

Article

Likelihood Inference for Copula Models Based on Left-Truncated and Competing Risks Data from Field Studies

Hirofumi Michimae ¹  and Takeshi Emura ^{2,3,*}

¹ Department of Clinical Medicine (Biostatistics), School of Pharmacy, Kitasato University, Tokyo 108-8641, Japan; michimaeh@pharm.kitasato-u.ac.jp

² Biostatistics Center, Kurume University, Kurume 830-0011, Japan

³ Research Center for Medical and Health Data Science, The Institute of Statistical Mathematics, Tokyo 190-8562, Japan

* Correspondence: takeshiemura@gmail.com

Abstract: Survival and reliability analyses deal with incomplete failure time data, such as censored and truncated data. Recently, the classical left-truncation scheme was generalized to analyze “field data”, defined as samples collected within a fixed period. However, existing competing risks models dealing with left-truncated field data are not flexible enough. We propose copula-based competing risks models for latent failure times, permitting a flexible parametric form. We formulate maximum likelihood estimation methods under the Weibull, lognormal, and gamma distributions for the latent failure times. We conduct simulations to check the performance of the proposed methods. We finally give a real data example. We provide the R code to reproduce the simulations and data analysis results.

Keywords: censoring; competing risk; left-truncation; lognormal distribution; multivariate survival analysis; Weibull distribution; reliability

MSC: 62H12; 62N01; 62N05; 62P30



Citation: Michimae, H.; Emura, T. Likelihood Inference for Copula Models Based on Left-Truncated and Competing Risks Data from Field Studies. *Mathematics* **2022**, *10*, 2163. <https://doi.org/10.3390/math10132163>

Academic Editor: Yuhlong Lio

Received: 22 April 2022

Accepted: 20 June 2022

Published: 21 June 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Survival analysis deals with incomplete failure time data, such as censored and truncated data [1–5]. *Left-truncation* arises if early failures occur before the data collection period and their failures are simply ignored. *Right-censoring* arises if late failures occur after the data collection period, where their exact times of failures are unascertainable. *Competing risks* arise if the failure event of interest becomes unavailable by the occurrence of other failure events.

Naïve statistical analyses of truncated, censored, and competing risks data lead to biased inference for the population of interest. Left-truncation, right-censoring, and competing risks are a classical topic in survival analysis that is widely recognized in both biostatistics [3] and reliability engineering [1–5].

Hong et al. [1] generalized the classical truncation scheme to a more complex scheme found in the “field data”, defined as samples collected within a fixed period (see Section 2). The well-known example is the failure time dataset for electric power transformers. Hong et al. [1] separated the samples into two parts: a truncated part (e.g., transformers installed before 1980) and an untruncated part (e.g., transformers installed after 1980), and then combined them into a single likelihood function. This type of truncated field data arises in many other practical applications, where there is a birth process (the installing process) of individuals and a certain data collection period of failure [6–9].

For left-truncated field reliability data, a variety of distributions have been considered for statistical inference. The inference methods for the lognormal and Weibull distributions were developed by [1,10–12]. The gamma and the generalized gamma distribution were

considered in [13,14]. Comparison of the numerical methods for maximum likelihood estimator were considered by [12,15]. Bayesian approaches for the lognormal, Weibull, and gamma distributions were considered by [2,16]. The Lehmann family of distributions was considered by [17] and the log-location-scale family of distributions was considered by [18]. The spline model was proposed by [7]. In this research topic, the most up-to-date review paper is [2].

Competing risks could occur together with left-truncation. Competing risks means that there are latent failure times to events, and only one of the failure times is observable. As competing risks are dependent, copula-based dependent risks models have been extensively utilized [19–25], which formulate the joint distribution of latent failure times. However, in analysis of left-truncated field data, copula-based dependent risks models have not been proposed yet.

Existing competing risks analyses for left-truncated field data are limited to the following models: the independent Weibull model with the common shape parameter [26], the Marshall–Olkin bivariate Rayleigh model [27], and the Marshall–Olkin bivariate Weibull model [4]. While these three models provide an important starting point, their models are somewhat specific for modeling latent failure times of competing risks. Therefore, there is much room for proposing more flexible models.

In this paper, we will develop copula-based models for dependent competing risks under left-truncation, permitting more flexible models for dependence structure and failure time distributions than the existing three models of [4,26,27]. Due to our main interest in applications to reliability analyses, we adopt the standard parametric distributions, including the Weibull, lognormal, and gamma distributions. These failure time distributions are commonly used in reliability studies, and also found in recent engineering applications [28–30]. While the proposed copula models permit a variety of copulas, we nonetheless focus on the Clayton copula for modeling dependence among failure times to facilitate computation.

The remainder of the paper is structured as follows. Section 2 provides the backgrounds and key concepts, such as field failure data and competing risks. Section 3 introduces copula models and develops likelihood functions for the Weibull, lognormal, gamma marginal models. Section 4 reports simulation studies to check the performance of the proposed methods. In Section 5, we analyze a dataset for illustration. Section 6 concludes the paper with future directions.

2. Left-Truncation and Competing Risks

We review the structure of left-truncation and competing risks for data collected from field studies.

Suppose that there are two mutually exclusive events for failure, named Event 1 and Event 2. Let (T_1, T_2) be a pair of failure times for Event 1 and Event 2, respectively. Under competing risks, the failure time is defined by $T_{12} \equiv \min(T_1, T_2)$, where only one of (T_1, T_2) is observable. For instance, if $T_1 < T_2$, the observable failure time is $T_{12} = T_1$, and the exact value of T_2 is unknown. The pair (T_1, T_2) is called “latent failure times” since only one of T_1 and T_2 is observable.

Figure 1 explains left-truncation in the data collection scheme, where the individuals were born *randomly* over time. However, we cannot obtain samples of all the individuals since the data collection starts only after time s (the starting time). If individuals are born and then fail before time s , they are discarded and not observable; see “Not observed” in Figure 1. This leads to biased sampling.

To define left-truncation mathematically, we let τ be the time from birth to time s (Figure 1). For individuals to be included in our samples, their failure time T_{12} must be greater than τ (Figure 1). Those who were born before s are defined as “truncated samples”, which are all conditional on $T_{12} \geq \tau$. The individuals born after s are always included in our samples. Therefore, our samples are divided into truncated ones (born before s) and untruncated ones (born after s). For untruncated samples, the sample is observed without any condition, and τ is undefined.

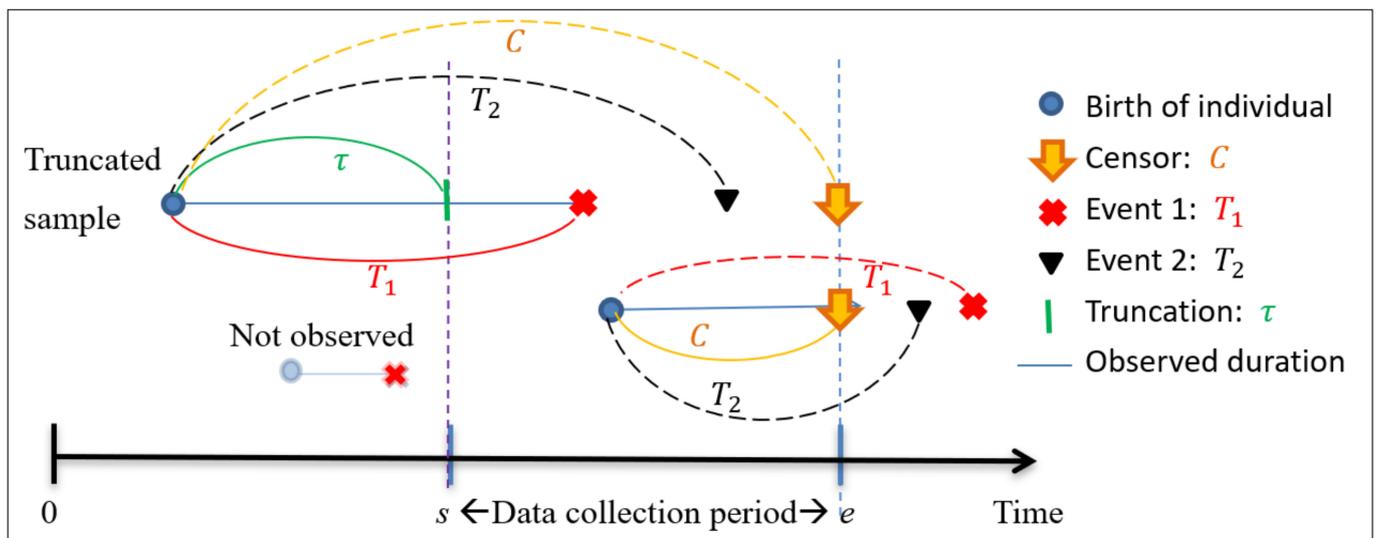


Figure 1. Left-truncated and competing risks data from a field study. Event 1 and Event 2 are subject to competing risks. Solid curves show observable times, and dashed curves show unobservable times. One can observe Event 1, Event 2, or Censoring, whichever occurs first, between the starting time (s) and the ending time (e). One cannot observe anything if Event 1 or Event 2 occurs before time s .

