1)$ is trivially satisfied because $\hat{\theta}$ is an M-estimator. Therefore we can conclude that $\hat{\theta} \to \theta_0$ when $n \to \infty$ by Theorem 5.7 in van der Vaart (2007). Proof of asymptotic normality of our parameter estimator: using the Taylor expansion at $\theta = \theta_0$ (assuming that θ_0 is the true parameter), our estimating equation can be written as:

$$0 = \sum_{i=1}^{n} g_n(X_i, \hat{\theta}) \approx \sum_{i=1}^{n} g_n(X_i, \theta_0) + \sum_{i=1}^{n} \frac{dg_n(X_i, \theta_0)}{d\theta}(\hat{\theta} - \theta_0),$$

where $g_n(X_i, \theta_0)$ is defined as:

$$\begin{aligned} g_n(X_i, \theta_0) &= \left\{ \hat{S}_1(X_i) - \exp\left(-\frac{X_i}{\hat{\mu}} \right) \right\} \left\{ \frac{d\hat{S}_1(X_i)}{d\theta} - \exp\left(-\frac{X_i}{\hat{\mu}} \right) \frac{X_i}{\hat{\mu}^2} \frac{d\hat{\mu}}{d\theta} \right\} \\ &= \hat{S}_1(X_i) \frac{d\hat{S}_1(X_i)}{d\theta} - \hat{S}_1(X_i) \exp\left(-\frac{X_i}{\hat{\mu}} \right) \frac{X_i}{\hat{\mu}^2} \frac{d\hat{\mu}}{d\theta} \\ &\quad - \exp\left(-\frac{X_i}{\hat{\mu}} \right) \frac{d\hat{S}_1(X_i)}{d\theta} + \exp\left(-\frac{2X_i}{\hat{\mu}} \right) \frac{X_i}{\hat{\mu}^2} \frac{d\hat{\mu}}{d\theta}. \end{aligned}$$

Therefore

$$\sqrt{n}(\hat{\theta} - \theta_0) = \frac{(\sqrt{n}/n) \sum_{i=1}^{n} g_n(X_i, \theta_0)}{(1/n) \sum_{i=1}^{n} -dg_n(X_i, \theta_0)/d\theta}.$$

Noticing the fact that when $\theta = \theta_0$, $S_1(X_i) - \exp(-X_i/\mu) = 0$ so that

$$\left\{ S_1(X_i) - \exp\left(-\frac{X_i}{\mu} \right) \right\} \left\{ \frac{dS_1(X_i)}{d\theta} - \exp\left(-\frac{X_i}{\mu} \right) \frac{X_i}{\mu^2} \frac{d\mu}{d\theta} \right\} = 0.$$

Therefore,

$$\sum_{i=1}^{n} g_n(X_i, \theta_0) = \text{Term 1} + \text{Term 2} + \text{Term 3} + \text{Term 4}.$$

Where

$$\begin{aligned} &\text{Term 1} \\ &= \sum_{i=1}^{n} \hat{S}_1(X_i) \frac{d\hat{S}_1(X_i)}{d\theta} - S_1(X_i) \frac{dS_1(X_i)}{d\theta} \\ &= \sum_{i=1}^{n} \hat{S}_1(X_i) \frac{d\hat{S}_1(X_i)}{d\theta} - \hat{S}_1(X_i) \frac{dS_1(X_i)}{d\theta} + \hat{S}_1(X_i) \frac{dS_1(X_i)}{d\theta} - S_1(X_i) \frac{dS_1(X_i)}{d\theta} \\ &= \sum_{i=1}^{n} \hat{S}_1(X_i) \left\{ \frac{d\hat{S}_1(X_i)}{d\theta} - \frac{dS_1(X_i)}{d\theta} \right\} + \frac{dS_1(X_i)}{d\theta} \left\{ \hat{S}_1(X_i) - S_1(X_i) \right\}. \end{aligned}$$

