Statistical Methods for
Comparing of Two Restricted Mean Survival Times
in the presence of Dependent Censoring

Yaeji Lee

The Graduate School
Yonsei University
Department of Biostatistics and Computing

Statistical Methods for Comparing of Two Restricted Mean Survival Times in the presence of Dependent Censoring

A Dissertation Submitted to the
Department of Biostatistics and Computing
and the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Biostatistics and Computing

Yaeji Lee

June 2024

This certifies that the Dissertation of ***Yaeji Lee*** is approved.

Chung Mo Nam: Thesis Supervisor

Inkyung Jung: Thesis Committee Member #1

Sohee Park: Thesis Committee Member #2

Jiwon Lee: Thesis Committee Member #3

Sejung Park: Thesis Committee Member #4

The Graduate School

Yonsei University

June 2024

Contents

List of Tables

Appendix

List of Figures

Abstract

Statistical Methods for

Comparing of Two Restricted Mean Survival Times

in the presence of Dependent Censoring

Survival analysis is a common method in clinical trials, often relying on the assumption of proportional hazards. When this assumption is violated, an alternative approach is the restricted mean survival time (RMST) method. RMST is a significant metric in survival analysis, offering an intuitive and interpretable measure of average survival time up to a specific point. Traditional survival analysis methods, assuming independent censoring, can be biased when this assumption is violated. Typically, the failure event and censoring time are positively correlated, necessitating the consideration of their joint distribution under dependent censoring.

The copula function provides a flexible tool for modeling the dependence between survival and censoring times. For analyzing survival under dependent censoring, the copula-graphic estimator is utilized. This study proposes new estimators that expand on copula-graphic and self-consistency estimators under the dependent censoring assumption. The objective is to adapt the copula method for estimating RMST with survival data that includes dependent censoring. We evaluate our proposed estimators through a series of simulations, examining bias, type I error, and power in RMST estimation under various scenarios of dependent censoring.

Simulation results confirmed that the type I error was generally well-controlled, and the proposed model demonstrated performance comparable to the true model in terms of power. The aim of this paper is to identify effective methods for estimating the difference in RMST between two groups when dependency exists. Previous research has not extensively explored copula-graphic estimators and self-consistency in the context of RMST, particularly regarding their statistical properties. This study contributes to the understanding of RMST estimation in data with dependent censoring and may further contribute to future studies.

Keywords: survival analysis, restricted mean survival time, dependent censoring, copula models, copula-graphic estimator, self-consistency estimator

Chapter 1

Introduction

Survival analysis is one of the most common analyses in clinical trials. In general, time-to-event analysis relies on the assumption of proportional hazards. However, quite frequently, we may find that the proportional hazards assumption is violated, especially in many immuno-oncology trials. When the proportional hazards assumption is violated, one of the alternative approaches is the restricted mean survival time (RMST) method. The RMST generally uses the Kaplan-Meier estimator to calculate and compare the area under the curve (AUC) for different treatment groups or different comparative groups.

Generally, event time and censoring time tend to be positively correlated. However, traditional methods of survival analysis critically rely on the independent censoring assumption: survival time and censoring time need to be statistically independent. This independence assumption is easily violated when patients drop out from the study due to the worsening of their health condition or are removed for transplantation (Staplin et al., 2015; Emura and Chen, 2018). In such cases, survival time is censored depending on their health status.

Consider a study investigating the survival times of patients with a particular disease who are treated with different medications. In this scenario, patients may drop out of the study or be lost to follow-up at different times. However, the likelihood of being censored may be influenced by factors such as the effectiveness of the treatment, the severity of the disease, or the patient's compliance with the medication regimen. For example, suppose a new, more effective medication is introduced midway through the study. Patients who are not responding well to the initial treatment may be more likely to switch to the new medication, leading to a higher probability of being censored at certain time points. Conversely, patients experiencing positive outcomes with the initial treatment may be less likely to drop out of the study, resulting in a lower probability of censoring. In this situation, the censoring times of patients are dependent on factors related to their treatment and disease progression. This dependent censoring complicates the analysis of survival data and requires careful consideration to ensure accurate estimation of survival probabilities and treatment effects.

However, the Kaplan-Meier estimator cannot consider dependent censoring, and there is no proposed method to estimate RMST and test for differences in RMST between two groups under dependent censoring. Therefore, it is necessary to consider how to test the differences in RMST under dependent censoring. The copula function is a popular model for modeling the dependency between survival and censoring times. For dependent censoring, the copula-graphic estimator was proposed to estimate the survival curve. Additionally, some methods based on self-consistency were proposed for dependent censoring. Prior research by Jiang, Hongyu, et al. (2005) has estimated the copula-graphic estimator as a self-consistency estimator in dependent censoring.

In this paper, we propose new estimators that expand upon the copula-graphic estimator and the self-consistency estimator under the assumption of dependent censoring. We evaluated our proposed estimators through a series of simulations. The purpose of this study is to adapt the copula method to estimate the difference between two RMSTs using survival data that includes dependent censoring. In Section 2, we briefly review the estimators of the survival function. We review the definition and basic statistical properties of RMST and summarize how to estimate it under the independent assumption in Section 3. In Section 4, we illustrate the issues of dependent censoring arising from medical research and provide the mathematical foundations of the copula models for applications to survival analysis under dependent censoring. In Section 5, we propose a set of estimators under dependent censoring based on the copula model, expanding on the RMST estimators. In Section 6, we evaluate the performance of the proposed estimators for censored survival data via a simulation study. The proposed estimator is applied to a real dataset, and the analysis results are reported in Section 7. The final section provides conclusions and discussion.

Chapter 2

Estimators of the Survival function

2.1 Kaplan-Meier Estimator

The commonly used survival function estimator, proposed by Kaplan and Meier (1958), is referred to as the Product-Limit estimator. The Kaplan-Meier estimator is a non-parametric statistic used to estimate the survival function from lifetime data. In other words, it provides a way to measure the proportion of subjects living for a certain amount of time after treatment or under certain conditions. It estimates the probability that a subject will survive beyond a certain time t. This estimator is defined as follows:

$$\hat{S}(t) = \begin{cases} 1 & \text{if } t \leq t_1, \\ \prod_{t_i \leq t} \left[1 - \frac{d_i}{n_i}\right] & \text{if } t_1 \leq t \cdot \end{cases}$$

where t_i are the distinct times when events occur, n_i is the number of subjects at risk just before time t_i and d_i is the number of events at time t_i.

The Kaplan-Meier method can manage right-censored data, where subjects exit the study before an event occurs or the study concludes before the event for some subjects. The Kaplan-Meier estimator generates a step function that changes its value only at each observed event time. Between these times, the function remains constant. The Product-

Limit estimator also uses a step function, jumping at each observed event time. The size of these jumps is influenced not only by the number of events at each time point t_i, but also by the arrangement of censored observations leading up t_i. This estimator is not defined beyond the maximum observation time t.

The variance of the Kaplan-Meier estimator is estimated using Greenwood's formula:

$$\hat{V}\left[\hat{S}(t)\right] = \hat{S}(t)^2 \sum_{T_i \leq t} \frac{d_i}{n_i(n_i - d_i)}.$$

The standard error of the Kaplan-Meier estimator is written by $\sqrt{\hat{V}\left[\hat{S}(t)\right]}$.

2.2 Self-consistency Estimator

A self-consistency estimator is an estimator that satisfies the property that if the estimator were known, the data would look as it does. In other words, the estimated distribution must be consistent with the observed data. If we had no censored observations, the estimator of the survival function at a time t can be straightforwardly defined as the proportion of observations which are larger than t as follows:

$$\hat{S}(t) = \frac{1}{n}\sum_{i=1}^{n} \phi(T_i)$$

where $\phi(T_i) = 1$ if $T_i > t$ and $\phi(T_i) = 0$, if $T_i \leq t$.

For right-censored data, a survival function estimator can be developed similarly by redefining the scoring function ϕ. Let $T_1, T_2, \dots, T_n$ be the observed times in the study. If T_i represents the time of death, we can definitively determine whether T_i is less than or greater that t. If T_i is a censored time that is greater than or equal to t, we know that the true death time must be larger than t because it exceeds T_i for this individual. For a censored observation less than t, we cannot determine if the corresponding death time is greater than t because it could occur between T_i and t. If we knew $S(t)$, we could estimate the probability of this censored observation being larger than t by $S(t)/S(T_i)$. Using these redefined scores, we will call an estimator $\hat{S}(t)$ a self-consistency estimator of S if

$$\hat{S}(t) = \frac{1}{n}\left[\sum_{T_i > t} \phi(T_i) + \sum_{\delta_i = 0, T_i \le t} \frac{\hat{S}(t)}{\hat{S}(T_i)}\right].$$

Self-consistency estimators are often derived using an iterative algorithm. This involves repeatedly updating the estimate based on the current estimate until convergence is achieved. The iterative process ensures that the final estimate is self-consistent. The Expectation-Maximization (EM) algorithm is a common framework used to find self-consistency estimators. The EM algorithm iteratively applies the Expectation step (E-step) and Maximization step (M-step) to refine the estimates. In the E-step, the expected value of the log-likelihood is computed given the current parameter estimates. In the M-step, the parameters are updated to maximize this expected log-likelihood. Self-consistency estimators typically converge to a stable solution after a sufficient number of iterations. The convergence is monitored by checking the change in estimates between successive iterations until it falls below a predefined threshold. Many self-consistency estimators, like the Turnbull estimator, are non-parametric, meaning they do not assume a specific parametric form for the underlying survival distribution. This flexibility makes them widely applicable. Self-consistency estimators provide a robust framework for dealing with complex data structures, particularly interval-censored data. They rely on iterative methods to ensure that the final estimates are consistent with the observed data, often utilizing algorithms like EM to achieve convergence. This methodology extends the flexibility and applicability of survival analysis beyond traditional right-censored data scenarios.

Chapter 3

Restricted Mean Survival Time

Restricted Mean Survival Time (RMST) is an important metric used in survival analysis. It serves as an alternative approach when dealing with survival time data, aiming to average survival times up to a specific time point. RMST is typically used in conjunction with traditional survival analysis methods like the Kaplan-Meier survival curve. One of the key advantages of RMST is its intuitiveness and interpretability. By directly measuring the average survival time up to a specific time, it is easier to interpret compared to other statistical methods. Additionally, RMST allows for comparing survival times while maintaining a constant observation period, which is useful for comparing outcomes in research or clinical trials. However, RMST comes with some limitations. The most significant limitation is that it only considers survival times up to a specific time point, potentially losing information from the entire survival curve. Additionally, RMST may be unstable with small sample sizes and may have limited ability to capture nonlinear effects. In summary, RMST is an important metric in survival analysis, offering intuitive and interpretable insights, but it has some limitations. While it facilitates comparison and interpretation of survival times up to a specific time point, it may lose information from the entire survival curve and can be unstable with small sample sizes.

3.1 Definition and Properties of RMST

Let t be a nonnegative random variable representing the time until an event occurs for a patient from a homogeneous population. Suppose γ is a specific time point of interest. Define $X(\gamma)$ be the minimum of t and γ, i.e., $X(\gamma) = \min(t, \gamma)$. The Restricted Mean Survival Time (RMST), or the mean survival time up to γ years, is defined as the expected value of $X(\gamma)$: $\mu(\gamma) = \mathrm{E}[X(\gamma)] = \mathrm{E}[\min(t, \gamma)]$.

For instance, if the RMST up to 4 years (i.e., $\gamma =$ 4 years) is 1 year, it indicates that a patient, on average, would survive for 1 year when followed for 4 years. The RMST up to γ represents the area under the survival curve from 0 to γ,

$$\mu(\gamma) = \int_0^{\gamma} S(t)\, dt,$$

where $S(t)$ is the survival function. The variance of the restricted survival time $X(\gamma)$ is

$$\mathrm{Var}[X(\gamma)] = 2\int_0^{\gamma} tS(t)\, dt - \left\{\int_0^{\gamma} S(t)\, dt\right\}^2$$

As mentioned above, the mean value (i.e., RMST) and the variance of the restricted survival time $X(\gamma)$ up to the specific time point γ are determined by the survival function. The estimation of RMST and its variance can essentially be derived from the estimated survival function.

3.2 Estimation of RMST

As discussed by Royston and Parmar (2011), several methods to estimate the RMST are available, including direct integration of Kaplan-Meier survival curves, a jackknife method, pseudo-value regression method, inverse probability of censoring weighting (IPCW) regression, conditional restricted mean survival time (CRMST), and a flexible parametric regression modeling. As the nonparametric Kaplan-Meier method is commonly used for estimating the survival curve, we outline the method of integrating the survival curve obtained by the Kaplan-Meier method. It is assumed that event times and censoring times are independent. Let $T_1 < T_2 < \cdots < T_D$ denote the distinct event times up to the specific time point γ, where $T_0 = 0$ and $T_{D+1} = \gamma$. For each $j = 1, \dots, D$, let n_j represents the size of the risk set just prior to T_j and let d_j represents the number of events occurring at time T_j. The RMST up to the specific time point γ is estimated as follows:

$$\hat{\mu}(\gamma) = \int_0^{\gamma} \hat{S}(t)\, dt = \sum_{j=0}^{D} (T_{j+1} - T_j)\hat{S}(T_j).$$

The variance of $\hat{\mu}(\gamma)$ is estimated as

$$\mathrm{Var}[\hat{\mu}(\gamma)] = \sum_{j=1}^{D} \left[\sum_{i=j}^{D} (T_{i+1} - T_i)\hat{S}(T_i) \right]^2 \frac{d_j}{n_j(n_j - d_j)},$$

using Greenwood's formula. A $100(1-\alpha)\%$ confidence interval for the mean is expressed by $\hat{\mu}(\gamma) \pm z_{1-\alpha/2}\sqrt{\hat{V}[\hat{\mu}_{\gamma}]}$.

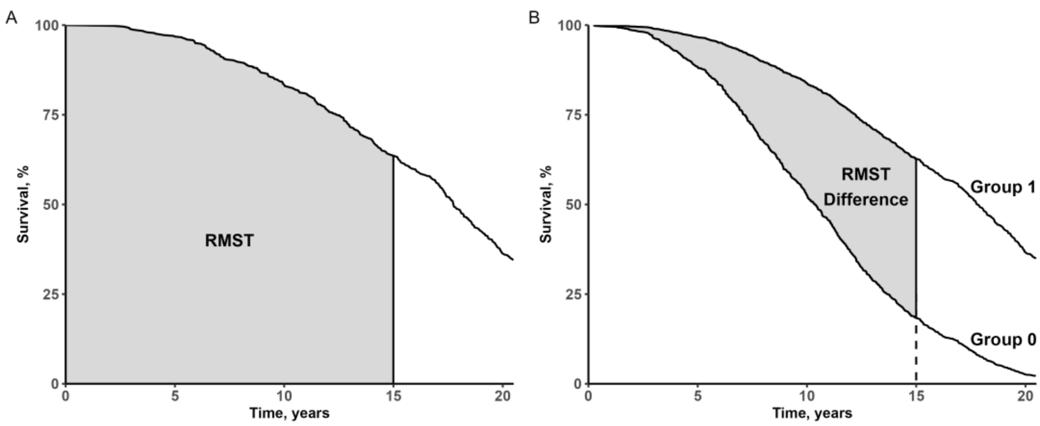


Figure 1: Illustrations of restricted mean survival time and its difference. (A) RMST is the integration of survival probability across a pre-specified time, which graphically is the area under the survival curve. (B) RMST difference represents the group separation of the survival analysis.

3.3 Comparison of the RMST between Two Groups

Let $\mu_g(\gamma)$ be the RMST for group g, where $g = 0$ represents the control group and $g = 1$ represents the treatment group. Let $\hat{\mu}_g(\gamma)$ be the estimated RMST, and let $\mathrm{Var}\left[\hat{\mu}_g(\gamma)\right]$ be the variance of $\hat{\mu}_g(\gamma)$. The difference of two RMST is written as:

$$\mathrm{RMST}(\delta) = \int_0^{\gamma} \left(\hat{S}_1(t) - \hat{S}_0(t)\right) dt = \hat{\mu}_1(\gamma) - \hat{\mu}_0(\gamma)$$

The $100(1-\alpha)\%$ confidence interval for the difference in RMST between the groups is estimated as:

$$\hat{\mu}_1(\gamma) - \hat{\mu}_0(\gamma) \pm z_{\frac{\alpha}{2}} \sqrt{\mathrm{Var}[\hat{\mu}_1(\gamma)] + \mathrm{Var}[\hat{\mu}_0(\gamma)]},$$

where z_α is the upper $100(1-\alpha)\%$ quantile of the standard normal distribution.

The null and alternative hypotheses to be tested are

$$H_0 : \mu_1(\gamma) - \mu_0(\gamma) = 0 \ \text{ versus } \ H_1 : \mu_1(\gamma) - \mu_0(\gamma) \neq 0.$$

Under the null hypothesis H_0, the test statistic, which is computed as

Asymptotically, S_D follows a standard normal distribution. Hence, the two-sided p-value is calculated as $2\{1 - \Phi(|S_D|)\}$, where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

Chapter 4

Dependent Censoring

Dependent censoring occurs when the relationship between censoring and survival time cannot be explained by observable covariates, indicating residual dependency that covariates do not address. To mitigate concerns about dependent censoring, it is advisable to collect a comprehensive set of covariates. For instance, late-stage cancer patients typically have shorter survival times and are more likely to drop out due to tumor progression, which establishes a positive correlation between survival and dropout times. Therefore, including cancer stage as a covariate can help achieve conditional independence between survival and dropout times.