To signify the difference between truncated and untruncated samples, we define the truncation indicator,

$$v = \begin{cases} 0, & \text{if the sample is truncated } (T_{12} \geq \tau \geq 0), \\ 1, & \text{if the sample is untruncated } (\tau = \text{undefined}). \end{cases}$$

There could be a random censoring time C that may censor the value of T_1 and T_2 , e.g., the duration of the data collection (Figure 1). If both Event 1 and Event 2 do not occur within the data collection period, one cannot exactly know T_1 and T_2 . One only knows that T_1 and T_2 are greater than the observed censoring time C that is defined as time from birth to e (the ending time).

In this sense, what we actually observe is $T = \min(T_1, T_2, C)$. The sample is censored if both $T_1 > C$ and $T_2 > C$ hold; otherwise, the sample is uncensored. We assume that two failures do not occur simultaneously, i.e., $T_1 = T_2$ never occurs.

Under competing risks, there are three possible cases: The first case is where Event 1 is observable by $T = T_1$ and $\{T_1 < T_2, T_1 < C\}$. The second case is where Event 2 is observable by $T = T_2$ and $\{T_2 < T_1, T_2 < C\}$. The last case is where neither Event 1 nor Event 2 is observable by $T = C$ and $\{C < T_1, C < T_2\}$. The first and last cases are shown in Figure 1.

In a dataset, there are n samples denoted as $\{(t_i, \tau_i, v_i); i = 1, 2, \dots, n\}$. In addition, as we know the event status of each individual, the samples are divided into three sets:

$$\begin{aligned} l_1 &= \{i : T_{1i} < T_{2i}, T_{1i} < C_i\}, \\ l_2 &= \{i : T_{2i} < T_{1i}, T_{2i} < C_i\}, \\ l_0 &= \{i : C_i < T_{1i}, C_i < T_{2i}\}. \end{aligned}$$

Since each sample belongs to one of the three sets, it holds that $l_1 \cup l_2 \cup l_0 = \{1, 2, \dots, n\}$.

3. Proposed Methods

In this section, we propose a copula-based model and a likelihood-based inference method.

3.1. Copula Model for Competing Risks

We use a copula model for bivariate competing risks as originally proposed by [19]. Let (T_1, T_2) be a pair of failure times whose marginal survival functions are $S_i(t) = P(T_i > t)$,

$i \in \{1, 2\}$. We assume that the failure times are continuous and hence have the density functions $f_i(t) = -dS_i(t)/dt, i \in \{1, 2\}$. To model the dependence of T_1 and T_2 , we assume a survival copula model [21]

$$P(T_1 > t_1, T_2 > t_2) = C_\theta(S_1(t_1), S_2(t_2)),$$

where C_θ is a parametric copula [31] with parameter θ .

Copulas are popular in statistical analyses for bivariate models involving competing risks [20–25,32–34]. There exist a number of copulas [31], and new copulas keep emerging [35,36]. Copulas are applied to a variety of fields, including reliability [37], ecology [38], meta-analysis [39,40], and econometrics [41]. However, to reduce the complexity of estimation, the copula has to be as simple as possible.

For computational simplicity, we particularly adopt the Clayton copula [31]

$$C_\theta(u, v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta}, \theta > 0,$$

where θ specifies dependence between T_1 and T_2 , and is transformed to Kendall’s tau: $\theta/(\theta + 2)$. Hence, T_1 and T_2 are positively dependent whose strength increases as θ gets large. If $\theta \rightarrow 0$, the Clayton copula reduces to $C(u, v) = uv$, the independent risks model of [26]. The Clayton copula is one of the most frequently used copulas in competing risks models [22,24] and other models [33,40–42] due to its remarkable mathematical tractability.

Other copulas might be considered. However, we do not suggest fitting many copulas since the true one may not be identifiable from the data. We use the Clayton copula, and then focus on the selection of θ .

3.2. Likelihood Function

We extend the likelihood function for the independent risks models [26] to the copula-based models. Let $\{(t_i, \tau_i, v_i); i = 1, 2, \dots, n\}$ be observed data as previously defined. Let (T_{1i}, T_{2i}) be a pair of latent failure times for Event 1 and Event 2, respectively. In addition, let C_i be independent and uninformative censoring times. Let $T_i = \min(T_{1i}, T_{2i}, C_i)$, and t_i be its observed value. For a sample without truncation ($v_i = 1$), the likelihood is divided into three cases:

- (i) $L_i = P(T_{1i} = t_i, t_i < T_{2i}, t_i < C_i)$ if $i \in l_1 = \{i : T_{1i} < T_{2i}, T_{1i} < C_i\}$ and $v_i = 1$,
- (ii) $L_i = P(T_{2i} = t_i, t_i < T_{1i}, t_i < C_i)$ if $i \in l_2 = \{i : T_{2i} < T_{1i}, T_{2i} < C_i\}$ and $v_i = 1$,
- (iii) $L_i = P(C_i = t_i, t_i < T_{1i}, t_i < T_{2i})$ if $i \in l_0 = \{i : C_i < T_{1i}, C_i < T_{2i}\}$ and $v_i = 1$,

For a sample with truncation ($v_i = 0$), the likelihood is divided into other three cases:

- (iv) $L_i = P(T_{1i} = t_i, t_i < T_{2i}, t_i < C_i | T_{1i} \geq \tau_i, T_{2i} \geq \tau_i, C_i \geq \tau_i)$ if $i \in l_1$ and $v_i = 0$,
- (v) $L_i = P(T_{2i} = t_i, t_i < T_{1i}, t_i < C_i | T_{1i} \geq \tau_i, T_{2i} \geq \tau_i, C_i \geq \tau_i)$ if $i \in l_2$ and $v_i = 0$,
- (vi) $L_i = P(C_i = t_i, t_i < T_{1i}, t_i < T_{2i} | T_{1i} \geq \tau_i, T_{2i} \geq \tau_i, C_i \geq \tau_i)$ if $i \in l_0$ and $v_i = 0$.

In the last three cases, the “ $C_i \geq \tau_i$ ” can be dropped. Combining all the six cases and ignoring the constant factors for censoring, the likelihood function is

$$\begin{aligned} L \equiv \prod_i L_i &= \prod_{i \in l_1} P(T_{1i} = t_i, t_i < T_{2i})^{v_i} P(T_{1i} = t_i, t_i < T_{2i} | T_{1i} \geq \tau_i, T_{2i} \geq \tau_i)^{1-v_i} \\ &\quad \times \prod_{i \in l_2} P(T_{2i} = t_i, t_i < T_{1i})^{v_i} P(T_{2i} = t_i, t_i < T_{1i} | T_{1i} \geq \tau_i, T_{2i} \geq \tau_i)^{1-v_i} \\ &\quad \times \prod_{i \in l_0} P(t_i < T_{1i}, t_i < T_{2i})^{v_i} P(t_i < T_{1i}, t_i < T_{2i} | T_{1i} \geq \tau_i, T_{2i} \geq \tau_i)^{1-v_i}. \end{aligned}$$

After some simplification, the likelihood function becomes

$$\begin{aligned} L &= \prod_{i \in l_1} P(T_{1i} = t_i, t_i < T_{2i}) \times \prod_{i \in l_2} P(T_{2i} = t_i, t_i < T_{1i}) \\ &\quad \times \prod_{i \in l_0} P(t_i < T_{1i}, t_i < T_{2i}) \prod_{v_i=0} P(T_{1i} \geq \tau_i, T_{2i} \geq \tau_i)^{-1}. \end{aligned}$$

To estimate parameters in the model, the maximum likelihood estimator (MLE) is adopted under parametric models for $S_1(\cdot)$ and $S_2(\cdot)$. We will consider commonly used parametric models for failure time distributions in reliability, including the Weibull, lognormal, and gamma distributions. Recent engineering applications of these distributions are found in [28–30]. The Weibull distribution will be detailed in Section 3.3 and the lognormal and gamma distributions in Appendices A and B.

3.3. Weibull Model

We set $S_1(t) = \exp(-\lambda_1 t^{\alpha_1})$ and $S_2(t) = \exp(-\lambda_2 t^{\alpha_2})$. Under the independent risks model $C(v, w) = vw$, the likelihood function is

$$L_{Weib}(\alpha_1, \alpha_2, \lambda_1, \lambda_2) = \prod_{i \in I_1} \left\{ \alpha_1 \lambda_1 t_i^{\alpha_1 - 1} \exp(-\lambda_1 t_i^{\alpha_1}) \exp(-\lambda_2 t_i^{\alpha_2}) \right\} \\ \times \prod_{i \in I_2} \left\{ \alpha_2 \lambda_2 t_i^{\alpha_2 - 1} \exp(-\lambda_2 t_i^{\alpha_2}) \exp(-\lambda_1 t_i^{\alpha_1}) \right\} \\ \times \prod_{i \in I_0} \left\{ \exp(-\lambda_1 t_i^{\alpha_1}) \exp(-\lambda_2 t_i^{\alpha_2}) \right\} \times \prod_{v_i=0} \left\{ \exp(-\lambda_1 \tau_i^{\alpha_1}) \exp(-\lambda_2 \tau_i^{\alpha_2}) \right\}^{-1}$$

If we set the common shape parameter $\alpha = \alpha_1 = \alpha_2$, the above likelihood is equivalent to the one derived in Kundu et al. [26]. We do not impose this assumption, and let $\alpha_1 \neq \alpha_2$ for generality.