Here

$$\hat{S}_1(t) = \psi_\theta \left\{ -\int_0^t \psi_\theta^{-1'}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{d\bar{N}(u)}{\bar{Y}(u)} \right\},$$

and

$$\begin{aligned}
&\frac{d\hat{S}_1(t)}{d\theta} \\
&= \psi'_\theta \left\{ -\int_0^t \psi_\theta^{-1'}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{d\bar{N}(u)}{\bar{Y}(u)} \right\} \left\{ -\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{d\bar{N}(u)}{\bar{Y}(u)} \right\} \\
&= \left\{ -\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{d\bar{N}(u)}{\bar{Y}(u)} \right\} \Big/ \psi_\theta^{-1'}\left[\psi_\theta \left\{ -\int_0^t \psi_\theta^{-1'}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{d\bar{N}(u)}{\bar{Y}(u)} \right\} \right] \\
&= \left\{ -\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{d\bar{N}(u)}{\bar{Y}(u)} \right\} \Big/ \psi_\theta^{-1'}[\hat{S}_1(t)],
\end{aligned}$$

hence

$$\begin{aligned}
&\frac{d\hat{S}_1(t)}{d\theta} - \frac{dS_1(t)}{d\theta} \\
&= \left\{ -\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{d\bar{N}(u)}{\bar{Y}(u)} \right\} \Big/ \psi_\theta^{-1'}[\hat{S}_1(t)] \\
&\quad - \left\{ -\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u) \right\} \Big/ \psi_\theta^{-1}[S_1(t)] \\
&= \left\{ -\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{d\bar{N}(u)}{\bar{Y}(u)} \right\} \Big/ \psi_\theta^{-1'}[\hat{S}_1(t)] \\
&\quad - \left\{ -\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u) \right\} \Big/ \psi_\theta^{-1'}[\hat{S}_1(t)] \\
&\quad + \left\{ -\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u) \right\} \Big/ \psi_\theta^{-1'}[\hat{S}_1(t)] \\
&\quad - \left\{ -\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u) \right\} \Big/ \psi_\theta^{-1}[S_1(t)] \\
&= \frac{1}{\psi_\theta^{-1'}[\hat{S}_1(t)]} \left[-\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{\Delta\bar{N}(u)}{\bar{Y}(u)} + \int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u) \right] \\
&\quad + \left\{ -\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u) \right\} \left[\frac{1}{\psi_\theta^{-1'}[\hat{S}_1(t)]} - \frac{1}{\psi_\theta^{-1'}[S_1(t)]} \right].
\end{aligned}$$

Mimicking the arguments given in Rivest and Wells (2001), we can show that

$$-\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{d\bar{N}(u)}{\bar{Y}(u)} + \int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u)$$
$$= -\frac{1}{n}\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\left\{\frac{\bar{Y}}{n}\right\}d\bar{M}(u) + \int_0^t \Big[-\frac{d\psi_\theta^{-1'}}{d\theta}\left\{\frac{\bar{Y}}{n}\right\}\frac{\bar{Y}(u)}{n}$$
$$+ \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u)\Big] d\Lambda^\sharp(u),$$

and also

$$-\frac{1}{\sqrt{n}}\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\left\{\frac{\bar{Y}(u)}{n}\right\}d\bar{M}(u) = -\frac{1}{\sqrt{n}}\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}d\bar{M}(u) + o_p(1).$$
$$-\frac{d\psi_\theta^{-1'}}{d\theta}\left\{\frac{\bar{Y}}{n}\right\}\frac{\bar{Y}(u)}{n} + \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u) = \Phi_1'(\pi(u))\left[\frac{\bar{Y}(u)}{n} - \pi(u)\right] + o_p(n^{-1/2})$$
$$= \Phi_1'(\pi(u))\frac{1}{n}\sum_{i=1}^n [I(X_i > u) - \pi(u)] + o_p(n^{-1/2}),$$

where Φ_1 is defined as $\Phi_1(s) = -s(d\psi_\theta^{-1'}/d\theta)(s)$. Combining these two results, we have

$$-\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{d\bar{N}(u)}{\bar{Y}(u)} + \int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u)$$
$$= \frac{1}{\sqrt{n}}\Big[-\frac{1}{\sqrt{n}}\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}d\bar{M}(u)$$
$$+ \frac{1}{\sqrt{n}}\int_0^t \sum_{i=1}^n \Phi_1'(\pi(u))\left[I(X_i > u) - \pi(u)\right] d\Lambda^\sharp(u) + o_p(1)\Big].$$

