Journal of Mathematics and Scientific Computing with Applications

E-ISSN : 2580-5754; P-ISSN : 2580-569X
Volume 7, Number 1, 2026
DOI: 10.53806/jmscowa.v7i1.1532
Page: 93–105

A COPULA-BASED SEMIPARAMETRIC BIVARIATE SURVIVAL MODEL WITH YANG–PRENTICE MARGINALS: ESTIMATION AND SIMULATION STUDY

Agnes Agatha Renaningtyas Itu[1], **Jerry Dwi Trijoyo Purnomo**[2], **Purhadi**[3]

[1,2,3]Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Indonesia

Article Info

Article history:
Received : May 29, 2026
Revised : June 15, 2026
Accepted : July 10, 2026

Keywords:
Archimedean Copula;
BFGS algorithm;
Bivariate survival;
Simulation study;
Yang–Prentice model.

ABSTRACT

This study proposed a semiparametric bivariate survival model that combines Yang–Prentice marginals with Archimedean copulas (Clayton, Frank, and Gumbel) to jointly address non-proportional hazards and dependence between paired survival times. Parameters were estimated via a two-stage inference-functions-for-margins approach with BFGS optimization, with performance evaluated through a Monte Carlo simulation study ($R = 500$ replications) across sample sizes of $n = 300$ and 500 and censoring rates of 10% and 40%. Results showed that baseline hazard parameters were estimated with near-zero bias across all scenarios, while regression coefficients exhibited increased variability under higher censoring rates. The Gumbel copula yielded the most stable dependence parameter estimates, the Frank copula produced the smallest bias, and the Clayton copula delivered balanced and consistent performance with well calibrated coverage probabilities. Overall, the proposed framework offers a numerically reliable basis for analyzing correlated, censored survival outcomes. The results imply that practitioners should choose the copula family according to whether lower-tail, symmetric, or upper-tail dependence best matches their data, since this choice directly affects the precision of the estimated dependence structure.

Corresponding Author:
Agnes Agatha Renaningtyas Itu,
Department of Statistics, Faculty of Science and Data Analytics,
Institut Teknologi Sepuluh Nopember,
Email: agnesagt216@gmail.com

1. INTRODUCTION

Survival analysis is a branch of statistics concerned with the time until the occurrence of a particular event, such as death, disease recurrence, or device failure. This analysis incorporates information from both individuals who experience the event and those whose observations are censored [1]. Among the various survival regression models,

the Cox proportional hazards (PH) model introduced by Cox has become the most widely used semiparametric model because it does not require a parametric specification of the baseline hazard function and allows the effects of covariates to be interpreted through hazard ratios [2]. Nevertheless, the Cox model is built upon two major assumptions, namely the independence of observations and proportional hazards over time. In many biomedical applications, both assumptions may be violated simultaneously. The independence assumption is not satisfied when two event times are observed from the same individual [3], [4]. Meanwhile, the proportional hazards assumption is violated when survival curves across treatment groups intersect over time, a phenomenon that often occurs when an intervention has adverse short-term effects but beneficial long-term effects, or vice versa. This violation frequently arises when a treatment or intervention produces opposing effects across different time horizons. For example, a surgical procedure may carry elevated short-term mortality risk due to operative complications, yet confer substantial long-term survival benefit compared to conservative management [5]. In such scenarios, the survival curves of the treatment and control groups cross over time, rendering the hazard ratio non-constant and invalidating the Cox proportional hazards model. Ignoring either of these conditions may lead to biased estimators and misleading statistical inference [6].

In addition, survival outcomes in many biomedical studies are naturally observed in pairs or clusters rather than as independent events, for instance, in ophthalmological research, the time to visual impairment in the left and right eyes of the same patient is inherently correlated, as both eyes share systemic risk factors such as blood pressure, diabetes severity, and genetic predisposition. To accommodate dependence between paired event times, copula-based bivariate survival models have increasingly been preferred over frailty models because copulas separate the marginal distributions from their dependence structure, allow explicit modeling of tail dependence, and provide Kendall's tau directly from the copula parameter [7], [8]. Within the Archimedean family, the Clayton, Frank, and Gumbel copulas are commonly employed because they respectively represent lower-tail, symmetric, and upper-tail dependence. For the marginal component, Yang & Prentice [5] introduced a regression model (YP) that includes the Cox proportional hazards (PH) and proportional odds (PO) models as special cases while allowing separate modeling of short-term and long-term covariate effects, making it well suited to data with the crossing survival curves described above. This model has since been extended in the univariate setting: Demarqui and Mayrink [9] developed a YP model with a piecewise exponential baseline hazard, while Oliveira et al. [10] further extended the framework through a more flexible semiparametric baseline hazard, enabling the capture of complex risk dynamics in survival data.

Several studies have combined copula-based bivariate survival models with different marginal distributions. Peres et al. [11] employed Weibull marginals with copulas for cure-fraction data; Mahara et al. [4] proposed a copula–Cox model for bivariate time-to-event data in cataract studies; whereas Miranda Filho and Demarqui [12] extended the framework by incorporating Yang–Prentice marginals. Nevertheless, studies focusing on the theoretical development of estimation algorithms and the evaluation of estimator performance under finite-sample conditions through systematic simulation studies remain relatively limited. Motivated by this research gap, the main contributions of this study are threefold: (i) deriving the joint likelihood function and a two-stage IFM estimation procedure for the copula-based bivariate Yang–Prentice model with a piecewise exponential baseline hazard; (ii) implementing the BFGS algorithm as a numerically stable optimization strategy for this estimation procedure; and (iii) conducting a systematic Monte Carlo simulation study that evaluates the finite-sample performance of the proposed estimators, in terms of bias and RMSE, across varying sample sizes, censoring rates, and copula families. Accordingly, this study addresses the following research question: how accurately and robustly can the copula-based bivariate Yang–Prentice model be estimated via the two-stage IFM procedure with BFGS optimization, and how does this performance vary across sample sizes, censoring rates, and Archimedean copula families (Clayton, Frank, and Gumbel).