Under the Clayton copula with parameter θ , the likelihood function is

$$L_{Weib}(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \theta) = \prod_{i \in I_1} \left\{ \alpha_1 \lambda_1 t_i^{\alpha_1 - 1} u_i^{-\theta} \left(u_i^{-\theta} + v_i^{-\theta} - 1 \right)^{-(1+1/\theta)} \right\} \\ \times \prod_{i \in I_2} \left\{ \alpha_2 \lambda_2 t_i^{\alpha_2 - 1} v_i^{-\theta} \left(u_i^{-\theta} + v_i^{-\theta} - 1 \right)^{-(1+1/\theta)} \right\} \\ \times \prod_{i \in I_0} \left(u_i^{-\theta} + v_i^{-\theta} - 1 \right)^{-1/\theta} \times \prod_{v_i=0} \left(x_i^{-\theta} + y_i^{-\theta} - 1 \right)^{1/\theta},$$

where $u_i = \exp(-\lambda_1 t_i^{\alpha_1})$, $v_i = \exp(-\lambda_2 t_i^{\alpha_2})$, $x_i = \exp(-\lambda_1 \tau_i^{\alpha_1})$, and $y_i = \exp(-\lambda_2 \tau_i^{\alpha_2})$. While the likelihood is undefined for $\theta = 0$, by letting $\theta \rightarrow +0$, the likelihood function reduces to the one under the independent risks model.

There are a few different ways to estimate $(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \theta)$ based on $L_{Weib}(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \theta)$. The first one is the MLE under an “assumed” value of θ , which is defined as

$$(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\lambda}_1, \hat{\lambda}_2 | \theta) = \operatorname{argmax}_{(\alpha_1, \alpha_2, \lambda_1, \lambda_2)} \{ \log L_{Weib}(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \theta) \}.$$

The MLE $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\lambda}_1, \hat{\lambda}_2)$ under the independence copula is regarded as $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\lambda}_1, \hat{\lambda}_2 | 0)$. The next one is the full MLE with the unknown θ , which is defined as

$$(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\theta}) = \operatorname{argmax}_{(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \theta)} \{ \log L_{Weib}(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \theta) \}.$$

In either case, one can maximize $\log L$ by applying an iteration algorithm. In R (<https://www.r-project.org/>), there are a variety of optimization functions. We suggest the “optim(.)” function (<https://stat.ethz.ch/R-manual/R-patched/library/stats/html/optim.html>, accessed on 19 June 2022). This function has the option “hessian=TRUE” to yield the Hessian matrix of $-\log L$. The standard error (SE) of the estimates can be computed by the diagonal components of the inverse of the Hessian matrix. Details can be seen from the R code given in Supplementary Materials.

We generally recommend using the full MLE if one wishes to estimate θ . However, the full MLE $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\theta})$ gives large variance since data have little information for θ . In addition, the Hessian matrix could not be invertible, making the SE unavailable. We conjecture that the regularity conditions for the MLE hold with the unknown θ (as shown in [25]), but fail to hold for the full MLE.

An alternative approach to this problem is the pseudo-MLE (PMLE). We first obtain preliminary estimates $(\hat{\alpha}_1^*, \hat{\alpha}_2^*, \hat{\lambda}_1^*, \hat{\lambda}_2^* | \theta^*)$ under an assumed value of θ^* , and obtain $\hat{\theta}^*$ that maximizes $\ell^*(\theta) = \log L_{Weib}(\hat{\alpha}_1^*, \hat{\alpha}_2^*, \hat{\lambda}_1^*, \hat{\lambda}_2^*, \theta)$. The final estimates $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\lambda}_1, \hat{\lambda}_2 | \hat{\theta}^*)$ are then obtained. In our simulation studies, we set $\theta^* = 2/9$ giving Kendall’s tau = 0.1. In practice, however, one may choose a few different values of θ^* , and then, choose the one that maximizes $\ell^*(\hat{\theta}^*)$.

Our numerical studies (Sections 4 and 5) will consider both known and unknown settings of θ .

4. Simulation Studies

We conducted simulation studies to check the performance of the proposed methods of Section 3. To this end, we compared the performance of fitting (i) the independence copula, (ii) the Clayton copula with the known parameter, (iii) the Clayton copula with the parameter estimated by the MLE, and (iv) the Clayton copula with the parameter estimated by the PMLE.

4.1. Simulation Settings

We generated latent failure times from the Weibull model: $T_{1i} \sim S_1(t_1) = \exp(-\lambda_1 t_1^{\alpha_1})$ and $T_{2i} \sim S_2(t_2) = \exp(-\lambda_2 t_2^{\alpha_2})$. Dependence structure for the two failure times is modeled by either the independence copula or the Clayton copula with $\theta = 2$ (Kendall’s tau = 0.5).

Mimicking the sampling scheme of Figure 1, we generated $\tau_i = s - B_i$, and $C_i = e - B_i$, where $B_i \sim \text{Unif}(0, s)$ is the birth time, s is the starting time, and e is the ending time. For truncated samples, we kept samples satisfying $\tau_i \leq t_i = \min(T_{1i}, T_{2i})$. This led to samples $\{(t_i, \tau_i, v_i); i = 1, 2, \dots, n\}$ with $v_i = 0$ for $n/2$ truncated samples and with $v_i = 1$ for $n/2$ untruncated samples.

Based on the samples, the MLE $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\lambda}_1, \hat{\lambda}_2)$ was computed by fitting the independent model or the Clayton copula with θ being known. In addition, the MLE and PMLE $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\theta})$ were computed by fitting the Clayton copula with θ being unknown (the PMLE wasso computed under an assumed value of $\theta^* = 2/9$ corresponding to Kendall’s tau = 0.1).

Based on 1000 repetitions, we calculated the mean, standard deviation (SD), and the mean squared error (MSE) of parameter estimates, and the average of the SEs. For instance, the MSE for α_1 is $E[(\hat{\alpha}_1 - \alpha_1)^2]$, where the expectation is taken by the average of 1000 repetitions.

The following parameter settings were considered to model different shapes of the hazard function (Figure 2):

- (a) **Decreasing hazard:** $(\alpha_i, \lambda_i) = (0.4, 0.1), i \in \{1, 2\}; s = 10; e = 13,$
- (b) **Constant hazard:** $(\alpha_i, \lambda_i) = (1.0, 0.4), i \in \{1, 2\}; s = 10; e = 13,$
- (c) **Increasing hazard:** $(\alpha_i, \lambda_i) = (1.5, 1.0), i \in \{1, 2\}; s = 3; e = 4.$

4.2. Simulation Results

Table 1 shows the means of parameter estimates. We see that the means of parameter estimates are nearly close to their true parameters when the copula is correctly specified. For instance, when the true copula is the Clayton copula with $\theta = 2$, accurate estimates are obtained by fitting the same copula. If the true copula is wrongly specified, the estimates are biased (e.g., by fitting the Clayton copula with $\theta = 2$ under the independence copula). This implies the importance of selecting the correct value of θ to estimate the parameter. If θ is treated as unknown and estimated by data, the biases of estimates can often be large. These biases exist for both MLE and PMLE methods for dealing with the unknown θ . This unavoidable bias comes from the well-known difficulty of estimating θ under competing risks.

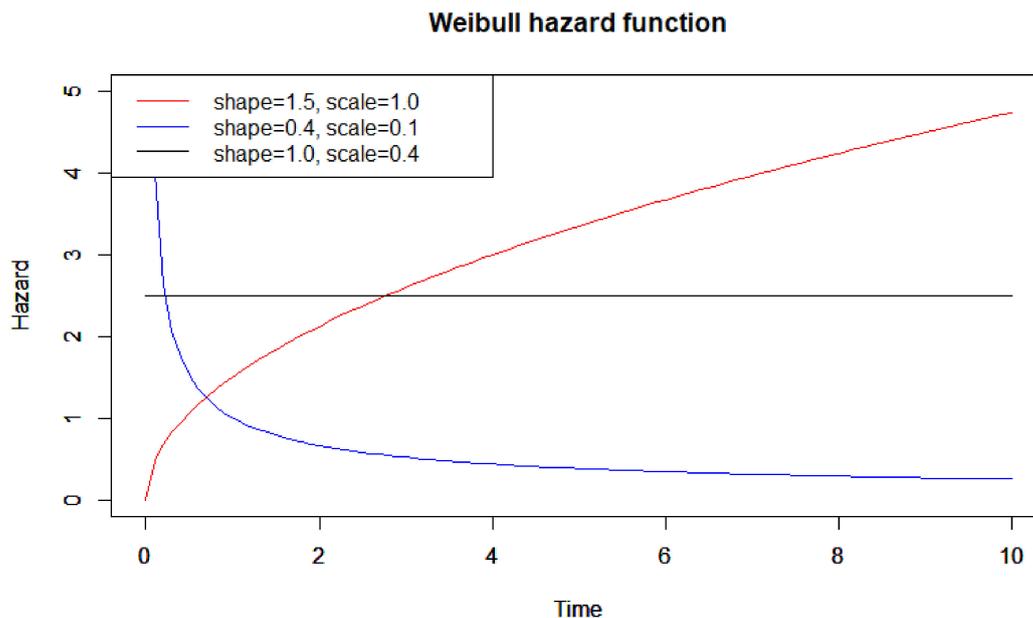


Figure 2. The shape parameter α_i and the scale parameter λ_i used in the simulation study for the Weibull model: $T_1 \sim S_1(t_1) = \exp(-\lambda_1 t_1^{\alpha_1})$ and $T_2 \sim S_2(t_2) = \exp(-\lambda_2 t_2^{\alpha_2})$.

