Modelling Competing Risks Data by a Bivariate Model with Singularity Originating from a Shock Model using the Lehmann Family of Distributions

Aakash Agrawal

University of California San Diego, California, USA; Email: aaa015@ucsd.edu

Ayon Ganguly

Indian Institute of Technology Guwahati, Assam, India; Email: aganguly@iitg.ac.in

Debanjan Mitra

Indian Institute of Management Udaipur, Rajasthan, India; Email: debanjan.mitra@iimu.ac.in

Abstract

In this article, a general family of bivariate distributions is used to model competing risks data with dependent causes. The general structure of competing risks data considered here includes ties. A comprehensive inferential framework for the proposed model for competing risks is presented, including maximum likelihood estimation and construction of confidence intervals. Model selection within the bivariate family of distributions for a given dependent competing risks data is discussed. A detailed Monte Carlo simulation study shows that the inferential methods provide quite reasonable results. Analysis of a real data from the Diabetic Retinopathy Study is carried out as an illustrative example.

Keywords: Lehmann family, Dependent competing risks, Singular distribution, Maximum likelihood estimator, Confidence interval, Bootstrap confidence interval, Model selection.

1 Introduction

In lifetime data analysis, competing risks scenarios arise when there are multiple event types for subjects under study, and occurrence of any one of the causes prevents the occurrence of the other relevant causes [12]. The causes, quite often, are dependent. For example, in cancer studies, dependent causes are commonly observed when the death of a subject may be the event of interest; death may occur due to the advancement of the particular type of cancer being studied, or some other cause(s) that are directly related to cancer such as the treatment of cancer or the complications that may arise thereafter. Apart from medical studies, competing risks models are widely studied in various domains including engineering and finance.

The earlier period literature on competing risks models mostly assumed independence among the causes, such as the latent failure time modelling approach or the cause-specific hazard modelling approach. The book by Crowder [5] gives a detailed account of classical models on this topic. However, naturally, a more general approach for modelling competing risks data is the one that does not assume indepdence among causes. The models based on cumulative incidence functions is of particular interest in this regard; see Kalbfleisch and Prentice [11]. Moeschberger and Klein [21] give an excellent review of the different methods used for modelling dependent competing risks data. A very popular approach is the Fine-Gray [9] model that considers a proportional hazards model for hazard derived from the cumulative incidence function.

An appealing approach for analyzing dependent competing risks data is to consider a suitable bivariate or multivariate probability distribution for joint modelling of the lifetimes corresponding to the causes. Copula-based models are commonly used for such purposes, in which lifetime distributions corresponding to the dependent causes are connected by an assumed copula from one of the copula families ([7], [15]). Recently, Michimae and Emura [20] have used the copula-based approach for competing risks data with left truncation and right censoring. Another recent work is by Shih and Emura [23] where the Farlie-Gumbel-Morgenstern (FGM) copula is used along with Burr III marginal distributions for dependent competing risks; see also the references therein. Lawless [14] developed the likelihood func-

tion for directly using a bivariate probability distribution as a model for competing risks data. Several researchers have used this approach very recently. Feizjavadian and Hashemi [8] have used the Marshall-Olkin bivariate Weibull (MOBW) distribution ([16]) to model dependent competing risks data under a progressive hybrid censoring scheme. Samanta and Kundu [22] have discussed Bayesian inference for the same model by using flexible Gamma-Dirichlet prior assumptions. Alqallaf and Kundu [1] have used a bivariate inverse generalized exponential distribution for the same problem. In a different direction, Bai et al. [3] have used different forms of bivariate exponential distributions to model dependent competing risks data in the context of step-stress experiments.

In parametric modelling, it is of utmost importance to select an appropriate model for a given data, as it leads to validity of the subsequent inference. A simple approach to the problem of model selection is to use a general parsimonious family of distributions for modelling purposes, and then to choose the model which is most appropriate within the family for a given data. This approach has been used by many researchers. For example, in the context of cure rate modelling, the Conway-Maxwell Poisson, and the generalized gamma distributions have been used to model the random number of competing causes and lifetimes, respectively; see [4] and the references therein.

In this article, we propose to model dependent competing risks data by using a general family of bivariate distributions. The bivariate family is constructed by using a well-known univariate family of distributions, known as the Lehmann family or the frailty parameter family ([18]), following the approach of Marshall and Olkin [16] assuming a shock model. The univariate Lehmann family contains the Weibull, Gompertz, and Lomax distributions as members of the family, all of which are quite well-known in the context of lifetime data. We call the bivariate family of distributions that we construct by using the univariate Lehmann family as the bivariate Lehmann family of distributions. The primary advantage of using this general family of bivariate distributions is its flexibility derived from its member distributions. Also, a prominent advantage of the proposed model compared to other recent models for competing risks data (for example, the copula-based models) is that it can structurally accommodate presence of ties in the data, where the failure times corresponding to the two causes are identical. For a given competing risks data, it is possible to choose the model

within the family that is most appropriate for the data, instead of fitting just one of the models which may not be the best one for the given data. In fact, the well-known MOBW distribution which was used by Feizjavadian and Hashemi [8] in this regard is a member of this bivariate family.

We develop likelihood inference for the bivariate family of distributions based on dependent competing risks data, and observe that computation of maximum likelihood estimates (MLEs) for the parameters is very convenient in this case. For a given data, it is naturally of interest to select the most appropriate model within the family. We carry out a study of model selection and observe that a simple likelihood-based approach is effective in choosing the suitable model for a given data within the family. This work, therefore, generalizes the works of researchers who have used the MOBW distribution to model dependent competing risks data. These are the main contributions of this paper.

The paper is organized as follows. The details of the construction of the bivariate family of distributions are given in Section 2. Likelihood inference for modelling dependent competing risks data by using the bivariate family is discussed in Section 3. This section also presents the construction of confidence intervals for model parameters by using approaches based on the Fisher information matrix and parametric bootstrap. A likelihood-based approach for model selection is presented in this section as well. In Section 4, the results and discussions of a detailed Monte Carlo simulation study are presented. The Monte Carlo study examines the performance of the MLEs, confidence intervals, and the model selection approach. Analysis of a real dataset is presented in Section 5. Finally, some concluding remarks are made in Section 6.

2 A General Family of Bivariate Distributions

2.1 The Lehmann family of distributions

The survival function of the Lehmann family or the frailty parameter family ([18]) is given by

$$S(t;\alpha,\lambda) = (S_0(t;\lambda))^{\alpha}, \quad t > 0, \tag{1}$$

where $S_0(\cdot;\lambda)$ is the baseline survival function depending on the parameter $\lambda\,(>0)$. Here, the power parameter $\alpha\,(>0)$ is sometimes called the frailty parameter ([18]). We assume that the baseline survival function $S_0(\cdot;\,\lambda)$ is absolutely continuous. The hazard rate function of the Lehmann family is given by

$$h(t;\alpha,\lambda) = \alpha h_0(t;\lambda),$$

where $h_0(t;\lambda) = -\frac{d}{dt}\log S_0(t:\lambda)$ is the hazard rate corresponding to the baseline distribution.

Special members of the Lehmann family are the Weibull, Gompertz, and Lomax distributions which are obtained for different choices of the baseline survival function $S_0(t;\lambda)$ in Eq.(1). In particular, Weibull distribution is obtained when

$$S_0(t;\lambda) = e^{-t^{\lambda}}, \quad t>0.$$

Gompertz distribution is obtained by choosing

$$S_0(t;\lambda) = e^{-(\exp(\lambda t)-1)}, \quad t>0,$$

and for Lomax distribution one chooses

$$S_0(t;\lambda) = \frac{1}{1+\lambda t}, \quad t>0.$$

With well-known lifetime distributions such as the Weibull, Gompertz, and Lomax models as its members, the Lehmann family of distributions is naturally of interest in lifetime data analysis.