2. RESEARCH METHOD

2.1. Yang-Prentice Marginal Model

Let T denote a non-negative random variable representing the survival time, and let be covariate. The Yang–Prentice (YP) model specifies the marginal survival function as:

$$S(t \mid \boldsymbol{x}) = \left[1 + \frac{\boldsymbol{\phi}^{(S)}}{\boldsymbol{\phi}^{(L)}} R_0(t)\right]^{-\phi^{(L)}} = \left[1 + \frac{\exp\left(\boldsymbol{x}^T \boldsymbol{\beta}^{(S)}\right)}{\exp\left(\boldsymbol{x}^T \boldsymbol{\beta}^{(L)}\right)} R_0(t)\right]^{-\exp\left(\boldsymbol{x}^T \boldsymbol{\beta}^{(L)}\right)}. \tag{1}$$

where:

$$R_0(t) = \frac{F_0(t)}{S_0(t)} = \frac{1 - S_0(t)}{S_0(t)}. \tag{2}$$

is the baseline odds function, $\boldsymbol{\beta}^{(S)}$ and $\boldsymbol{\beta}^{(L)}$ are vectors of short-term and long-term regression coefficients. The corresponding marginal hazard function is

$$h\left(t \mid \boldsymbol{x}\right) = h_0\left(t\right) \frac{\exp\left(\boldsymbol{x}^T\boldsymbol{\beta}^{(S)}\right) \exp\left(\boldsymbol{x}^T\boldsymbol{\beta}^{(L)}\right)}{\exp\left(\boldsymbol{x}^T\boldsymbol{\beta}^{(S)}\right) F_0\left(t\right) + \exp\left(\boldsymbol{x}^T\boldsymbol{\beta}^{(L)}\right) S_0\left(t\right)}. \tag{3}$$

The mathematical properties of this hazard function indicate that the hazard ratio in this model is not constant. As time approaches $t \to 0$, the hazard ratio approaches short-term effect, and as time approaches $t \to \infty$, the hazard ratio approaches long-term effect. A crossing between the hazard and survival functions occurs when the two regression coefficients have opposite signs.

To maximize semiparametric flexibility, the baseline distribution function can be estimated using a Piecewise Exponential (PE) model. The Piecewise Exponential (PE) model uses a time grid partition on the horizontal axis denoted as $E = a_0, a_1, \ldots, a_m$. The time axis is divided into m intervals based on the cut-off points $0 = a_0 < a_1 < \cdots < a_m < \infty$. In this model, the hazard function is assumed to be constant within each adjacent time interval, so that hazard values may differ between intervals but remain constant within the same interval. With this assumption, the PE model provides a flexible approach to the shape of the baseline hazard without requiring the specification of a specific parametric form overall. The hazard function is estimated to be constant over each interval using the following formula

$$h_{0k}(t) = \gamma_k, \quad \text{for } t \in (a_{k-1}, a_k). \tag{4}$$

where $(a_{k-1},\ a_k)\,,\ k = 1, \ldots, m$ is a set of m intervals formed by the time grid E, and $\boldsymbol{\gamma} = [\gamma_1, \cdots, \gamma_m]^T$ is the baseline parameter to be estimated. The cumulative hazard function is obtained by summing the hazard rate per unit time:

$$H_{0k}(t) = \sum_{k=1}^{m} \gamma_k (t_k - a_{k-1}). \tag{5}$$

2.2. Copula-Based Bivariate Survival Model

Copula models are a statistical approach used to model the dependence structure among random variables independently of their marginal distributions. Conceptually, copulas allow researchers to combine multiple marginal distributions into a multivariate distribution with a specific dependence structure. This core idea is introduced through Sklar's Theorem, which states that for any multivariate distribution with a joint cumulative distribution function $S(t_1, t_2)$, it can be expressed in terms of the marginal survival functions $S_1(t_1)$ and $S_2(t_2)$ using a survival copula function as follows

$$S(t_1, t_2) = C_\theta[S_1(t_1), S_2(t_2)] \tag{6}$$

where C_θ is a copula with dependence parameter θ and $S_1(t_1), S_2(t_2)$ is the marginal survival function specified by Equation (1).

This study focuses on the Archimedean copula family, which is defined by the additive generator function ψ and its pseudo-inverse $\psi^{-1}(t)$ as follows

$$C_\theta[S_1(t_1), S_2(t_2)] = \psi[\psi^{-1}(S_1(t_1)) + \psi^{-1}(S_2(t_2))] \tag{7}$$

Some popular models that belong to the Archimedean copula class with a single dependence parameter θ are Clayton, Frank, and Gumble. To measure the degree of dependence between two random variables, Kendall's τ is used. Kendall's τ measures the strength of the relationship based on the concordance and discordance of pairs of observations. The value of τ ranges from $-1 \leq \tau \leq 1$, $\tau = 1$ indicates perfect positive dependence, $\tau = -1$ indicates perfect negative dependence, and $\tau = 0$ indicates no dependence. The formulas for the generator function, the copula function, and the τ value for each copula are shown in Table 1.

Table 1: Generator function and the τ value for each copula

Copula	$\boldsymbol{\psi(t)}$	**Kendall's** τ
Clayton	$(1+\theta t)^{-1/\theta},\ \theta > 0$	$\tau = \dfrac{\theta}{\theta+2}$
Frank	$-\dfrac{1}{\theta}\log\left(1-\left(1-e^{-\theta}\right)e^{-t}\right),\ \theta \neq 0$	$\tau = 1+\dfrac{4}{\theta}\left[D_1(\theta)-1\right]$
Gumbel	$\exp\left(-t^{1/\theta}\right),\ \theta \geq 1$	$\tau = \dfrac{\theta-1}{\theta}$

2.3. Steps for Obtaining a Parameter Estimator

The parameter estimates are obtained using the Two-Stage Estimation method, also known as Inference Functions for Margins (IFM) [13]. This approach breaks down the high-dimensional optimization space into two separate steps, significantly improving numerical stability. The steps are as follows:

a. Define the Yang–Prentice marginal model according to Equation (1).
b. Derive the Copula–Yang Prentice joint survival function according to Equation (6).
c. Formulating the joint likelihood function.
d. Determine the joint log-likelihood function.
e. Estimate the marginal parameters of the Yang–Prentice model.
f. Estimate the copula parameter θ using the log-likelihood based on the estimated marginal results.

2.4. Broyden-Fletcher-Goldfarb-Shanno (BFGS) Algorithm

Broyden-Fletcher-Goldfarb-Shanno (BFGS) is a numerical optimization algorithm that belongs to the class of quasi-Newton methods for maximizing or minimizing differentiable objective functions. This algorithm was first developed independently by Broyden, Fletcher, Goldfarb, and Shanno. BFGS is widely used in the estimation of complex statistical models due to its ability to efficiently converge parameters without requiring explicit computation of the Hessian matrix [14]. The BFGS algorithm works through a series of iterative steps to find the parameter values that maximize the log-likelihood function.