Table 1. The means (biases) of parameter estimates based on 1000 simulation runs.

True Model	Fitted Model	Event 1		Event 2	
		$\alpha_1 = 0.4$	$\lambda_1 = 0.1$	$\alpha_2 = 0.4$	$\lambda_2 = 0.1$
Indep. ($\theta = 0$)	Indep. ($\theta = 0$)	0.406 (0.006)	0.099 (−0.001)	0.403 (0.003)	0.100 (−0.000)
	Clayton ($\theta = 2$)	0.441 (0.041)	0.113 (0.011)	0.437 (0.037)	0.112 (0.012)
	Clayton ($\theta =$ MLE)	0.457 (0.057)	0.149 (0.049)	0.455 (0.055)	0.149 (0.049)
	Clayton ($\theta =$ PMLE)	0.412 (0.012)	0.101 (0.001)	0.408 (0.008)	0.101 (0.001)
Clayton ($\theta = 2$)	Indep. ($\theta = 0$)	0.375 (−0.025)	0.090 (−0.010)	0.379 (−0.021)	0.089 (−0.011)
	Clayton ($\theta = 2$)	0.404 (0.004)	0.100 (−0.000)	0.408 (0.008)	0.100 (−0.000)
	Clayton ($\theta =$ MLE)	0.426 (0.026)	0.132 (0.032)	0.428 (0.028)	0.132 (0.032)
	Clayton ($\theta =$ PMLE)	0.382 (−0.018)	0.092 (−0.008)	0.385 (−0.015)	0.091 (−0.009)
Indep. ($\theta = 0$)	Indep. ($\theta = 0$)	1.003 (0.003)	0.398 (−0.002)	1.001 (0.001)	0.401 (0.001)
	Clayton ($\theta = 2$)	1.133 (0.133)	0.557 (0.157)	1.131 (0.131)	0.559 (0.159)
	Clayton ($\theta =$ MLE)	1.082 (0.082)	0.630 (0.230)	1.081 (0.081)	0.631 (0.231)
	Clayton ($\theta =$ PMLE)	1.109 (0.109)	0.425 (0.025)	1.107 (0.107)	0.428 (0.028)
Clayton ($\theta = 2$)	Indep. ($\theta = 0$)	0.869 (−0.131)	0.298 (−0.102)	0.869 (−0.131)	0.300 (−0.101)
	Clayton ($\theta = 2$)	1.002 (0.002)	0.399 (−0.001)	1.001 (0.001)	0.400 (0.000)
	Clayton ($\theta =$ MLE)	0.952 (−0.048)	0.470 (0.070)	0.951 (−0.049)	0.471 (0.071)
	Clayton ($\theta =$ PMLE)	0.934 (−0.066)	0.313 (−0.087)	0.934 (−0.066)	0.314 (−0.086)
Indep. ($\theta = 0$)	Indep. ($\theta = 0$)	1.500 (−0.000)	1.002 (0.002)	1.501 (0.001)	1.001 (0.001)
	Clayton ($\theta = 2$)	1.684 (0.184)	1.575 (0.575)	1.685 (0.185)	1.574 (0.574)
	Clayton ($\theta =$ MLE)	1.642 (0.142)	1.225 (0.225)	1.643 (0.143)	1.223 (0.223)
	Clayton ($\theta =$ PMLE)	1.671 (0.171)	1.193 (0.193)	1.672 (0.172)	1.191 (0.191)
Clayton ($\theta = 2$)	Indep. ($\theta = 0$)	1.315 (−0.185)	0.666 (−0.334)	1.316 (−0.184)	0.665 (−0.335)
	Clayton ($\theta = 2$)	1.502 (0.001)	1.002 (0.002)	1.503 (0.003)	1.002 (0.002)
	Clayton ($\theta =$ MLE)	1.496 (−0.004)	0.991 (−0.009)	1.496 (−0.004)	0.990 (−0.010)
	Clayton ($\theta =$ PMLE)	1.426 (−0.073)	0.747 (−0.025)	1.426 (−0.073)	0.746 (−0.253)

NOTE: The simulation is based on simulated samples of $\{(t_i, \tau_i, v_i); i = 1, 2, \dots, 1000\}$, consisting of 500 truncated samples and another 500 untruncated samples. The Clayton copula with $\theta = 2$ yields Kendall's tau = 0.5 for two failure times.

Table 2 shows the MSEs for the estimates. The smallest MSEs are attained when the true copula equals the fitted one. This agrees with the pattern of biases (Table 1),

again implying the importance of selecting the correct value of θ . The MSEs remain large when estimating θ by the MLE or PMLE. Again, this problem comes from the difficulty of estimating θ under competing risks.

Table 2. The MSEs of parameter estimates based on 1000 simulation runs.

True Model	Fitted Model	Event 1		Event 2	
		$\alpha_1 = 0.4$	$\lambda_1 = 0.1$	$\alpha_2 = 0.4$	$\lambda_2 = 0.1$
Independent ($\theta = 0$)	Indep. ($\theta = 0$)	0.00186	0.00017	0.00170	0.00017
	Clayton ($\theta = 2$)	0.00374	0.00035	0.00335	0.00036
	Clayton ($\theta = \text{MLE}$)	0.00517	0.00320	0.00498	0.00321
	Clayton ($\theta = \text{PMLE}$)	0.00204	0.00018	0.00184	0.00017
Clayton ($\theta = 2$)	Indep. ($\theta = 0$)	0.00229	0.00025	0.00221	0.00028
	Clayton ($\theta = 2$)	0.00195	0.00018	0.00207	0.00020
	Clayton ($\theta = \text{MLE}$)	0.00261	0.00178	0.00274	0.00178
	Clayton ($\theta = \text{PMLE}$)	0.00210	0.00022	0.00204	0.00025
Independent ($\theta = 0$)	Indep. ($\theta = 0$)	0.00111	0.00093	0.00111	0.00087
	Clayton ($\theta = 2$)	0.01847	0.02621	0.01799	0.02673
	Clayton ($\theta = \text{MLE}$)	0.00115	0.07285	0.01124	0.07301
	Clayton ($\theta = \text{PMLE}$)	0.01323	0.00175	0.01266	0.00184
Clayton ($\theta = 2$)	Indep. ($\theta = 0$)	0.01831	0.01095	0.01814	0.01066
	Clayton ($\theta = 2$)	0.00102	0.00090	0.00088	0.00081
	Clayton ($\theta = \text{MLE}$)	0.00629	0.01394	0.00622	0.01385
	Clayton ($\theta = \text{PMLE}$)	0.00575	0.00828	0.00560	0.00803
Independent ($\theta = 0$)	Indep. ($\theta = 0$)	0.00229	0.00346	0.00244	0.00314
	Clayton ($\theta = 2$)	0.03547	0.33575	0.03570	0.33428
	Clayton ($\theta = \text{MLE}$)	0.02966	0.08040	0.02991	0.08363
	Clayton ($\theta = \text{PMLE}$)	0.03181	0.04223	0.03254	0.04108
Clayton ($\theta = 2$)	Indep. ($\theta = 0$)	0.03659	0.11357	0.03642	0.11415
	Clayton ($\theta = 2$)	0.00193	0.00300	0.00179	0.00297
	Clayton ($\theta = \text{MLE}$)	0.00294	0.00875	0.00282	0.00867
	Clayton ($\theta = \text{PMLE}$)	0.00830	0.00664	0.00808	0.06680

NOTE: The simulation is based on simulated samples of $\{(t_i, \tau_i, v_i); i = 1, 2, \dots, 1000\}$, consisting of 500 truncated samples and another 500 untruncated samples. The Clayton copula with $\theta = 2$ yields Kendall's tau = 0.5 for two failure times.

Table 3 shows how the MSEs reduce as the sample size increases (only the results for estimating α_1 are shown as the results for other parameters exhibiting similar patterns). When the true copula's parameter is correctly specified, the MSEs vanish toward zero with increased samples. This implies the consistency of the parameter estimates. When the copula parameter is wrongly chosen, the MSEs reduce very slowly or do not vanish toward zero. When the copula parameter is estimated, the MSEs do not always reduce with increased samples. This indicates the inconsistency of the parameter estimates when the copula parameter is wrongly specified or estimated.