If censoring involves dropout or withdrawal due to worsening symptoms, it can introduce bias in statistical analysis. This form of dropout, known as informative dropout, is among several reasons for censoring. Broadly speaking, when the time of an event of interest is censored due to a mechanism associated with the event itself, this is termed dependent censoring. Most standard survival analysis methods yield unbiased results assuming independent censoring. Hence, careful consideration is required in survival analysis when censoring is not independent.

In cancer follow-up studies, survival times may be censored due to dropout from tumor progression, treatment toxicity, or initiation of new treatments, among other factors.

Consequently, overall survival and censoring times may be positively correlated, as patients often pass away shortly after dropout. Informative censoring of this nature can adversely impact data analysis. For example, many terminally ill patients drop out of clinical trials to receive home care, potentially leading to missed observable deaths. Consequently, Kaplan-Meier survival curves that treat such patients as censored may exhibit upward bias.

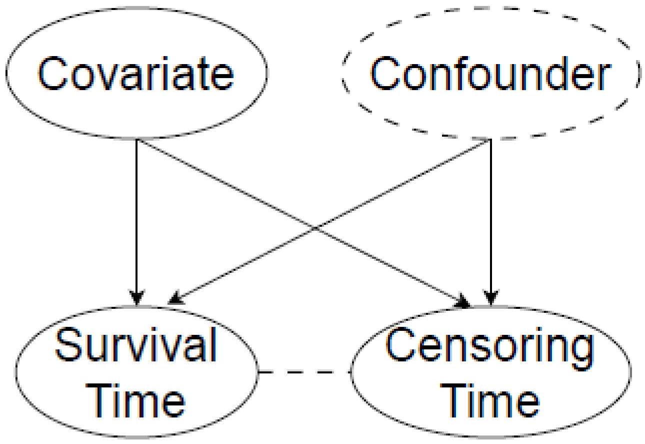


Figure 2: Illustration of dependent censoring in survival analysis. Solid and dashed nodes denote observed and hidden variables, respectively.

4.1 Copula Models

In this section, we introduce a mathematical background to bivariate copula models and a likelihood-based inference method. Let T is survival time and C is censoring time. In addition, let $S_T(t) = P(T > t)$ and $S_C(c) = P(C > c)$ are the marginal survival functions. To model the dependence of T and C, we assume a survival copula model

$$P(T > t, C > c) = C_\theta\big(S_T(t), S_C(c)\big),$$

where C_θ is a parametric copula with parameter θ describes the degree of dependency between T and C. A bivariate copula is defined as a bivariate distribution whose marginal distribution is the uniform distribution on $[0,1]$. Let a bivariate copula, $C_\theta : [0,1]^2 \mapsto [0,1]$, is indexed by a parameter θ. By the definition, any bivariate copula should be satisfying the following conditions

$$(C1)\ \ C_\theta(u,0) = C_\theta(0,v) = 0,\ \ C_\theta(u,1) = u,\ \text{and}\ \ C_\theta(1,v) = v$$

$$\text{for}\ \ 0 \le u \le 1\ \ \text{and}\ \ 0 \le v \le 1.$$

$$(C2)\ \ C_\theta(u_2,v_2) - C_\theta(u_2,v_1) - C_\theta(u_1,v_2) + C_\theta(u_1,v_1) \ge 0$$

$$\text{for}\ \ 0 \le u_1 \le u_2 \le 1\ \ \text{and}\ \ 0 \le v_1 \le v_2 \le 1.$$

$(C1)$ requires the two marginal uniform distributions and $(C2)$ requires that C_θ produces a probability mass on the rectangular region $[u_1, u_2] \times [v_1, v_2]$.

For a copula C_θ, we can consider a pair of random variables (V, W) such that

$P(V \leq u, W \leq v) = C_\theta(u, v)$. If one defines a pair of random variables (T, C) by setting $T = S_T^{-1}(V)$ and $C = S_C^{-1}(W)$, its bivariate survival function satisfies $P(T > t, C > c) = C_\theta\{S_T(t), S_C(c)\}$.

There are some copulas meet conditions $(C1)$ and $(C2)$:

(a) the independent copula is

$$C(u, v) = uv.$$

(b) the Clayton copula by Clayton (1978) is

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

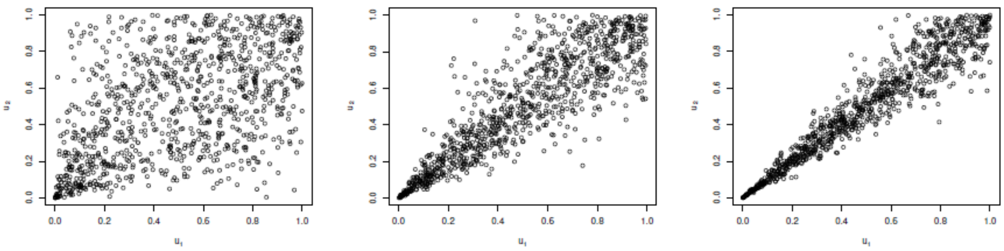
(c) the Gumbel copula by Gumbel (1960) is

$$C_\theta(u, v) = \exp\left(-\left((-\log u)^\theta + (-\log v)^\theta\right)^{\frac{1}{\theta}}\right), \qquad \theta \geq 1.$$

(d) the Frank copula by Frank (1979) is

$$C_\theta(u, v) = -\frac{1}{\theta}\log\left\{1 + \frac{\left(e^{-\theta u} - 1\right)\left(e^{-\theta v} - 1\right)}{e^{-\theta} - 1}\right\}, \qquad \theta \neq 0.$$

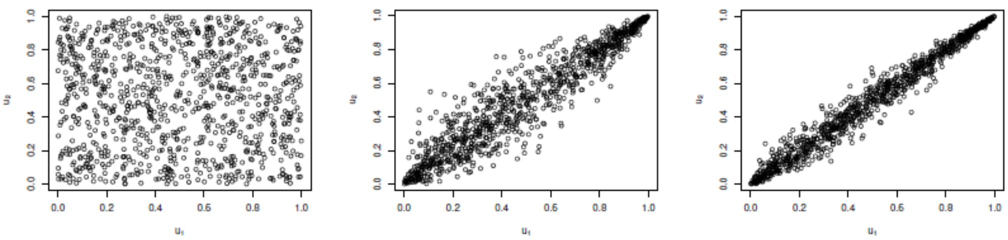
By Tovar Cuevas et al. (2019), the Clayton copula function models a highly dependent asymmetric data structure with the left tail indicating that the cloud is expanding.



(a) $\theta = 1$ (b) $\theta = 5$ (c) $\theta = 10$

Figure 3. Scatter plot of data under the Clayton copula with different θ.

The Gumbel copula is useful for modeling data structures that have a strong dependency on the upper tail and a weak dependency on the lower tail, where we expect the upper data to be strongly correlated and the lower data to be weakly correlated.



(a) $\theta = 1$ (b) $\theta = 5$ (c) $\theta = 10$

Figure 4. Scatter plot of data under the Gumbel copula with different θ.

An Archimedean copula is defined as

$$C_\theta(u, v) = \phi_\theta^{-1}\{\phi_\theta(u) + \phi_\theta(v)\},$$

where $\phi_\theta : [0,1] \longmapsto [0, \infty]$ is called a generator of the copula that is continuous and strictly decreasing function from $\phi_\theta(0) > 0$ to $\phi_\theta(1) = 0$. If $\phi_\theta(0) \equiv \lim_{t \to 0} \phi_\theta(t) = \infty$, the generator is called a strict generator and has the inverse function ϕ_θ^{-1}: $[0,\infty] \longmapsto [0,1]$. The Clayton, Gumbel, and Frank copulas have a strict generator.

Let (V, W) be a pair of random variables that satisfy $P(V \leq u, W \leq v) = C_\theta(u, v)$. To measure of dependence between V and W, Kendall's tau is defined as

$$\tau_\theta = Pr\{(V_2 - V_1)(V_2 - V_1) > 0\} - Pr\{(V_2 - V_1)(V_2 - V_1) < 0\},$$

where (V_1, V_2) and (W_1, W_2) also have the same distribution as (V, W). It can be expressed that

$$\tau_\theta = 4 \int_0^1 \int_0^1 C_\theta(u, v) C_\theta(du, dv) - 1 = 4 \int_0^1 \int_0^1 C_\theta(u, v) C_\theta^{[1,1]}(u, v) du dv - 1,$$

where $C_\theta^{[1,1]}(u, v) = \frac{\partial^2}{\partial u \partial v} C_\theta(u, v)$.

Table 1 summarizes τ_θ for copulas and τ_θ increases with $\tau_\theta \to 1$ as $\theta \to \infty$.

Table 1. Examples of copulas

Copula	Range of θ	Generator: ϕ_θ	Kendall's tau: τ_θ
Clayton	$(0,\ \infty)$	$\theta^{-1}(t^{-\theta} - 1)$	$\theta/(\theta + 2)$
Gumbel	$[1, \infty)$	$(-\log t)^\theta$	$1 - 1/\theta$
Frank	$(-\infty, \infty)$	$-log\left(\frac{e^{-\theta t} - 1}{e^{-\theta} - 1}\right)$	$1 - \frac{4}{\theta}\left(1 - \frac{1}{\theta}\int_0^\theta \frac{t}{e^t - 1} dt\right)$

4.2 Copula-Graphic (CG) Estimator

Zheng and Klein (1995) introduced the concept of using copulas in survival data analysis under dependent censoring. They examined a bivariate distribution function for survival and censoring times with fully specified copula forms, including parameter values, to ensure model identifiability. They used a copula-graphic (CG) estimator to estimate the marginal survival function under this copula assumption. The CG estimator's survival function closely resembles that of the Kaplan-Meier estimator, effectively reducing to a Kaplan-Meier estimator when independence copulas are assumed. In practice, the CG estimator is typically calculated using one of the Archimedean copulas. Rivest and Wells (2001) simplified the CG estimator's expression for Archimedean copulas. Today, CG estimators are crucial tools in survival analysis with dependent censoring (Braekers and Veraverbeke 2005; Emura and Chen 2018).

Under dependent censoring, Kaplan-Meier estimator may introduce biased information but a survival curve calculated form CG estimator gives unbiased information if copula function between death and censoring time is rightly specified. We introduce the CG estimator proposed by Rivest and Wells (2001). Consider random variables defied as T is survival time and C is censoring time and an Archimedean copula model

$$P(T > t, C > c) = \phi_\theta^{-1}[\phi_\theta\{S_T(t)\} + \phi_\theta\{S_C(t)\}],$$

where $\phi_\theta: [0,1] \longmapsto [0,\infty]$ is generator function, which is strictly decreasing and continuous from $\phi_\theta(0) = \infty$ to $\phi_\theta(1) = 0$; $S_T(t) = P(T > t)$ and $S_C(c) = P(C >$

c) are the marginal survival functions.

Let $(t_i, \delta_i), i = 1, \ldots, n,$ be survival data without covariates, where $t_i = \min\{T_i, C_i\}$, $\delta_i = I(T_i \le C_i)$, $I(\cdot)$ is the indicator function. All the observed times are assumed to distinct ($t_i \ne t_j$ whenever $i \ne j$). The CG estimator is defined as

$$\hat{S}_T(t) = \phi_\theta^{-1}\left[\sum_{t_i \le t, \delta_i = 1} \phi_\theta\left(\frac{n_i - 1}{n}\right) - \phi_\theta\left(\frac{n_i}{n}\right)\right], \qquad 0 \le t \le \max_i(t_i)$$

where $n_i = \sum_{\ell=1}^{n} I(t_\ell \ge t_i)$ is the number at risk at time t_i; $\hat{S}_T(t) = 1$ if no death occurs up to time t; $\hat{S}_T(t)$ is undefined for $t > \max_i(t_i)$.

The derivation of the CG estimator can be obtained as follows. Assume that $S_T(t)$ is decreasing step function with jumps at death times. Then, $\delta_i = 1$ implies $S_T(t_i) \ne S_T(t_i - dt)$ and $S_C(t_i) = S_C(t_i - dt)$. Let's set $t = c = t_i$ in $P(T > t, C > c) = \phi_\theta^{-1}[\phi_\theta\{S_T(t)\} + \phi_\theta\{S_C(t)\}]$, we have

$$\phi_\theta\, P(T > t_i, C > t_i) = \phi_\theta\{S_T(t_i)\} + \phi_\theta\{S_C(t_i)\}.$$

In the left-side of the preceding equation, estimate $P(T > t_i, C > t_i)$ by $(n_i - 1)/n$, where $n_i - 1 = \sum_{\ell=1}^{n} I(t_\ell > t_i)$ is the number of survivors at time t_i. Accordingly,

$$\phi_\theta\left(\frac{n_i - 1}{n}\right) = \phi_\theta\{S_T(t_i)\} + \phi_\theta\{S_C(t_i)\}, \qquad \delta_i = 1.$$

Meanwhile, we set $t = c = t_i - dt$ in equation $P(T > t, C > c) =$

$\phi_\theta^{-1}[\phi_\theta\{S_T(t)\}+\phi_\theta\{S_C(t)\}]$ and then estimate $P(T>t_i-dt, C>t_i-dt)$ by n_i/n.

$$\phi_\theta\left(\frac{n_i}{n}\right)=\phi_\theta\{S_T(t_i-dt)\}+\phi_\theta\{S_C(t_i)\},\quad \delta_i=1.$$

The result in the system of difference equation is

$$\phi_\theta\left(\frac{n_i-1}{n}\right)-\phi_\theta\left(\frac{n_i}{n}\right)=\phi_\theta\{S_T(t_i)\}-\phi_\theta\{S_T(t_i-dt)\},\quad \delta_i=1.$$

When t_i is the smallest, we can impose the constraint that $S_T(t_i-dt)=1$. Then, the solution of different equations is

$$\begin{aligned}\phi_\theta\{S_T(t)\}&=\sum_{t_i\le t,\delta_i=1}[\phi_\theta\{S_T(t_i)\}-\phi_\theta\{S_T(t_i-dt)\}]\\&=\sum_{t_i\le t,\delta_i=1}\left[\phi_\theta\left(\frac{n_i-1}{n}\right)-\phi_\theta\left(\frac{n_i}{n}\right)\right],\end{aligned}$$

which is equivalent to the CG estimator.

When $\phi_\theta(t)=-\log(t)$ under independence copula, the CG estimator is same to the Kaplan-Meier estimator and given by $\phi_\theta(t)=(t^{-\theta}-1)/\theta$ for $\theta>0$ under Clayton copula, the CG estimator is

$$\hat{S}_T(t)=\left[1+\sum_{t_i\le t,\delta_i=1}\left\{\left(\frac{n_i-1}{n}\right)^{-\theta}-\left(\frac{n_i}{n}\right)^{-\theta}\right\}\right]^{-\frac{1}{\theta}}.$$

Chapter 5

Proposed Method

We proposed some methods for comparison two RMSTs to deal with dependent censoring. To estimate survival function under dependent censoring, we proposed two survival function estimation method based on modified CG estimator and self-consistency method. There are two situations when comparing two RMSTs under dependent censoring: the assumed copula situation and the situation where the copula assumption is violated. When the copula assumption is violated, it can be categorized as either a copula type misspecified or an association parameter misspecified. For the assumed copula situation, we proposed a method to compare two RMSTs based on modified CG estimator. When the copula assumption is violated, we proposed methods to compare two RMSTs based on modified CG estimator and self-consistency estimator.

5.1 Modified Copula-Graphic Estimator

The CG estimator can be used to estimate RMST under dependent censoring. When test the difference between two RMSTs, we need an estimate of the variance of each RMST. There is no closed form for the RMST variance estimate for the CG estimator, and the RMST variance estimate can be obtained via a bootstrap method. When using the bootstrap method for variance estimation, a tied data can be generated with approximately 63.2% ties. The book by Chernick and LaBudde (2011, p.199) states the following result about bootstrap resamples: If the sample size is large and we generate many bootstrap samples, we will find that on average, approximately 36.8% of the original observations will be missing from the individual bootstrap samples. Another way to look at this is that for any particular observation, approximately 36.8% of the bootstrap samples will not contain it. However, the CG estimator can be used for the data without ties. Thus, we have proposed a modified CG (MCG) estimator to consider tied data as following formula for summarized survival data for distinct time $(t_i, d_i),\ i = 1, \dots, I$

$$\hat{S}_T(t) = \phi_\theta^{-1}\left[\sum_{t_i \le t, d_i \ne 0} \phi_\theta\left(\frac{n_i - d_i}{n}\right) - \phi_\theta\left(\frac{n_i}{n}\right)\right], 0 \le t \le \max_i(t_i)$$

where n is the number of data, $n_i = \sum_{\ell=1}^{L} I(t_\ell \ge t_i)$ is the number at risk at time t_i, $d_i = \sum_{\ell=1}^{L} I(t_\ell = t_i)$ is the number of events at time t_i, and L is the number of distinct times.