2.2 Construction of the bivariate family of distributions

Let U_0, U_1, and U_2 denote three independent random variables with probability distributions specified by the survival functions

$$P(U_i > u_i) = S(u_i; \lambda, \alpha_i) = (S_0(u_i; \lambda))^{\alpha_i}, \quad u_i > 0, \quad i = 0, 1, 2. \tag{2}$$

Define $X = \min\{U_0, U_1\}$ and $Y = \min\{U_0, U_2\}$. Then, the joint survival function of X and Y is given by

$$\begin{aligned} S_{X,Y}(x, y) &= P(\min\{U_0, U_1\} \geq x, \min\{U_0, U_2\} \geq y) \\ &= \begin{cases} S^{(1)}(x, y), & 0 < x < y < \infty \\ S^{(2)}(x, y), & 0 < y < x < \infty \\ S^{(3)}(x), & 0 < x = y < \infty \end{cases} \end{aligned} \tag{3}$$

where

$$\begin{aligned} S^{(1)}(x, y) &= (S_0(y; \lambda))^{\alpha_0+\alpha_2} (S_0(x; \lambda))^{\alpha_1}, \\ S^{(2)}(x, y) &= (S_0(x; \lambda))^{\alpha_0+\alpha_1} (S_0(y; \lambda))^{\alpha_2}, \\ S^{(3)}(x) &= (S_0(x; \lambda))^{\alpha_0+\alpha_1+\alpha_2}. \end{aligned}$$

Note that for the Lehmann family of distributions, λ is the parameter involved in the baseline distribution, whereas α is the power parameter. By using a common λ for the underlying random variables U_0, U_1 and U_2, we essentially assume that the baseline parameter for the lifetimes X and Y corresponding to the two competing causes are the same. The power parameter, which is assumed to be different for different variables, i.e., α_0, α_1 and α_2 for U_0, U_1 and U_2, respectively, brings the difference in their distributions. For similar models discussed in statistical literature, this assumption is not new. For example, Kundu and Gupta [13] made a similar assumption while constructing a class of bivariate generalized exponential distributions. Very recently, Martínez-Flórez et al. [19] used the same assumption for proposing a general class of bivariate proportional hazard distributions.

From the joint survival function of X and Y, their joint probability density function (JPDF) can be worked out. For $0 < x < y < \infty$, the JPDF of X and Y is given by

$$f^{(1)}(x,\, y) = \frac{\partial^2}{\partial x \partial y} S^{(1)}(x,\, y) = \alpha_1 \left(\alpha_0 + \alpha_2\right) \left(S_0\left(y; \lambda\right)\right)^{\alpha_0+\alpha_2-1} \left(S_0\left(x; \lambda\right)\right)^{\alpha_1-1} f\left(x; \lambda\right) f\left(y; \lambda\right),$$

where $f(\cdot;\, \lambda)$ is the probability density function corresponding to the survival function $S_0(\cdot;\, \lambda)$. For $0 < y < x < \infty$, the JPDF of X and Y is

$$f^{(2)}(x,\, y) = \frac{\partial^2}{\partial x \partial y} S^{(2)}(x,\, y) = \alpha_2 \left(\alpha_0 + \alpha_1\right) \left(S_0\left(y; \lambda\right)\right)^{\alpha_2-1} \left(S_0\left(x; \lambda\right)\right)^{\alpha_0+\alpha_1-1} f\left(x; \lambda\right) f\left(y; \lambda\right).$$

Therefore, we can calculate

$$\begin{aligned} P\left(X < Y\right) &= \int_0^\infty \int_0^y \alpha_1 \left(\alpha_0 + \alpha_2\right) S^{\alpha_0+\alpha_2-1}\left(y; \lambda\right) S^{\alpha_1-1}\left(x; \lambda\right) f\left(x; \lambda\right) f\left(y; \lambda\right) dxdy \\ &= \frac{\alpha_1}{\alpha_0 + \alpha_1 + \alpha_2}. \end{aligned}$$

Similarly, we obtain

$$P\left(Y < X\right) = \frac{\alpha_2}{\alpha_0 + \alpha_1 + \alpha_2}.$$

From here, it is straightforward to see that

$$P\left(X = Y\right) = 1 - P\left(X < Y\right) - P\left(Y < X\right) = \frac{\alpha_0}{\alpha_0 + \alpha_1 + \alpha_2}. \tag{4}$$

This implies that the joint distribution of X and Y has a singular component on the straight line $x = y$. The PDF of the singular part can be obtained as follows:

$$f^{(3)}\left(x\right) = P\left(X = Y\right) \left[-\frac{\partial}{\partial x} S^{(3)}(x)\right] = \alpha_0 (S_0\left(x; \lambda\right))^{\alpha_0+\alpha_1+\alpha_2-1} f\left(x; \lambda\right).$$

Summarizing the above, the JPDF of X and Y is given by

$$f_{X,Y}(x,\, y) \;\; = \begin{cases} f^{(1)}(x,\, y) & \text{if } 0 < x < y < \infty \\ f^{(2)}(x,\, y) & \text{if } 0 < y < x < \infty \\ f^{(3)}(x) & \text{if } 0 < x = y < \infty \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

For convenience of reference, we say that (X, Y) has a distribution in bivariate Lehmann family if the joint survival function of X and Y is given by Eq.(3) or, equivalently, the JPDF is given by Eq.(5), and we denote it as

$$(X, Y) \sim BVF(\alpha_0, \alpha_1, \alpha_2, \lambda),$$

for brevity. Clearly, by choosing different baseline survival functions in the distributions of U_is in Eq.(2), we arrive at different bivariate models - bivariate Weibull, bivariate Gompertz, and bivariate Lomax. Plots of the bivariate densities are presented in Figure 1. It may be mentioned here that a similar family of bivariate distributions was considered by Shoaee [24] for theoretical derivations. But Shoaee [24] did not consider applying the family of distributions to model dependent competing risks data.

Here, it is instructive to discuss measures of association such as the Pearson's correlation coefficient and the Kendall's τ between the marginal lifetimes. For this bivariate model, Pearson's correlation coefficient between the marginal lifetimes cannot be explicitly derived due to analytical intractability. Moreover, for the Lomax distribution, the expectation does not exist when the shape parameter is less than or equal to 1. Therefore, for Lomax baselines in this bivariate model, the Pearson's correlation coefficient does not exist for some values of the parameters. However, the Kendall's τ between the marginal lifetimes can be explicitly derived for this bivariate model, as follows.

The Kendall's τ between two random variables X and Y is defined by

$$\tau_{X,Y} = 2P\left((X_1 - X_2)\,(Y_1 - Y_2) > 0\right) - 1 = 4P\left(X_1 > X_2,\, Y_1 > Y_2\right) - 1$$

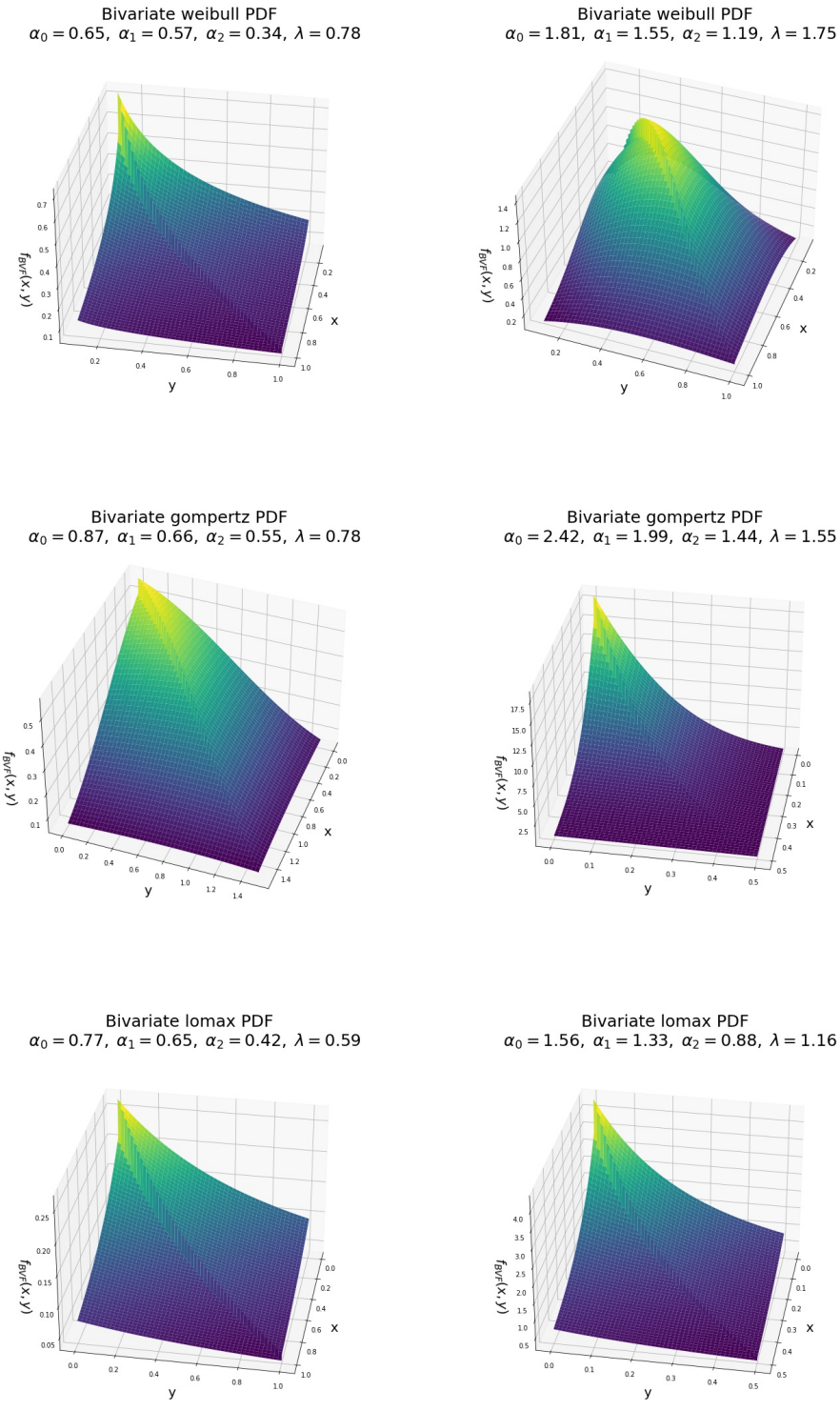

Figure 1: Surface plots of different members of the $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$ distribution where the density is in the y-axis: bivariate Weibull, bivariate Gompertz, and bivariate Lomax (top to bottom).