Table 4 compares the SDs of the parameter estimates with the average of the SEs. The SDs and SEs are very close to each other, indicating that the SEs capture the true variations of the estimates. Hence, the SEs calculated by the proposed method (by the Hessian matrix from the "optim(.)" R function) are highly reliable. However, if θ is treated as unknown and estimated by data, there are some simulation runs where the SEs cannot be computed due to the calculation problem of the Hessian caused by some errors in the "optim(.)" function. Hence, we could not present the results in this case. We suspect that the regularity conditions for the Hessian matrix (to be negative definite) may not hold in this case.

Table 3. The MSEs of the estimates of α_1 under various sample sizes n based on 1000 runs.

True Par.	True Model	Fitted Model	$n = 500$	$n = 1000$	$n = 1500$	$n = 2000$
$\alpha_1 = 0.40$	Indep. ($\theta = 0$)	Indep. ($\theta = 0$)	0.00369	0.00186	0.00114	0.00096
		Clayton ($\theta = 2$)	0.00614	0.00374	0.00261	0.00240
		Clayton ($\theta = \text{MLE}$)	0.00706	0.00517	0.00404	0.00417
		Clayton ($\theta = \text{PMLE}$)	0.00412	0.00204	0.00122	0.00104
	Clayton ($\theta = 2$)	Indep. ($\theta = 0$)	0.00394	0.00229	0.00174	0.00145
		Clayton ($\theta = 2$)	0.00404	0.00195	0.00134	0.00094
		Clayton ($\theta = \text{MLE}$)	0.00408	0.00261	0.00215	0.00180
		Clayton ($\theta = \text{PMLE}$)	0.00382	0.00209	0.00153	0.00122
$\alpha_1 = 0.10$	Indep. ($\theta = 0$)	Indep. ($\theta = 0$)	0.00239	0.00111	0.00073	0.00053
		Clayton ($\theta = 2$)	0.01950	0.01847	0.01738	0.01743
		Clayton ($\theta = \text{MLE}$)	0.01295	0.00115	0.00955	0.00927
		Clayton ($\theta = \text{PMLE}$)	0.00147	0.01323	0.01204	0.01179
	Clayton ($\theta = 2$)	Indep. ($\theta = 0$)	0.01967	0.01831	0.01818	0.01800
		Clayton ($\theta = 2$)	0.00195	0.00102	0.00065	0.00047
		Clayton ($\theta = \text{MLE}$)	0.00678	0.00629	0.00728	0.00720
		Clayton ($\theta = \text{PMLE}$)	0.00717	0.00575	0.00543	0.00523
$\alpha_1 = 0.15$	Indep. ($\theta = 0$)	Indep. ($\theta = 0$)	0.00457	0.00229	0.00161	0.00110
		Clayton ($\theta = 2$)	0.03838	0.03547	0.03474	0.03482
		Clayton ($\theta = \text{MLE}$)	0.03402	0.02966	0.02840	0.02907
		Clayton ($\theta = \text{PMLE}$)	0.03612	0.03181	0.03141	0.03109
	Clayton ($\theta = 2$)	Indep. ($\theta = 0$)	0.03807	0.03659	0.03572	0.03527
		Clayton ($\theta = 2$)	0.00394	0.00193	0.00127	0.00091
		Clayton ($\theta = \text{MLE}$)	0.00605	0.00294	0.00179	0.00129
		Clayton ($\theta = \text{PMLE}$)	0.01083	0.008300.	0.00729	0.00676

NOTE: The simulation is based on 1000 runs. The Clayton copula uses $\theta = 2$ (Kendall's tau = 0.5).

Table 4. The standard derivation (SD) and the average of SEs based on 1000 simulation runs.

True Model	Fitted Model	$\alpha_1 = 0.4$		$\lambda_1 = 0.1$		$\alpha_2 = 0.4$		$\lambda_2 = 0.1$	
		SD	SE	SD	SE	SD	SE	SD	SE
Independent ($\theta = 0$)	Independent ($\theta = 0$)	0.043	0.042	0.013	0.013	0.041	0.042	0.013	0.013
	Clayton ($\theta = 2$)	0.046	0.044	0.015	0.014	0.044	0.044	0.014	0.014
	Clayton ($\theta = \text{PMLE}$)	0.044	0.043	0.013	0.013	0.042	0.042	0.013	0.013
Clayton ($\theta = 2$)	Independent ($\theta = 0$)	0.041	0.043	0.012	0.012	0.042	0.043	0.013	0.012
	Clayton ($\theta = 2$)	0.044	0.044	0.013	0.014	0.045	0.044	0.014	0.014
	Clayton ($\theta = \text{PMLE}$)	0.042	0.043	0.012	0.013	0.043	0.043	0.013	0.013
		$\alpha_1 = 1.0$		$\lambda_1 = 0.4$		$\alpha_2 = 1.0$		$\lambda_2 = 0.4$	
		SD	SE	SD	SE	SD	SE	SD	SE
Independent ($\theta = 0$)	Independent ($\theta = 0$)	0.033	0.033	0.030	0.029	0.033	0.033	0.029	0.029
	Clayton ($\theta = 2$)	0.029	0.028	0.038	0.035	0.028	0.028	0.036	0.035
	Clayton ($\theta = \text{PMLE}$)	0.036	0.034	0.034	0.031	0.036	0.034	0.033	0.031
Clayton ($\theta = 2$)	Independent ($\theta = 0$)	0.035	0.035	0.024	0.024	0.033	0.035	0.024	0.024
	Clayton ($\theta = 2$)	0.032	0.031	0.030	0.029	0.030	0.031	0.029	0.029
	Clayton ($\theta = \text{PMLE}$)	0.037	0.036	0.026	0.025	0.035	0.036	0.025	0.025
		$\alpha_1 = 1.5$		$\lambda_1 = 1.0$		$\alpha_2 = 1.5$		$\lambda_2 = 1.0$	
		SD	SE	SD	SE	SD	SE	SD	SE
Independent ($\theta = 0$)	Independent ($\theta = 0$)	0.048	0.049	0.059	0.058	0.049	0.049	0.056	0.058
	Clayton ($\theta = 2$)	0.040	0.040	0.075	0.076	0.040	0.040	0.073	0.076
	Clayton ($\theta = \text{PMLE}$)	0.051	0.050	0.071	0.067	0.053	0.050	0.068	0.067
Clayton ($\theta = 2$)	Independent ($\theta = 0$)	0.049	0.050	0.043	0.043	0.049	0.050	0.042	0.043
	Clayton ($\theta = 2$)	0.044	0.043	0.055	0.055	0.042	0.043	0.055	0.055
	Clayton ($\theta = \text{PMLE}$)	0.054	0.051	0.050	0.048	0.052	0.051	0.049	0.048

NOTE: The simulation is based on simulated samples of $\{(t_i, \tau_i, v_i); i = 1, 2, \dots, 1000\}$, consisting of 500 truncated samples and another 500 untruncated samples. The Clayton copula with $\theta = 2$ yields Kendall's tau = 0.5 for two failure times.

5. Data Analysis

This section illustrates the proposed method by a data example.

Table 5 shows the artificial dataset created by Kundu et al. [26], who mimicked the setting of electric power transformers that were randomly installed from 1960 to 2008 [1]. The dataset consists of $n = 100$ samples (electric power transformers) that are observed from the starting year ($s = 1980$) to the ending year ($e = 2008$). However, some transformers installed before 1980 are subject to left-truncation. The 30 transformers are truncated ($v_i = 0$; installed before 1980) and the other 70 transformers are untruncated ($v_i = 1$; installed after 1980). Observed events are three types of “Event” (=1 for Event 1; =2 for Event 2; =0 for censoring) in the 5th column of Table 5. The number of events is 14 for Event 1, 33 for Event 2, and 53 for censoring. The data seem to be a simulated dataset from the independent Weibull with the common shape parameter as considered in [26], though the details were not explained. We therefore will fit the Weibull failure time models.

Table 5. Artificial data from Kundu et al. [26], which consist of $n = 100$ power transformers observed from the starting year ($s = 1980$) to the ending year ($e = 2008$). B_i is the installation year. E_i is either failure year or censoring year (2008). The 30 transformers are truncated ($v_i = 0$) and other 70 transformers are untruncated ($v_i = 1$). Observed events are shown in Event (=0 for censoring; =1 for Event 1; =2 for Event 2).

i (Index)	B_i (Install)	E_i (End)	v_i (Truncation)	Event	$\tau_i = 1980 - B_i$ (Truncation Time)	$t_i = E_i - B_i$ (Failure Time)
1	1961	1996	0	2	19	35
2	1964	1985	0	1	16	21
3	1962	2007	0	2	18	45
4	1962	1986	0	2	18	24
5	1961	1992	0	2	19	31
:	:	:	:	:	:	:
30	1963	1994	0	1	17	31
31	1987	2008	1	0	Undefined	21
32	1980	2008	1	0	Undefined	28
33	1988	2008	1	0	Undefined	20
34	1985	2008	1	0	Undefined	23
:	:	:	:	:	:	:
100	1989	2008	1	0	Undefined	19

For the i -th transformer, let (T_{1i}, T_{2i}) be a pair of latent failure times for Event 1 and Event 2, respectively, for $i = 1, 2, \dots, 100$. Following [26], we scaled the failure times (in days) by dividing them by 100 to avoid many decimal places in parameter estimates.

