

On estimation and influence diagnostics for a Bivariate Promotion Lifetime Model Based on the FGM Copula: A Fully Bayesian Computation

A.K. SUZUKI*, F. LOUZADA and V.G. CANCHO

Received on November 28, 2012 / Accepted on November 1, 2013

ABSTRACT. In this paper we propose a bivariate long-term model based on the Farlie-Gumbel-Morgens-tern copula to model, where the marginals are assumed to be long-term promotion time structured. The proposed model allows for the presence of censored data and covariates. For inferential purpose a Bayesian approach via Markov Chain Monte Carlo is considered. Further, some discussions on the model selection criteria are given. In order to examine outlying and influential observations, we present a Bayesian case deletion influence diagnostics based on the Kullback-Leibler divergence. The newly developed procedures are illustrated on artificial and real data.

Keywords: Bayesian approach, case deletion influence diagnostics, copula modeling, long-term survival.

1 INTRODUCTION

In survival and reliability analysis we can observe two lifetimes for a same patient or equipment. For instance, in the medical area we can have interest in studying the lifetimes of matched human organs, as kidneys and eyes, and double recurrence of a certain disease. In industrial applications this type of data can occur for systems whose duration times depend on the durability of two components. For instance, the damage of dual generators in a power plant or the lifetime of motors in a twin-engine airplane. The examples set up above are illustrations of bivariate survival data. In literature, an extensive list of papers on modeling multivariate survival data can be found. For instance, we can cite [1, 27, 24] amongst others.

Bivariate survival data in general are correlated and the study of that dependence has been focus of many researches. The most popular approach is the frailty models in that one or more random effects are included in the model in order to model the dependence between the observations (see for instance, [6, 46, 25, 34, 47, 24]). In this case, the marginal times are conditionally independent given the frailty variable. Details on early developments of the such models may be seen in [6]. [40] considered a comparative study between the frailty models and the bivariate models based on the Exponential and Weibull distributions for bivariate survival data according to Bayesian criteria.

*Corresponding author: Adriano Kamimura Suzuki
Department of Applied Mathematics and Statistics, University of São Paulo, Brazil. E-mails: suzuki@icmc.usp.br; louzada@icmc.usp.br; garibay@icmc.usp.br

Recently, the copula models (see for instance, [16, 45, 33, 11, 10, 15, 49, 29]) have become a popular tools to model the dependence of multivariate data, especially in biological areas, actuarial sciences and finances. [35, 22] show some relationships between frailty models and copula models.

A copula is a function that connects the marginal distributions to restore the joint distribution. Different copula functions represent different dependence structures between variables ([33]). By comparing to the joint distribution approach, a copula model is a more convenient tool in studying the dependence structure. Indeed it is more flexible in applications, since, when the scatter of the data does not fit any known family of joint distributions, it may be difficult to specify the joint distribution. Using copulas, however, we can first estimate the marginal distributions and then estimate the copula. Another advantage of the copula modeling is its relatively mathematical simplicity. Also, it is possible to build a variety of dependence structures based on parametric or non-parametric models for the marginal distributions.

In survival analysis, models based on copulas are considered in [26, 34, 41, 23]. Particularly, from a Bayesian perspective, [39] considered an application of the Archimedian copula family for modeling the dependence of bivariate survival. In their analysis, the author considered a Weibull distribution for the marginal distributions of the bivariate lifetime components.

A difficulty arises if a part of the population is not susceptible to the event of interest. For instance, in bivariate clinical studies a population can respond favorably to a treatment, being considered cured. Considering univariate lifetimes, models which assumes part of the population is cured have been widely developed and are usually called long term survival models. Perhaps the most popular type of cure rate model is the mixture model introduced by [3, 2]. In this model, it is assumed that a certain proportion of the patients, say p , are cured, in the sense that they do not present the event of interest during a long period of time and can be seen as to be immune to the cause of failure under study. Later on, the literature on mixture long-term model is extensive and and interested readers can refer to [31, 37, 38], amongst others.

Long-term promotion time model ([48, 7]) assume that the number of causes or risks of a particular event of interest follows a Poisson distribution with mean θ . If the number of these cases follows a Bernoulli distribution with parameter p , we have the mixture model ([3, 2]), in other words, there is a single cause of occurrence. In this paper, we consider a copula model based on the Farlie-Gumbel-Morgenstern (FGM) distribution (see, [12]) assuming promotion time models as marginal distributions.

The motivation for considering the FGM copula is modeling bivariate data with a weak dependence between marginal times. Also, the FGM copula can measure both positive and negative dependence, while the usual copulas, such as the Archimedian copula family, can only measure positive one. Then, the main objective of the paper is to present the FGM long-term bivariate survival copula model assuming promotion time models as marginal distributions, and to develop diagnostic measures from a Bayesian perspective based on the Kullback-Leibler (K-L) divergence to the proposed modeling as proposed by [36]. For each individual lifetime variable

we consider a long-term Weibull promotion time models. For inferential purpose joint estimation was performed and a sampling based approach via Markov Chain Monte Carlo (MCMC) was considered. A simulation study and an application on real data illustrates our approach. We compare the methodology propose with the Positive stable frailty (PSF) model ([25]).