The derivation of the modified CG estimator can be obtained by a similar process to

CG estimator. The $d_i > 0$ implies that, in the left-side of the equation mentioned above, estimate $\Pr(\mathrm{T} > t_i, \mathrm{C} > t_i)$ by $(n_i - d_i)/n$, where $n_i - d_i = \sum_{\ell=1}^{n} I(t_\ell > t_i)$ is the number of survivors at time t_i. Accordingly,

$$\phi_\theta\left(\frac{n_i - d_i}{n}\right) = \phi_\theta\{S_T(t_i)\} + \phi_\theta\{S_C(t_i)\}, \qquad \delta_i = 1.$$

The result in the system of difference equations is

$$\phi_\theta\left(\frac{n_i - d_i}{n}\right) - \phi_\theta\left(\frac{n_i}{n}\right) = \phi_\theta\{S_T(t_i)\} - \phi_\theta\{S_T(t_i - dt)\}, \qquad \delta_i = 1.$$

When t_i is the smallest, we can impose the constraint that $S_T(t_i - dt) = 1$. Then, the solution of different equations is

$$\phi_\theta\{S_T(t)\} = \sum_{t_i \le t, \delta_i = 1} [\phi_\theta\{S_T(t_i)\} - \phi_\theta\{S_T(t_i - dt)\}]$$

$$= \sum_{t_i \le t, \delta_i = 1} \left[\phi_\theta\left(\frac{n_i - 1}{n}\right) - \phi_\theta\left(\frac{n_i}{n}\right)\right].$$

5.2 Self-consistency Estimator for Dependent Censoring

Prior research by Jiang, Hongyu, et al (2005) has estimated copula graphic estimator as self-consistency estimator in dependent censoring. For dependent censoring survival data, the self-consistency estimator can be defined in non-parametric approaches. Survival function for event time T and censored time C can be written using copula function as follows:

$$S_T(t) = \frac{1}{n}\left[\sum_{T_i>t}\phi(T_i) + \sum_{\delta_i=0,T_i\leq t}\frac{P(T>t|C=T_i)}{P(T>T_i|C=T_i)}\right]$$

$$= \frac{1}{n}\left[\sum_{T_i>t}\phi(T_i) + \sum_{\delta_i=0,T_i\leq t}\frac{S_T(t|C=T_i)}{S_T(T_i|C=T_i)}\right]$$

$$= \frac{1}{n}\left[\sum_{T_i>t}\phi(T_i) + \sum_{\delta_i=0,T_i\leq t}\frac{\dot{C}_\theta\big(v(T_i),u(t)\big)}{\dot{C}_\theta\big(v(T_i),u(T_i)\big)}\right]$$

and

$$S_C(c) = \frac{1}{n}\left[\sum_{C_i>c}\phi(C_i) + \sum_{\delta_i=1,C_i\leq c}\frac{P(C>c|T=C_i)}{P(C>C_i|T=C_i)}\right]$$

$$= \frac{1}{n}\left[\sum_{C_i>c}\phi(C_i) + \sum_{\delta_i=1,C_i\leq c}\frac{S_C(c|T=C_i)}{S_C(C_i|T=C_i)}\right]$$

$$= \frac{1}{n}\left[\sum_{C_i > c} \phi(C_i) + \sum_{\delta_i = 1, C_i \le c} \frac{\dot{C}_\theta\big(u(C_i), v(c)\big)}{\dot{C}_\theta\big(u(C_i), v(C_i)\big)}\right]$$

where $u,\ v$ are survival functions for $T,\ C$, respectively and $\dot{C}_\theta(x, y) = \frac{\partial}{\partial x} C_\theta(x, y)$.

For non-parametric self-consistency estimator, the survival functions $S_T(t)$ and $S_C(c)$ can be estimated iteratively until they converge by updating the estimates of $S_T(t)$ and $S_C(c)$ alternately. The initial values of $S_T(t)$ and $S_C(c)$ can be used as estimated survival probabilities under independent censoring assumption. In this study, we only consider three copula types (Clayton, Gumbel, Frank) and two marginal distributions (exponential, Weibull) for event and censored times.

5.3 Marginal Survival Estimation

Previous studies (Zheng and Klein, 1994, 1995; Huang and Zhang, 2008; Chen, 2010) have shown that the proposed model is robust to the functional form of the copula (e.g., whether the underlying copula is Clayton, Gumbel, etc.) but sensitive to the assumed level of association (e.g., the copula parameter value that directly corresponds to Kendall's tau).

The authors noted in their simulation study results that a critical requirement for a good estimate of the marginal survival function is a reasonable guess of the strength of association between T and C, rather than the functional form of the copula.

The Kaplan-Meier estimator is biased in estimating the survival curve when there is a dependency between survival time and censoring time. Survival probability tends to be overestimated. However, the MCG model is not biased significantly even if the copula type is misspecified if the association parameter is set to the same as the data generated. However, even if the copula type is the same, it tends to be underestimated or overestimated if the association parameter is different.

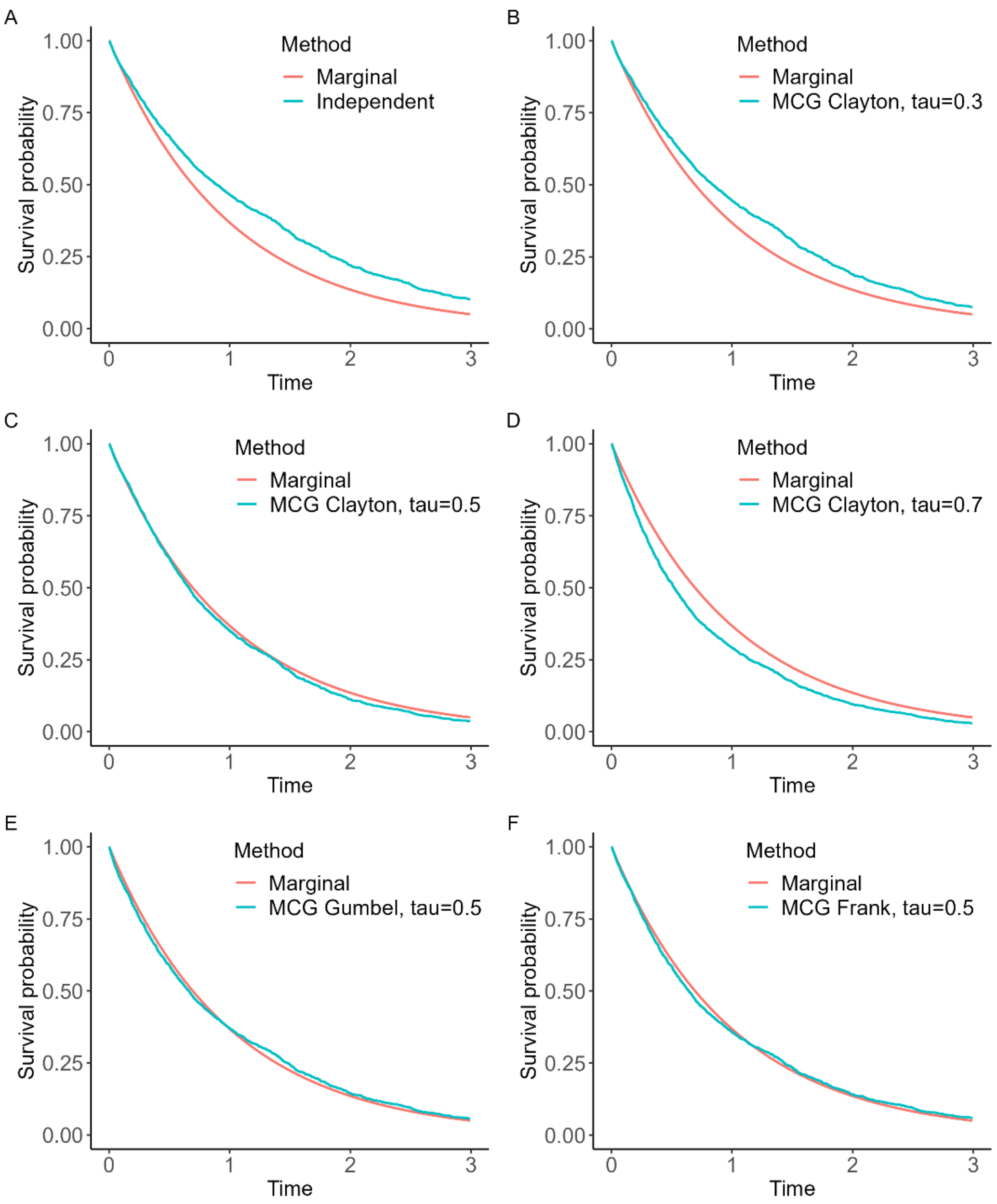


Figure 5. Survival curve estimated by modified Copula-Graphic (MCG) estimators according to Copula type and association parameter.

5.4 Comparison of the RMST for Dependent Censoring

Previous research on copula graphic estimators and self-consistency in the context of RMST has not been conducted, and particularly, studies focusing on their statistical properties (Bias, Type I error) have been absent. We proposed methods to compare two RMSTs under dependent censoring by considering two situations: assumed copula situation and the situation where the copula assumption is violated.

In assumed copula situation, we can estimate the difference of RMST between group 1 and group 0 until γ using the MCG estimator as follows:

$$\widehat{D}_{MCG}(\gamma) = \int_0^{\gamma} \left[\hat{S}_1(\cdot) - \hat{S}_0(\cdot)\right] dt$$

where $\hat{S}(\cdot)$ is the MCG estimator.

In situation that copula assumption is violated, two problems can happen: a copula type misspecified or an association parameter misspecified. It is known to have weak effect for the misspecified copula type and strong effect for the misspecified association parameter on the bias for RMST estimates. Thus, we proposed ensemble method to estimate RMST for pre-specified copula type and association parameters. The weights can be calculated by log-likelihood for copula models under dependent censoring. The difference of RMST between group 1 and group 0 until γ as follows:

Using modified copula graphic estimator under clayton copula,

$$T1 = \sum_{k=1}^{K} w_k \int_0^{\gamma} \left[\hat{S}_{1,k}(t) - \hat{S}_{0,k}(t)\right] dt$$

Using self-consistency estimator under clayton copula,

$$T2 = \sum_{k=1}^{K} w_k \int_0^{\gamma} \left[\tilde{S}_{1,k}(t) - \tilde{S}_{0,k}(t)\right] dt$$

where w_k is association parameter weight, k is the index for association parameter, and $\hat{S}(\cdot)$ can be estimated using the MCG estimator and SC estimator.

The estimated difference of RMST $\widehat{D}(\gamma)$ obtained by proposed methods can be approximated by a normal distribution under null hypothesis ($D = 0$) as follow:

$$\widehat{D}(\gamma) \sim N\left(0, Var\left(\widehat{D}(\gamma)\right)\right)$$

where $Var\left(\widehat{D}(\gamma)\right)$ is the bootstrap variance and test statistic is defined as

$$T = \frac{\widehat{D}(\gamma)}{\sqrt{Var\left(\widehat{D}(\gamma)\right)}}.$$

Chapter 6

Simulation Study

6.1 Simulation Setting

In this section, we conduct various simulation studies to the performance of the proposed new estimators evaluate for testing difference between two RMST by varying dependency, number of samples, type of copulas, association parameters, survival distribution, censoring distribution, and censoring rate. Assuming that the survival distributions of the two groups are the same, the difference of RMST between the two groups was confirmed by performing simulations. The number of samples per group was set to 30, 50, and 100, and both data without dependence censoring and data with dependence censoring were examined. The marginal survival distributions were considered to be the exponential and Weibull distribution as following probability density function:

Exponential distribution:

$$f(x) = \lambda e^{-\lambda x}, \qquad x \geq 0$$

Weibull distribution:

$$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, \qquad x \geq 0$$

The parameters of marginal survival distribution are presented in Table 2 under the null and alternative hypotheses for the difference in RMST between Group A and Group B.

Table 2. Parameters of marginal survival distribution

Marginal survival distribution	Null hypothesis		Alternative hypothesis	
	Group A	Group B	Group A	Group B
Exponential	$\lambda = 0.2$		$\lambda = 0.2$	$\lambda = 0.4$
Weibull	$k = 2, \lambda = 5$		$k = 2, \lambda = 5$	$k = 2, \lambda = 3.5$

We considered three types of copula types: Clayton, Gumbel, and Frank, and set the tau values of the association parameter Kendall's tau to 0.3, 0.5, and 0.7. In this paper, we ensembled three RMST estimators of the Clayton copula with association parameter Kendall's tau of 0.3, 0.5, and 0.7. We only considered Clayton copula because there was no significant difference in RMST estimators according to copula types. We assumed the censoring distribution is the same for each group. The censoring rates considered were 30, 50, and 70%. Since it is dependence data, we need to calculate a parameter for censoring rate to take this into account. It should be obtained through the expression censoring percentage below, and for the convenience of calculation, the value was calculated under the empirical distribution with $n = 100{,}000$.

$$\text{Censoring percentage} = \mathrm{P(X > C)} = \int_0^\infty \int_c^\infty f(x,c)\,dxdc$$

$$= \int_0^\infty \int_c^\infty f(x)g(c)\ddot{C}(S_X(x), S_C(c))\,dxdc$$

Proposed and independent methods were evaluated in terms of type I error, power, and bias. γ is set to 80th percentile of the population's survival time. All analyses used in the simulation were analyzed using R version 4.3.0. Each scenario was repeated 1000 times independently.

6.2 Results

The results of type I error under independent censoring assumption are presented in Tables 3 and are similar to those in assumed copula. The simulation results for type I error under the Clayton copula are presented in Tables 4. The type I error remained constant in both the independent and proposed methods, with no significant deviation for simulation settings of the number of sample size and association parameter Kendall's tau. Under the Gumbel and Frank copula, the results of type I error showed similar to results of Clayton copula (Table 5-6). When the copula type was misspecified, the results of type I error are showed similar to assumed copula, with no significant increase in type I errors.

Tables 7 present the simulation results of power under the Clayton copula according to simulation settings of the number of sample size and association parameter Kendall's tau. The proposed methods were showed overall higher power than independent method in all simulation settings. The decrease in power was lower than the proposed method when the censoring rate increased. Similar trends in power were showed for Gumbel and Frank copula type assumption to results of Clayton copula in Tables 8-9. When the copula type was misspecified, there was no significant difference in power compared to the results of assumed copula.

Table 3. The results of type I error for independent censoring survival data generated with exponential distribution according to the sample size (n) and censoring rate.

n	Censoring rate	Independent	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Clayton Ensemble I	SC Clayton (τ=True)	SC Clayton Ensemble I
	0.3	0.060	0.063	0.064	0.065	0.064	0.050	0.050
30	0.5	0.060	0.061	0.061	0.055	0.061	0.040	0.050
	0.7	0.066	0.070	0.074	0.063	0.068	0.040	0.040
	0.3	0.059	0.056	0.060	0.065	0.060	0.020	0.020
50	0.5	0.054	0.054	0.059	0.057	0.058	0.020	0.020
	0.7	0.060	0.065	0.068	0.055	0.067	0.030	0.030
	0.3	0.050	0.046	0.049	0.054	0.045	0.020	0.020
100	0.5	0.047	0.053	0.050	0.046	0.054	0.040	0.020
	0.7	0.047	0.045	0.049	0.047	0.048	0.060	0.050

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table 3. The results of type I error for independent censoring survival data generated with exponential distribution according to the sample size (n) and censoring rate.

n	Censoring rate	MCG Gumbel ($\tau = 0.3$)	MCG Gumbel ($\tau = 0.5$)	MCG Gumbel ($\tau = 0.7$)	MCG Frank ($\tau = 0.3$)	MCG Frank ($\tau = 0.5$)	MCG Frank ($\tau = 0.7$)
	0.3	0.060	0.063	0.065	0.061	0.062	0.064
30	0.5	0.059	0.058	0.055	0.061	0.054	0.054
	0.7	0.064	0.058	0.050	0.065	0.059	0.051
	0.3	0.056	0.058	0.062	0.055	0.058	0.062
50	0.5	0.052	0.055	0.056	0.054	0.054	0.056
	0.7	0.060	0.057	0.051	0.064	0.066	0.049
	0.3	0.046	0.051	0.054	0.050	0.049	0.057
100	0.5	0.046	0.050	0.050	0.051	0.052	0.048
	0.7	0.046	0.045	0.048	0.046	0.048	0.046

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table 4. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.066	0.067	0.068	0.067	0.064	0.067	0.063	0.065	0.066
	0.5	0.075	0.072	0.068	0.070	0.067	0.063	0.071	0.068	0.062
	0.7	0.071	0.065	0.066	0.063	0.063	0.061	0.067	0.065	0.064
50	0.3	0.066	0.066	0.065	0.057	0.059	0.065	0.061	0.062	0.061
	0.5	0.062	0.065	0.068	0.061	0.067	0.065	0.065	0.069	0.069
	0.7	0.066	0.067	0.068	0.061	0.062	0.068	0.062	0.065	0.065
100	0.3	0.064	0.061	0.059	0.059	0.062	0.058	0.060	0.058	0.053
	0.5	0.056	0.058	0.058	0.056	0.058	0.057	0.055	0.059	0.056
	0.7	0.060	0.058	0.058	0.055	0.055	0.052	0.057	0.055	0.055

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table 4. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.060	0.069	0.068	0.070	0.070	0.070
	0.5	0.072	0.069	0.069	0.070	0.060	0.060
	0.7	0.073	0.065	0.064	0.070	0.060	0.060
50	0.3	0.064	0.065	0.066	0.070	0.080	0.080
	0.5	0.058	0.066	0.065	0.100	0.110	0.110
	0.7	0.059	0.068	0.068	0.100	0.110	0.100
100	0.3	0.056	0.062	0.060	0.080	0.080	0.080
	0.5	0.053	0.056	0.056	0.060	0.060	0.060
	0.7	0.060	0.056	0.058	0.080	0.080	0.060