Therefore, for this bivariate family of distribution, we have

$$\begin{aligned}
P\left(X_1 > X_2,\, Y_1 > Y_2\right) &= \int\int_{\mathbb{R}^2} P\left(X_1 > x,\, Y_1 > y\right) f_{X_2,\, Y_2}(x,\, y) dxdy \\
&= \alpha_1(\alpha_0+\alpha_2) \int\int_{0<x<y<\infty} S_0^{2(\alpha_0+\alpha_2)-1}(y) S_0^{2\alpha_1-1}(x) f_0(x) f_0(y) dxdy \\
&\quad + \alpha_2(\alpha_0+\alpha_1) \int\int_{0<y<x<\infty} S_0^{2(\alpha_0+\alpha_1)-1}(x) S_0^{2\alpha_2-1}(y) f_0(x) f_0(y) dxdy \\
&\quad + \alpha_0 \int_0^\infty S_0^{2(\alpha_0+\alpha_1+\alpha 2)-1}(x) f_0(x) dx \\
&= \frac{\alpha_1}{2} \int_0^\infty S_0^{2(\alpha_0+\alpha_1+\alpha_2)-1}(x) f_0(x) dx + \frac{\alpha_2}{2} \int_0^\infty S_0^{2(\alpha_0+\alpha_1+\alpha_2)-1} f_0(y) dy \\
&\quad + \frac{\alpha_0}{2(\alpha_1+\alpha_2+\alpha_3)} \\
&= \frac{2\alpha_0+\alpha_1+\alpha_2}{4(\alpha_0+\alpha_1+\alpha_2)}.
\end{aligned}$$

Thus, the Kendall's τ between X and Y is

$$\tau_{X,\,Y} = \frac{2\alpha_0+\alpha_1+\alpha_2}{\alpha_0+\alpha_1+\alpha_2} - 1 = \frac{\alpha_0}{\alpha_0+\alpha_1+\alpha_2}.$$

Obviously, the Kendall's τ does not depend upon the baseline distribution $S_0(\cdot)$.

3 Modelling dependent competing risks data

3.1 Dependent competing risks data

Consider two competing causes in a lifetime study involving n subjects. The two causes can influence each other. Suppose X and Y are random variables denoting the lifetimes of a subject under the two causes, and the observed lifetime is $T = \min(X, Y)$. It is quite possible in reality that a subject fails from both the causes simultaneously. For example, in complex medical studies, two causes may get activated simultaneously and be responsible for the event of interest (e.g. death). Therefore, we consider a general set up where we allow ties in the competing risks data.

The form of observed data is

$$Data = \{(t_i, \delta_i),\, i = 1,\, 2,\, \ldots,\, n\}\,,$$

where t_i is the lifetime of the $i-$th unit, and δ_i is the indicator variable that gives information about the failure mode of the subjects, as follows:

$$\delta_i = \begin{cases} 0, & X_i = Y_i \\ 1, & X_i < Y_i \\ 2, & X_i > Y_i \\ 3, & X_i > C, Y_i > C, \end{cases} \tag{6}$$

where C is the right censoring time for a censored lifetime when $\delta_i = 3$. For convenience of exposition, define index sets A_k that contains subjects with indicator variable $\delta = k$, , $k = 0,\, 1,\, 2,\, 3,\, i.e.$,

$$A_k = \{i : \delta_i = k,\, i = 1,\, 2,\, \ldots,\, n\}\,.$$

Let $m_k = |A_k|$, $k = 0, ..., 3$ (i.e., the number of observations of the sets). Of course, $\sum_{k=0}^{3} m_k = n$.

3.2 Maximum likelihood estimation

The likelihood function is constructed considering contributions of subjects according to their failure status. Note that

(a) When $\delta = 1$, the contribution is $-\frac{\partial}{\partial x} S_{X,Y}(x,\, y)\Big|_{x=t,y=t}$,

(b) When $\delta = 2$, the contribution is $\frac{\partial}{\partial y} S_{X,Y}(x,\, y)\Big|_{x=t,y=t}$,

(c) When $\delta = 0$ (i.e., a tie between the causes), the contribution is $f^{(3)}(t)$, and finally

(d) When $\delta = 3$, the contribution is $S^{(3)}\,(t, t)$. Combining these cases, the likelihood function

without the multiplicative constant is given by

$$L(\boldsymbol{\theta}) \propto \prod_{i \in A_1} \left[-\frac{\partial}{\partial x} S^{(1)}(x,y)\Big|_{x=t_i, y=t_i} \right] \prod_{i \in A_2} \left[-\frac{\partial}{\partial y} S^{(2)}(x,y)\Big|_{x=t_i, y=t_i} \right] \times \prod_{i \in A_0} \left[f^{(3)}(t_i) \right] \prod_{i \in A_3} \left[S^{(3)}(t_i) \right], \tag{7}$$

where $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \alpha_2, \lambda)$. Then, the corresponding log-likelihood function without the additive constant is

$$\log L(\boldsymbol{\theta}) = \sum_{k=0}^{2} m_k \log \alpha_k + (\alpha_0 + \alpha_1 + \alpha_2) + \sum_{i=1}^{n} \log S(t_i; \lambda) + \sum_{i \in A_0 \cup A_1 \cup A_2} \log \frac{f(t_i; \lambda)}{S(t_i; \lambda)}. \tag{8}$$

Equating the partial derivative of the log-likelihood function with respect to α_k to zero, we obtain

$$\alpha_k = \frac{-m_k}{\sum_{i=1}^{n} \log S(t_i; \lambda)}, \quad k = 0, 1, 2. \tag{9}$$

Substituting α_k in Eq.(8), the profile log-likelihood in λ is obtained as

$$p(\lambda) = -(m_0 + m_1 + m_2) \log \left(-\sum_{i=1}^{n} \log S(t_i; \lambda) \right) + \sum_{i \in A_0 \cup A_1 \cup A_2} \log \frac{f(t_i; \lambda)}{S(t_i; \lambda)}. \tag{10}$$

The above derivations greatly simplifies the optimization of the log-likelihood function in this case, reducing it to a problem of a one-dimensional optimization. The profile log-likelihood in Eq.(10) can be maximized to get the MLE $\widehat{\lambda}$ of λ, which can then be plugged into Eq.(9) to obtain MLEs $\widehat{\alpha}_0$, $\widehat{\alpha}_1$, $\widehat{\alpha}_2$ of $\alpha_0, \alpha_1, \alpha_2$, respectively. For optimizing Eq.(10), any routine one-dimensional optimizer from a standard statistical software may be used.

3.3 Asymptotic confidence intervals

The asymptotic variance of the MLEs can be estimated by using the observed Fisher information matrix, which is the negative of the hessian of the log-likelihood function. That is, the observed Fisher information matrix $\boldsymbol{I}(\boldsymbol{\theta})$ is

$$\boldsymbol{I}(\boldsymbol{\theta}) = -\nabla^2 (\log L(\boldsymbol{\theta})).$$

Due to asymptotic normality of MLEs, the distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})$ may be approximated by a $N_4(\mathbf{0}, \boldsymbol{I}^{-1}(\widehat{\boldsymbol{\theta}}))$ distribution for large values of sample sizes. Therefore, estimated asymptotic variances of $\widehat{\alpha}_0, \widehat{\alpha}_1, \widehat{\alpha}_2$ and $\widehat{\lambda}$ are given by the diagonal elements of $\boldsymbol{I}^{-1}(\widehat{\boldsymbol{\theta}})$. Using these information, asymptotic 95% confidence intervals (CIs) for the parameters can be easily constructed; for example,

$$\widehat{\alpha}_0 \pm 1.96\sqrt{\widehat{Var(\widehat{\alpha}_0)}},$$

is an asymptotic 95% CI for α_0.