Similarly, we have

$$\left\{-\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u)\right\}\left[\frac{1}{\psi_\theta^{-1'}[\hat{S}_1(t)]} - \frac{1}{\psi_\theta^{-1'}[S_1(t)]}\right]$$
$$= \left\{-\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u)\right\}\frac{\psi_\theta^{-1'}(S_1(t))}{[\psi_\theta^{-1'}(S_1(t))]^2}[\hat{S}_1(t) - S_1(t)]$$
$$= \left\{-\int_0^t \frac{d\psi_\theta^{-1'}}{d\theta}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u)\right\}\frac{\psi_\theta^{-1'}(S_1(t))}{[\psi_\theta^{-1'}(S_1(t))]^2}\frac{1}{[\psi_\theta^{-1'}(S_1(t))]}$$
$$\times \left\{ -\int_0^t \psi_\theta^{-1'}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{d\bar{N}(u)}{\bar{Y}(u)} + \int_0^t \psi_\theta^{-1'}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u)\right\}$$

$$= \left\{ - \int_0^t \frac{d\psi_\theta^{-1'}}{d\theta} \{\pi(u)\} \pi(u) d\Lambda^\sharp(u) \right\} \frac{\psi_\theta^{-1'}(S_1(t))}{[\psi_\theta^{-1'}(S_1(t))]^3}$$

$$\times \frac{1}{\sqrt{n}} \Bigg[- \frac{1}{\sqrt{n}} \int_0^t \psi_\theta^{-1'} \{\pi(u)\} d\bar{M}(u)$$

$$+ \frac{1}{\sqrt{n}} \int_0^t \sum_{i=1}^n \Phi'(\pi(u)) \left[I(X_i > u) - \pi(u) \right] d\Lambda^\sharp(u) + o_p(1) \Bigg],$$

where $\Phi(s) = -s\psi_\theta^{-1'}(s)$. Therefore,

$$\sqrt{n} \left[\frac{d\hat{S}_1(X_i)}{d\theta} - \frac{dS_1(X_i)}{d\theta} \right]$$

$$= \frac{1}{\sqrt{n}} \frac{1}{\psi_\theta^{-1'}[S_1(X_i)]} \int_0^{X_i} - \frac{d\psi_\theta^{-1'}}{d\theta} \{\pi(u)\} d\bar{M}(u)$$

$$+ \frac{1}{\sqrt{n}} \frac{1}{\psi_\theta^{-1'}[S_1(X_i)]} \int_0^{X_i} \sum_{j=1}^n \Phi_1'(\pi(u)) \left[I(X_j > u) - \pi(u) \right] d\Lambda^\sharp(u)$$

$$+ \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ - \int_0^{X_i} \frac{d\psi_\theta^{-1'}}{d\theta} \{\pi(u)\} \pi(u) d\Lambda^\sharp(u) \right\} \frac{\psi_\theta^{-1'}(S_1(X_i))}{[\psi_\theta^{-1'}(S_1(X_i))]^3}$$

$$\times \left[\int_0^{X_i} -\psi_\theta^{-1'} \{\pi(u)\} dM_j(u) + \int_0^{X_i} \Phi'(\pi(u)) \left[I(X_j > u) - \pi(u) \right] d\Lambda^\sharp(u) \right]$$

$$+ o_p(1)$$

$$= \frac{1}{\sqrt{n}} \sum_{j=1}^n \Bigg[\frac{1}{\psi_\theta^{-1'}[S_1(X_i)]} \int_0^{X_i} -\psi_\theta^{-1'} \{\pi(u)\} dM_j(u)$$

$$+ \frac{1}{\psi_\theta^{-1'}[S_1(X_i)]} \int_0^{X_i} \Phi_1'(\pi(u)) \left[I(X_j > u) - \pi(u) \right] d\Lambda^\sharp(u) \Bigg]$$