2.5. Simulation Study Design

It should be noted at the outset that the evaluation in this study is based exclusively on simulated data, no real bivariate survival dataset is analyzed here, and this scope should be regarded as a limitation of the study. This simulation study aims to evaluate the performance of the Two-Stage Inference Functions for Margins (IFM) estimation with MLE and BFGS iteration method in a bivariate semiparametric survival model that integrates the Yang-Prentice (YP) model as a marginal distribution with the Archimedean copulas (Clayton, Frank, and Gumbel). The simulation design was constructed to resemble a bivariate survival setting in which the two survival times are correlated through an Archimedean copula and the marginal distributions follow the semiparametric Yang–Prentice model. The performance of the estimator was assessed under different sample sizes, censoring proportions, dependence levels, and copula families in order to examine the robustness and stability of the proposed estimation procedure. The censoring time C was generated independently from a Uniform(0, c) distribution, where the upper bound c was calibrated through pilot simulation to achieve the target censoring rates of approximately 10% and 40%. For each simulation scenario, synthetic bivariate survival data were repeatedly generated, the model parameters were estimated using the proposed BFGS optimization algorithm. the estimation results were summarized using several performance measures such as bias and root mean squared error (RMSE), standard error (SE), and Coverage Probability (CP). The detailed simulation settings are shown in Table 2.

Table 2: Simulation settings

Component	**Description**
Sample size	300 and 500
Replication	500
Marginal parameters	$\beta_1^{(S)} = 0.6,\ \beta_1^{(L)} = -0.6$ $\beta_2^{(S)} = 0.4,\ \beta_2^{(L)} = -0.3$
Time grids	$m = 3$
Baseline rate	$\gamma_{11} = 0.01,\ \gamma_{12} = 0.02,\ \gamma_{13} = 0.04$

Continued on next page

Table 2 continued from previous page

Component	Description
	$\gamma_{21} = 0.01,\ \gamma_{22} = 0.03,\ \gamma_{23} = 0.05$
Dependence parameter	$\theta_{\text{Clayton}} = 2,\ \tau_{\text{Clayton}} = 0.5$
	$\theta_{\text{Frank}} = 5.7,\ \tau_{\text{Frank}} = 0.55$
	$\theta_{\text{Gumbel}} = 2,\ \tau_{\text{Gumbel}} = 0.5$
Covariate	$X_1 \sim \text{Bernoulli}(0.5)$ generated independently for each subject
Censoring	$C \sim \text{Uniform}(0, c)$, where c is chosen to yield approximately 10% and 40% censoring rates
Convergence criterion	$\varepsilon = 10^{-6}$

The estimation results were summarized using several performance measures such as bias and root mean squared error (RMSE). Let $\boldsymbol{\eta} = [\theta, (\beta_1^{(S)}), \beta_2^{(S)}), \beta_1^{(L)}), \beta_2^{(L)}), (\boldsymbol{\gamma}_1)^T, (\boldsymbol{\gamma}_2)^T]^T$, is the vector of all parameters to be estimated, the bias, RMSE, SE, and CP in the estimation of parameters are given by,

$$\widehat{\text{Bias}}(\hat{\boldsymbol{\eta}}) = \frac{1}{R}\sum_{r=1}^{R}\left(\hat{\boldsymbol{\eta}}^{(r)} - \boldsymbol{\eta}\right). \tag{8}$$

$$\widehat{\text{RMSE}}(\hat{\boldsymbol{\eta}}) = \sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(\hat{\boldsymbol{\eta}}^{(r)} - \boldsymbol{\eta}\right)^2}. \tag{9}$$

$$\widehat{\text{SE}}(\hat{\boldsymbol{\eta}}) = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(\hat{\boldsymbol{\eta}}^{(r)} - \bar{\hat{\boldsymbol{\eta}}}\right)^2}. \tag{10}$$

where $\hat{\boldsymbol{\eta}}^{(r)}$ denote the estimated parameter of η and R is the number of replications.

$$\text{CP}(\%) = \frac{1}{R}\sum_{r=1}^{R} I_r \times 100\%, \qquad I_r = \begin{cases} 1, & \text{if } L_r \le \theta \le U_r, \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

where L_r is and U_r represent the lower and upper bounds of the confidence interval for the $r-th$ simulation replication, respectively.

The simulation settings in Table 2 were chosen to reflect realistic applied survival conditions rather than arbitrary values. The sample sizes ($n = 300$ and 500) correspond to typical cohort sizes encountered in clinical and epidemiological bivariate survival studies. The marginal regression coefficients were set with opposing signs for the short- and long-term effects (e.g., $\beta_1^{(S)} = 0.6$ and $\beta_1^{(L)} = -0.6$) specifically to induce a crossing hazard pattern, so that the simulation directly tests the model's ability to recover non-proportional hazard effects. The baseline hazard values (γ) were selected to produce moderate, monotonically increasing piecewise hazards consistent with typical disease-progression patterns. The copula dependence parameters were calibrated so that Kendall's τ was approximately equal (0.5–0.55) across the three copula families, ensuring that differences in estimator performance reflect the copula's functional form rather than differences in the underlying strength of dependence. The 10% and 40% censoring rates were chosen to represent light and heavy censoring regimes commonly observed in real bivariate survival data, allowing the robustness of the estimation procedure to be assessed under realistic data-loss conditions.

3. RESULT AND ANALYSIS

In survival analysis, the true survival time is denoted by t. However, in practical applications, survival data are commonly subject to censoring, such that the true survival time may not be fully observed for all individuals. Consequently, the observed time used in the estimation process is defined as $y = min(T, C)$, where C represents the censoring time. Accordingly, the survival function in the model implementation is expressed as $S(y)$, as the likelihood construction and parameter estimation are based on the observed survival times.

3.1. Parameter Estimation

3.1.1 Define the Yang-Prentice marginal model

Let T_{ij} denote the random survival time variable for the $i-th$ subject $i = 1, 2, \ldots, n$ at the $j-th$ margin $j = 1, 2$. The marginal survival function is defined as follows:

$$S(y_{ij} \mid x_{ij}) = \left[1 + \frac{\exp\left(x_{ij}^{\top}\beta_j^{(S)}\right)}{\exp\left(x_{ij}^{\top}\beta_j^{(L)}\right)} R_{0j}(y_{ij})\right]^{-\exp\left(x_{ij}^{\top}\beta_j^{(L)}\right)} \tag{12}$$