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table 5. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	MCG Clayton ($\tau = 0.3$)	MCG Clayton ($\tau = 0.5$)	MCG Clayton ($\tau = 0.7$)	MCG Gumbel ($\tau = 0.3$)	MCG Gumbel ($\tau = 0.5$)	MCG Gumbel ($\tau = 0.7$)	MCG Frank ($\tau = 0.3$)	MCG Frank ($\tau = 0.5$)	MCG Frank ($\tau = 0.7$)
30	0.3	0.056	0.056	0.053	0.051	0.050	0.048	0.051	0.048	0.048
	0.5	0.054	0.053	0.054	0.052	0.050	0.048	0.055	0.052	0.049
	0.7	0.055	0.058	0.051	0.056	0.050	0.047	0.055	0.052	0.046
50	0.3	0.072	0.066	0.058	0.073	0.072	0.063	0.072	0.064	0.059
	0.5	0.067	0.064	0.066	0.067	0.061	0.062	0.067	0.062	0.061
	0.7	0.061	0.065	0.062	0.062	0.060	0.057	0.062	0.059	0.058
100	0.3	0.055	0.056	0.055	0.054	0.052	0.058	0.056	0.054	0.056
	0.5	0.053	0.052	0.059	0.056	0.055	0.062	0.057	0.058	0.062
	0.7	0.050	0.051	0.053	0.052	0.055	0.056	0.049	0.052	0.054

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table 5. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.049	0.057	0.053	0.060	0.050	0.060
	0.5	0.057	0.053	0.053	0.040	0.030	0.030
	0.7	0.057	0.052	0.058	0.040	0.040	0.040
50	0.3	0.071	0.071	0.065	0.090	0.070	0.060
	0.5	0.067	0.066	0.065	0.060	0.050	0.060
	0.7	0.068	0.061	0.064	0.060	0.040	0.050
100	0.3	0.059	0.054	0.054	0.050	0.040	0.060
	0.5	0.060	0.053	0.052	0.070	0.050	0.060
	0.7	0.052	0.053	0.051	0.040	0.050	0.050

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table 6. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	MCG Clayton ($\tau = 0.3$)	MCG Clayton ($\tau = 0.5$)	MCG Clayton ($\tau = 0.7$)	MCG Gumbel ($\tau = 0.3$)	MCG Gumbel ($\tau = 0.5$)	MCG Gumbel ($\tau = 0.7$)	MCG Frank ($\tau = 0.3$)	MCG Frank ($\tau = 0.5$)	MCG Frank ($\tau = 0.7$)
30	0.3	0.074	0.074	0.064	0.068	0.070	0.065	0.065	0.069	0.062
	0.5	0.072	0.069	0.056	0.069	0.061	0.059	0.068	0.061	0.056
	0.7	0.062	0.061	0.059	0.053	0.055	0.054	0.058	0.058	0.054
50	0.3	0.060	0.062	0.053	0.055	0.058	0.053	0.059	0.060	0.054
	0.5	0.067	0.065	0.059	0.063	0.057	0.055	0.063	0.059	0.055
	0.7	0.069	0.072	0.061	0.067	0.069	0.063	0.068	0.064	0.064
100	0.3	0.056	0.058	0.058	0.056	0.057	0.056	0.054	0.054	0.058
	0.5	0.047	0.049	0.057	0.047	0.047	0.053	0.047	0.048	0.057
	0.7	0.048	0.050	0.055	0.050	0.047	0.050	0.051	0.048	0.055

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table 6. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.071	0.072	0.073	0.080	0.080	0.070
	0.5	0.061	0.069	0.069	0.080	0.080	0.080
	0.7	0.064	0.062	0.060	0.080	0.080	0.080
50	0.3	0.066	0.059	0.061	0.060	0.050	0.060
	0.5	0.063	0.062	0.062	0.070	0.060	0.070
	0.7	0.063	0.067	0.072	0.080	0.100	0.090
100	0.3	0.050	0.061	0.058	0.070	0.080	0.080
	0.5	0.050	0.049	0.049	0.090	0.080	0.090
	0.7	0.049	0.051	0.049	0.100	0.090	0.090

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

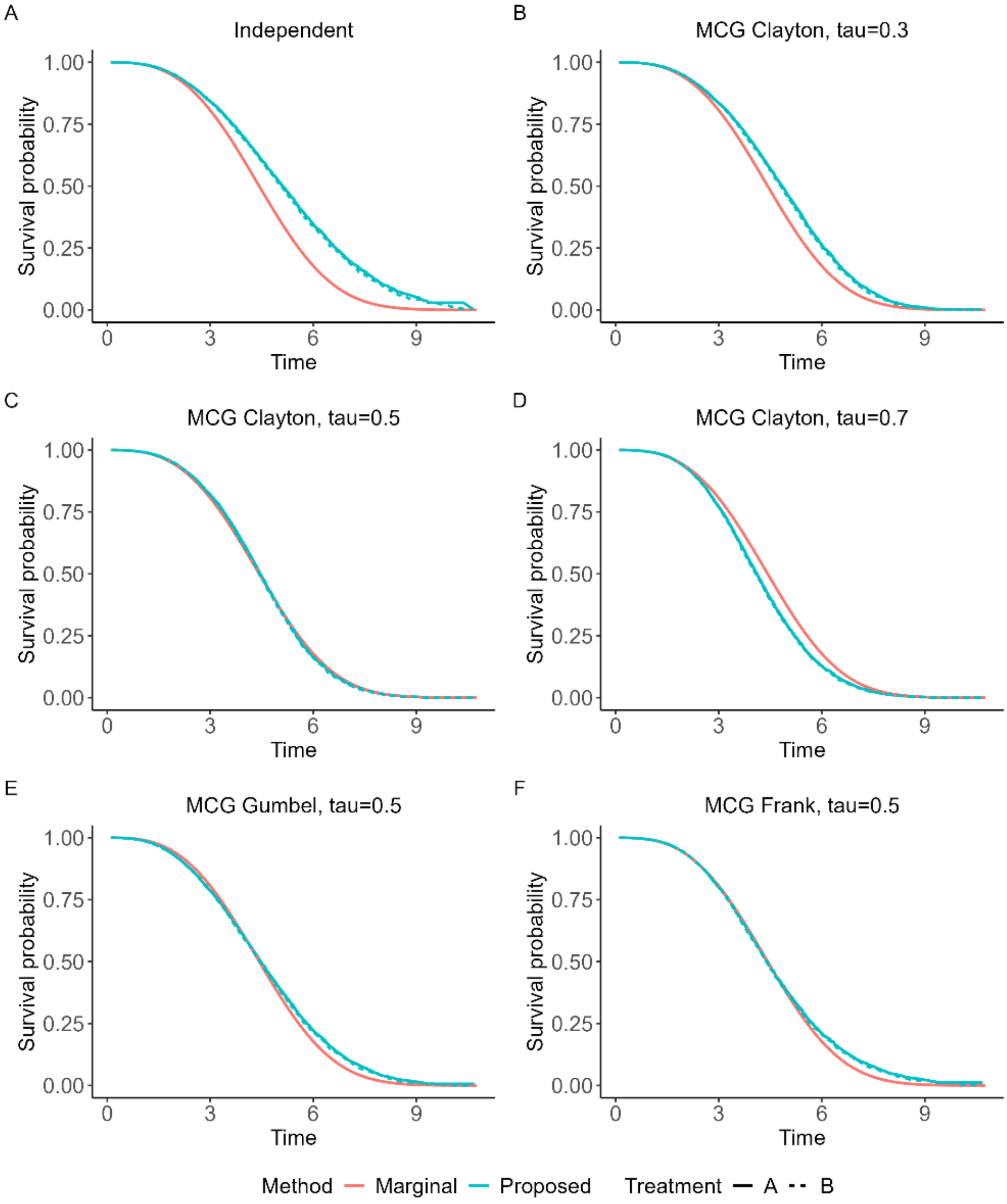
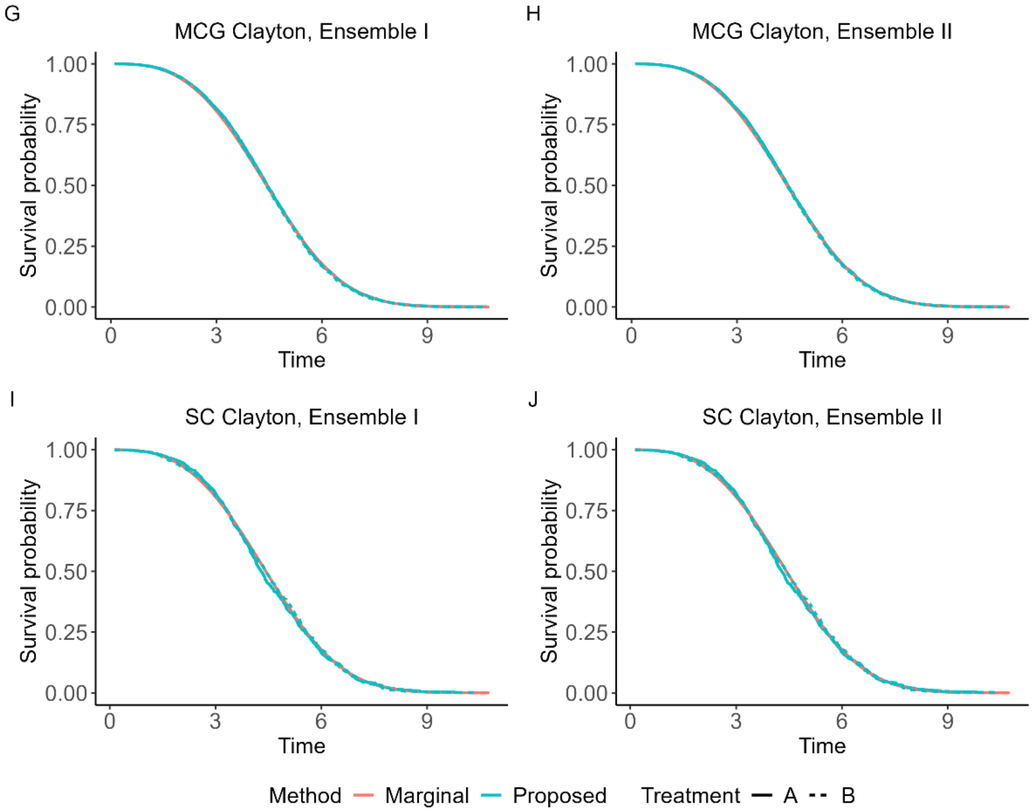


Figure 6. Survival curve estimated by Kaplan-Meier curve (Independent), modified Copula-Graphic (MCG) estimators and Self-consistency (SC) estimators according to Copula type and association parameter under the null hypothesis.



(continuous) Figure 6. Survival curve estimated by Kaplan-Meier curve (Independent), modified Copula-Graphic (MCG) estimators and Self-consistency (SC) estimators according to Copula type and association parameter under the null hypothesis.

Table 7. The results of power for dependent censoring survival data generated with bivariate exponential distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.510	0.556	0.583	0.476	0.522	0.572	0.496	0.536	0.569
	0.5	0.539	0.591	0.619	0.510	0.563	0.605	0.521	0.570	0.607
	0.7	0.543	0.600	0.625	0.517	0.569	0.614	0.531	0.579	0.614
50	0.3	0.721	0.791	0.825	0.691	0.763	0.816	0.707	0.776	0.818
	0.5	0.734	0.803	0.836	0.711	0.774	0.829	0.726	0.790	0.831
	0.7	0.754	0.804	0.838	0.726	0.797	0.830	0.742	0.792	0.824
100	0.3	0.958	0.981	0.987	0.951	0.975	0.987	0.959	0.977	0.986
	0.5	0.971	0.988	0.991	0.959	0.985	0.991	0.966	0.989	0.991
	0.7	0.974	0.989	0.991	0.965	0.984	0.992	0.969	0.987	0.991

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table 7. The results of power for dependent censoring survival data generated with bivariate exponential distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.407	0.534	0.554	0.530	0.530	0.530
	0.5	0.447	0.587	0.587	0.570	0.580	0.570
	0.7	0.442	0.613	0.591	0.590	0.620	0.580
50	0.3	0.596	0.761	0.790	0.660	0.710	0.720
	0.5	0.618	0.803	0.803	0.790	0.780	0.790
	0.7	0.636	0.819	0.802	0.790	0.760	0.730
100	0.3	0.879	0.974	0.981	0.940	0.960	0.980
	0.5	0.896	0.988	0.988	0.990	0.990	0.990
	0.7	0.911	0.991	0.989	0.990	0.990	0.990

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table 8. The results of power for dependent censoring survival data generated with bivariate exponential distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.458	0.506	0.530	0.425	0.470	0.521	0.443	0.484	0.513
	0.5	0.425	0.466	0.495	0.408	0.441	0.474	0.418	0.448	0.477
	0.7	0.447	0.494	0.503	0.417	0.464	0.491	0.431	0.475	0.493
50	0.3	0.612	0.674	0.715	0.588	0.645	0.698	0.605	0.654	0.703
	0.5	0.591	0.644	0.663	0.567	0.599	0.634	0.586	0.617	0.639
	0.7	0.652	0.707	0.734	0.612	0.663	0.707	0.627	0.685	0.712
100	0.3	0.835	0.926	0.999	0.810	0.883	0.966	0.825	0.907	0.987
	0.5	0.858	0.960	1.000	0.816	0.899	0.986	0.846	0.925	1.000
	0.7	0.908	0.949	0.966	0.890	0.932	0.960	0.899	0.938	0.967

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table 8. The results of power for dependent censoring survival data generated with bivariate exponential distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.352	0.479	0.502	0.390	0.400	0.450
	0.5	0.336	0.462	0.461	0.400	0.400	0.400
	0.7	0.347	0.503	0.492	0.470	0.460	0.470
50	0.3	0.502	0.641	0.669	0.580	0.600	0.670
	0.5	0.504	0.636	0.636	0.630	0.630	0.620
	0.7	0.564	0.721	0.704	0.710	0.710	0.680
100	0.3	0.702	0.879	0.929	0.940	0.940	0.960
	0.5	0.728	0.961	0.961	0.950	0.950	0.950
	0.7	0.804	0.959	0.947	0.970	0.970	0.960

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table 9. The results of power for dependent censoring survival data generated with bivariate exponential distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.448	0.488	0.507	0.427	0.465	0.497	0.438	0.472	0.500
	0.5	0.452	0.496	0.505	0.432	0.468	0.499	0.446	0.473	0.495
	0.7	0.477	0.519	0.532	0.454	0.488	0.514	0.470	0.499	0.513
50	0.3	0.651	0.717	0.749	0.626	0.695	0.733	0.639	0.705	0.741
	0.5	0.651	0.717	0.748	0.624	0.676	0.722	0.636	0.689	0.729
	0.7	0.680	0.735	0.763	0.644	0.696	0.749	0.659	0.709	0.750
100	0.3	0.922	0.956	0.970	0.905	0.947	0.966	0.916	0.951	0.969
	0.5	0.938	0.965	0.974	0.918	0.954	0.971	0.928	0.958	0.971
	0.7	0.939	0.972	0.980	0.929	0.963	0.975	0.938	0.967	0.974

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table 9. The results of power for dependent censoring survival data generated with bivariate exponential distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.365	0.470	0.488	0.400	0.400	0.420
	0.5	0.378	0.493	0.493	0.420	0.410	0.420
	0.7	0.395	0.525	0.514	0.470	0.470	0.450
50	0.3	0.518	0.687	0.714	0.600	0.650	0.710
	0.5	0.510	0.716	0.716	0.660	0.660	0.650
	0.7	0.542	0.750	0.729	0.720	0.710	0.700
100	0.3	0.819	0.942	0.955	0.920	0.920	0.950
	0.5	0.828	0.965	0.965	0.950	0.950	0.950
	0.7	0.851	0.977	0.971	0.960	0.960	0.950