Sometimes, confidence intervals constructed using the observed Fisher information matrix may not perform well in the sense that the coverage probability of such intervals may be lower than the nominal confidence level. This could be due to the fact that the observed Fisher information matrix depends on the observed data, and hence is subjected to variability of the observed data to a large extent. An alternative method of constructing confidence intervals is the bootstrap approach [10]. Although both parametric and nonparametric versions of the bootstrap approach can be used, here we focus on the parametric bootstrap approach for constructing asymptotic confidence intervals. The parametric bootstrap samples are generated using the mechanism involving the three underlying independent variables U_0, U_1, U_2 (as described in Section 2.2) utilizing maximum likelihood estimates of the parameters. To implement parametric bootstrap, we use the following algorithm:

Algorithm:

1. For a given data of size n, obtain MLE $\widehat{\boldsymbol{\theta}}$ of parameter $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4) = (\alpha_0, \alpha_1, \alpha_3, \lambda)$
2. Using $\widehat{\boldsymbol{\theta}}$, generate a data of the same size n from the assumed model
3. Based on the generated data, obtain MLE $\widehat{\boldsymbol{\theta}}^*$
4. Repeat steps 2-3 B times, to get B bootstrap estimates $\widehat{\boldsymbol{\theta}}^*_1$, $\widehat{\boldsymbol{\theta}}^*_2$, ..., $\widehat{\boldsymbol{\theta}}^*_B$
5. To construct bootstrap confidence interval for θ_i, arrange $\widehat{\theta}^*_{i1}$, ..., $\widehat{\theta}^*_{iB}$ (the ith components of $\widehat{\boldsymbol{\theta}}^*_1$, $\widehat{\boldsymbol{\theta}}^*_2$, ..., $\widehat{\boldsymbol{\theta}}^*_B$, respectively) in the ascending order to get $\widehat{\theta}^*_{i(1)} < \widehat{\theta}^*_{i(2)} < ... < \widehat{\theta}^*_{i(B)}$, $i = 1, 2, 3, 4$
6. A $100(1-\alpha)\%$ parametric bootstrap confidence interval of θ_i is given by $\left(\widehat{\theta}^*_{i([B\frac{\alpha}{2}])}, \widehat{\theta}^*_{i([B(1-\frac{\alpha}{2})])}\right)$.

3.4 Model Selection

Marshall et al. [17] used a simple yet powerful approach to investigate whether data can correctly identify their parent distribution in the context of univariate distributions. Following them, we use a model selection procedure which can be conveniently used for dependent competing risks data.

Suppose $\mathcal{M}_1$,...,$\mathcal{M}_l$ are l candidate models for a competing risks data with dependent causes. The candidate models may or may not be nested. Let the vector of parameters associated with these models be $\boldsymbol{\theta}_1$, ..., $\boldsymbol{\theta}_l$, respectively, and that they all are of the same dimension; denote the MLEs of the parameter vectors by $\widehat{\boldsymbol{\theta}}_1$, ..., $\widehat{\boldsymbol{\theta}}_l$, respectively. Then, for the given data, the most suitable model among the candidate models is $\mathcal{M}^*$ if

$$\widehat{L}(\mathcal{M}^*) = Max\{\widehat{L}_{\mathcal{M}_1}(\widehat{\boldsymbol{\theta}}_1), \widehat{L}_{\mathcal{M}_2}(\widehat{\boldsymbol{\theta}}_2), ..., \widehat{L}_{\mathcal{M}_l}(\widehat{\boldsymbol{\theta}}_l)\}, \tag{11}$$

where $\widehat{L}_{\mathcal{M}_j}(\widehat{\boldsymbol{\theta}}_j)$ is the maximized likelihood function evaluated at the MLE for the j-th candidate model, $j = 1, ..., l$, and $\widehat{L}(\mathcal{M}^*)$ is the maximum of all those maximized likelihoods.

When the parameter $\boldsymbol{\theta}_1$, ..., $\boldsymbol{\theta}_l$ are not of the same dimension, instead of using the values of maximized likelihoods for model selection, one may use Akaike's information criterion ([2]) in Eq.(11). The Akaike's information criterion adjusts for the different number of parameters in a model by adding a penalty term to the maximized likelihood.

Within the $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$ family, there are three candidate models: the bivariate Weibull, bivariate Gompertz, and bivariate Lomax distributions. For a given dependent competing risks data, we fit all the candidate models of the $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$ family and select the one with the largest maximized likelihood value as the most suitable model for the data. Note that this procedure is expected to offer a final model that indeed is most suitable for a given dependent competing risks data, as the $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$ family is very rich, containing the bivariate Weibull, Gompertz, and Lomax as special cases.

4 Simulation Study

The goal of the simulation study, performed by using the R software, here is twofold. First, it is to assess the performance of the MLEs and the confidence intervals for the proposed model. Secondly, it is to examine the performance of the model selection approach. We have observed through simulations that all the methods of inference proposed here perform satisfactorily. The detailed results of the numerical experiments are presented in this section.

4.1 Performance of MLEs and confidence intervals

Three sample sizes are used: $n = 100$, 200, and 400. Along with complete data, right censored data with roughly 20% and 40% censoring are also used. Tables 1 - 3 present the results of the numerical experiments for the bivariate Weibull, bivariate Gompertz, and bivariate Lomax models respectively.

Performance of the MLEs are assessed by using relative mean squared error (MSE) and relative bias, as defined below for one of the parameters, say α_0:

$$\text{Relative MSE}(\alpha_0) = \frac{\text{MSE}(\widehat{\alpha}_0)}{(\alpha_0^*)^2},$$

$$\text{Relative Bias}(\alpha_0) = \frac{\text{Bias}(\widehat{\alpha}_0)}{\alpha_0^*}.$$

where the true value of α_0 is α_0^*. For the confidence intervals, the coverage probabilities are estimated by the Monte Carlo probabilities of including the true parameter value.

We observe that the MLEs, for all the models, have quite reasonable bias and MSE. As expected, the bias and MSE reduce with increasing sample size. Also, censoring has adverse effects on bias and MSE of MLEs for all the models: the higher the censoring, the more the MLEs suffer. Relatively, inference for bivariate Gompertz distribution based on dependent competing risks data seems to be the most affected due to censoring. Inference for bivariate Lomax is also affected to some extent. However, for censored data, the models are less affected for larger sample sizes compared to smaller sample sizes. Censoring seems to have the least effect on inference for the bivariate Weibull distribution among the three models.

The confidence intervals also render satisfactory performance with respect to average

length and coverage probability. The coverage probabilities for both types of confidence intervals are quite close to the nominal confidence level of 95%. And it is also noteworthy that even with an increase in sample size, the confidence intervals can retain their coverage probability although their average lengths reduce. The confidence intervals from the parametric bootstrap approach, in a relative sense, has a slightly lower coverage probability than that for the observed Fisher information matrix. For construction of each bootstrap confidence interval, 250 bootstrap samples have been used. In summary, we can conclude that the proposed method to obtain MLEs and confidence intervals for the model parameters of the $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$ family of distributions based on dependent competing risks data performs quite well.

4.2 A study of model selection

The simulation study on model selection for dependent competing risks data is carried out within the bivariate family $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$. Here, we consider three models for comparison, *viz.*, the bivariate Weibull, bivariate Gompertz, and bivariate Lomax distributions. The general approach is the following: we generate dependent competing risks data from a parent distribution belonging to the family $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$, and then fit all the candidate models to the generated data. The set of candidate models includes the parent distribution along with other model(s) from the family. We use both two- and three-model settings in the model selection study. These steps of data generation and model fitting are repeated a large number of times, to calculate the Monte Carlo probability for each model of being selected as the model of choice for a given dependent competing risks data.

There are two motivating factors for the study on model selection: first, the direct goal is to see whether the likelihood-based approach of Marshall et al. [17], originally used for univariate models, is successful in identifying the parent distribution in case of dependent competing risks data as well. In case it is successful, this approach can then be used as an effective tool for model selection in case of parametric modelling of dependent competing risks data. Secondly, as a tangential topic, this study will also indicate the relative richness of the bivariate distributions of the members of the family $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$, in the sense of accommodating dependent competing risks data. Figures 2 and 3 give the results

of the simulation study on model selection with the two-model and three-model settings, respectively.