$$+ \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ - \int_0^{X_i} \frac{d\psi_\theta^{-1'}}{d\theta} \{\pi(u)\} \pi(u) d\Lambda^\sharp(u) \right\} \frac{\psi_\theta^{-1'}(S_1(X_i))}{[\psi_\theta^{-1'}(S_1(X_i))]^3}$$

$$\times \left[\int_0^{X_i} -\psi_\theta^{-1'} \{\pi(u)\} dM_j(u) + \int_0^{X_i} \Phi'(\pi(u)) \left[I(X_j > u) - \pi(u) \right] d\Lambda^\sharp(u) \right]$$

$$+ o_p(1),$$

and

$$\sqrt{n}[\hat{S}_1(X_i) - S_1(X_i)]$$

$$
\begin{aligned}
&= \frac{1}{[\psi_\theta^{-1\prime}(S_1(X_i))]}\Bigg\{ -\int_0^{X_i} \psi_\theta^{-1\prime}\{\hat{\pi}(u)\}\hat{\pi}(u)\frac{d\bar{N}(u)}{\bar{Y}(u)} \\
&\quad + \int_0^{X_i} \psi_\theta^{-1\prime}\{\pi(u)\}\pi(u)d\Lambda^\sharp(u)\Bigg\} \\
&= \frac{1}{\sqrt{n}}\frac{1}{[\psi_\theta^{-1\prime}(S_1(X_i))]}\Bigg[\int_0^{X_i} -\psi_\theta^{-1\prime}\{\pi(u)\}d\bar{M}(u) \\
&\quad + \frac{1}{\sqrt{n}}\int_0^{X_i}\sum_{j=1}^n \Phi'(\pi(u))\left[I(X_j > u) - \pi(u)\right]d\Lambda^\sharp(u)\Bigg] + o_p(1) \\
&= \frac{1}{\sqrt{n}}\frac{1}{[\psi_\theta^{-1\prime}(S_1(X_i))]}\sum_{j=1}^n\Bigg[\int_0^{X_i} -\psi_\theta^{-1\prime}\{\pi(u)\}dM_j(u) \\
&\quad + \int_0^{X_i} \Phi'(\pi(u))\left[I(X_j > u) - \pi(u)\right]d\Lambda^\sharp(u)\Bigg] + o_p(1).
\end{aligned}
$$

Hence $(\sqrt{n}/n)$Term 1 can be expressed as $\sqrt{n}$ times a linear combination of order 2 U statistics, therefore it is asymptotically normal. Now

$$
\begin{aligned}
\text{Term 2 } &= -\sum_{i=1}^n \hat{S}_1(X_i)\exp\left(-\frac{X_i}{\hat{\mu}}\right)\frac{X_i}{\hat{\mu}^2}\frac{d\hat{\mu}}{d\theta} + S_1(X_i)\exp\left(-\frac{X_i}{\mu}\right)\frac{X_i}{\mu^2}\frac{d\mu}{d\theta}. \\
&= -\sum_{i=1}^n \exp\left(-\frac{X_i}{\mu}\right)\frac{X_i}{\mu^2}\frac{d\mu}{d\theta}[\hat{S}_1(X_i) - S_1(X_i)] \\
&\quad + S_1(X_i)\left[\exp\left(-\frac{X_i}{\hat{\mu}}\right)\frac{X_i}{\hat{\mu}^2}\frac{d\hat{\mu}}{d\theta} - \exp\left(-\frac{X_i}{\mu}\right)\frac{X_i}{\mu^2}\frac{d\mu}{d\theta}\right] + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= -\sum_{i=1}^n \exp\left(-\frac{X_i}{\mu}\right)\frac{X_i}{\mu^2}\frac{d\mu}{d\theta}[\hat{S}_1(X_i) - S_1(X_i)] \\
&\quad + S_1(X_i)\frac{d\mu}{d\theta}\left[\exp\left(-\frac{X_i}{\hat{\mu}}\right)\frac{X_i}{\hat{\mu}^2} - \exp\left(-\frac{X_i}{\mu}\right)\frac{X_i}{\mu^2}\right] \\
&\quad + S_1(X_i)\exp\left(-\frac{X_i}{\mu}\right)\frac{X_i}{\mu^2}\left[\frac{d\hat{\mu}}{d\theta} - \frac{d\mu}{d\theta}\right] + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= -\sum_{i=1}^n \exp\left(-\frac{X_i}{\mu}\right)\frac{X_i}{\mu^2}\frac{d\mu}{d\theta}[\hat{S}_1(X_i) - S_1(X_i)] \\
&\quad + S_1(X_i)\frac{d\mu}{d\theta}\left[-\frac{X_i^2}{\mu^2}\exp\left(-\frac{X_i}{\mu}\right) + \frac{2X_i}{\mu}\exp\left(-\frac{X_i}{\mu}\right)\right]\left\{\frac{1}{\hat{\mu}} - \frac{1}{\mu}\right\} \\
&\quad + S_1(X_i)\exp\left(-\frac{X_i}{\mu}\right)\frac{X_i}{\mu^2}\left[\frac{d\hat{\mu}}{d\theta} - \frac{d\mu}{d\theta}\right] + o_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}
$$