For each subject i at margin $j(j = 1, 2)$, the individual likelihood contribution depends on the censoring indicator δ_{ij}. If $\delta_{ij} = 1$, the contribution is given by the probability density function $f(y_{ij})$, if $\delta_{ij} = 0$, the contribution is given by the survival function $S(y_{ij})$. The marginal likelihood contribution is written as:

$$L_i = f(y_{ij} \mid x_{ij})^{\delta_{ij}} S(y_{ij} \mid x_{ij})^{1-\delta_{ij}} \tag{13}$$

known $f(y) = h(y)S(y)$, and the hazard function defined as

$$h(y_{ij} \mid x_{ij}) = h_{0j}(y_{ij}) \frac{\exp\left(x_{ij}^{\top}\beta_j^{(S)}\right) \exp\left(x_{ij}^{\top}\beta_j^{(L)}\right)}{\exp\left(x_{ij}^{\top}\beta_j^{(S)}\right) F_{0j}(y_{ij}) + \exp\left(x_{ij}^{\top}\beta_j^{(L)}\right) S_{0j}(y_{ij})}. \tag{14}$$

the likelihood expression can be rewritten as

$$L_i = \{h(y_{ij} \mid x_{ij})S(y_{ij} \mid x_{ij})\}^{\delta_{ij}} \{S(y_{ij} \mid x_{ij})\}^{1-\delta_{ij}} = \{h(y_{ij} \mid x_{ij})\}^{\delta_{ij}} S(y_{ij} \mid x_{ij}). \tag{15}$$

where the marginal survival function $S(y_{ij} \mid x_{ij})$ follows the Yang–Prentice model in Equation (12), and the marginal hazard function $h(y_{ij} \mid x_{ij})$ follows Equation (14).

3.1.2 Derive the joint survival function Copula-Yang Prentice

Based on Sklar's theorem, the joint bivariate survival function is expressed through the survival copula function C_θ. Using the characteristics of the Archimedean copula family, the joint survival function can be written as:

$$C_\theta\left(S_1(y_{i1}), S_2(y_{i2})\right) = \psi_\theta\left[\psi_\theta^{-1}\left(S_1(y_{i1})\right) + \psi_\theta^{-1}\left(S_2(y_{i2})\right)\right]. \tag{16}$$

or equivalently,

$$C_\theta\left(S(y_{i1} \mid x_{i1}), S(y_{i2} \mid x_{i2})\right) = \psi_\theta\left\{\sum_{j=1}^{2} \psi_\theta^{-1}\left(S(y_{ij} \mid x_{ij})\right)\right\}. \tag{17}$$

3.1.3 Formulating the joint likelihood function

The likelihood for bivariate data is obtained by differentiating the joint survival function with respect to the time variables y_{ij} associated with observed events. Let $d_i = \delta_{i1} + \delta_{i2}$ denote the total number of events for subject i. Then, the likelihood contribution of subject i is:

$$L_i = \frac{\partial^{d_i}}{\partial y_{i1}^{\delta_{i1}} \partial y_{i2}^{\delta_{i2}}} C_\theta\left(S(y_{i1} \mid x_{i1}), S(y_{i2} \mid x_{i2})\right), \tag{18}$$

Applying the chain rule to the copula function with respect to y_{ij}, yields:

$$L_i = \left[\frac{\partial^{d_i} C_\theta\left(S(y_{i1} \mid x_{i1}), S(y_{i2} \mid x_{i2})\right)}{\partial y_{i1}^{\delta_{i1}} \partial y_{i2}^{\delta_{i2}}}\right] \prod_{j=1}^{2} \left[f_j(y_{ij} \mid x_{ij})\right]^{\delta_{ij}}. \tag{19}$$

Combining all likelihood contributions from the η individuals, the joint likelihood function for all parameters η is obtained as follows:

$$L(\eta) = \prod_{i=1}^{n} \left(\frac{\partial^{d_i}}{\partial \prod_{j=1}^{2} (y_{ij})^{\delta_{ij}}} \psi_\theta \left\{ \sum_{j=1}^{2} \psi_\theta^{-1} \left[S(y_{ij} \mid x_{ij}) \right] \right\} \times \prod_{j=1}^{2} \left\{ h(y_{ij} \mid x_{ij}) S(y_{ij} \mid x_{ij}) \right\}^{\delta_{ij}} \right). \tag{20}$$

where $\boldsymbol{\eta} = [\theta, (\beta_1^{(S)}), (\beta_2^{(S)}), (\beta_1^{(L)}), (\beta_2^{(L)}), (\boldsymbol{\gamma}_1)^T, (\boldsymbol{\gamma}_2)^T]^T$ is the vector of all parameters to be estimated.

3.1.4 Determining the log likelihood function

The log-likelihood function is given by:

$$\ell L(\boldsymbol{\eta}) = \sum_{i=1}^{n} \left[\log \left(\frac{\partial^{d_i}}{\prod_{j=1}^{2} \partial y_{ij}^{\delta_{ij}}} \psi_\theta \left\{ \sum_{j=1}^{2} \psi_\theta^{-1} \left[S(y_{ij} \mid x_{ij}) \right] \right\} \right) + \sum_{j=1}^{2} \delta_{ij} \log h(y_{ij} \mid x_{ij}) + \sum_{j=1}^{2} \delta_{ij} \log S(y_{ij} \mid x_{ij}) \right]. \tag{21}$$

3.1.5 Estimate the marginal parameters of the Yang–Prentice model

Estimate the parameters of the marginal model using the likelihood function in Equation (17). The log-likelihood function is given by:

$$\ell = \sum_{i=1}^{n} \sum_{j=1}^{2} \left[\delta_{ij} \log h(y_{ij} \mid x_{ij}) + \log S(y_{ij} \mid x_{ij}) \right]. \tag{22}$$

Compute the first derivative of the marginal log-likelihood function with respect to each estimated parameter and set it equal to zero to obtain the score equations. Since the resulting score equations do not have closed-form solutions, the marginal parameter estimation process is continued using the BFGS (Broyden–Fletcher–Goldfarb–Shanno) numerical iteration method.

3.1.6 Estimate the copula parameter θ using the log-likelihood obtained by substituting the estimated marginal results into the joint log-likelihood function.

The log-likelihood is given by:

$$\ell(\theta, \hat{\eta}_1, \hat{\eta}_2) = \sum_{i=1}^{n} \log \left[\frac{\partial^{d_i}}{\partial y_{i1}^{\delta_{i1}} \partial y_{i2}^{\delta_{i2}}} \psi_\theta \left\{ \psi_\theta^{-1} \left(\hat{S}_1(y_{i1}) \right) + \psi_\theta^{-1} \left(\hat{S}_2(y_{i2}) \right) \right\} \right] \times \left\{ \hat{h}(y_{i1} \mid x_{i1}) \hat{S}(y_{i1} \mid x_{i1}) \right\}^{\delta_{i1}} \left\{ \hat{h}(y_{i2} \mid x_{i2}) \hat{S}(y_{i2} \mid x_{i2}) \right\}^{\delta_{i2}}. \tag{23}$$

If the copula parameter estimation equations do not have closed-form solutions, the copula parameter estimation process is continued using the BFGS numerical iteration method. The BFGS numerical iteration procedure yields the estimated parameter $\hat{\theta}$.