In a two-model setting, it is clear that the parent distribution will always be selected as the model of choice when the sample size is large. For small to moderate sample sizes, however, there are some interesting observations. Clearly, when the parent is the bivariate Weibull, bivariate Gompertz is a weak choice with its probability of being selected dropping below 20% when the sample size exceeds 100. However, for bivariate Gompertz as the parent, bivariate Weibull has a significant chance of being selected; even for sample size 150, the probability is above 20%. The bivariate Lomax is always a very weak choice when data are generated from either the bivariate Weibull or Gompertz. Similarly, when bivariate Lomax is the parent, selection probabilities for both the bivariate Weibull and Gompertz drop rapidly with increasing sample size.

From the above observations on the two-model comparisons within the family $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$, we can conclude that the likelihood-based model selection approach can identify the parent model correctly, with increasing confidence as the sample size increases. Moreover, we can also conclude that for modelling dependent competing risks data, the bivariate Weibull and the bivariate Gompertz distributions seem to be far apart in nature compared to the bivariate Lomax distribution, although all of them belong to the same family. Finally, we can also perhaps conclude that between bivariate Weibull and Gompertz, the former seems to be stronger in the sense that it fits well to dependent competing risks data generated from the latter, especially for small to moderate sample sizes.

In the three-model setting, the bivariate Lomax has almost no chance of being selected when the parent is either bivariate Weibull or bivariate Gompertz. In contrast, bivariate Weibull and Gompertz give a fair competition to each other: each of them having a significant probability of being selected when the other is the parent. It is also clear that bivariate Weibull is a richer model than the bivariate Gompertz, in the sense that the former's probability of being selected ($\sim$20%) as double the latter's probability of being selected ($\sim$10%), even when the sample is large (300). Moreover, when the parent is the bivariate Lomax, bivariate Weibull remains a strong choice throughout, with its selection probability about 25% even for large sample size. Nonetheless, it is also clear that the likelihood-based approach

eventually picks the parent distribution of the dependent competing risks data correctly as sample size increases.

In summary, it is quite clear that for dependent competing risks data, the likelihood-based approach is successful in identifying the parent distribution correctly, in both two-model and three-model settings. Also, we can conclude that the bivariate Weibull is the richest model in this family $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$ to model this data structure, as we have observed that apart from being the best choice for data generated from itself, it is a good choice for data generated from the other two parents as well. In view of the above discussions, it is clear that likelihood-based model selection approach is quite successful for dependent competing risks data, and can be used for selecting the appropriate model for a given data.

5 Analysis of Diabetic Retinopathy data

The Diabetic Retinopathy Study conducted by the National Eye Institute, United Sates of America was a clinical trial to evaluate laser-based treatment for patients with proliferative diabetic retinopathy. In this study, a total of 1758 patients were enrolled during the period 1972 to 1975. At enrollment, the minimum best corrected visual acuity in each eye was 20/100 for each of the study subjects. For each subject, one of the eyes was given a laser-based treatment, and the other eye was left untreated. Csörgö and Welsh [6] reported the uncensored part of the data for white male patients who received argon laser treatment which was one of three types of laser-based treatment given to the enrolled patients. For the data reported in Csörgö and Welsh [6], failure of an eye, measured in days, was defined as the first time the best corrected visual acuity was below 5/200. For more details regarding the data, refer to Csörgö and Welsh [6].

The diabetic retinopathy data can be looked at as a competing risks data, as analysed by Feijzavidian and Hashemi [8] in the following way. Let X_i and Y_i denote the time to failure for the treated and the untreated eye, respectively for the i-th patient. Define $T_i = Min(X_i, Y_i)$ as the time to blindness. Also, it is easy to define an indicator variable ν_i such

that

$$\nu_i = \begin{cases} 0, & \text{two eyes fail at the same time} \\ 1, & \text{the treated eye fails first} \\ 2, & \text{the untreated eye fails first.} \end{cases}$$

Moreover, the dependent competing risks model based on the family $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$ proposed here is an ideal candidate for this data, as there are ties in the data. Out of total 71 observations, there were 28 cases where the treated eye failed first, 33 cases where the untreated eye failed first, and 10 cases where failure occurred in both eyes at the same time.

The members of the family $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$ were fitted to this data following the proposed approach; the results of model fitting for the three models are given in Table 4. Note that the bivariate Weibull model is the most appropriate model for this data, as it has the largest maximized likelihood (-319.82) among the candidate models. The MLEs of parameters do not exist for the bivariate Lomax model based on this data. Indeed, from the plots of the profile log-likelihood $p(\lambda)$ in the parameter λ given in Figure 4, we observe that the profile log-likelihood in case of bivariate Lomax distribution based on the Diabetic Retinopathy data is a monotonic decreasing function.

It is of course of interest to see how closely the bivariate Weibull model fits to the data, compared to the nonparametric Kaplan-Meier survival curve. The plot, ignoring the cause of failure, is given in Figure 5. The Kaplan-Meier curve estimates the survival function of time to blindness, regardless of the treatment status of eyes. For the parametric curve, we consider the distribution of minimum of X and Y, i.e., time to blindness regardless of the treatment status of eyes. Here, essentially we are looking at the distribution of $Z = min(U_0, U_1, U_2)$. The distribution function of Z can be easily derived as follows:

$$P(Z \leq z) = 1 - \prod_{i=0}^{2} P(U_i > z) = 1 - e^{-(\alpha_0+\alpha_1+\alpha_2)z^{\lambda}}.$$

It is clear that the two survival curves (parametric and nonparametric) are quite close.

6 Conclusion

In this article, a general approach for modelling dependent competing risks data with a general bivariate family of distributions is presented, including the details such the construction of the bivariate family of distributions, its use in modelling dependent competing risks data, likelihood inference, and a model selection approach. Through a detailed Monte Carlo simulation study, it is observed that all the proposed methods of inference in this paper perform quite well. Analysis of a real data is presented as an illustration.

In summary, this work provides a comprehensive inferential framework for modelling dependent competing risks data with ties using a general family of bivariate distributions. The model and inferential framework, due to their general and comprehensive nature, are expected to accommodate dependent competing risks data arising from different spheres of science.

A natural generalization of our work would be to extend the model to accommodate multiple causes of failure. For this, a multivariate Lehmann family needs to be constructed. However, this will not be a trivial extension. Marshall and Olkin [16] outlined construction of a multivariate exponential distribution, and observed that the ensuing analytical treatment was quite complicated due to existence of singular distributions in lower dimensions. More research is necessary to extend the bivariate Lehmann family that is considered here to a multivariate family, and then to use the family to model competing risks data with multiple causes.

Data Availability Statement

The real data used in this paper is publicly available. See Sec 5 for relevant details.

Declaration of Conflict of Interests

The Authors declare that there is no conflict of interest.

Acknowledgments

The authors express their sincere thanks to the anonymous reviewers who made many constructive comments which has significantly improved an earlier version of this paper.

References

[1] Alqallaf, F. A., Kundu, D. (2022) A bivariate inverse generalized exponential distribution and its applications in dependent competing risks model. *Communications in Statistics - Simulation and Computation, 51*, 7019–7036.

[2] Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transactions on Automatic Control, 19*, 716–723.

[3] Bai, X., Shi, Y., Ng, H. K. T. (2022). Statistical inference of Type-I progressively censored step-stress accelerated life test with dependent competing risks. *Communications in Statistics - Theory and Methods, 51*, 3077–3103.

[4] Balakrishnan, N., Pal, S. (2015). An EM algorithm for the estimation of parameters of a flexible cure rate model with generalized gamma lifetime and model discrimination using likelihood- and information-based methods. *Computational Statistics, 30*, 151–189.

[5] Crowder M. (2001). *Classical Competing Risks Model*. Chapman & Hall, New York.

[6] Csörgö, S., Welsh, A. H. (1985). Testing for exponential and Marshall-Olkin distributions. *Technical Report No. 222*, Department of Statistics, Stanford University.

[7] Escarela, G., Carriere, J. F. (2003). Fitting competing risks with an assumed copula. *Statistical Methods in Medical Research, 12*, 333–349.

[8] Feizjavadian, S. H., Hashemi, R. (2015). Analysis of dependent competing risks in the presence of progressive hybrid censoring using Marshall–Olkin bivariate Weibull distribution. *Computational Statistics and Data Analysis, 82*, 19–34.

[9] Fine, J. P., and Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association, 94*, 496–509.

[10] Davison, A. C., and Hinkley, D. V. (1997). *Bootstrap Methods and Their Application*. Cambridge University Press: Cambridge. `https://doi.org/10.1017/CBO9780511802843`

[11] Kalbfleisch, J. D., and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*, (2nd ed.), John Wiley & Sons, Hoboken, New Jersey.