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

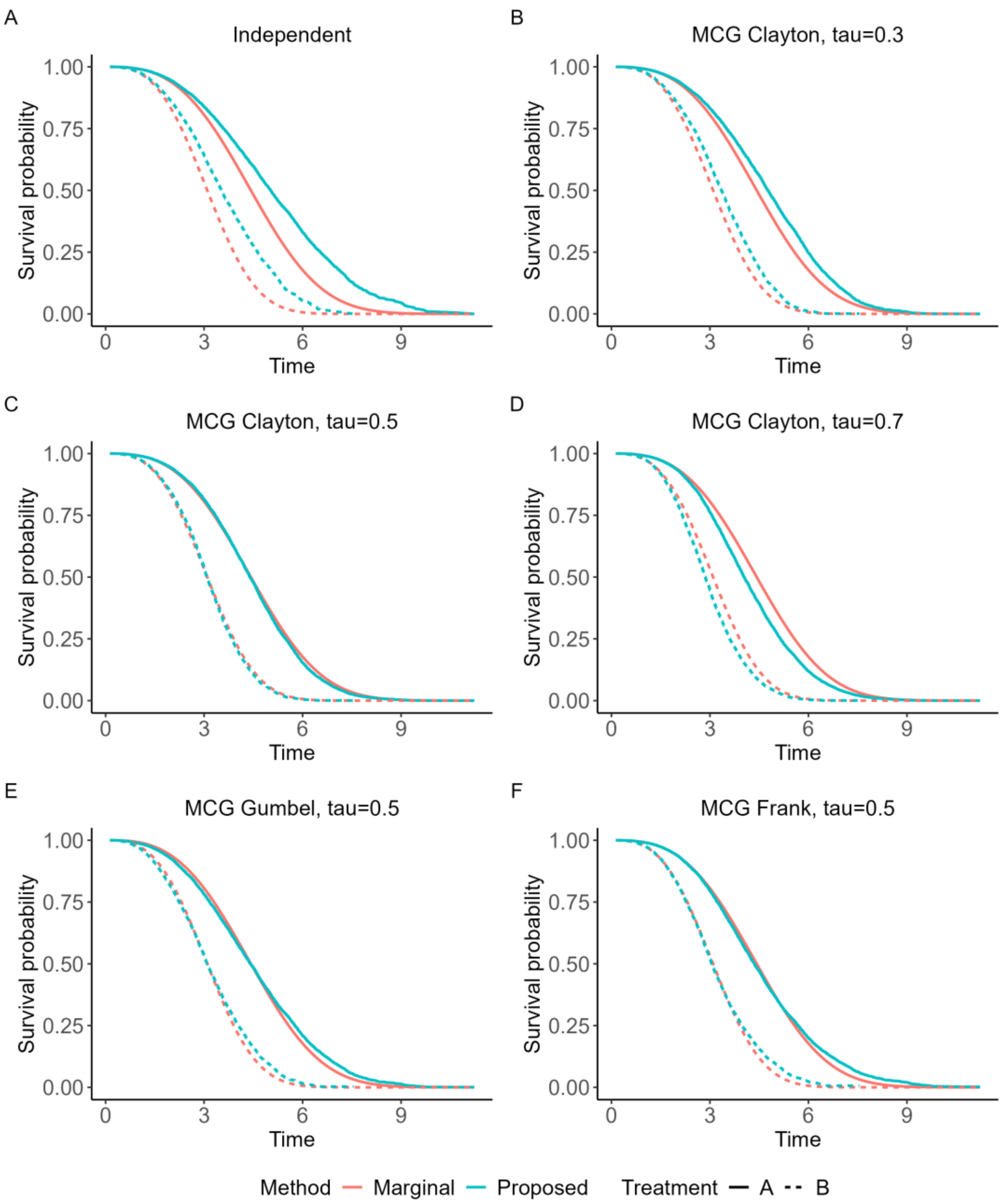
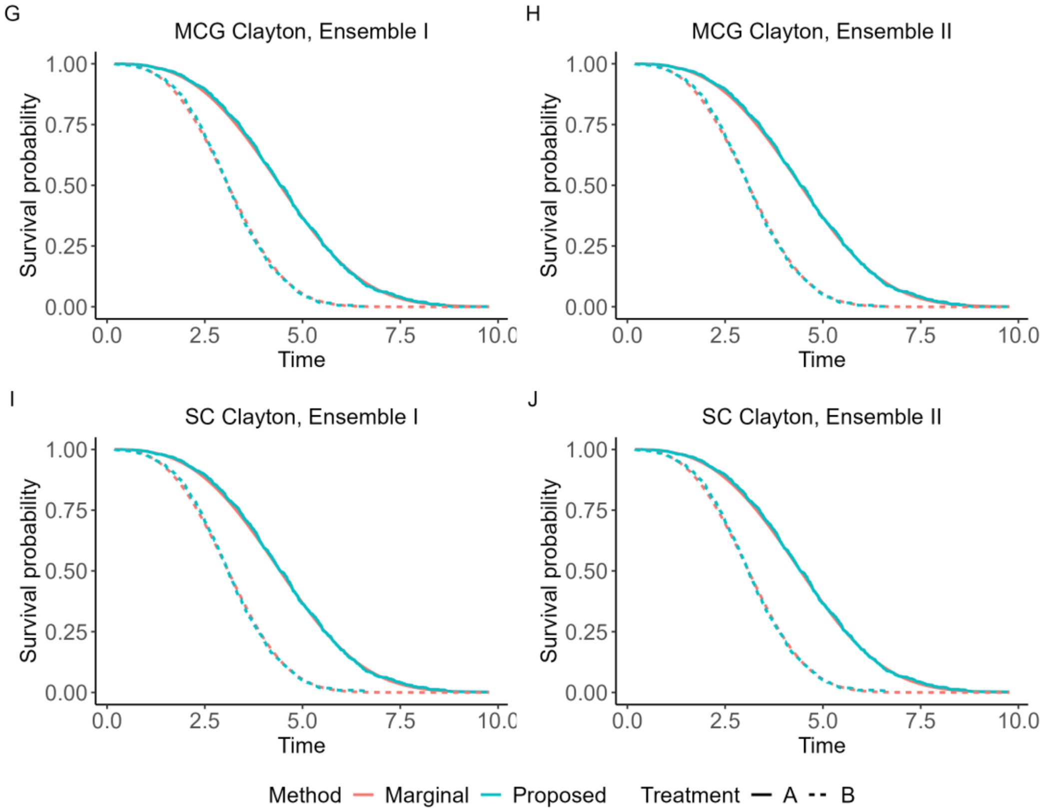


Figure 7. Survival curve estimated by Kaplan-Meier curve (Independent), modified Copula-Graphic (MCG) estimators and Self-consistency (SC) estimators according to Copula type and association parameter under the alternative hypothesis.



(continuous) Figure 7. Survival curve estimated by Kaplan-Meier curve (Independent), modified Copula-Graphic (MCG) estimators and Self-consistency (SC) estimators according to Copula type and association parameter under the alternative hypothesis.

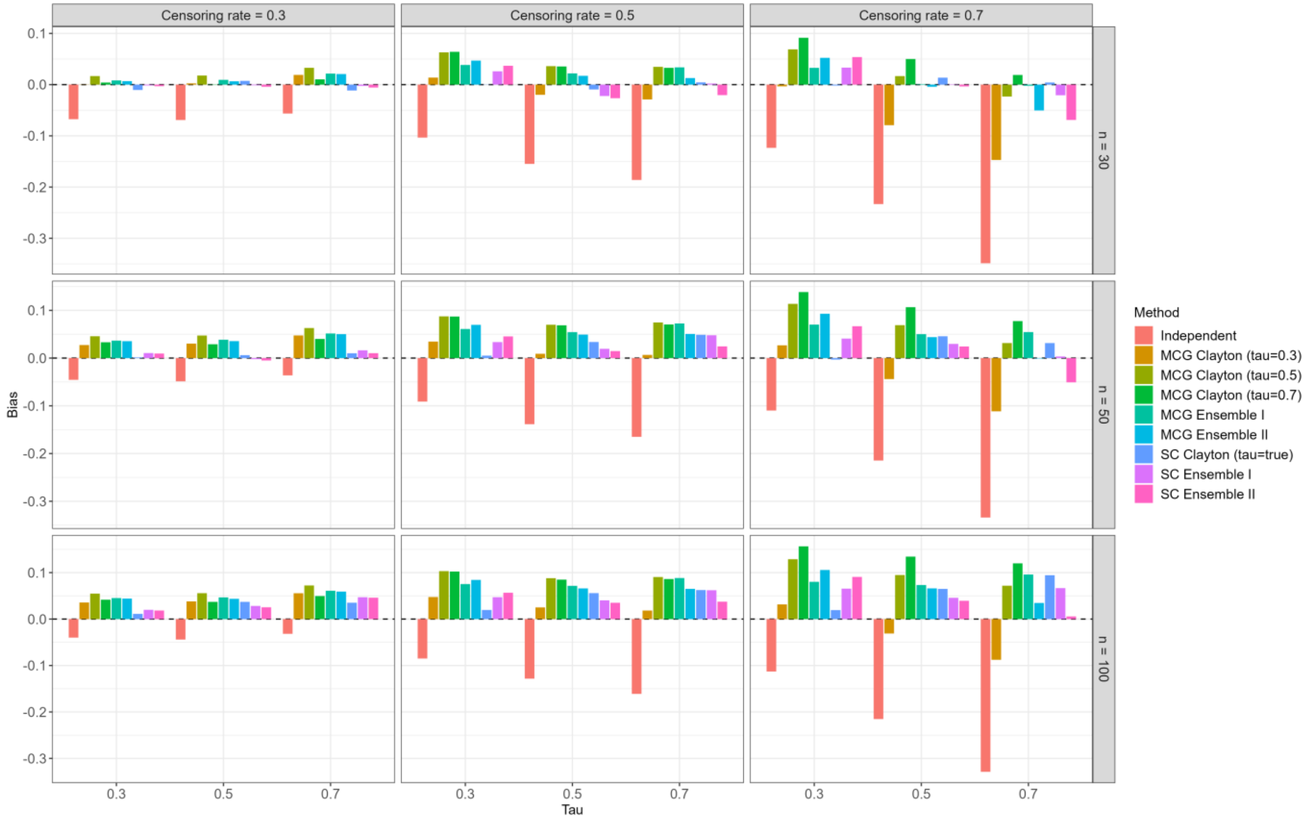


Figure 8. The results of bias for dependent censoring survival data generated with bivariate exponential distribution under the Clayton copula according to the censoring rate and association parameter Kendall's tau (τ).

Chapter 7

Real Data Analysis

7.1 Data on Kidney Transplant Patients

In this section, we used real data examples to illustrate the performance of the proposed method. We applied the proposed estimator to a dataset on the time to death of 863 kidney transplant patients. All patients underwent their transplants at The Ohio State University Transplant Center between 1982 and 1992. The maximum follow-up time was 9.47 years. Patients were censored if they moved away from Columbus (lost to follow-up) or if they were alive on June 30, 1992.

The age of patients at the time of transplant ranged from 9.5 months to 74.5 years, with an average age of 42.8 years. The sample included 524 males and 339 females, with 712 white patients and 151 black patients. Out of the 863 patients, 140 (16.2%) experienced failure events (deaths), while 723 (83.8%) were censored by the end of the study. Specifically, 87 males (16.6%), 53 females (15.6%), 112 white patients (15.7%), and 28 black patients (18.5%) died before the study concluded.

Figure 9, 10 represents the survival function estimated by the product-limit estimator and proposed estimator according to gender and race. We apply the proposed method to directly calculate the difference of RMST and investigate the impact of gender and race on RMST among patients with kidney transplant.

The difference of RMST at τ = 3, 5 and 7 years were analyzed. Table 10, 11 summarizes the results. Overall, the covariate effects show similar trends across all models. Toward the end of the survival curve, we confirm that the survival probability of the proposed models decreases. There was a trend for female to survive longer to a given time point γ compared to male, but it was not significant. Similarly, in race, there were no models and time points showing significant differences, but when the difference in RMST up to 3 years was identified, the trend was different according to the model.

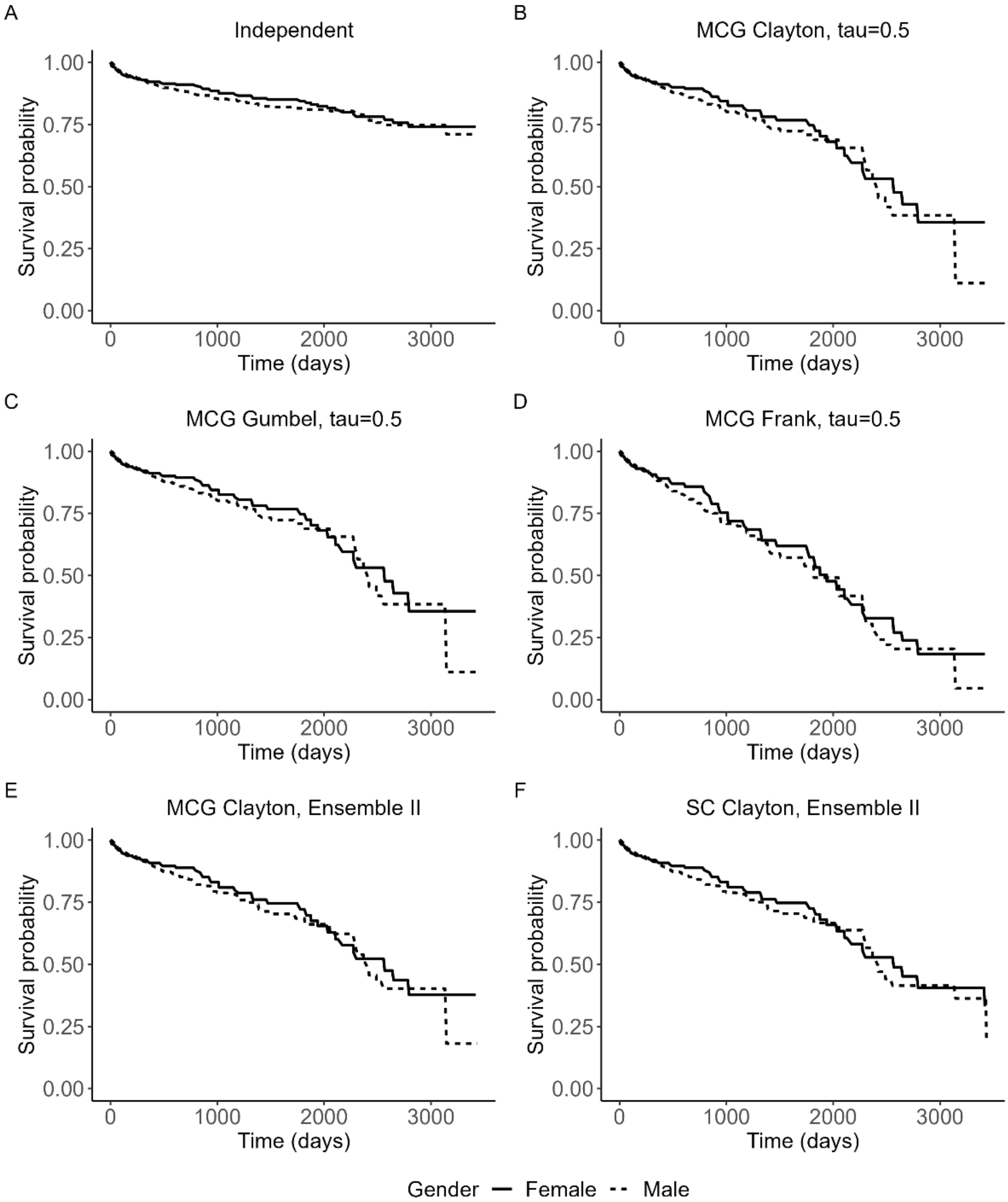


Figure 9. Estimator of survival function using prevailing method and proposed estimator. Product-limit estimator (A), MCG Clayton, tau=0.5 (B), MCG Gumbel, tau=0.5 (C), MCG Frank, tau=0.5 (D), MCG Clayton, Ensemble II (E) and Self-consistency Clayton, Ensemble II (F) for patients by gender (solid line for female; dashed line for male)

Table 10. The difference of RMST by gender with 95% confidence intervals (CIs) and p-values at various values of γ (years)

γ	Independent		MCG Clayton, tau=0.3		MCG Clayton, tau=0.5		MCG Clayton, tau=0.7		MCG Gumbel, tau=0.3	
	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value
3	15.91 (-22.91, 54.73)	0.422	18.59 (-22.52, 59.69)	0.375	21.91 (-23.03, 66.85)	0.339	26.71 (-26.70, 80.11)	0.327	21.20 (-20.06, 62.47)	0.314
5	36.70 (-37.71, 111.12)	0.334	42.89 (-40.05, 125.84)	0.311	49.80 (-46.22, 145.82)	0.309	55.01 (-59.17, 169.18)	0.345	44.85 (-35.44, 125.15)	0.274
7	42.96 (-72.31, 158.22)	0.465	45.52 (-91.38, 182.42)	0.515	51.42 (-111.63, 214.48)	0.537	64.86 (-107.98, 237.71)	0.462	45.95 (-79.90, 171.79)	0.474

γ	MCG Gumbel, tau=0.5		MCG Gumbel, tau=0.7		MCG Frank, tau=0.3		MCG Frank, tau=0.5		MCG Frank, tau=0.7	
	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value
3	23.47 (-20.44, 67.37)	0.295	22.16 (-25.11, 69.43)	0.358	20.66 (-22.27, 63.58)	0.346	24.98 (-23.32, 73.27)	0.311	26.22 (-28.93, 81.37)	0.351
5	45.77 (-39.63, 131.17)	0.294	39.94 (-49.72, 129.60)	0.383	45.57 (-40.69, 131.82)	0.300	50.06 (-47.68, 147.80)	0.315	46.41 (-57.14, 149.96)	0.380
7	40.63 (-91.61, 172.88)	0.547	33.25 (-100.52, 167.02)	0.626	44.41 (-92.83, 181.65)	0.526	41.81 (-107.79, 191.41)	0.584	40.65 (-105.18, 186.48)	0.585

γ	MCG Ensemble II		SC Clayton, tau=0.3		SC Clayton, tau=0.5		SC Clayton, tau=0.7		SC Ensemble II	
	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value
3	22.40 (-23.54, 68.34)	0.339	18.59 (-23.99, 61.16)	0.392	21.93 (-23.48, 67.35)	0.344	27.05 (-24.51, 78.62)	0.304	22.52 (-22.74, 67.79)	0.329
5	49.23 (-46.77, 145.24)	0.315	42.93 (-39.55, 125.40)	0.308	50.02 (-43.52, 143.57)	0.295	55.87 (-55.47, 167.22)	0.325	49.61 (-45.01, 144.22)	0.304
7	53.94 (-100.69, 208.56)	0.494	44.29 (-91.10, 179.69)	0.521	48.16 (-111.96, 208.29)	0.556	51.14 (-125.65, 227.93)	0.571	47.86 (-110.73, 206.46)	0.554