Let $\boldsymbol{\eta}_1 = [(\beta_1^{(S)}, \beta_1^{(L)}, \gamma_1^{(T)}]^T$ and $\boldsymbol{\eta} = [(\beta_2^{(S)}, \beta_2^{(L)}, \gamma_2^{(T)}]^T$ is the vector of all parameters marginal to be estimated, the BFGS iteration numerical algorithm can be carried out through the following steps.

Step 1. Determine the initial parameter value $\boldsymbol{\eta}_j^0$ as the starting point of the iterative process.

Step 2. Initialize the identity matrix $\boldsymbol{M}_0 = \boldsymbol{I}$ as the initial approximation of the inverse Hessian matrix.

Step 3. Define the convergence criterion as $\varepsilon = 10^{-6}$. The optimization was additionally capped at a maximum of 100 BFGS iterations per replication.

Step 4. Compute the gradient of the ln-likelihood function:

$$\nabla l(\boldsymbol{\eta}_j) = \left[\frac{\partial l}{\partial \beta_j^{(S)}}, \frac{\partial l}{\partial \beta_j^{(L)}}, \frac{\partial l}{\partial \boldsymbol{\gamma}_j} \right]^\top \quad for \;\; j = \; 1, \; 2 \tag{24}$$

Step 5. Determine the search direction at the z-th iteration using:

$$\boldsymbol{b}^z = -\boldsymbol{M}^z \nabla \ell \left(\boldsymbol{\eta}_j^z \right) \tag{25}$$

Step 6. ompute the step length λ_z such that it satisfies the Wolfe conditions:

$$\ell(\boldsymbol{\eta}_j^z + \lambda_z \boldsymbol{b}^z) \leq \ell(\boldsymbol{\eta}_j^z) + \mathrm{c}_1\, \lambda_z \nabla \ell(\eta_j^z)^\top \boldsymbol{b}^z \tag{26}$$

where λ_z is the step length and $c_1 \in (0, 1)$, and

$$\nabla \ell(\boldsymbol{\eta}_j^z + \lambda_z \boldsymbol{b}^z)^\top \boldsymbol{b}^z \geq \mathrm{c}_2\, \nabla \ell(\boldsymbol{\eta}_j^z)^\top \boldsymbol{b}^z \tag{27}$$

where $c_2 \in (0, 1)$ and $c_1 < c_2$. if the Wolfe line search failed to identify an acceptable step at a given iteration, a backtracking step-halving fallback was applied, and a replication was flagged as non-converged if the tolerance was not met within the iteration cap.

Step 7. Update the parameter values using:

$$\boldsymbol{\eta}_j^{z+1} = \boldsymbol{\eta}_j^z + \lambda_z \boldsymbol{b}^z \tag{28}$$

Step 8. Compute the difference between the updated parameter values and the previous parameter values:

$$\boldsymbol{w}^z = \boldsymbol{\eta}_j^{z+1} - \boldsymbol{\eta}_j^z \tag{29}$$

Step 9. Compute the difference between the gradient at the updated parameter values and the gradient at the previous parameter values:

$$\boldsymbol{v}^z = \nabla \ell(\boldsymbol{\eta}_j^{z+1}) - \nabla \ell(\boldsymbol{\eta}_j^z) \tag{30}$$

Step 10. Update the inverse Hessian approximation matrix using the parameter and gradient changes:

$$\boldsymbol{M}^{z+1} = (\boldsymbol{I} - \frac{\boldsymbol{w}^Z (\boldsymbol{v}^Z)^T}{(\boldsymbol{v}^Z)^T \boldsymbol{w}^Z}) \boldsymbol{M_2} (I - \frac{\boldsymbol{v}^Z (\boldsymbol{w}^Z)^T}{(\boldsymbol{v}^Z)^T w^Z}) + \frac{\boldsymbol{w}^Z (w^Z)^T}{(v^Z)^T v^Z} \tag{31}$$

Step 11. Check whether the gradient norm at the current parameter value is sufficiently small:

$$\parallel \nabla \ell(\boldsymbol{\eta}_j^{z+1}) \parallel < \varepsilon \tag{32}$$

Step 12. Check whether the change in parameter values between successive iterations is sufficiently small:

$$\parallel \boldsymbol{\eta}_j^{z+1} - \boldsymbol{\eta}_z \parallel < \varepsilon \tag{33}$$

Step 13. Stop the iteration process when at least one convergence criterion is satisfied. The BFGS numerical iteration procedure yields the estimated parameters $\hat{\boldsymbol{\eta}}_{\boldsymbol{j}} = [\hat{\beta}_j^{(s)^T} \quad \hat{\beta}_j^{(L)^T} \quad \hat{\boldsymbol{\gamma}}_j^T]^T$ for $j = 1, 2$.

The BFGS algorithm for copulas follows the same steps as in Equations 24 until Equations 33, with the matrices and vectors adjusted to scalars

3.2. Simulation Result

The overall results of the simulation study show that the two-stage IFM estimation method using MLE and the BFGS iterative algorithm is capable of providing sufficiently accurate estimates for most model parameters across all tested simulation scenarios. The performance of the estimator was evaluated based on 500 Monte Carlo replications using four primary performance measures namely bias and Root Mean Square Error (RMSE), Standard Error (SE), and Coverage Probability (CP) across combinations of sample sizes, censor levels, and three Archimedean copula families (Clayton, Frank, and Gumbel). The algorithm's convergence rate reached 100% across all scenarios for all three copulas, indicating the numerical stability of the applied BFGS optimization procedure. The baseline hazard parameter (γ) for all three copulas exhibited a very small bias approaching zero and a low RMSE across all scenarios. This indicates that the semiparametric component of the model with a piecewise exponential baseline is able to capture the baseline hazard pattern well [9]. The marginal regression parameters (β) also exhibit relatively small bias across all copulas, although they show increased variability as the censoring proportion rises from 10% to 40%. This is consistent with the theory of likelihood inference for censored survival data, which states that the effective information in the data decreases as censoring increases [6]. Details of the simulation results are presented in Table 3, Table 4, and Table 5.