We postulate the model

$$P(T_{1i} > t_1, T_{2i} > t_2) = C_\theta(S_1(t_1), S_2(t_2)),$$

where $C_\theta(v, w) = (v^{-\theta} + w^{-\theta} - 1)^{-1/\theta}$, $S_1(t_1) = \exp(-\lambda_1 t_1^{\alpha_1})$ and $S_2(t_2) = \exp(-\lambda_2 t_2^{\alpha_2})$. We estimated the parameters $\alpha_1, \lambda_1, \alpha_2,$ and λ_2 under the assumed value of $\theta = 0, 0.5, 2,$ or 8 . We also estimated the parameters without giving the value of θ (all parameters were estimated together) by the MLE or the PMLE.

Table 6 compares the estimates for $(\alpha_1, \lambda_1, \alpha_2, \lambda_2)$ based on different values for θ . Under the assumed value of $\theta = 0$, the estimates are very close to the estimates obtained by Kundu et al. [26]. This is reasonable since the proposed model with $\theta = 0$ and that of Kundu et al. [26] are both the independent Weibull models. Note that the model of Kundu et al. [26] imposed the common shape parameter $\alpha_1 = \alpha_2$ while our proposed model did not. Estimates $\alpha_1 = 2.82$ and $\alpha_2 = 2.79$ show that the common shape parameter seems to be valid for these data. For the assumed values of $\theta \in \{0.5, 2, 8\}$, the estimates deviate from the results of Kundu et al. [26]. The results under $\theta \in \{0.5, 2, 8\}$ should not

be taken since their log-likelihood values are lower than the one under $\theta = 0$. The PMLE under the assumed values of $\theta^* \in \{0.5, 2, 8\}$ improves the log-likelihood value, yet none of them reaches the log-likelihood value under $\theta = 0$. Indeed, the log-likelihood value under $\theta = 0$ almost reaches the largest log-likelihood value under $\hat{\theta} = 0.004$ that comes from the MLE. In conclusion, our sensitivity analysis by the fixed values and estimated value of θ extremely confirms adequacy of the independent risks model for this dataset.

Table 6. The estimates (SEs in parenthesis) for the Clayton copula model with the Weibull failure times based on the dataset from Kundu et al. [26]. The Clayton copula with fixed θ or estimated θ are fitted.

θ	$\hat{\theta}$	$\hat{\alpha}_1$; Shape	$\hat{\lambda}_1$; Scale	$\hat{\alpha}_2$; Shape	$\hat{\lambda}_2$; Scale	logL
$\theta = 0^*$	-	2.82 (0.61)	6.93 (5.24)	2.79 (0.39)	15.77 (7.73)	-8.984
$\theta = 0.5$	-	3.18 (0.65)	12.86 (10.56)	2.88 (0.40)	18.74 (9.29)	-9.024
$\theta = 2$	-	3.68 (0.57)	33.12 (22.90)	2.93 (0.38)	21.93 (10.19)	-9.389
$\theta = 8$	-	3.39 (0.40)	35.97 (16.21)	2.92 (0.35)	24.65 (10.64)	-9.239
MLE	0.004 (0.11)	2.82 (0.60)	6.96 (5.16)	2.79 (0.39)	15.79 (7.70)	-8.983
PMLE ($\theta^* = 0.5$)	0.402	3.12 (0.65)	11.57 (9.42)	2.87 (0.40)	18.31 (9.10)	-9.008
PMLE ($\theta^* = 2$)	1.853	3.66 (0.59)	31.57 (22.44)	2.93 (0.38)	21.73 (10.15)	-9.384
PMLE ($\theta^* = 8$)	8.179	3.38 (0.40)	35.77 (16.07)	2.92 (0.35)	24.67 (10.63)	-9.238
Results of [26]	-	2.80 (-)	6.76 (-)	2.80 (-)	15.93 (-)	-

* $\theta = 0$ means the independence model.

6. Conclusions and Future Work

In this paper, we propose a copula-based model for dependent competing risks in the presence of left-truncation. The copula-based competing risks model permits more flexible failure time distributions and dependence structure than the existing competing risks model [4,26,27]. Parametric likelihood-based inference methods are then formulated, motivated by the practically important applications in field reliability analyses.

Below, we list several problems to be considered or resolved in the future.

First, different failure time models can be examined besides the Weibull, lognormal, and gamma models considered in this paper. The candidate distributions could be the extreme-value distributions [43], the generalized Pareto distribution [44], and the generalized Bilal distribution [45]. The Fréchet (type II) and Gumbel (type I) distributions are of particular interest in the extreme-value distributions [43]. In addition, the model may include a covariate (e.g., an acceleration factor) associated with failure time as in the Weibull regression [1]. In addition to this usual failure time regression model, a conditional copula regression model is of another interest [46]. Cure fractions and latent variables may also be considered to be incorporated into the failure time model [47].

Second, the prediction analysis and the test design may be explored in addition to the estimation method considered in this paper. The prediction of remaining lifetime was considered under left-truncated and right-censored field data without competing risks [1,17]. This risk prediction at a censored time point may be informative for engineers to make decisions on future maintenance plans. However, it is not trivial to extend this prediction method under competing risks. This risk prediction scheme is equivalent to dynamic prediction [48–53]. Moreover, the accelerated life test plans may be considered including the case of one-shot devices with competing risks [54,55].

Third, a goodness-of-fit test and/or graphical model-diagnostic method for the proposed model may be developed. A possible strategy is to apply a distance measure between the empirical sub-distribution function and the model-based one. This strategy was adopted for right-censored competing risks data analyses [23] without left-truncation. In order to handle left-truncation, further methodological and numerical works are necessary.

Fourth, Bayesian methods can be considered as a competitor to the MLE methods of this paper. The choice of the marginal prior distributions could follow [2,16], while the

bounded uniform prior could be recommended for the copula prior [56]. It is of interest to see if the difficulty of estimating the copula parameter is resolved by Bayesian methods.

Finally, the limited availability of open datasets could be resolved by encouraging engineers to collect field failure data. Due to the difficulty of obtaining real datasets in the literature, we applied our methods to the artificial dataset created by Kundu et al. [26], who considered the setting of electric power transformers. Dataset can also be searched from other applications of survival analysis, such as those for medical research.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/math10132163/s1>, Supplementary Materials of this paper include “R.code.zip”, including the R code to reproduce the simulations and data analysis results.

Author Contributions: Conceptualization, H.M. and T.E.; methodology, H.M. and T.E.; writing—original draft preparation, H.M. and T.E.; writing—review and editing, H.M. and T.E. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by JSPS KAKENHI (JP21K12127) and (JP22K11948).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the numerical results of the paper are reproduced by the R code in Supplementary Materials.

Acknowledgments: We thank the special issue editor and five reviewers for their helpful comments that improved the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Likelihood for the Gamma Model

We set the marginal survival functions

$$S_1(t) = 1 - \frac{\int_0^{t/\eta_1} u^{\beta_1-1} e^{-u} du}{\Gamma(\beta_1)}, \quad S_2(t) = 1 - \frac{\int_0^{t/\eta_2} u^{\beta_2-1} e^{-u} du}{\Gamma(\beta_2)},$$

Under the independent risks model $C(v, w) = vw$, the likelihood function is

$$\begin{aligned} L_{Gamma}(\beta_1, \beta_2, \eta_1, \eta_2) &= \prod_{i \in I_1} \left[\frac{1}{\eta_1^{\beta_1} \Gamma(\beta_1)} t_i^{\beta_1-1} \exp\left(-\frac{t_i}{\eta_1}\right) \left\{ 1 - \frac{\int_0^{t_i/\eta_2} u^{\beta_2-1} e^{-u} du}{\Gamma(\beta_2)} \right\} \right] \\ &\times \prod_{i \in I_2} \left[\frac{1}{\eta_2^{\beta_2} \Gamma(\beta_2)} t_i^{\beta_2-1} \exp\left(-\frac{t_i}{\eta_2}\right) \left\{ 1 - \frac{\int_0^{t_i/\eta_1} u^{\beta_1-1} e^{-u} du}{\Gamma(\beta_1)} \right\} \right] \\ &\times \prod_{i \in I_0} \left[\left\{ 1 - \frac{\int_0^{t_i/\eta_1} u^{\beta_1-1} e^{-u} du}{\Gamma(\beta_1)} \right\} \left\{ 1 - \frac{\int_0^{t_i/\eta_2} u^{\beta_2-1} e^{-u} du}{\Gamma(\beta_2)} \right\} \right] \\ &\times \prod_{v_i=0} \left[\left\{ 1 - \frac{\int_0^{t_i/\eta_1} u^{\beta_1-1} e^{-u} du}{\Gamma(\beta_1)} \right\} \left\{ 1 - \frac{\int_0^{t_i/\eta_2} u^{\beta_2-1} e^{-u} du}{\Gamma(\beta_2)} \right\} \right]^{-1}. \end{aligned}$$