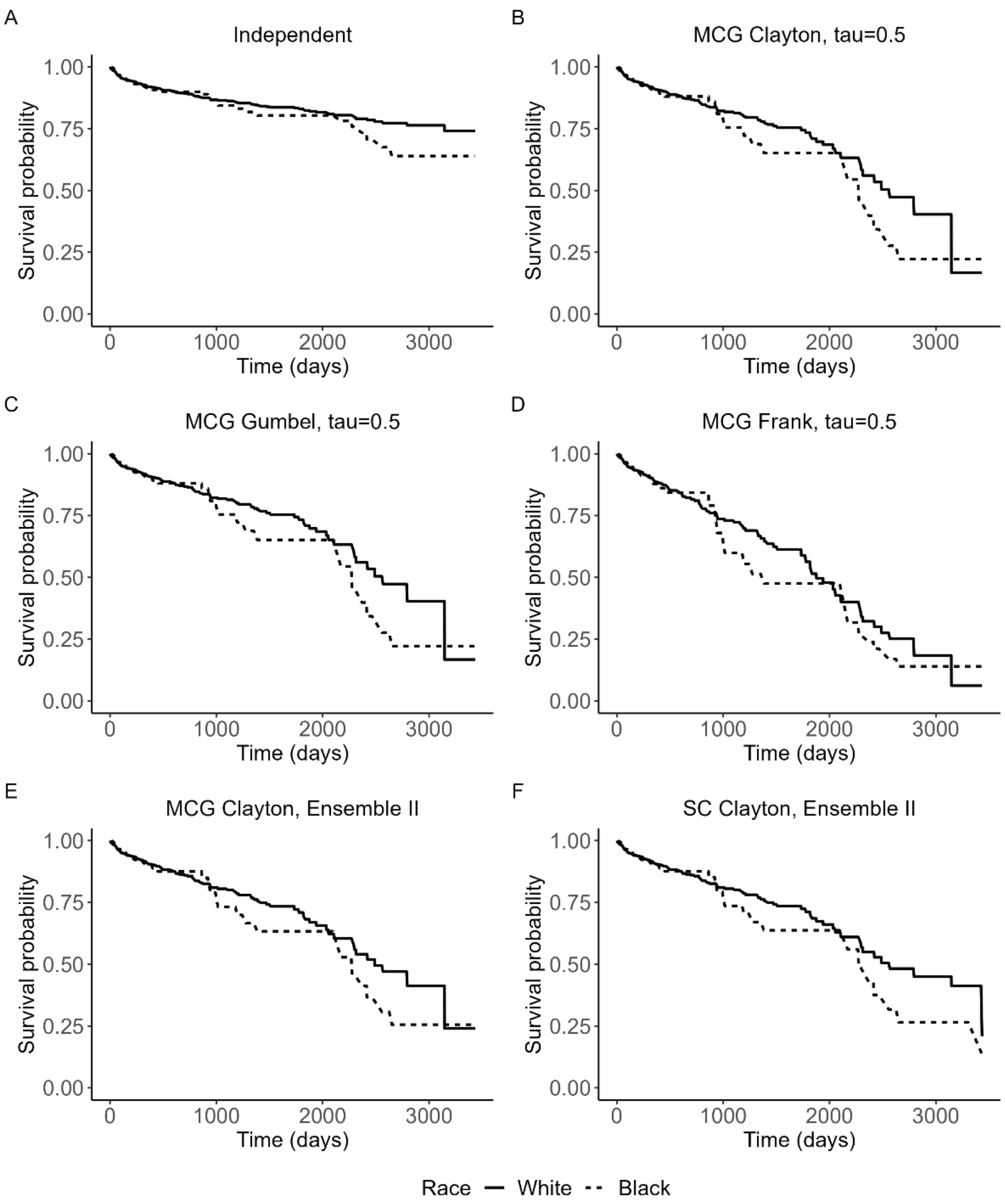


Figure 10. Estimator of survival function using prevailing method and proposed estimator. Product-limit estimator (A), MCG Clayton, tau=0.5 (B), MCG Gumbel, tau=0.5 (C), MCG Frank, tau=0.5 (D), MCG Clayton, Ensemble II (E) and Self-consistency Clayton, Ensemble II (F) for patients by race (solid line for black; dashed line for white)

Table 11. The difference of RMST by race with 95% confidence intervals (CIs) and p-values at various values of γ (years)

γ	Independent		MCG Clayton, tau=0.3		MCG Clayton, tau=0.5		MCG Clayton, tau=0.7		MCG Gumbel, tau=0.3	
	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value
3	-0.06 (-50.42, 50.30)	0.998	0.54 (-53.17, 54.26)	0.984	2.15 (-56.78, 61.07)	0.943	6.23 (-63.08, 75.54)	0.860	-2.34 (-55.66, 50.98)	0.932
5	21.57 (-77.95, 121.08)	0.671	40.20 (-77.25, 157.65)	0.502	70.15 (-68.99, 209.29)	0.323	106.37 (-46.33, 259.07)	0.172	32.69 (-76.86, 142.24)	0.559
7	44.40 (-112.15, 200.96)	0.578	86.63 (-105.00, 278.26)	0.376	131.25 (-93.24, 355.73)	0.252	141.23 (-84.16, 366.61)	0.219	66.95 (-104.93, 238.83)	0.445

γ	MCG Gumbel, tau=0.5		MCG Gumbel, tau=0.7		MCG Frank, tau=0.3		MCG Frank, tau=0.5		MCG Frank, tau=0.7	
	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value
3	-1.66 (-58.33, 55.00)	0.954	1.51 (-59.53, 62.54)	0.961	0.32 (-55.55, 56.20)	0.991	1.66 (-60.89, 64.21)	0.959	4.04 (-66.59, 74.67)	0.911
5	49.30 (-66.96, 165.55)	0.406	71.12 (-49.05, 191.28)	0.246	43.49 (-75.81, 162.80)	0.475	68.07 (-64.06, 200.20)	0.313	86.04 (-47.83, 219.90)	0.208
7	92.44 (-86.96, 271.84)	0.313	113.65 (-64.34, 291.65)	0.211	84.54 (-103.76, 272.84)	0.379	115.71 (-84.69, 316.10)	0.258	122.09 (-66.97, 311.14)	0.206

γ	MCG Ensemble II		SC Clayton, tau=0.3		SC Clayton, tau=0.5		SC Clayton, tau=0.7		SC Ensemble II	
	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value
3	2.97 (-57.01, 62.96)	0.923	0.42 (-52.40, 53.24)	0.988	1.56 (-56.53, 59.65)	0.958	3.84 (-63.14, 70.81)	0.911	1.94 (-55.74, 59.61)	0.948
5	72.24 (-61.93, 206.41)	0.291	39.35 (-76.07, 154.78)	0.504	66.79 (-69.17, 202.76)	0.336	101.06 (-52.24, 254.35)	0.196	69.07 (-62.06, 200.19)	0.302
7	119.70 (-90.37, 329.77)	0.264	83.75 (-102.27, 269.78)	0.378	122.41 (-95.68, 340.50)	0.271	130.33 (-101.34, 362.00)	0.270	112.17 (-103.15, 327.48)	0.307

7.2 Data on Tongue Cancer Patients

The data come from a study examining the effects of ploidy on the prognosis of patients with oral cancers (Sickle-Santanello et al., 1988; Klein & Moeschberger, 1997). Patients were chosen based on the availability of paraffin-embedded cancerous tissue samples collected during surgery. Follow-up survival data was obtained on each patient. The tissue samples were analyzed with a flow cytometer to identify if the tumor had an aneuploid (abnormal) or diploid (normal) DNA profile, as outlined by Sickle-Santanello et al. (1988).

Our goal is to calculate the difference in RMST. Among the 80 patients, 52 (65.0%) had an aneuploid tumor, and 28 (35.0%) had a diploid tumor. Out of the 80 patients, 53 (66.3%) experienced deaths and 27 (33.7%) were censored by the end of the study. Specifically, 31 patients with aneuploid tumors (59.6%) and 22 patients with diploid tumors (78.6%) died before the study concluded. The maximum follow-up time was 7.67 years, with a median follow-up time of 1.4 years.

Figure 11 shows the survival function estimated by the product-limit estimator and the proposed estimator among tongue cancer patients with aneuploid and diploid tumors. We applied the proposed method to directly calculate the difference in RMST and examine the impact of tumor ploidy on RMST among patients with tongue cancer.

The difference of RMST at τ = 50, 100 and 150 weeks were analyzed. Table 12 summarizes the results. Overall, the covariate effects show similar trends across all models.

In the models with MCG Gumbel (tau=0.7) and MCG Frank (tau=0.7), patients with aneuploid tumors are significantly associated with longer average survival times at all values of γ=50. Patients with aneuploid tumors survived longer up to the specified time point γ compared to those with diploid tumors. Specifically, aneuploid tumors are associated with an increase in survival time by 8.14 weeks (95% CI: 0.12-16.16) on average during the next 50 weeks post-diagnosis, using the MCG Gumbel (tau=0.7) model, and by 8.16 weeks (95% CI: 0.08-16.24), using the MCG Frank (tau=0.7) model. Other models showed similar but non-significant values and trends.

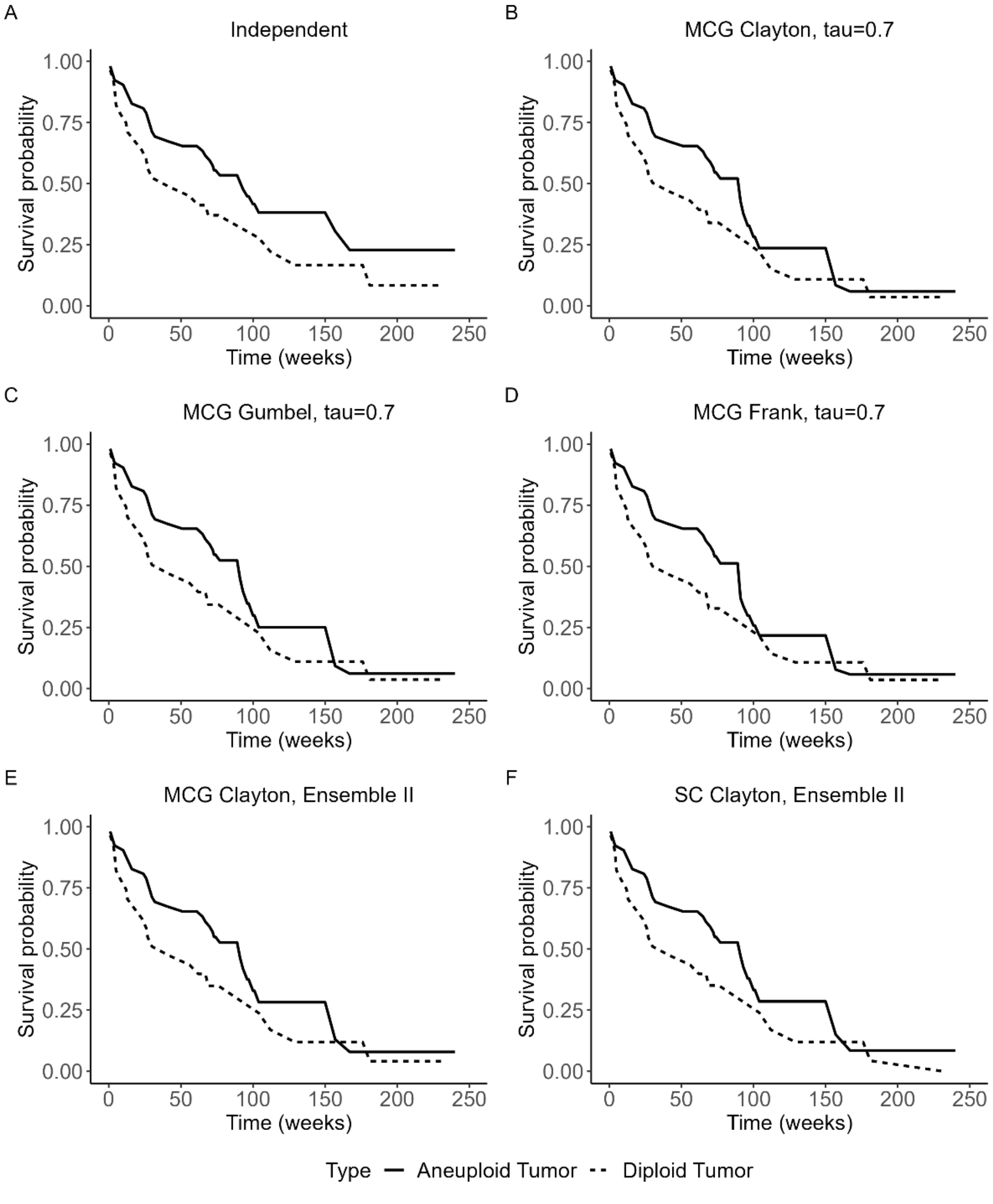


Figure 11. Estimator of survival function using prevailing method and proposed estimator. Product-limit estimator (A), MCG Clayton, tau=0.7 (B), MCG Gumbel, tau=0.7 (C), MCG Frank, tau=0.7 (D), MCG Clayton, Ensemble II (E) and Self-consistency Clayton, Ensemble II (F) for patients by tumor type (solid line for aneuploid; dashed line for diploid)

Table 12. The difference of RMST by tumor type with 95% confidence intervals (CIs) and p-values at various values of γ (weeks)

γ	Independent		MCG Clayton, tau=0.3		MCG Clayton, tau=0.5		MCG Clayton, tau=0.7		MCG Gumbel, tau=0.3	
	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value
50	7.36 (-1.05, 15.76)	0.086	7.58 (-0.51, 15.67)	0.066	7.79 (-0.32, 15.90)	0.060	8.06 (-0.06, 16.18)	0.052	7.62 (-0.39, 15.63)	0.062
100	16.02 (-2.16, 34.19)	0.084	16.47 (-1.16, 34.10)	0.067	16.73 (-0.85, 34.31)	0.062	16.84 (-0.59, 34.27)	0.058	16.53 (-0.88, 33.95)	0.063
150	24.25 (-1.93, 50.42)	0.069	23.63 (-1.52, 48.79)	0.066	22.35 (-2.21, 46.92)	0.075	20.86 (-2.76, 44.48)	0.083	24.17 (-0.66, 49.00)	0.056

γ	MCG Gumbel, tau=0.5		MCG Gumbel, tau=0.7		MCG Frank, tau=0.3		MCG Frank, tau=0.5		MCG Frank, tau=0.7	
	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value
50	7.86 (-0.14, 15.86)	0.054	8.14 (0.12, 16.16)	0.047	7.66 (-0.40, 15.72)	0.063	7.90 (-0.16, 15.97)	0.055	8.16 (0.08, 16.24)	0.048
100	16.87 (-0.43, 34.18)	0.056	17.04 (-0.16, 34.25)	0.052	16.59 (-0.87, 34.06)	0.063	16.88 (-0.46, 34.23)	0.056	16.95 (-0.26, 34.15)	0.054
150	23.64 (-0.61, 47.90)	0.056	22.31 (-1.22, 45.85)	0.063	23.85 (-0.79, 48.49)	0.058	23.03 (-0.89, 46.94)	0.059	21.56 (-1.54, 44.66)	0.067

γ	MCG Ensemble II		SC Clayton, tau=0.3		SC Clayton, tau=0.5		SC Clayton, tau=0.7		SC Ensemble II	
	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value	RMST (95% CI)	p-value
50	7.81 (-0.29, 15.91)	0.059	7.57 (-0.52, 15.67)	0.067	7.78 (-0.34, 15.90)	0.060	8.05 (-0.09, 16.18)	0.052	7.80 (-0.63, 16.23)	0.070
100	16.68 (-0.85, 34.20)	0.062	16.46 (-1.17, 34.09)	0.067	16.73 (-0.88, 34.33)	0.063	16.85 (-0.64, 34.34)	0.059	16.68 (-1.60, 34.95)	0.074
150	22.28 (-2.11, 46.67)	0.073	23.68 (-1.68, 49.04)	0.067	22.51 (-2.33, 47.35)	0.076	21.07 (-2.84, 44.98)	0.084	22.42 (-2.87, 47.71)	0.082

Chapter 8

Conclusion and Discussion

The paper aims to develop a method for accurately estimating the difference in Restricted Mean Survival Time (RMST) between two groups when there is a dependency. The challenge lies in identifying the association parameter in data with dependent censoring. To address this, the authors propose providing an appropriate association range. They found that using modified Copula-Graphic (MCG) or Self-consistency (SC) models reduces bias in the survival curve compared to the Kaplan-Meier model, which assumes independence. The study emphasizes the importance of fitting the association parameters well over the copula type and proposes a weighted sum form with association parameters. Simulations confirmed that the method satisfies the type I error rate and performs similarly to the true model. The results indicate that the association parameter significantly impacts the model, even when the copula type is misspecified. The paper suggests expanding future models to include both the association parameter and copula type and highlights the need for theoretical development due to the self-consistency conversion rate problem.