Table 3: Bias, RMSE, SE, and CP of Parameter Estimates under Clayton Copula

Censor	Parameter	$N = 300$				$N = 500$			
		Bias	RMSE	SE	CP (%)	Bias	RMSE	SE	CP (%)
10%	$\beta_1^{(S)}$	-0.0331	0.2657	0.0118	94.2	-0.0671	0.2204	0.0094	94.2
	$\beta_1^{(L)}$	0.0579	0.1817	0.0077	94.4	0.0779	0.1563	0.0061	91.0
	$\beta_2^{(S)}$	-0.0059	0.2380	0.0107	95.8	-0.0166	0.1859	0.0083	94.2
	$\beta_2^{(L)}$	0.0350	0.2159	0.0095	94.8	0.0319	0.1592	0.0070	93.6
	γ_{11}	0.0002	0.0020	0.0001	94.8	0.0002	0.0015	0.0001	94.4
	γ_{12}	0.0002	0.0025	0.0001	95.4	-0.0001	0.0018	0.0001	95.2
	γ_{13}	0.0016	0.0055	0.0002	95.2	0.0008	0.0037	0.0002	93.4
	γ_{21}	0.0000	0.0019	0.0001	95.2	0.0001	0.0014	0.0001	94.8
	γ_{22}	0.0000	0.0032	0.0001	95.8	-0.0001	0.0025	0.0001	95.6
	γ_{23}	0.0011	0.0073	0.0003	95.2	0.0003	0.0049	0.0002	96.0
	θ	-0.0240	0.2097	0.0093	95.6	-0.0188	0.1662	0.0074	95.2
	Conv. rate	100%				100%			
	Avg. time (s)	1.8146 s				4.5463 s			
40%	$\beta_1^{(S)}$	-0.0035	0.2909	0.0130	94.2	-0.0162	0.2241	0.0100	95.0
	$\beta_1^{(L)}$	0.0337	0.2654	0.0118	94.8	0.0351	0.2206	0.0097	94.8
	$\beta_2^{(S)}$	-0.0027	0.2688	0.0120	95.4	-0.0094	0.2104	0.0094	94.8
	$\beta_2^{(L)}$	0.0439	0.2917	0.0129	94.8	0.0135	0.2150	0.0096	95.4
	γ_{11}	0.0001	0.0020	0.0001	95.4	0.0002	0.0016	0.0001	95.6
	γ_{12}	0.0000	0.0027	0.0001	93.8	0.0001	0.0022	0.0001	94.8
	γ_{13}	0.0012	0.0077	0.0003	95.0	0.0009	0.0059	0.0003	93.8
	γ_{21}	0.0000	0.0019	0.0001	95.4	0.0001	0.0016	0.0001	95.4
	γ_{22}	0.0001	0.0037	0.0002	94.4	0.0004	0.0031	0.0001	95.6
	γ_{23}	0.0011	0.0108	0.0005	96.2	0.0011	0.0076	0.0003	95.8
	θ	-0.0169	0.2130	0.0095	95.6	-0.0151	0.1771	0.0079	95.2
	Conv. rate	100%				100%			
	Avg. time (s)	2.0722 s				4.5204 s			

In the context of estimating the dependence parameter θ), the simulation results reveal clear differences in performance among copula families. The Gumbel copula performed best in reducing variability, as evidenced by achieving the smallest RMSE values (0.0901 to 0.1517) and the narrowest standard errors across all sample sizes. On the other hand, although the Frank copula proved capable of producing the least parameter bias in some scenarios, it exhibits high susceptibility to variability in dependence estimation, with the RMSE spiking as high as 0.6813 in small samples. The Clayton copula serves as a balanced alternative, providing moderate and consistent performance in estimating θ. Furthermore, the dependence parameter θ for all three copulas yields Coverage Probability (CP) values within a narrow range of 94.6% to 95.6%. These results are very close to the nominal 95% level, which convincingly confirms that the confidence intervals generated by this estimation procedure are very well calibrated and valid for statistical inference purposes.

Table 4: Bias, RMSE, SE, and CP of Parameter Estimates under Frank Copula

Censor	Parameter	$N = 300$				$N = 500$			
		Bias	RMSE	SE	CP (%)	Bias	RMSE	SE	CP (%)
10%	$\beta_1^{(S)}$	-0.0675	0.2545	0.0110	93.2	-0.0453	0.2131	0.0093	95.4
	$\beta_1^{(L)}$	0.0749	0.1840	0.0075	93.8	0.0701	0.1570	0.0063	91.6
	$\beta_2^{(S)}$	-0.0090	0.2344	0.0105	93.8	0.0060	0.1845	0.0083	94.8
	$\beta_2^{(L)}$	0.0111	0.1907	0.0085	95.2	0.0119	0.1531	0.0068	95.2
	γ_{11}	0.0004	0.0018	0.0001	93.6	0.0002	0.0015	0.0001	95.0

Continued on next page

Table 4 – Continued from previous page

Censor	Parameter	$N = 300$				$N = 500$			
		Bias	RMSE	SE	CP (%)	Bias	RMSE	SE	CP (%)
40%	γ_{12}	0.0002	0.0024	0.0001	95.0	0.0000	0.0020	0.0001	95.4
	γ_{13}	0.0013	0.0053	0.0002	93.4	0.0011	0.0042	0.0002	93.2
	γ_{21}	0.0001	0.0018	0.0001	93.8	0.0000	0.0015	0.0001	94.8
	γ_{22}	0.0004	0.0034	0.0001	94.0	-0.0001	0.0025	0.0001	95.6
	γ_{23}	0.0017	0.0076	0.0003	94.0	0.0011	0.0055	0.0002	94.4
	θ	-0.0252	0.4888	0.0219	94.8	-0.0456	0.3815	0.0170	94.8
	Conv. rate		100%				100%		
	Avg. time (s)		2.8005 s				4.4708 s		
	$\beta_1^{(S)}$	-0.0042	0.3076	0.0138	95.4	-0.0098	0.2299	0.0103	95.2
	$\beta_1^{(L)}$	0.0586	0.2924	0.0128	94.8	0.0355	0.2095	0.0092	94.4
	$\beta_2^{(S)}$	0.0222	0.2670	0.0119	94.2	-0.0121	0.2059	0.0092	95.2
	$\beta_2^{(L)}$	0.0199	0.2945	0.0132	94.0	0.0288	0.2141	0.0095	95.2
	γ_{11}	0.0001	0.0020	0.0001	94.4	0.0001	0.0016	0.0001	96.4
	γ_{12}	0.0002	0.0028	0.0001	95.4	0.0001	0.0024	0.0001	95.0
	γ_{13}	0.0006	0.0075	0.0003	95.2	0.0008	0.0056	0.0002	94.4
	γ_{21}	-0.0001	0.0019	0.0001	95.0	0.0000	0.0015	0.0001	94.6
	γ_{22}	0.0002	0.0038	0.0002	94.4	0.0001	0.0030	0.0001	95.8
	γ_{23}	0.0016	0.0105	0.0005	94.8	0.0007	0.0083	0.0004	94.0
	θ	-0.0082	0.6813	0.0305	95.0	-0.0340	0.5061	0.0226	95.4
	Conv. rate		100%				100%		
	Avg. time (s)		2.7841 s				4.4932 s		