[12] Klein, J. P. (2006). Modelling competing risks in cancer studies. *Statistics in Medicine, 25*, 1014–1034.

[13] Kundu, D., and Gupta, R. (2009). Bivariate generalized exponential distribution, *Journal of Multivariate Analysis, 10*, 581–593.

[14] Lawless, J. F. (1982). *Statistical Models and Methods for Lifetimes Data*. John Wiley & Sons, New York.

[15] Lo, S. M. S., Wilke, R. A. (2010). A copula model for dependent competing risks. *Journal of the Royal Statistical Society, Series C, 59*, 359–376.

[16] Marshall, A. W., Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association, 62*, 30–44.

[17] Marshall, A. W., Meza, J. C., Olkin, I. (2001). Can data recognize its parent distribution? *Journal of Computational and Graphical Statistics, 10*, 555–580.

[18] Marshall, A. W., & Olkin I. (2007). *Life Distributions; Structure of Nonparametric, Semiparametric, and Parametric Families*, Springer, New York.

[19] Martínez-Flórez, G., Barrera-Causil, C., and Lemonte, A. J. (2022). Power Families of Bivariate Proportional Hazard Models, *Mathematics, 10*, 4410.

[20] Michimae, H., Emura, T. (2022). Likelihood inference for copula models based on left-truncated and competing risks data from field studies. *Mathematics, 10*, 2163.

[21] Moeschberger, M. L., & Klein, J. P. (1995). Statistical methods for dependent competing risks. *Lifetime Data Analysis, 1*, 195–204.

[22] Samanta, D., Kundu, D. (2021). Bayesian inference of a dependent competing risk data. *Journal of Statistical Computation and Simulation, 91*, 3069–3086.

[23] Shih, J. H., Emura, T. (2018). Likelihood-based inference for bivariate latent failure time models with competing risks under the generalized FGM copula. *Computational Statistics 33*, 1293–1323.

[24] Shoaee, S. (2020). On a new class of bivariate survival distributions based on the model of dependent lives and its generalization. *Applications and Applied Mathematics: An International Journal, 15*, 801–819.

Table 1: Performance of the MLEs and the CIs of the parameters of bivariate Weibull ($\alpha_0 = 1.34$, $\alpha_1 = 1.17$, $\alpha_2 = 0.86$, $\lambda = 0.91$) distribution based on dependent competing risks data.

Complete Data							
		Point Estimate		95% CI (Asymptotic)		95% CI (Bootstrapping)	
Sample Size	Para	Relative MSE	Relative Bias	Average Length	Coverage Percentage	Average Length	Coverage Percentage
100	α_0	0.031	0.028	0.930	0.954	0.788	0.894
	α_1	0.039	0.041	0.867	0.951	0.744	0.876
	α_2	0.049	0.044	0.730	0.949	0.622	0.892
	λ	0.007	0.018	0.284	0.934	0.241	0.874
200	α_0	0.014	0.012	0.641	0.955	0.545	0.880
	α_1	0.016	0.013	0.594	0.958	0.500	0.906
	α_2	0.023	0.012	0.501	0.940	0.421	0.884
	λ	0.003	0.006	0.198	0.940	0.165	0.896
400	α_0	0.007	0.004	0.449	0.953	0.374	0.894
	α_1	0.008	0.003	0.416	0.945	0.347	0.902
	α_2	0.011	0.010	0.352	0.944	0.294	0.884
	λ	0.001	0.004	0.139	0.950	0.116	0.902
Data with 20% censoring							
		Point Estimate		95% CI (Asymptotic)		95% CI (Bootstrapping)	
Sample Size	Para	Relative MSE	Relative Bias	Average Length	Coverage Percentage	Average Length	Coverage Percentage
100	α_0	0.046	0.031	1.065	0.944	0.912	0.900
	α_1	0.047	0.021	0.973	0.954	0.844	0.898
	α_2	0.060	0.027	0.811	0.947	0.700	0.898
	λ	0.009	0.014	0.329	0.945	0.278	0.878
200	α_0	0.021	0.015	0.742	0.951	0.625	0.896
	α_1	0.023	0.009	0.679	0.929	0.569	0.874
	α_2	0.030	0.018	0.567	0.941	0.474	0.910
	λ	0.004	0.007	0.231	0.948	0.193	0.876
400	α_0	0.009	0.006	0.519	0.956	0.432	0.882
	α_1	0.010	0.007	0.478	0.954	0.397	0.892
	α_2	0.013	0.000	0.395	0.945	0.326	0.918
	λ	0.002	0.003	0.163	0.942	0.135	0.902
Data with 40% censoring							
		Point Estimate		95% CI (Asymptotic)		95% CI (Bootstrapping)	
Sample Size	Para	Relative MSE	Relative Bias	Average Length	Coverage Percentage	Average Length	Coverage Percentage
100	α_0	0.063	0.039	1.342	0.948	1.158	0.898
	α_1	0.077	0.038	1.221	0.942	1.049	0.888
	α_2	0.080	0.019	0.982	0.943	0.849	0.872
	λ	0.011	0.012	0.391	0.956	0.328	0.894
200	α_0	0.033	0.024	0.932	0.951	0.786	0.894
	α_1	0.035	0.017	0.846	0.945	0.714	0.878
	α_2	0.042	0.021	0.692	0.946	0.582	0.898
	λ	0.005	0.006	0.275	0.952	0.230	0.900
400	α_0	0.015	0.008	0.649	0.952	0.545	0.900
	α_1	0.017	0.007	0.591	0.942	0.479	0.902
	α_2	0.021	0.003	0.481	0.945	0.405	0.890
	λ	0.002	0.003	0.194	0.956	0.161	0.898

Table 2: Performance of the MLEs and the CIs of the parameters of bivariate Gompertz ($\alpha_0 = 1.13$, $\alpha_1 = 0.96$, $\alpha_2 = 0.79$, $\lambda = 1.05$) distribution based on dependent competing risks data.

Complete data							
		Point Estimate		95% CI (Asymptotic)		95% CI (Bootstrapping)	
Sample Size	Para	Relative MSE	Relative Bias	Average Length	Coverage Percentage	Average Length	Coverage Percentage
100	α_0	0.227	-0.029	2.800	0.845	2.110	0.854
	α_1	0.212	-0.042	2.352	0.833	1.827	0.854
	α_2	0.230	-0.040	1.959	0.838	1.478	0.862
	λ	0.211	0.182	1.926	0.955	1.563	0.824
200	α_0	0.170	0.023	2.089	0.895	1.686	0.902
	α_1	0.192	0.037	1.814	0.886	1.448	0.906
	α_2	0.187	0.030	1.492	0.880	1.218	0.868
	λ	0.100	0.074	1.333	0.963	1.093	0.884
400	α_0	0.091	0.023	1.416	0.914	1.185	0.884
	α_1	0.092	0.017	1.204	0.914	1.013	0.890
	α_2	0.093	0.022	1.003	0.909	0.836	0.892
	λ	0.047	0.033	0.932	0.952	0.776	0.886
Data with 20% censoring							
		Point Estimate		95% CI (Asymptotic)		95% CI (Bootstrapping)	
Sample Size	Para	Relative MSE	Relative Bias	Average Length	Coverage Percentage	Average Length	Coverage Percentage
100	α_0	2.047	0.323	11.67	0.817	3.618	0.884
	α_1	2.338	0.354	10.32	0.809	3.180	0.882
	α_2	2.095	0.327	8.154	0.813	2.589	0.900
	λ	0.439	0.199	3.119	0.972	2.098	0.876
200	α_0	1.376	0.273	6.945	0.872	3.406	0.918
	α_1	1.533	0.288	6.074	0.873	2.965	0.930
	α_2	1.456	0.266	4.891	0.866	2.383	0.918
	λ	0.233	0.090	2.206	0.969	1.578	0.928
400	α_0	0.901	0.248	4.140	0.899	3.120	0.904
	α_1	0.874	0.239	3.488	0.902	2.619	0.908
	α_2	0.930	0.251	2.921	0.897	2.192	0.886
	λ	0.133	0.007	1.562	0.960	1.173	0.902
Data with 40% censoring							
		Point Estimate		95% CI (Asymptotic)		95% CI (Bootstrapping)	
Sample Size	Para	Relative MSE	Relative Bias	Average Length	Coverage Percentage	Average Length	Coverage Percentage
100	α_0	2.123	0.178	20.30	0.731	2.791	0.814
	α_1	2.139	0.172	17.09	0.730	2.640	0.810
	α_2	2.254	0.189	14.32	0.732	2.152	0.798
	λ	1.430	0.654	5.673	0.964	3.657	0.802
200	α_0	1.924	0.266	14.00	0.821	3.646	0.870
	α_1	2.077	0.260	12.04	0.808	3.168	0.874
	α_2	2.045	0.255	9.845	0.806	2.566	0.868
	λ	0.648	0.315	4.014	0.969	2.596	0.870
400	α_0	1.486	0.253	8.810	0.843	3.600	0.918
	α_1	1.403	0.247	7.367	0.833	3.049	0.918
	α_2	1.391	0.241	6.031	0.839	2.547	0.918
	λ	0.353	0.169	2.844	0.968	1.973	0.922

Table 3: Performance of the MLEs and the CIs of the parameters of bivariate Lomax ($\alpha_0 = 0.85$, $\alpha_1 = 0.57$, $\alpha_2 = 0.74$, $\lambda = 0.69$) distribution based on dependent competing risks data.