The study successfully demonstrates that addressing dependent censoring in survival analysis by focusing on the accurate estimation of association parameters enhances model performance. The proposed method, which uses a weighted sum form with association parameters, outperforms traditional methods that assume independence, such as the

Kaplan-Meier model. The simulation results validate the robustness and reliability of the proposed method, even with copula type misspecification, underscoring the critical importance of the association parameter.

This research has significant implications for survival analysis, particularly in fields like medical research where dependent censoring is prevalent. By reducing bias in survival curve estimation, the proposed method can lead to more accurate and reliable assessments of treatment effects and patient outcomes. Future research should explore integrating both association parameters and copula types to further enhance model accuracy. Additionally, practical applications using real-world datasets will be essential for further validation and refinement of the method. Addressing the balance between model complexity and computational efficiency will remain a key consideration, and ongoing advancements in statistical methodologies will likely continue to improve the applicability and precision of these models. The identified need for theoretical development to address the self-consistency conversion rate problem also presents an important area for future investigation.

Reference

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. Journal of the American statistical association, 53(282), 457-481.

Royston, P., & Parmar, M. K. (2011). The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. Statistics in medicine, 30(19), 2409-2421.

Hasegawa, T., Misawa, S., Nakagawa, S., Tanaka, S., Tanase, T., Ugai, H., ... & JPMA Task Force Members. (2020). Restricted mean survival time as a summary measure of time-to-event outcome. Pharmaceutical statistics, 19(4), 436-453.

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(1), 141-151.

Gumbel, E. J. (1960). Bivariate exponential distributions. Journal of the American Statistical Association, 55(292), 698-707.

Frank, A. G. (1979). Dependent accumulation (Vol. 492). NYU Press.

Peres, M. V. D. O., de Oliveira, R. P., Achcar, J. A., & Martinez, E. Z. (2020). The Bivariate defective Gompertz distribution based on Clayton Copula with applications to medical data. arXiv preprint arXiv:2012.07824.

Emura, T., & Chen, Y. H. (2018). Analysis of survival data with dependent censoring: copula-based approaches (Vol. 450). Singapore: Springer.

Zheng, M., & Klein, J. P. (1995). Estimates of marginal survival for dependent competing risks based on an assumed copula. Biometrika, 82(1), 127-138.

Rivest, L. P., & Wells, M. T. (2001). A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. Journal of Multivariate Analysis, 79(1), 138-155.

Braekers, R., & Veraverbeke, N. (2005). A copula-graphic estimator for the conditional survival function under dependent censoring. Canadian Journal of Statistics, 33(3), 429-447.

Sun, J., Chernick, M. R., & LaBudde, R. A. (2011). A bootstrap test for comparing two variances: simulation of size and power in small samples. Journal of biopharmaceutical statistics, 21(6), 1079-1093.

Klein, J. P., Lee, S. C., & Moeschberger, M. L. (1990). A partially parametric estimator of survival in the presence of randomly censored data. Biometrics, 795-811.

Klein, John P., and Melvin L. Moeschberger. Survival analysis: techniques for censored and truncated data. Vol. 1230. New York: Springer, 2003.

Staplin, N. D., et al. "Dependent censoring in piecewise exponential survival models." Statistical Methods in Medical Research 24.3 (2015): 325-341.

Emura, Takeshi, and Yi-Hau Chen. Analysis of survival data with dependent censoring: copula-based approaches. Vol. 450. Singapore: Springer, 2018.

Hyun-Soo Zhang. A Unified Parametric Approach to the Estimation of Dependence and Marginal Distributions in Bivariate Competing Risks Survival Data. Diss. Yonsei University, 2023.

Choa Yun. Estimation of censored medical cost with dependent censoring using copula method. Diss. Yonsei University, 2023.

Jiang, Hongyu, et al. "Pseudo self-consistent estimation of a copula model with informative censoring." Scandinavian Journal of Statistics 32.1 (2005): 1-20.

Zhang, Weijia, Chun Kai Ling, and Xuanhui Zhang. "Deep Copula-Based Survival Analysis for Dependent Censoring with Identifiability Guarantees." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 38. No. 18. 2024.

Appendix

Table A1. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 30%.

n	Tau	MCG Clayton ($\tau = 0.3$)	MCG Clayton ($\tau = 0.5$)	MCG Clayton ($\tau = 0.7$)	MCG Gumbel ($\tau = 0.3$)	MCG Gumbel ($\tau = 0.5$)	MCG Gumbel ($\tau = 0.7$)	MCG Frank ($\tau = 0.3$)	MCG Frank ($\tau = 0.5$)	MCG Frank ($\tau = 0.7$)
30	0.3	0.072	0.071	0.074	0.069	0.065	0.068	0.071	0.066	0.068
	0.5	0.072	0.067	0.064	0.070	0.065	0.067	0.072	0.063	0.067
	0.7	0.072	0.065	0.065	0.066	0.066	0.066	0.069	0.063	0.063
50	0.3	0.073	0.070	0.072	0.071	0.066	0.073	0.071	0.070	0.074
	0.5	0.075	0.068	0.068	0.073	0.069	0.070	0.074	0.068	0.072
	0.7	0.075	0.072	0.069	0.074	0.073	0.069	0.077	0.073	0.067
100	0.3	0.062	0.062	0.060	0.059	0.056	0.058	0.062	0.057	0.055
	0.5	0.064	0.059	0.064	0.058	0.055	0.059	0.062	0.056	0.060
	0.7	0.063	0.057	0.058	0.060	0.054	0.055	0.061	0.053	0.053

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A1. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 30%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.070	0.072	0.071	0.130	0.130	0.100
	0.5	0.077	0.067	0.067	0.100	0.090	0.090
	0.7	0.071	0.062	0.065	0.070	0.080	0.080
50	0.3	0.069	0.071	0.070	0.110	0.110	0.100
	0.5	0.073	0.069	0.069	0.100	0.110	0.110
	0.7	0.072	0.072	0.072	0.090	0.090	0.090
100	0.3	0.064	0.062	0.060	0.090	0.090	0.090
	0.5	0.063	0.059	0.059	0.100	0.100	0.100
	0.7	0.060	0.056	0.058	0.100	0.100	0.100

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A2. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 70%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.069	0.073	0.059	0.057	0.054	0.050	0.062	0.062	0.051
	0.5	0.079	0.078	0.069	0.070	0.060	0.052	0.074	0.069	0.056
	0.7	0.066	0.066	0.068	0.060	0.054	0.052	0.069	0.057	0.056
50	0.3	0.057	0.062	0.067	0.052	0.054	0.052	0.056	0.055	0.054
	0.5	0.062	0.062	0.067	0.057	0.056	0.058	0.060	0.055	0.056
	0.7	0.060	0.065	0.061	0.051	0.057	0.054	0.057	0.060	0.054
100	0.3	0.050	0.053	0.049	0.045	0.046	0.047	0.050	0.049	0.046
	0.5	0.057	0.061	0.051	0.053	0.058	0.051	0.060	0.061	0.049
	0.7	0.044	0.051	0.049	0.040	0.047	0.049	0.045	0.047	0.050

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A2. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 70%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.068	0.071	0.067	0.070	0.070	0.070
	0.5	0.075	0.080	0.077	0.090	0.090	0.090
	0.7	0.062	0.066	0.063	0.060	0.070	0.070
50	0.3	0.060	0.060	0.060	0.080	0.080	0.080
	0.5	0.052	0.061	0.061	0.090	0.080	0.080
	0.7	0.052	0.064	0.064	0.070	0.080	0.080
100	0.3	0.053	0.053	0.051	0.050	0.040	0.050
	0.5	0.054	0.062	0.061	0.080	0.070	0.080
	0.7	0.044	0.050	0.046	0.090	0.090	0.080

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A3. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 30%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.049	0.051	0.055	0.049	0.051	0.051	0.047	0.049	0.051
	0.5	0.051	0.054	0.057	0.052	0.049	0.054	0.051	0.049	0.054
	0.7	0.056	0.055	0.061	0.052	0.053	0.056	0.053	0.054	0.062
50	0.3	0.055	0.057	0.054	0.063	0.058	0.054	0.058	0.057	0.055
	0.5	0.058	0.052	0.053	0.059	0.052	0.053	0.058	0.050	0.054
	0.7	0.061	0.055	0.057	0.059	0.058	0.054	0.060	0.059	0.054
100	0.3	0.059	0.064	0.061	0.056	0.060	0.059	0.056	0.061	0.061
	0.5	0.062	0.064	0.062	0.061	0.063	0.061	0.062	0.060	0.063
	0.7	0.055	0.060	0.064	0.058	0.058	0.063	0.061	0.060	0.067

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A3. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 30%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.054	0.051	0.051	0.070	0.060	0.070
	0.5	0.055	0.054	0.053	0.060	0.060	0.070
	0.7	0.053	0.061	0.054	0.050	0.040	0.050
50	0.3	0.058	0.056	0.057	0.060	0.080	0.070
	0.5	0.062	0.052	0.052	0.050	0.050	0.050
	0.7	0.063	0.053	0.057	0.050	0.050	0.050
100	0.3	0.057	0.060	0.063	0.080	0.080	0.090
	0.5	0.065	0.064	0.064	0.060	0.070	0.070
	0.7	0.050	0.062	0.059	0.050	0.050	0.040

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A4. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 70%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.057	0.063	0.058	0.058	0.049	0.050	0.061	0.055	0.052
	0.5	0.064	0.062	0.054	0.053	0.052	0.048	0.060	0.054	0.049
	0.7	0.063	0.062	0.054	0.054	0.055	0.047	0.062	0.059	0.050
50	0.3	0.072	0.067	0.061	0.069	0.061	0.052	0.069	0.061	0.051
	0.5	0.069	0.060	0.048	0.066	0.061	0.047	0.070	0.060	0.048
	0.7	0.065	0.061	0.054	0.065	0.054	0.052	0.065	0.055	0.047
100	0.3	0.063	0.061	0.057	0.060	0.060	0.066	0.063	0.057	0.060
	0.5	0.059	0.062	0.055	0.059	0.063	0.059	0.059	0.060	0.054
	0.7	0.065	0.067	0.060	0.064	0.062	0.054	0.067	0.062	0.050

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A4. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 70%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.051	0.064	0.061	0.070	0.070	0.070
	0.5	0.054	0.060	0.060	0.090	0.080	0.080
	0.7	0.064	0.060	0.059	0.060	0.050	0.060
50	0.3	0.072	0.068	0.064	0.070	0.050	0.050
	0.5	0.070	0.061	0.061	0.080	0.070	0.060
	0.7	0.070	0.059	0.059	0.060	0.070	0.070
100	0.3	0.061	0.060	0.061	0.060	0.050	0.040
	0.5	0.056	0.059	0.057	0.050	0.050	0.050
	0.7	0.061	0.063	0.066	0.080	0.090	0.090

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A5. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 30%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.084	0.078	0.077	0.085	0.074	0.075	0.083	0.074	0.075
	0.5	0.073	0.070	0.066	0.071	0.064	0.066	0.071	0.064	0.066
	0.7	0.063	0.066	0.065	0.063	0.061	0.061	0.064	0.060	0.061
50	0.3	0.059	0.064	0.060	0.056	0.060	0.062	0.060	0.065	0.063
	0.5	0.066	0.062	0.062	0.058	0.057	0.057	0.059	0.058	0.058
	0.7	0.056	0.061	0.060	0.056	0.056	0.061	0.057	0.061	0.062
100	0.3	0.049	0.053	0.057	0.047	0.056	0.055	0.047	0.052	0.052
	0.5	0.052	0.057	0.057	0.054	0.057	0.055	0.053	0.056	0.054
	0.7	0.053	0.047	0.053	0.049	0.052	0.052	0.054	0.048	0.052

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A5. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 30%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.079	0.082	0.078	0.090	0.100	0.090
	0.5	0.072	0.071	0.071	0.100	0.100	0.100
	0.7	0.074	0.066	0.065	0.100	0.100	0.100
50	0.3	0.062	0.065	0.063	0.070	0.060	0.060
	0.5	0.067	0.063	0.063	0.050	0.060	0.060
	0.7	0.065	0.060	0.062	0.080	0.080	0.070
100	0.3	0.050	0.051	0.053	0.080	0.090	0.100
	0.5	0.054	0.057	0.057	0.090	0.090	0.100
	0.7	0.055	0.050	0.050	0.060	0.080	0.080

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A6. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 70%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
	0.3	0.080	0.088	0.079	0.073	0.065	0.066	0.079	0.072	0.066
30	0.5	0.079	0.088	0.078	0.066	0.072	0.064	0.074	0.080	0.067
	0.7	0.076	0.080	0.085	0.069	0.066	0.068	0.077	0.071	0.074
	0.3	0.051	0.060	0.061	0.047	0.051	0.054	0.051	0.055	0.053
50	0.5	0.060	0.066	0.067	0.050	0.063	0.060	0.057	0.060	0.061
	0.7	0.048	0.059	0.057	0.040	0.049	0.047	0.049	0.055	0.047
	0.3	0.050	0.054	0.050	0.050	0.048	0.052	0.052	0.048	0.050
100	0.5	0.050	0.058	0.057	0.047	0.051	0.054	0.048	0.055	0.051
	0.7	0.050	0.056	0.050	0.047	0.052	0.050	0.050	0.050	0.051

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A6. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 70%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.058	0.084	0.082	0.120	0.110	0.150
	0.5	0.065	0.084	0.084	0.100	0.120	0.110
	0.7	0.070	0.083	0.078	0.090	0.080	0.090
50	0.3	0.055	0.054	0.058	0.050	0.060	0.060
	0.5	0.051	0.065	0.063	0.090	0.090	0.090
	0.7	0.045	0.060	0.060	0.070	0.080	0.070
100	0.3	0.054	0.050	0.051	0.060	0.050	0.060
	0.5	0.045	0.057	0.057	0.070	0.080	0.080
	0.7	0.052	0.053	0.053	0.060	0.060	0.050

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A7. The results of type I error for independent censoring survival data generated with weibull distribution according to the sample size (n) and censoring rate.

n	Censoring rate	Independent	MCG Clayton ($\tau = 0.3$)	MCG Clayton ($\tau = 0.5$)	MCG Clayton ($\tau = 0.7$)	MCG Clayton Ensemble I	SC Clayton (τ=True)	SC Clayton Ensemble I
30	0.3	0.063	0.061	0.065	0.064	0.061	0.100	0.100
	0.5	0.073	0.077	0.074	0.070	0.077	0.100	0.110
	0.7	0.078	0.075	0.075	0.078	0.075	0.060	0.060
50	0.3	0.063	0.063	0.060	0.064	0.061	0.060	0.060
	0.5	0.066	0.066	0.061	0.058	0.063	0.070	0.060
	0.7	0.063	0.063	0.063	0.062	0.068	0.100	0.100
100	0.3	0.053	0.049	0.044	0.042	0.046	0.050	0.050
	0.5	0.052	0.052	0.046	0.039	0.049	0.060	0.050
	0.7	0.044	0.039	0.048	0.042	0.043	0.060	0.060

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A7. The results of type I error for independent censoring survival data generated with weibull distribution according to the sample size (n) and censoring rate.

n	Censoring rate	MCG Gumbel ($\tau = 0.3$)	MCG Gumbel ($\tau = 0.5$)	MCG Gumbel ($\tau = 0.7$)	MCG Frank ($\tau = 0.3$)	MCG Frank ($\tau = 0.5$)	MCG Frank ($\tau = 0.7$)
30	0.3	0.063	0.061	0.062	0.060	0.062	0.066
	0.5	0.075	0.068	0.069	0.076	0.073	0.072
	0.7	0.066	0.061	0.060	0.072	0.071	0.067
50	0.3	0.061	0.059	0.058	0.060	0.057	0.059
	0.5	0.060	0.053	0.057	0.063	0.057	0.055
	0.7	0.054	0.057	0.056	0.062	0.059	0.058
100	0.3	0.048	0.044	0.041	0.049	0.047	0.040
	0.5	0.044	0.047	0.042	0.049	0.047	0.038
	0.7	0.038	0.038	0.039	0.041	0.041	0.042

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A8. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 30%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.077	0.073	0.071	0.078	0.071	0.070	0.076	0.070	0.071
	0.5	0.074	0.073	0.070	0.073	0.075	0.072	0.072	0.073	0.073
	0.7	0.075	0.072	0.067	0.071	0.070	0.068	0.067	0.067	0.068
50	0.3	0.079	0.080	0.077	0.080	0.077	0.076	0.082	0.072	0.072
	0.5	0.075	0.075	0.072	0.071	0.077	0.074	0.077	0.077	0.074
	0.7	0.080	0.079	0.076	0.078	0.076	0.076	0.075	0.078	0.072
100	0.3	0.058	0.057	0.056	0.057	0.057	0.056	0.060	0.059	0.058
	0.5	0.056	0.057	0.057	0.056	0.055	0.056	0.061	0.058	0.059
	0.7	0.059	0.059	0.055	0.062	0.056	0.054	0.061	0.060	0.054

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A8. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 30%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.076	0.076	0.073	0.073	0.075	0.072
	0.5	0.077	0.072	0.072	0.073	0.071	0.073
	0.7	0.075	0.067	0.072	0.066	0.071	0.069
50	0.3	0.078	0.080	0.082	0.110	0.110	0.110
	0.5	0.073	0.075	0.077	0.100	0.100	0.100
	0.7	0.074	0.076	0.080	0.100	0.090	0.080
100	0.3	0.058	0.057	0.057	0.090	0.090	0.080
	0.5	0.057	0.057	0.057	0.090	0.090	0.090
	0.7	0.057	0.059	0.059	0.090	0.090	0.090