Noticing the fact that

$$\begin{aligned}\sqrt{n}\Big\{\frac{1}{\hat{\mu}}-\frac{1}{\mu}\Big\} &= \frac{\sqrt{n}}{n}\sum_{j=1}^{n}\Big\{\frac{-\log(\hat{S}(X_j))}{X_j}-E\Big(\frac{-\log(S(X))}{X}\Big)\Big\}\\ &= \frac{\sqrt{n}}{n}\sum_{j=1}^{n}\Big[\frac{-\log(\hat{S}(X_j))}{X_j}-\frac{-\log(S(X_j))}{X_j}\\ &\quad+\frac{-\log(S(X_j))}{X_j}-E\Big(\frac{-\log(S(X))}{X}\Big)\Big]\\ &= \frac{\sqrt{n}}{n}\sum_{j=1}^{n}\Big[\frac{-\hat{S}(X_j)}{S(X_j)X_j}-\frac{-S(X_j)}{S(X_j)X_j}\\ &\quad+\frac{-\log(S(X_j))}{X_j}-E\Big(\frac{-\log(S(X))}{X}\Big)+o_p(1)\Big]\\ &= \frac{1}{n^{3/2}}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\int_0^{X_j}-\psi_\theta^{-1\prime}\{\pi(u)\}dM_k(u)}{\psi_\theta^{-1\prime}(S_1(X_j))S_1(X_j)X_j}\\ &\quad+\frac{1}{n^{3/2}}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\int_0^{X_j}\Phi'(\pi(u))\,[I(X_k>u)-\pi(u)]\,d\Lambda^\sharp(u)]}{\psi_\theta^{-1\prime}(S_1(X_j))S(X_j)X_j}\\ &\quad+\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\Big[\frac{-\log(S(X_j))}{X_j}-E\Big(\frac{-\log(S(X))}{X}\Big)\Big]+o_p(1),\end{aligned}$$

and

$$\begin{aligned}&\sqrt{n}\Big(\frac{d\hat{\mu}}{d\theta}-\frac{d\mu}{d\theta}\Big) = \frac{\sqrt{n}}{\hat{\mu}^2}\frac{1}{n}\sum_{j=1}^{n}\Big[\frac{-d\hat{S}(X_j)/d\theta}{\hat{S}(X_j)X_j}-E\Big(-\frac{dS(X)/d\theta}{S(X)X}\Big)\Big]\\ &= \frac{1}{\hat{\mu}^2}\frac{\sqrt{n}}{n}\sum_{j=1}^{n}\Big[\frac{-d\hat{S}(X_j)/d\theta}{\hat{S}(X_j)X_j}+\frac{d\hat{S}(X_j)/d\theta}{S(X_j)X_j}-\frac{d\hat{S}(X_j)/d\theta}{S(X_j)X_j}\\ &\quad+\frac{dS(X_j)/d\theta}{S(X_j)X_j}-\frac{dS(X_j)/d\theta}{S(X_j)X_j}+E\Big(\frac{dS(X)/d\theta}{S(X)X}\Big)\Big]\\ &= \frac{1}{\hat{\mu}^2}\frac{\sqrt{n}}{n}\sum_{j=1}^{n}\Big[\frac{-dS(X_j)/d\theta}{X_j}\Big(\frac{1}{\hat{S}(X_j)}-\frac{1}{S(X_j)}\Big)\\ &\quad-\frac{1}{S(X_j)X_j}\Big(\frac{d\hat{S}(X_j)}{d\theta}-\frac{dS(X_j)}{d\theta}\Big)-\frac{dS(X_j)/d\theta}{S(X_j)X_j}+E\Big(\frac{dS(X)/d\theta}{S(X)X}\Big)\Big]\\ &= \frac{1}{\mu^2}\frac{\sqrt{n}}{n}\sum_{j=1}^{n}\Big[\frac{-[dS(X_j)/d\theta]^2}{S(X_j)^2X_j}(\hat{S}(X_j)-S(X_j))\end{aligned}$$