Complete data							
		Point Estimate		95% CI (Asymptotic)		95% CI (Bootstrapping)	
Sample Size	Para	Relative MSE	Relative Bias	Average Length	Coverage Percentage	Average Length	Coverage Percentage
100	α_0	0.793	0.222	2.034	0.945	2.238	0.884
	α_1	0.742	0.210	1.383	0.952	1.544	0.872
	α_2	0.743	0.214	1.754	0.945	1.965	0.886
	λ	0.218	0.174	1.234	0.912	1.010	0.894
200	α_0	0.129	0.096	0.990	0.968	0.983	0.860
	α_1	0.132	0.102	0.694	0.957	0.680	0.892
	α_2	0.120	0.094	0.869	0.964	0.871	0.884
	λ	0.093	-0.014	0.843	0.926	0.705	0.862
400	α_0	0.039	0.045	0.622	0.956	0.580	0.904
	α_1	0.048	0.047	0.438	0.955	0.405	0.872
	α_2	0.042	0.044	0.549	0.963	0.515	0.888
	λ	0.052	-0.005	0.594	0.930	0.488	0.870
Data with 20% censoring							
		Point Estimate		95% CI (Asymptotic)		95% CI (Bootstrapping)	
Sample Size	Para	Relative MSE	Relative Bias	Average Length	Coverage Percentage	Average Length	Coverage Percentage
100	α_0	9.741	0.417	5.996	0.894	2.319	0.940
	α_1	10.84	0.434	4.107	0.889	1.601	0.944
	α_2	15.86	0.459	5.287	0.903	1.991	0.930
	λ	0.579	0.208	2.170	0.976	1.829	0.928
200	α_0	0.900	0.186	2.361	0.915	1.752	0.924
	α_1	0.790	0.186	1.603	0.914	1.191	0.920
	α_2	0.862	0.179	2.047	0.914	1.500	0.916
	λ	0.276	0.073	1.442	0.962	1.198	0.920
400	α_0	0.128	0.063	1.166	0.915	1.121	0.896
	α_1	0.139	0.072	0.804	0.939	0.761	0.878
	α_2	0.132	0.068	1.026	0.930	0.979	0.904
	λ	0.127	0.041	0.996	0.965	0.832	0.902
Data with 40% censoring							
		Point Estimate		95% CI (Asymptotic)		95% CI (Bootstrapping)	
Sample Size	Para	Relative MSE	Relative Bias	Average Length	Coverage Percentage	Average Length	Coverage Percentage
100	α_0	9.929	0.387	9.482	0.820	2.513	0.910
	α_1	13.02	0.400	6.287	0.837	1.751	0.912
	α_2	11.17	0.410	8.461	0.829	2.263	0.902
	λ	1.852	0.548	3.445	0.986	2.996	0.912
200	α_0	2.470	0.238	5.212	0.881	2.292	0.938
	α_1	2.673	0.241	3.580	0.877	1.545	0.928
	α_2	2.483	0.233	4.538	0.873	2.005	0.934
	λ	0.599	0.258	2.203	0.983	1.788	0.932
400	α_0	0.309	0.092	2.979	0.901	1.851	0.930
	α_1	0.317	0.091	1.403	0.902	1.247	0.928
	α_2	0.339	0.089	1.825	0.893	1.625	0.926
	λ	0.261	0.119	1.482	0.982	1.172	0.932

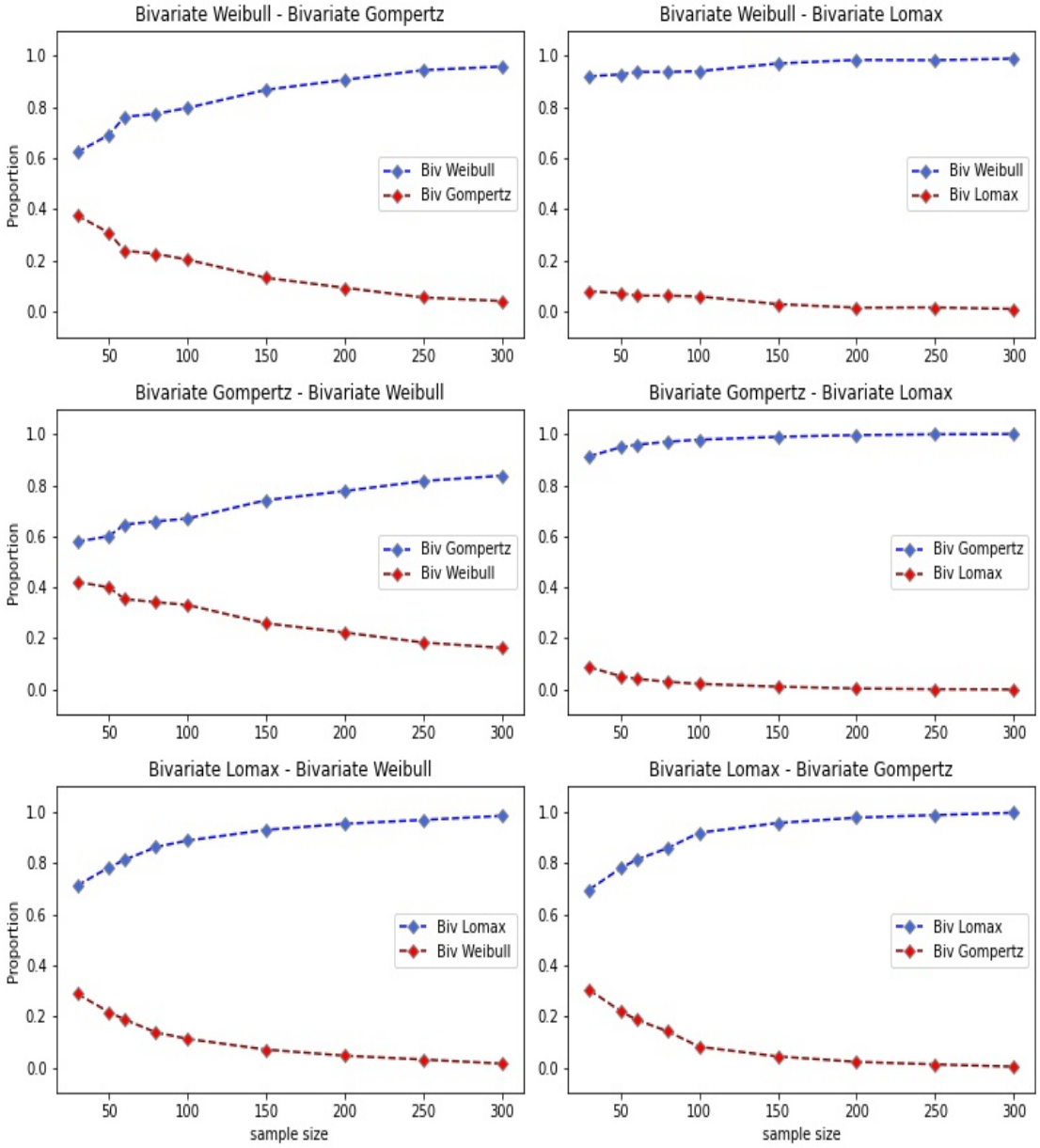

Figure 2: Empirical probability of model selection in the 2-model setting; for each plot, the "AA-BB" caption above the plot indicates AA as the parent distribution, and BB as the model fitted to the data other than the parent.