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A9. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.060	0.066	0.066	0.063	0.060	0.063	0.062	0.059	0.062
	0.5	0.074	0.078	0.076	0.073	0.071	0.068	0.067	0.071	0.069
	0.7	0.069	0.066	0.066	0.064	0.062	0.064	0.067	0.062	0.062
50	0.3	0.068	0.065	0.064	0.064	0.064	0.065	0.066	0.061	0.064
	0.5	0.066	0.063	0.071	0.065	0.066	0.075	0.064	0.064	0.071
	0.7	0.074	0.069	0.065	0.070	0.064	0.068	0.065	0.062	0.065
100	0.3	0.058	0.059	0.058	0.054	0.057	0.054	0.058	0.056	0.057
	0.5	0.056	0.057	0.057	0.057	0.059	0.057	0.059	0.056	0.056
	0.7	0.063	0.058	0.060	0.061	0.058	0.058	0.061	0.058	0.056

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A9. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.061	0.069	0.067	0.059	0.064	0.064
	0.5	0.069	0.077	0.077	0.077	0.074	0.077
	0.7	0.073	0.065	0.066	0.063	0.063	0.065
50	0.3	0.068	0.066	0.064	0.050	0.060	0.070
	0.5	0.062	0.063	0.063	0.100	0.100	0.100
	0.7	0.061	0.068	0.067	0.120	0.120	0.130
100	0.3	0.059	0.056	0.058	0.070	0.060	0.070
	0.5	0.062	0.057	0.057	0.070	0.060	0.070
	0.7	0.064	0.061	0.057	0.080	0.070	0.070

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A10. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 70%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.074	0.068	0.063	0.072	0.064	0.058	0.065	0.063	0.058
	0.5	0.072	0.074	0.066	0.072	0.066	0.057	0.064	0.055	0.057
	0.7	0.068	0.070	0.077	0.066	0.063	0.065	0.067	0.064	0.065
50	0.3	0.058	0.060	0.066	0.059	0.063	0.065	0.059	0.064	0.061
	0.5	0.056	0.059	0.062	0.053	0.058	0.057	0.055	0.055	0.057
	0.7	0.054	0.063	0.057	0.054	0.059	0.059	0.054	0.060	0.059
100	0.3	0.054	0.053	0.048	0.058	0.054	0.045	0.055	0.051	0.047
	0.5	0.064	0.058	0.049	0.062	0.056	0.051	0.059	0.058	0.053
	0.7	0.049	0.050	0.053	0.049	0.052	0.052	0.048	0.054	0.050

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A10. The results of type I error for dependent censoring survival data generated with bivariate exponential distribution under the Clayton copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 70%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.070	0.071	0.068	0.074	0.067	0.070
	0.5	0.070	0.073	0.072	0.072	0.071	0.070
	0.7	0.064	0.073	0.069	0.065	0.065	0.066
50	0.3	0.062	0.061	0.063	0.060	0.063	0.062
	0.5	0.056	0.059	0.058	0.080	0.080	0.080
	0.7	0.055	0.062	0.063	0.070	0.090	0.100
100	0.3	0.057	0.056	0.054	0.040	0.060	0.060
	0.5	0.061	0.058	0.058	0.070	0.060	0.070
	0.7	0.050	0.054	0.052	0.080	0.090	0.100

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A11. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 30%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.054	0.050	0.053	0.057	0.052	0.055	0.055	0.052	0.054
	0.5	0.059	0.059	0.059	0.057	0.052	0.054	0.057	0.056	0.053
	0.7	0.051	0.055	0.059	0.049	0.053	0.056	0.050	0.056	0.057
50	0.3	0.059	0.062	0.061	0.059	0.063	0.063	0.060	0.064	0.062
	0.5	0.060	0.064	0.068	0.062	0.065	0.061	0.057	0.067	0.067
	0.7	0.055	0.061	0.065	0.053	0.061	0.064	0.055	0.063	0.066
100	0.3	0.067	0.068	0.067	0.064	0.066	0.068	0.070	0.066	0.070
	0.5	0.068	0.066	0.065	0.068	0.070	0.063	0.067	0.069	0.064
	0.7	0.055	0.065	0.070	0.063	0.065	0.064	0.061	0.067	0.065

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A11. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 30%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.048	0.056	0.050	0.080	0.080	0.090
	0.5	0.062	0.059	0.059	0.060	0.055	0.055
	0.7	0.058	0.061	0.055	0.070	0.050	0.050
50	0.3	0.058	0.062	0.062	0.060	0.060	0.060
	0.5	0.062	0.065	0.065	0.100	0.090	0.090
	0.7	0.060	0.066	0.060	0.060	0.060	0.070
100	0.3	0.059	0.067	0.069	0.080	0.090	0.090
	0.5	0.067	0.067	0.067	0.070	0.060	0.070
	0.7	0.052	0.070	0.065	0.070	0.070	0.070

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A12. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.047	0.051	0.059	0.042	0.047	0.052	0.045	0.053	0.052
	0.5	0.055	0.057	0.052	0.052	0.053	0.050	0.053	0.055	0.050
	0.7	0.057	0.056	0.051	0.050	0.055	0.049	0.051	0.055	0.053
50	0.3	0.073	0.068	0.064	0.071	0.068	0.066	0.074	0.065	0.065
	0.5	0.073	0.077	0.070	0.074	0.075	0.065	0.076	0.078	0.070
	0.7	0.073	0.067	0.066	0.067	0.064	0.067	0.066	0.067	0.064
100	0.3	0.062	0.064	0.058	0.058	0.059	0.067	0.063	0.061	0.058
	0.5	0.062	0.055	0.058	0.058	0.061	0.058	0.061	0.054	0.058
	0.7	0.055	0.056	0.058	0.056	0.056	0.061	0.057	0.056	0.058

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A12. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.052	0.047	0.053	0.050	0.050	0.050
	0.5	0.060	0.057	0.057	0.054	0.054	0.053
	0.7	0.061	0.057	0.055	0.050	0.050	0.040
50	0.3	0.074	0.070	0.068	0.080	0.080	0.070
	0.5	0.074	0.076	0.075	0.090	0.080	0.070
	0.7	0.066	0.068	0.068	0.050	0.070	0.070
100	0.3	0.063	0.063	0.061	0.080	0.060	0.060
	0.5	0.061	0.055	0.055	0.060	0.050	0.050
	0.7	0.055	0.058	0.055	0.040	0.050	0.050

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A13. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 70%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
	0.3	0.060	0.064	0.061	0.054	0.050	0.047	0.059	0.058	0.053
30	0.5	0.050	0.051	0.056	0.043	0.043	0.038	0.048	0.049	0.044
	0.7	0.065	0.066	0.063	0.057	0.056	0.050	0.063	0.058	0.050
	0.3	0.069	0.067	0.062	0.068	0.066	0.058	0.066	0.068	0.059
50	0.5	0.064	0.065	0.055	0.067	0.067	0.056	0.067	0.063	0.052
	0.7	0.066	0.063	0.057	0.068	0.060	0.059	0.066	0.061	0.059
	0.3	0.064	0.064	0.059	0.061	0.055	0.064	0.064	0.061	0.061
100	0.5	0.064	0.059	0.052	0.060	0.061	0.052	0.067	0.059	0.056
	0.7	0.062	0.061	0.056	0.063	0.060	0.060	0.063	0.059	0.057

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A13. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Gumbel copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 70%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.050	0.060	0.062	0.060	0.070	0.070
	0.5	0.045	0.049	0.049	0.053	0.047	0.049
	0.7	0.059	0.064	0.061	0.090	0.070	0.070
50	0.3	0.072	0.064	0.066	0.080	0.080	0.090
	0.5	0.063	0.063	0.064	0.080	0.050	0.060
	0.7	0.066	0.060	0.061	0.040	0.050	0.070
100	0.3	0.067	0.062	0.064	0.080	0.070	0.070
	0.5	0.058	0.059	0.059	0.060	0.060	0.050
	0.7	0.056	0.062	0.059	0.080	0.080	0.080

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A14. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 30%.

n	Tau	MCG Clayton ($\tau = 0.3$)	MCG Clayton ($\tau = 0.5$)	MCG Clayton ($\tau = 0.7$)	MCG Gumbel ($\tau = 0.3$)	MCG Gumbel ($\tau = 0.5$)	MCG Gumbel ($\tau = 0.7$)	MCG Frank ($\tau = 0.3$)	MCG Frank ($\tau = 0.5$)	MCG Frank ($\tau = 0.7$)
30	0.3	0.080	0.079	0.075	0.082	0.075	0.074	0.079	0.075	0.072
	0.5	0.079	0.076	0.075	0.076	0.071	0.076	0.072	0.073	0.075
	0.7	0.068	0.067	0.069	0.062	0.063	0.069	0.063	0.066	0.068
50	0.3	0.060	0.062	0.066	0.061	0.063	0.067	0.065	0.066	0.067
	0.5	0.071	0.072	0.069	0.064	0.065	0.066	0.068	0.069	0.069
	0.7	0.050	0.060	0.062	0.057	0.062	0.065	0.053	0.062	0.061
100	0.3	0.054	0.060	0.062	0.054	0.058	0.056	0.054	0.060	0.057
	0.5	0.060	0.058	0.057	0.059	0.058	0.059	0.059	0.059	0.061
	0.7	0.055	0.051	0.057	0.055	0.052	0.053	0.056	0.054	0.060

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A14. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 30%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.077	0.079	0.077	0.090	0.090	0.090
	0.5	0.073	0.077	0.077	0.090	0.080	0.070
	0.7	0.074	0.068	0.066	0.080	0.090	0.090
50	0.3	0.059	0.060	0.063	0.110	0.080	0.080
	0.5	0.063	0.071	0.071	0.090	0.100	0.100
	0.7	0.059	0.061	0.060	0.090	0.080	0.090
100	0.3	0.050	0.057	0.060	0.080	0.090	0.110
	0.5	0.058	0.059	0.059	0.090	0.090	0.090
	0.7	0.050	0.055	0.053	0.070	0.070	0.070

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A15. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	MCG Clayton (τ = 0.3)	MCG Clayton (τ = 0.5)	MCG Clayton (τ = 0.7)	MCG Gumbel (τ = 0.3)	MCG Gumbel (τ = 0.5)	MCG Gumbel (τ = 0.7)	MCG Frank (τ = 0.3)	MCG Frank (τ = 0.5)	MCG Frank (τ = 0.7)
30	0.3	0.069	0.072	0.066	0.060	0.067	0.063	0.067	0.068	0.064
	0.5	0.068	0.067	0.064	0.060	0.063	0.058	0.065	0.065	0.061
	0.7	0.059	0.063	0.064	0.051	0.056	0.057	0.057	0.059	0.059
50	0.3	0.068	0.063	0.065	0.065	0.061	0.062	0.066	0.061	0.068
	0.5	0.056	0.061	0.053	0.058	0.060	0.061	0.059	0.061	0.054
	0.7	0.068	0.066	0.066	0.067	0.064	0.066	0.068	0.062	0.064
100	0.3	0.055	0.059	0.062	0.059	0.063	0.061	0.058	0.060	0.061
	0.5	0.054	0.056	0.058	0.052	0.055	0.059	0.053	0.053	0.062
	0.7	0.046	0.050	0.055	0.048	0.050	0.049	0.049	0.050	0.055

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A15. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 50%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
30	0.3	0.070	0.070	0.071	0.060	0.060	0.060
	0.5	0.058	0.067	0.067	0.060	0.060	0.060
	0.7	0.065	0.064	0.062	0.080	0.080	0.080
50	0.3	0.068	0.064	0.064	0.070	0.080	0.060
	0.5	0.060	0.061	0.061	0.100	0.090	0.100
	0.7	0.067	0.065	0.067	0.090	0.090	0.090
100	0.3	0.050	0.057	0.061	0.070	0.080	0.090
	0.5	0.053	0.056	0.056	0.100	0.090	0.090
	0.7	0.044	0.055	0.051	0.090	0.100	0.080

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

Table A16. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 70%.

n	Tau	MCG Clayton ($\tau = 0.3$)	MCG Clayton ($\tau = 0.5$)	MCG Clayton ($\tau = 0.7$)	MCG Gumbel ($\tau = 0.3$)	MCG Gumbel ($\tau = 0.5$)	MCG Gumbel ($\tau = 0.7$)	MCG Frank ($\tau = 0.3$)	MCG Frank ($\tau = 0.5$)	MCG Frank ($\tau = 0.7$)
30	0.3	0.068	0.075	0.069	0.061	0.054	0.062	0.068	0.065	0.062
	0.5	0.072	0.080	0.068	0.061	0.069	0.063	0.066	0.073	0.064
	0.7	0.076	0.080	0.076	0.069	0.063	0.060	0.076	0.067	0.066
50	0.3	0.057	0.057	0.054	0.051	0.056	0.058	0.055	0.057	0.050
	0.5	0.055	0.066	0.064	0.055	0.063	0.063	0.059	0.061	0.059
	0.7	0.059	0.058	0.054	0.048	0.051	0.055	0.054	0.056	0.054
100	0.3	0.053	0.054	0.057	0.045	0.057	0.057	0.050	0.056	0.057
	0.5	0.054	0.056	0.056	0.043	0.049	0.056	0.050	0.051	0.050
	0.7	0.045	0.050	0.053	0.045	0.047	0.053	0.048	0.047	0.053

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

(continuous) Table A16. The results of type I error for dependent censoring survival data generated with bivariate weibull distribution under the Frank copula according to the sample size (n) and association parameter Kendall's tau (τ) when censoring rate is 70%.

n	Tau	Independent	MCG Clayton Ensemble I	MCG Clayton Ensemble II	SC Clayton (τ=True)	SC Clayton Ensemble I	SC Clayton Ensemble II
	0.3	0.055	0.074	0.072	0.100	0.120	0.110
30	0.5	0.066	0.076	0.074	0.110	0.110	0.110
	0.7	0.073	0.077	0.073	0.070	0.080	0.090
	0.3	0.052	0.059	0.057	0.070	0.070	0.050
50	0.5	0.052	0.066	0.067	0.090	0.100	0.070
	0.7	0.049	0.056	0.059	0.060	0.070	0.060
	0.3	0.053	0.052	0.051	0.060	0.050	0.050
100	0.5	0.046	0.055	0.054	0.070	0.070	0.070
	0.7	0.046	0.052	0.051	0.070	0.070	0.070

MCG: modified Copula-Graphic estimator; SC: self-consistency estimator; τ: association parameter Kendall's tau;

국문요약

의존적 중도절단이 존재하는 경우

두 그룹의 제한된 평균 생존 시간 비교를 위한 통계적 방법

생존 분석은 임상 시험에서 흔히 사용되는 방법으로, 대부분의 생존 분석 방법은 비례 위험 가정을 전제로 한다. 그러나 이 가정이 위배될 경우, 제한된 평균 생존 시간 방법이 대안이 될 수 있다. 제한된 평균 생존 시간은 특정 시점까지의 평균 생존 시간을 측정하는 직관적이고 해석하기 쉬운 중요한 지표이다. 독립적 중도절단 가정이 위배되면 전통적인 생존 분석 방법은 편향될 수 있으며, 일반적으로 사건과 중도절단 시간은 양의 상관관계를 보인다. 따라서 의존적 중도절단 하에서 생존 시간과 중도절단 시간의 결합 분포를 고려할 필요가 있다.

코퓰라 함수는 생존 시간과 중도절단 시간 간의 의존성을 모델링하는 유연하고 유망한 도구이다. 의존적 중도절단 하에서의 생존 분석을 위해 우리는 코퓰라-그래픽 추정량을 사용하였다. 본 연구에서는 의존적 중도절단 가정 하에서 코퓰라-그래픽 및 자기 일관성 추정량을 확장한 새로운 추정량을 제안하였다. 본 연구의 목적은 의존적 중도절단을 포함한 생존 데이터를 사용하여

제한된 평균 생존 시간을 추정하기 위해 코퓰라 방법을 적용하는 것이다. 우리는 다양한 시나리오를 가정한 시뮬레이션 연구를 통해 제안된 추정량의 편향, 제1종 오류 및 검정력을 평가하였다.

시뮬레이션 결과, 전체적인 시뮬레이션 설정 하에서 제1종 오류가 잘 충족되었고, 검정력을 확인한 결과 제안된 모델이 실제 모델과 유사한 성능을 보이는 것으로 확인되었다. 본 논문의 목적은 의존성 중도절단 자료가 있을 때 두 그룹 간 제한된 평균 생존 시간 차이를 잘 추정하는 방법을 찾는 것이다. 이전 연구에서는 제한된 평균 생존 시간의 맥락에서 코퓰라-그래픽 추정량과 자기 일관성 추정량에 대한 연구가 거의 이루어지지 않았으며, 특히 이들의 통계적 특성에 중점을 둔 연구는 부족했다. 따라서, 본 연구는 의존성 중도절단이 있는 데이터에서 제한된 평균 생존 시간을 추정하는 방법에 기여할 것이며, 향후 관련 연구에도 기여할 것이다.

핵심되는 말: 생존 분석, 제한된 평균 생존 시간, 의존성 중도절단, 코퓰라 모델, 코퓰라-그래픽 추정량, 자기 일관성 추정량.