$$-\frac{1}{S(X_j)X_j}\left(\frac{d\hat{S}(X_j)}{d\theta}-\frac{dS(X_j)}{d\theta}\right)-\frac{dS(X_j)/d\theta}{S(X_j)X_j}+E\left(\frac{dS(X)/d\theta}{S(X)X}\right)\Bigg]+o_p(1).$$

From above expression, we can easily show that $(\sqrt{n}/n)$Term 2 can be expressed as $\sqrt{n}$ times a linear combination of order 2 or order 3 U statistics following the same way we have used for proving the asymptotic normality of $(\sqrt{n}/n)$Term 1, hence $(\sqrt{n}/n)$Term 2 is asymptotically normal. Similar arguments can be applied to prove the asymptotic normality of $(\sqrt{n}/n)$Term 3 and $(\sqrt{n}/n)$Term 4. We can conclude that our parameter estimator is asymptotically normal based on above derivations.

References

Carrière, J. (1995). Removing cancer when it is correlated with other causes of death. *Biome. J.* **37**, 339–350.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141–151.

Escarela, G. and Carrierè, J. (2003). Fitting competing risks with an assumed copula. *Stat. Methods Med. Res.* **12**, 333–349.

Fine, J. P., Jiang, H. and Chappell, R. (2001). On semi-competing risks data. *Biometrika* **88**, 907–919.

Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis.* John Wiley & Sons, Inc, Hoboken.

Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika* **74**, 549–555.

Heckman, J. J. and Honoré, B. E. (1989). The identifiability of the competing risks model. *Biometrika* **76**, 325–330.

Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika* **73**, 671–678.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53**, 457–481.

Klein, J and Moeschberger, M. (1997). *Survival Analysis: Techniques for Censored and Truncated Data.* Springer, New York.

Lakhal, L., Rivest, L.-P. and Abdous, B. (2008). Estimating survival and association in a semi-competing risks model. *Biometrics* **64**, 180–188.

Nelsen, R. B. (2007). *An Introduction to Copulas.* Springer Science & Business Media, New York.

Rivest, L.-P. and Wells M. T. (2001). A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. *J. Multivariate Anal.* **79**, 138–155.

Tsiatis, A. (1975). A nonidentifiability of the problem of competing risks. *Proc. Natl. Acad. Sci. USA* **72**, 20–22.

van der Vaart, A. W. (2007). *Asymptotic Statistics.* Cambridge University Press, Cambridge.

Wang, A. (2012). On the nonidentifiability property of Archimedean copula models under dependent censoring. *Statist. Probab.* **82**, 621–625.

Wang, A. (2014). Properties of the marginal survival functions for dependent censored data under an assumed archimedean copula. *J. Multivariate Anal.* **129**, 57–68.

Wang, A., Chandra, K., Xu, R. and Sun, J. (2015). The identifiability of dependent competing risks models induced by bivariate frailty models. *Scand. J. Statist.* **42**, 427–437.

Zeng, D., Lin, D. and Lin, X. (2008). Semiparametric transformation models with random effects for clustered failure time data. *Statist. Sinica* **18**, 355–377.

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

Antai Wang

Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, NJ 07102, USA.

E-mail: aw224@njit.edu

(Received February 2020; accepted August 2021)