Table 4: Estimates of parameters for the different members of the family $BVF(\alpha_0, \alpha_1, \alpha_2, \lambda)$ based on the Diabetic Retinopathy Data

Member	Maximized Likelihood	Parameter Estimates $(\widehat{\alpha}_0, \widehat{\alpha}_1, \widehat{\alpha}_2, \widehat{\lambda})$
Bivariate Weibull	-319.82	0.066, 0.185, 0.218, 1.558
Bivariate Gompertz	-323.10	0.140, 0.393, 0.463, 0.412
Bivariate Lomax	NA	NA

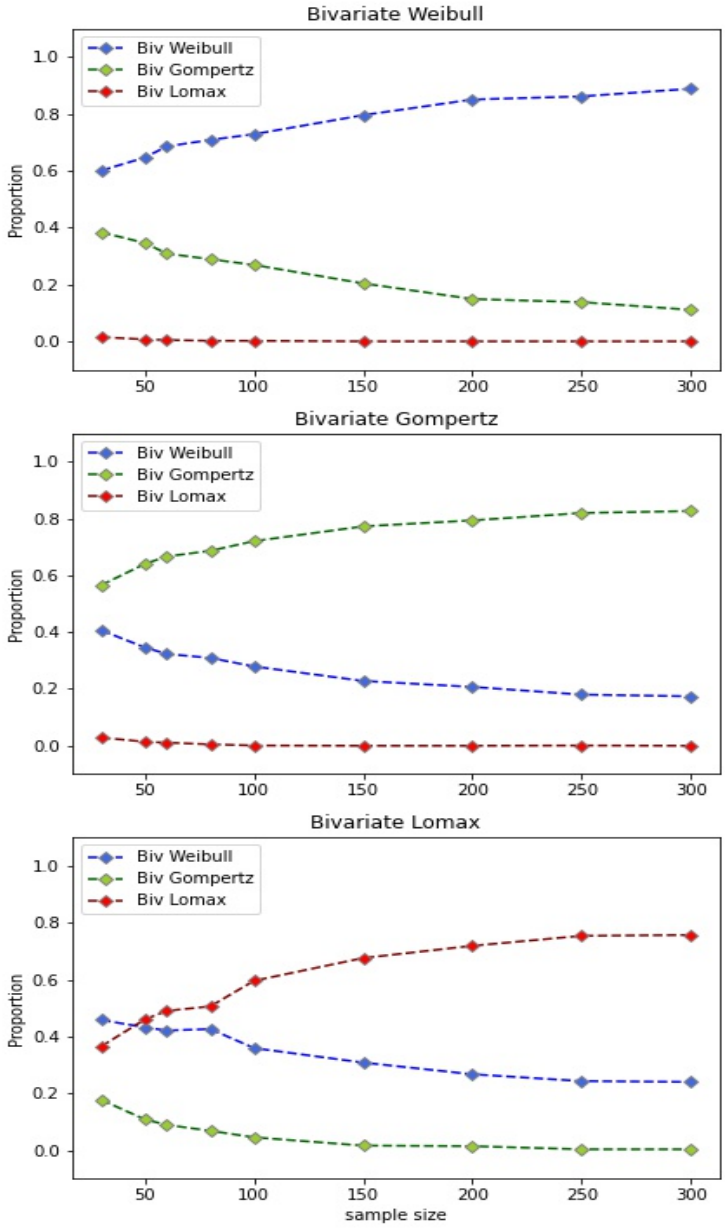

Figure 3: Empirical probability of model selection in the 3-model setting; the parent distributions for generating the data are indicated in captions above each plot.

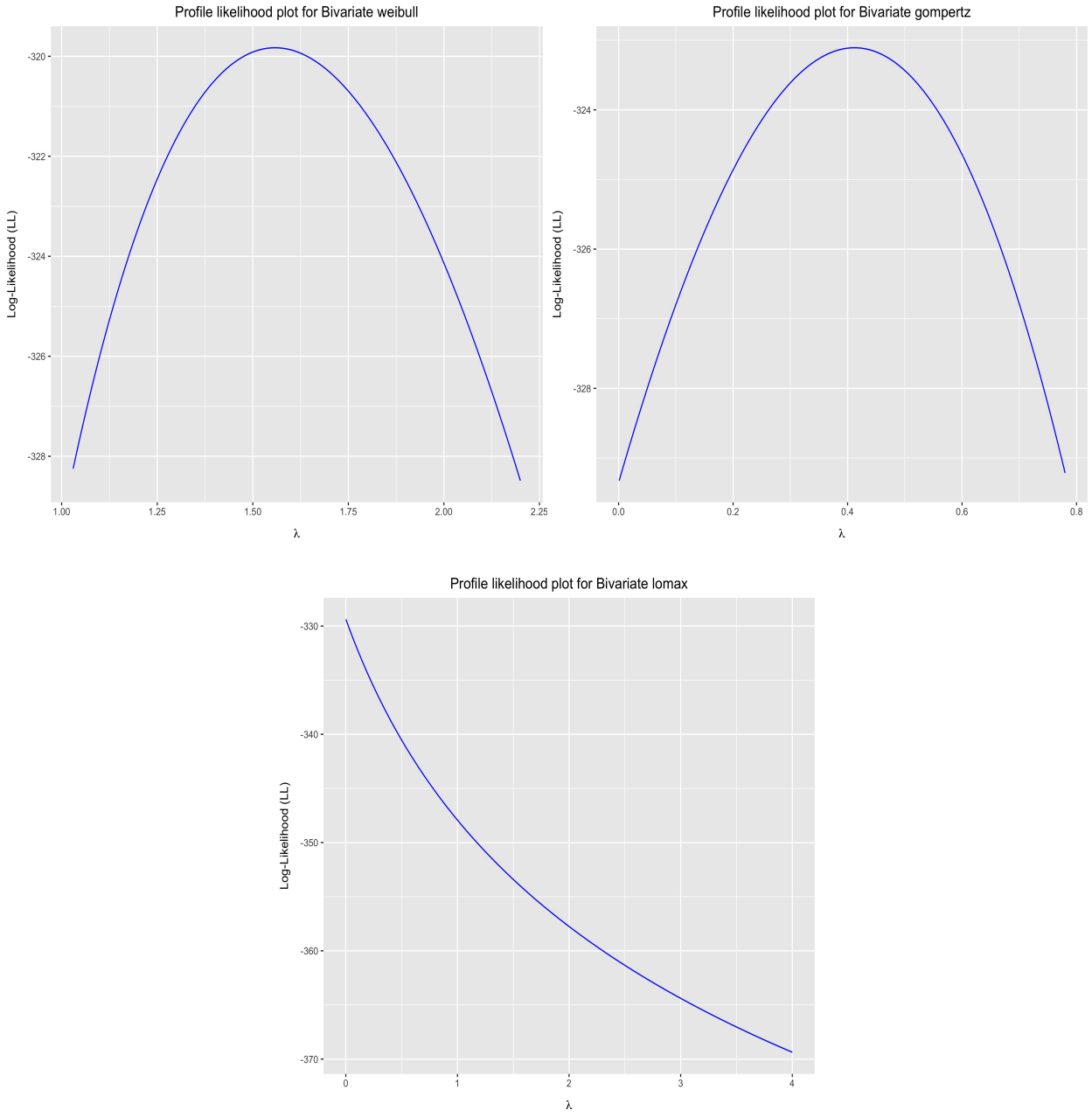

Figure 4: Plot of the profile log-likelihood $p(\lambda)$ for the candidate models based on the Diabetic Retinopathy Data

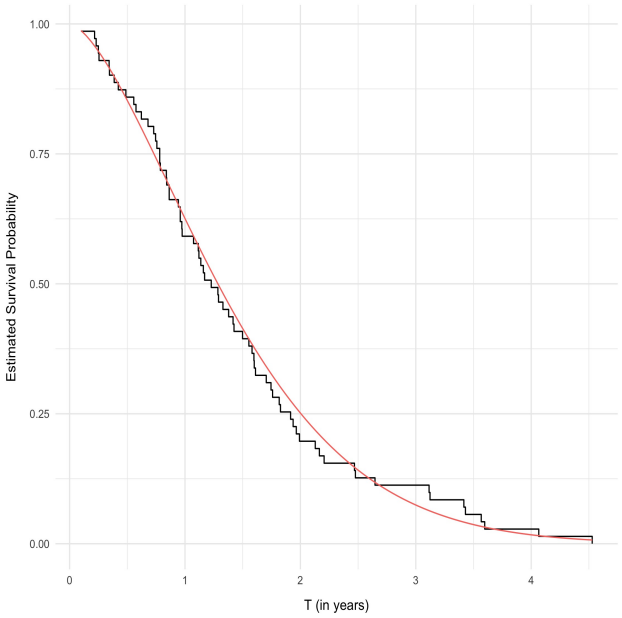

Figure 5: Comparison of the fitted bivariate Weibull survival function (the smooth curve) with the Kaplan-Meier curve (the step curve) for the Diabetic Retinopathy Data