Under the Clayton copula, the likelihood function is

$$\begin{aligned} L_{Gamma}(\beta_1, \beta_2, \eta_1, \eta_2, \theta) &= \prod_{i \in I_1} \left\{ \frac{1}{\eta_1^{\beta_1} \Gamma(\beta_1)} t_i^{\beta_1-1} \exp\left(-\frac{t_i}{\eta_1}\right) u_i^{-(\theta+1)} \left(u_i^{-\theta} + v_i^{-\theta} - 1\right)^{-(1+\frac{1}{\theta})} \right\} \\ &\times \prod_{i \in I_2} \left\{ \frac{1}{\eta_2^{\beta_2} \Gamma(\beta_2)} t_i^{\beta_2-1} \exp\left(-\frac{t_i}{\eta_2}\right) v_i^{-(\theta+1)} \left(u_i^{-\theta} + v_i^{-\theta} - 1\right)^{-(1+\frac{1}{\theta})} \right\} \\ &\times \prod_{i \in I_0} \left(u_i^{-\theta} + v_i^{-\theta} - 1\right)^{-\frac{1}{\theta}} \times \prod_{v_i=0} \left(x_i^{-\theta} + y_i^{-\theta} - 1\right)^{\frac{1}{\theta}}. \end{aligned}$$

where

$$u_i = S_1(t_i) = 1 - \frac{\int_0^{t_i} u^{\beta_1-1} e^{-u} du}{\Gamma(\beta_1)}, \quad v_i = S_2(t_i) = 1 - \frac{\int_0^{t_i} u^{\beta_2-1} e^{-u} du}{\Gamma(\beta_2)},$$

$$x_i = S_1(\tau_i) = 1 - \frac{\int_0^{\tau_i} u^{\beta_1-1} e^{-u} du}{\Gamma(\beta_1)}, \quad y_i = S_2(\tau_i) = 1 - \frac{\int_0^{\tau_i} u^{\beta_2-1} e^{-u} du}{\Gamma(\beta_2)}.$$

Appendix B. Likelihood for the Lognormal Model

We set the marginal survival functions

$$S_1(t) = 1 - \Phi\left(\frac{\log t - \mu_1}{\sigma_1}\right), \quad S_2(t) = 1 - \Phi\left(\frac{\log t - \mu_2}{\sigma_2}\right),$$

where Φ is the cdf of the standard normal distribution. Under the independent risks model $C(v, w) = vw$, the likelihood function is

$$L_{\text{lognorm}}(\mu_1, \mu_2, \sigma_1, \sigma_2) = \prod_{i \in I_1} \left[\frac{1}{\sigma_1 t_i \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \left(\frac{\log t_i - \mu_1}{\sigma_1}\right)^2\right\} \left\{1 - \Phi\left(\frac{\log t_i - \mu_2}{\sigma_2}\right)\right\} \right]$$

$$\times \prod_{i \in I_2} \left[\frac{1}{\sigma_2 t_i \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \left(\frac{\log t_i - \mu_2}{\sigma_2}\right)^2\right\} \left\{1 - \Phi\left(\frac{\log t_i - \mu_1}{\sigma_1}\right)\right\} \right]$$

$$\times \prod_{i \in I_0} \left[\left\{1 - \Phi\left(\frac{\log t_i - \mu_1}{\sigma_1}\right)\right\} \left\{1 - \Phi\left(\frac{\log t_i - \mu_2}{\sigma_2}\right)\right\} \right]$$

$$\times \prod_{v_i=0} \left[\left\{1 - \Phi\left(\frac{\log \tau_i - \mu_1}{\sigma_1}\right)\right\} \left\{1 - \Phi\left(\frac{\log \tau_i - \mu_2}{\sigma_2}\right)\right\} \right]^{-1},$$

Under the Clayton copula, the likelihood function is

$$L_{\text{lognorm}}(\mu_1, \mu_2, \sigma_1, \sigma_2, \theta) = \prod_{i \in I_1} \left\{ \frac{u_i^{-(\theta+1)} (u_i^{-\theta} + v_i^{-\theta} - 1)^{-(1+\frac{1}{\theta})}}{\sigma_1 t_i \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \left(\frac{\log t_i - \mu_1}{\sigma_1}\right)^2\right\} \right\}$$

$$\times \prod_{i \in I_2} \left\{ \frac{v_i^{-(\theta+1)} (u_i^{-\theta} + v_i^{-\theta} - 1)^{-(1+\frac{1}{\theta})}}{\sigma_2 t_i \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \left(\frac{\log t_i - \mu_2}{\sigma_2}\right)^2\right\} \right\}$$

$$\times \prod_{i \in I_0} \left(u_i^{-\theta} + v_i^{-\theta} - 1 \right)^{-\frac{1}{\theta}} \times \prod_{v_i=0} \left(x_i^{-\theta} + y_i^{-\theta} - 1 \right)^{\frac{1}{\theta}}.$$

where

$$u_i = 1 - \Phi\left(\frac{\log t_i - \mu_1}{\sigma_1}\right), \quad v_i = 1 - \Phi\left(\frac{\log t_i - \mu_2}{\sigma_2}\right),$$

$$x_i = 1 - \Phi\left(\frac{\log \tau_i - \mu_1}{\sigma_1}\right), \quad y_i = 1 - \Phi\left(\frac{\log \tau_i - \mu_2}{\sigma_2}\right).$$

References

- Hong, Y.; Meeker, W.Q.; McCalley, J.D. Prediction of remaining life of power transformers based on left truncated and right censored lifetime data. *Ann. Appl. Stat.* **2009**, *3*, 857–879. [\[CrossRef\]](#)
- Emura, T.; Michimae, H. Left-truncated and right-censored field failure data: Review of parametric analysis for reliability. *Qual. Reliab. Eng. Int.* **2022**. [\[CrossRef\]](#)
- Klein, J.P.; Moeschberger, M.L. *Survival Analysis Techniques for Censored and Truncated Data*; Springer: New York, NY, USA, 2003.
- Wang, L.; Lio, Y.; Tripathi, Y.M.; Dey, S.; Zhang, F. Inference of dependent left-truncated and right-censored competing risks data from a general bivariate class of inverse exponentiated distributions. *Statistics* **2022**, *56*, 347–374. [\[CrossRef\]](#)
- Lawless, J.F. *Statistical Models and Methods for Lifetime Data*, 2nd ed.; Wiley & Sons: Hoboken, NJ, USA, 2003.
- Li, Q.; Li, D.; Huang, B.; Jiang, Z.E.M.; Ma, J. Failure analysis for truncated and fully censored lifetime data with a hierarchical grid algorithm. *IEEE Access* **2020**, *8*, 34468–34480. [\[CrossRef\]](#)
- Jiang, W.; Ye, Z.; Zhao, X. Reliability estimation from left-truncated and right-censored data using splines. *Stat Sin.* **2020**, *30*, 845–875.
- Dörre, A. Semiparametric likelihood inference for heterogeneous survival data under double truncation based on a Poisson birth process. *Jpn. J. Stat. Data Sci.* **2021**, *4*, 1203–1226. [\[CrossRef\]](#)