Based on the simulation studies conducted, the superior performance of the Gumbel copula can be clearly distinguished between purely empirical findings and its theoretical justification. Empirically, the results of the Monte Carlo experiments in this study consistently show that the Gumbel copula produces the most stable and precise estimator of the dependence parameter (θ), characterized by a sharp decrease in RMSE as the sample size increases and strong robustness against an increase in the sensor percentage of up to 40%. These empirical findings are theoretically justified because the Gumbel copula is specifically designed to capture upper-tail dependence, a mathematical structure in which correlations among survival variables are concentrated on events occurring at late observation times (toward the end of the follow-up period). This theoretical characteristic is particularly valuable in the context of medical modelling, such as the progression of chronic diseases, where late-occurring events are often the focal point of analysis but are prone to being truncated by study termination (right-censoring).

The theoretical strength of the Gumbel copula in handling censored data is further supported by the classical literature on dependency modelling. Joe [8] explains in his study that dependency information located in the upper tail of the distribution is relatively better protected from information loss due to right-censoring compared to dependency in the lower tail, which is the focus of the Clayton copula. Furthermore, mathematically, the structure of the Gumbel function yields a more regular likelihood surface geometry. This regularity, according to the optimization theory of Nocedal and Wright [14], greatly facilitates gradient-based algorithms such as BFGS in converging to the global optimum without getting trapped in excessive variability. The stability of the dependency estimator achieved by the Gumbel distribution also reaffirms the asymptotic proof by Shih and Louis [15] regarding the consistency of survival-based copula estimators, and echoes the conclusions of Miranda Filho and Demarqui [12] and Peres et al. [11] that the alignment between the theoretical structure of dependence and the natural censoring properties of the data is a key determinant of the quality of statistical inference.

Table 5: Bias, RMSE, SE, and CP of Parameter Estimates under Gumbel Copula

Censor	Parameter	$N = 300$				$N = 500$			
		Bias	RMSE	SE	CP (%)	Bias	RMSE	SE	CP (%)
10%	$\beta_1^{(S)}$	-0.0627	0.2731	0.0119	94.0	-0.0598	0.2121	0.0091	94.2
	$\beta_1^{(L)}$	0.0600	0.1844	0.0078	93.2	0.0644	0.1495	0.0060	92.2
	$\beta_2^{(S)}$	-0.0206	0.2465	0.0110	95.0	-0.0066	0.1908	0.0085	93.4
	$\beta_2^{(L)}$	0.0284	0.2079	0.0092	94.6	0.0151	0.1480	0.0066	95.4
	γ_{11}	0.0004	0.0019	0.0001	94.6	0.0002	0.0015	0.0001	95.4
	γ_{12}	0.0005	0.0025	0.0001	93.6	0.0002	0.0018	0.0001	96.4
	γ_{13}	0.0018	0.0055	0.0002	93.4	0.0014	0.0040	0.0002	93.8
	γ_{21}	0.0004	0.0020	0.0001	93.8	0.0000	0.0014	0.0001	95.2
	γ_{22}	0.0005	0.0034	0.0001	95.4	0.0002	0.0025	0.0001	95.6
	γ_{23}	0.0010	0.0069	0.0003	94.4	0.0008	0.0051	0.0002	95.0
	θ	-0.0183	0.1204	0.0053	95.0	-0.0234	0.0901	0.0039	94.6
	Conv. rate		100%				100%		
	Avg. time (s)		3.0948 s				4.5353 s		
40%	$\beta_1^{(S)}$	-0.0220	0.2871	0.0128	95.2	0.0044	0.2392	0.0107	95.0
	$\beta_1^{(L)}$	0.0575	0.2817	0.0123	95.4	0.0293	0.2013	0.0089	95.2
	$\beta_2^{(S)}$	0.0093	0.2723	0.0122	95.4	0.0002	0.2165	0.0097	94.8
	$\beta_2^{(L)}$	0.0332	0.3026	0.0135	96.2	0.0297	0.2219	0.0098	94.2
	γ_{11}	0.0002	0.0021	0.0001	94.0	0.0001	0.0017	0.0001	95.4
	γ_{12}	0.0002	0.0028	0.0001	95.0	-0.0001	0.0023	0.0001	95.2
	γ_{13}	0.0014	0.0076	0.0003	95.8	0.0009	0.0056	0.0002	94.2
	γ_{21}	0.0000	0.0019	0.0001	94.6	0.0001	0.0016	0.0001	93.6
	γ_{22}	0.0004	0.0040	0.0002	94.2	0.0001	0.0028	0.0001	96.0
	γ_{23}	0.0014	0.0108	0.0005	93.8	0.0011	0.0083	0.0004	95.2
	θ	-0.0029	0.1517	0.0068	94.8	-0.0139	0.1199	0.0053	95.2
	Conv. rate		100%				100%		
	Avg. time (s)		3.0892 s				3.3054 s		

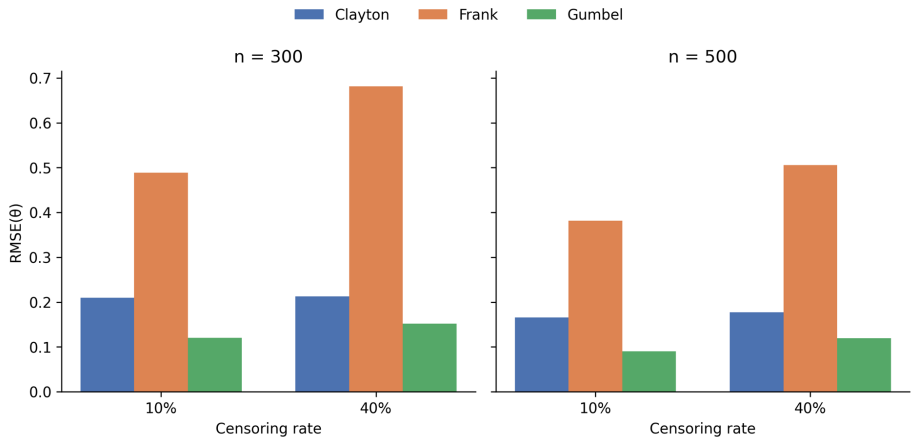

Figure 1. RMSE of the dependence parameter θ across copula families (Clayton, Frank, Gumbel)