9. Dörre, A.; Huang, C.Y.; Tseng, Y.K.; Emura, T. Likelihood-based analysis of doubly-truncated data under the location-scale and AFT model. *Comp. Stat.* **2021**, *36*, 375–408. [[CrossRef](#)]
10. Balakrishnan, N.; Mitra, D. Likelihood inference for lognormal data with left truncation and right censoring with an illustration. *J. Stat. Plan Inf.* **2011**, *141*, 3536–3553. [[CrossRef](#)]
11. Balakrishnan, N.; Mitra, D. Left truncated and right censored Weibull data and likelihood inference with an illustration. *Comp. Stat. Data Anal.* **2012**, *56*, 4011–4025. [[CrossRef](#)]
12. Balakrishnan, N.; Mitra, D. Some further issues concerning likelihood inference for left truncated and right censored lognormal data. *Comm. Stat. Simul Comp.* **2014**, *43*, 400–416. [[CrossRef](#)]
13. Balakrishnan, N.; Mitra, D. Likelihood inference based on left truncated and right censored data from a gamma distribution. *IEEE Trans. Reliab.* **2013**, *62*, 679–688. [[CrossRef](#)]
14. Balakrishnan, N.; Mitra, D. EM-based likelihood inference for some lifetime distributions based on left truncated and right censored data and associated model discrimination. *S. Afr. Stat. J.* **2014**, *48*, 125–171.
15. Emura, T.; Shiu, S. Estimation and model selection for left-truncated and right-censored lifetime data with application to electric power transformers analysis. *Commun. Stat. Simul.* **2016**, *45*, 3171–3189.
16. Ranjan, R.; Sen, R.; Upadhyay, S.K. Bayes analysis of some important lifetime models using MCMC based approaches when the observations are left truncated and right censored. *Reliab. Eng. Syst. Saf.* **2021**, *214*, 107747. [[CrossRef](#)]
17. Mitra, D.; Kundu, D.; Balakrishnan, N. Likelihood analysis and stochastic EM algorithm for left truncated right censored data and associated model selection from the Lehmann family of life distributions. *Jpn. J. Stat. Data Sci.* **2021**, *4*, 1019–1048. [[CrossRef](#)]
18. Mitra, D.; Balakrishnan, N. Statistical inference based on left truncated and interval censored data from log-location-scale family of distributions. *Comm. Stat. Simul. Comp.* **2021**, *50*, 1073–1093. [[CrossRef](#)]
19. Zheng, M.; Klein, J.P. A self-consistent estimator of marginal survival functions based on dependent competing risk data and an assumed copula. *Commun. Stat. Theory Methods* **1994**, *23*, 2299–2311. [[CrossRef](#)]
20. Zheng, M.; Klein, J.P. Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika* **1995**, *82*, 127–138. [[CrossRef](#)]
21. Escarela, G.; Carriere, J.F. Fitting competing risks with an assumed copula. *Stat. Methods Med. Res.* **2003**, *12*, 333–349. [[CrossRef](#)]
22. Emura, T.; Shih, J.H.; Ha, I.D.; Wilke, R.A. Comparison of the marginal hazard model and the sub-distribution hazard model for competing risks under an assumed copula. *Stat. Methods Med. Res.* **2020**, *29*, 2307–2327. [[CrossRef](#)]
23. Wang, Y.C.; Emura, T.; Fan, T.H.; Lo, S.M.; Wilke, R.A. Likelihood-based inference for a frailty-copula model based on competing risks failure time data. *Qual. Reliab. Eng. Int.* **2020**, *36*, 1622–1638. [[CrossRef](#)]
24. De Uña-Álvarez, J.; Veraverbeke, N. Copula-graphic estimation with left-truncated and right-censored data. *Statistics* **2017**, *51*, 387–403. [[CrossRef](#)]
25. Shih, J.H.; Emura, T. Likelihood-based inference for bivariate latent failure time models with competing risks under the generalized FGM copula. *Comput. Stat.* **2018**, *33*, 1293–1323. [[CrossRef](#)]
26. Kundu, D.; Mitra, D.; Ganguly, A. Analysis of left truncated and right censored competing risks data. *Comp. Stat. Data Anal.* **2017**, *108*, 12–26. [[CrossRef](#)]
27. Wu, K.; Wang, L.; Yan, L.; Lio, Y. Statistical inference of left truncated and right censored data from Marshall–Olkin bivariate Rayleigh distribution. *Mathematics* **2021**, *9*, 2703. [[CrossRef](#)]
28. Shuto, S.; Amemiya, T. Sequential Bayesian inference for Weibull distribution parameters with initial hyperparameter optimization for system reliability estimation. *Reliab. Eng. Syst. Saf.* **2022**, *224*, 108516. [[CrossRef](#)]
29. Bouwmeester, J.; Menicucci, A.; Gill, E.K.A. Improving CubeSat reliability: Subsystem redundancy or improved testing? *Reliab. Eng. Syst. Saf.* **2022**, *220*, 108288. [[CrossRef](#)]
30. Wu, C.W.; Lee, A.H.; Liu, S.W. A repetitive group sampling plan based on the lifetime performance index under gamma distribution. *Qual. Reliab. Eng. Int.* **2022**, *38*, 2049–2064. [[CrossRef](#)]
31. Nelsen, R.B. *An Introduction to Copulas*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2006.
32. Shih, J.H.; Emura, T. Bivariate dependence measures and bivariate competing risks models under the generalized FGM copula. *Stat. Pap.* **2019**, *60*, 1101–1118. [[CrossRef](#)]
33. Peng, M.; Xiang, L.; Wang, S. Semiparametric regression analysis of clustered survival data with semi-competing risks. *Comp. Stat. Data Anal.* **2018**, *124*, 53–70. [[CrossRef](#)]
34. Wu, M.; Shi, Y.; Zhang, C. Statistical analysis of dependent competing risks model in accelerated life testing under progressively hybrid censoring using copula function. *Commun. Stat. Simul. Comput.* **2017**, *46*, 4004–4017. [[CrossRef](#)]
35. Chesneau, C. Theoretical study of some angle parameter trigonometric copulas. *Modelling* **2022**, *3*, 140–163. [[CrossRef](#)]
36. Susam, S.O. A multi-parameter Generalized Farlie-Gumbel-Morgenstern bivariate copula family via Bernstein polynomial. *Hacet. J. Math. Stat.* **2022**, *51*, 618–631. [[CrossRef](#)]
37. Ota, S.; Kimura, M. Effective estimation algorithm for parameters of multivariate Farlie–Gumbel–Morgenstern copula. *Jpn. J. Stat. Data Sci.* **2021**, *4*, 1049–1078. [[CrossRef](#)]
38. Ghosh, S.; Sheppard, L.W.; Holder, M.T.; Loecke, T.D.; Reid, P.C.; Bever, J.D.; Reuman, D.C. Copulas and their potential for ecology. In *Advances in Ecological Research*; Academic Press: Cambridge, MA, USA, 2020; Volume 62, pp. 409–468.
39. Shih, J.H.; Konno, Y.; Chang, Y.T.; Emura, T. Estimation of a common mean vector in bivariate meta-analysis under the FGM copula. *Statistics* **2019**, *53*, 673–695. [[CrossRef](#)]

40. Shih, J.H.; Konno, Y.; Chang, Y.T.; Emura, T. Copula-based estimation methods for a common mean vector for bivariate meta-analyses. *Symmetry* **2022**, *14*, 186. [[CrossRef](#)]
41. Zhuang, H.; Diao, L.; Grace, Y.Y. A Bayesian nonparametric mixture model for grouping dependence structures and selecting copula functions. *Econom. Stat.* **2022**, *22*, 172–189. [[CrossRef](#)]
42. Emura, T.; Hsu, J.H. Estimation of the Mann–Whitney effect in the two-sample problem under dependent censoring. *Comp. Stat. Data Anal.* **2020**, *150*, 106990. [[CrossRef](#)]
43. Alves, M.I.F.; Neves, C. Extreme Value Distributions. *Int. Encycl. Stat. Sci.* **2011**, *2*, 493–496.
44. Grimshaw, S.D. Computing maximum likelihood estimates for the generalized Pareto distribution. *Technometrics* **1993**, *35*, 185–191. [[CrossRef](#)]
45. Akhter, Z.; Almetwally, E.M.; Chesneau, C. On the Generalized Bilal Distribution: Some Properties and Estimation under Ranked Set Sampling. *Axioms* **2022**, *11*, 173. [[CrossRef](#)]
46. Emura, T.; Sofeu, C.L.; Rondeau, V. Conditional copula models for correlated survival endpoints: Individual patient data meta-analysis of randomized controlled trials. *Stat. Method Med. Res.* **2021**, *30*, 2634–2650. [[CrossRef](#)] [[PubMed](#)]
47. Yang, Q.; He, H.; Lu, B.; Song, X. Mixture additive hazards cure model with latent variables: Application to corporate default data. *Comp. Stat. Data Anal.* **2022**, *167*, 107365. [[CrossRef](#)]
48. Emura, T.; Nakatochi, M.; Matsui, S.; Michimae, H.; Rondeau, V. Personalized dynamic prediction of death according to tumour progression and high-dimensional genetic factors: Meta-analysis with a joint model. *Stat. Method Med. Res.* **2018**, *27*, 2842–2858. [[CrossRef](#)] [[PubMed](#)]
49. Emura, T.; Michimae, H.; Matsui, S. Dynamic risk prediction via a joint frailty-copula model and IPD meta-analysis: Building web applications. *Entropy* **2022**, *24*, 589. [[CrossRef](#)] [[PubMed](#)]
50. Kawakami, R.; Michimae, H.; Lin, Y.H. Assessing the numerical integration of dynamic prediction formulas using the exact expressions under the joint frailty-copula model. *Jpn. J. Stat. Data Sci.* **2021**, *4*, 1293–1321. [[CrossRef](#)]
51. Noughabi, M.S.; Kayid, M. Bivariate quantile residual life: A characterization theorem and statistical properties. *Stat. Pap.* **2019**, *60*, 2001–2012. [[CrossRef](#)]
52. Kayid, M.; Shafaei Noughabi, M.; Abouammoh, A.M. A nonparametric estimator of bivariate quantile residual life model with application to tumor recurrence data set. *J. Classificat.* **2020**, *37*, 237–253. [[CrossRef](#)]
53. Zheng, R.; Najafi, S.; Zhang, Y. A recursive method for the health assessment of systems using the proportional hazards model. *Reliab. Eng. Syst. Saf.* **2022**, *221*, 108379. [[CrossRef](#)]
54. Meeker, W.Q.; Hahn, G.J. A comparison of accelerated test plans to estimate the survival probability at a design stress. *Technometrics* **1978**, *20*, 245–247. [[CrossRef](#)]
55. Ling, M.H. Optimal constant-stress accelerated life test plans for one-shot devices with components having exponential lifetimes under gamma frailty models. *Mathematics* **2022**, *10*, 840. [[CrossRef](#)]
56. Michimae, H.; Emura, T. Bayesian ridge estimators based on copula-based joint prior distributions for regression coefficients. *Comput. Stat.* **2022**. [[CrossRef](#)]