As summarized visually in Figure 1, the RMSE of the dependence parameter θ is consistently lowest for the Gumbel copula and highest for the Frank copula across both sample sizes and censoring rates, while the Clayton copula falls in between. RMSE also decreases for all three copulas as the sample size increases from $n = 300$ to $n = 500$. The overall findings of this study conclusively confirm that the choice of copula family is a crucial factor that significantly influences the quality of dependency inference in censored data. The two-stage semiparametric estimation framework, which integrates the Yang-Prentice marginal with BFGS numerical optimization, has been shown to provide a reliable and robust foundation. In particular, the advantages of the Gumbel copula make it the

most recommended choice for modelling applications that prioritize estimation stability for long-term dependencies, such as in the analysis of chronic disease survival data where events are often prone to being censored at the end of the observation period. Ultimately, the flexibility of this proposed framework offers a comprehensive analytical tool for modelling real bivariate survival data, simultaneously capable of accommodating complex risk dynamics resulting from violations of the proportional hazards assumption and precisely mapping dependency structures.

These findings should nonetheless be interpreted in light of several design choices in the simulation study that may affect their generalizability. First, the copula dependence parameters were calibrated so that Kendall's tau was approximately equal (0.5-0.55) across the three copula families. The relative ranking of copulas observed here, particularly the advantage of the Gumbel copula, may not hold under substantially weaker or stronger dependence levels and warrants further investigation. Second, the piecewise-exponential baseline hazard was specified with a fixed time grid ($m = 3$), a coarser or finer grid could alter the accuracy of the baseline hazard estimates and, in turn, the estimated regression and dependence parameters. Third, the covariate was generated as a single Bernoulli(0.5) variable, results may differ under continuous covariates, multiple covariates, or unbalanced covariate distributions. Future simulation work should therefore explore a wider range of dependence strengths, time-grid specifications, and covariate structures to assess the robustness of the copula-ranking conclusions reached here.

4. CONCLUSION

This study proposes a semiparametric bivariate survival model based on the Archimedean copula (Clayton, Frank, and Gumbel) with a Yang–Prentice marginal distribution, estimated using the two-stage IFM method and the BFGS algorithm. Simulation results show that the proposed estimation procedure achieves a 100% convergence rate across all scenarios, confirming the numerical stability of the BFGS algorithm. The baseline hazard parameter (γ) is estimated with high accuracy and bias approaching zero, while the regression parameter (β) exhibits increased variability as the proportion of censored data rises, a pattern consistent with likelihood inference theory for censored data. In terms of dependence parameter estimation, the Gumbel copula performed best with the smallest RMSE, followed by the Clayton copula, while the Frank copula produced the smallest bias but a larger RMSE due to its symmetric dependence structure. The coverage probability of the dependence parameter θ approached the nominal 95% level (94.6%–95.6%) across all scenarios for all three copulas, whereas CP values for the regression and baseline hazard parameters showed wider variation, ranging from approximately 91% to 96.4% depending on the parameter and scenario. It should be noted that these conclusions are based solely on Monte Carlo simulation scenarios, validation on real bivariate survival data remains necessary before the proposed framework can be generalized to applied settings.

In practical terms, these findings provide guidance on the selection of copulas for modeling censored bivariate survival data. The Gumbel copula is recommended when the stability of dependence parameter estimates is a top priority, the Frank copula is suitable for data with symmetric dependence, and the Clayton copula offers a balanced and consistent alternative, with moderate RMSE and well calibrated coverage probabilities across all censoring scenarios. The proposed model has the potential to be applied to real bivariate survival data, such as the analysis of failure times for two organs in the same patient or clinical studies with two correlated endpoints.

REFERENCES

[1] D. Collett, *Modelling Survival Data in Medical Research*. Boca Raton: Chapman & Hall/CRC, 2023.

[2] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 34, no. 2, pp. 187–202, 1972.

[3] J. R. Wilianto, R. Ruhiyat, and H. Sumarno, "Pemodelan waktu hilangnya penglihatan penderita retinopati diabetik dengan beberapa model survival," *Limits: Journal of Mathematics and Its Applications*, vol. 21, no. 2, pp. 139–166, 2024.

[4] D. O. Mahara, S. W. Purnami, and S. Andari, "Copula-based cox regression to modelling bivariate time-to-event data," in *Proceedings of the 2024 16th International Conference on Machine Learning and Computing*, 2024, pp. 658–664.

[5] S. Yang and R. L. Prentice, "Semiparametric analysis of short-term and long-term hazard ratios with two-sample survival data," *Biometrika*, vol. 92, no. 1, pp. 1–17, 2005.

[6] D. G. Kleinbaum and M. Klein, *Survival Analysis: A Self-Learning Text*, 3rd ed. New York: Springer, 2012.

[7] R. B. Nelsen, *An Introduction to Copulas*, 2nd ed. New York: Springer, 2006.

[8] H. Joe, *Dependence Modeling with Copulas*. Boca Raton: Chapman & Hall/CRC, 2014.

[9] F. N. Demarqui and V. D. Mayrink, "Yang and prentice model with piecewise exponential baseline distribution for modeling lifetime data with crossing survival curves," *Brazilian Journal of Probability and Statistics*, vol. 35, no. 1, pp. 172–186, 2021.

[10] C. H. X. Oliveira, F. N. Demarqui, and V. D. Mayrink, "A class of semiparametric yang and prentice frailty models," *arXiv*, 2024, https://doi.org/10.48550/arXiv.2403.07650.

[11] M. V. Oliveira Peres, J. A. Achcar, and E. Z. Martinez, "Bivariate lifetime models in presence of cure fraction: A comparative study with many different copula functions," *Heliyon*, vol. 6, no. 6, p. e03961, 2020.

[12] W. dos Reis Miranda Filho and F. N. Demarqui, "A class of semiparametric models for bivariate survival data," *Lifetime Data Analysis*, vol. 31, no. 1, pp. 102–125, 2025.

[13] H. Joe and J. J. Xu, "The estimation method of inference functions for margins for multivariate models," University of British Columbia, Vancouver, Tech. Rep., 1996.

[14] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York: Springer, 2006.

[15] J. H. Shih and T. A. Louis, "Inferences on the association parameter in copula models for bivariate survival data," *Biometrics*, vol. 51, no. 4, pp. 1384–1399, 1995.