The paper is organized as follows. Section 2 presents a survival bivariate model by considering a FGM copula distribution. Section 3 presents the Bayesian inferential procedure as well as some discussions on the model selection criteria are given. Section 4 a Bayesian case deletion influence diagnostics based on the K-L divergence is presented. Section 5 presents the results of a simulation study and re-analyses the dataset on the efficacy of photocoagulation treatment for proliferative retinopathy on diabetic patients reported by The Diabetic Retinopathy Study ([44]) comparing the methodology propose with the Positive stable frailty (PSF) model ([25]). Final comments in Section 6 conclude the paper.

2 THE BIVARIATE SURVIVAL MODEL BASED ON FGM COPULA

Long-term survivors models have been widely used for fitting data where some individuals may never suffer the cause of failure under study. In this type of modeling, it is assumed that, due to some unobserved prognostic factors, a certain fraction π of the population is immune to the cause of failure under study or a long-term survivor. The survivor function for the entire population can be written as

$$S_{pop}(t) = \exp\{-\theta[1 - S_0(t)]\}. \quad (2.1)$$

The long-term survivors cannot be identified but we can infer their presence in a data set if many of the largest times are censored. Common choices for $S_0(t)$ are the Gompertz, Exponential and Weibull distributions. For this model, the cure rate $\pi = \exp\{-\theta\}$. When $S_0(t)$ has an Exponential and Weibull distributions, we have the so called exponential promotion time and Weibull promotion time models, respectively.

The model (2.1) can be rewritten as follows (see, [7]),

$$S_{pop}(t) = \exp\{-\theta\} + (1 - \exp\{-\theta\})S^*(t) \quad (2.2)$$

where

$$S^*(t) = \frac{\exp\{-\theta[1 - S_0(t)]\} - \exp\{-\theta\}}{1 - \exp\{-\theta\}},$$

obtaining a mathematical relationship between the model mixture model ([3], [2]).

In order to define a copula we first suppose that C_ϕ is a distribution function with density function c_ϕ on $[0, 1]^2$ for $\phi \in \mathcal{R}$. Then, let (T_1, T_2) denote the paired failure times, S_{pop_j} and f_{pop_j} denoting, respectively, the marginal long-term survival functions and the marginal long-term density function of T_j , $j = 1, 2$. Consider (T_1, T_2) comes from the C_ϕ copula for some ϕ then the joint survival and density function of (T_1, T_2) are given

$$S_{pop}(t_1, t_2) = C_\phi(S_{pop1}(t_1), S_{pop2}(t_2)), t_1, t_2 > 0, \quad (2.3)$$

and

$$f_{pop}(t_1, t_2) = c_\phi(S_{pop1}(t_1), S_{pop2}(t_2))f_{pop1}(t_1)f_{pop2}(t_2), t_1, t_2 > 0, \tag{2.4}$$

respectively. Notice that the marginal distributions and the dependence structure can be visualized separately and this dependence structure is represented by a copula. Following [12], considering the Farlie-Gumbel-Morgenstern distribution, hereafter FGM copula, we have the copula distribution function given as

$$C_\phi(u, v) = uv[1 + \phi(1 - u)(1 - v)], \tag{2.5}$$

where $0 \leq u, v \leq 1$ e $-1 \leq \phi \leq 1$.

For the FGM copula, C_ϕ given in (2.5), the Kendall's tau ([33]) is given by

$$\tau_\phi = \frac{2\phi}{9}, \tag{2.6}$$

for $\phi = 0$, τ_ϕ is equals 0.

Other dependence measures between the variables T_1 and T_2 can be found in [33]. For instance, if the random variables T_1 and T_2 have Weibull distribution with density function

$$f(t_j) = \alpha_j \lambda_j t_j^{\alpha_j - 1} \exp\{-\lambda_j t_j^{\alpha_j}\}, j = 1, 2, \tag{2.7}$$

it can be proved that the Pearson's correlation coefficient between T_1 and T_2 , is given by

$$\rho = \frac{\phi \left(1 - 2^{-\frac{1}{\alpha_1}}\right) \left(1 - 2^{-\frac{1}{\alpha_2}}\right)}{\sqrt{\left(\frac{\Gamma(\frac{2}{\alpha_1} + 1)}{\Gamma(\frac{1}{\alpha_1} + 1)^2} - 1\right) \left(\frac{\Gamma(\frac{2}{\alpha_2} + 1)}{\Gamma(\frac{1}{\alpha_2} + 1)^2} - 1\right)}}, \tag{2.8}$$

observe when $\phi = 0$ the correlation coefficient, ρ is equals 0. When T_j has Exponential distribution, with parameter $\alpha_j = 1, j = 1, 2$, the correlation coefficient is given by, $\rho = \phi/4$.

Consider (T_1, T_2) comes from the FGM copula (2.5) then the joint long-term survival of (T_1, T_2) are given by

$$S_{pop}(t_1, t_2) = S_{pop1}(t_1)S_{pop2}(t_2)(1 + \phi(1 - S_{pop1}(t_1))(1 - S_{pop2}(t_2))), \tag{2.9}$$

where ϕ parameter measures the intensity of the dependence between the lifetimes. Observe that when $\phi = 0, S_{pop}(t_1, t_2) = S_{pop1}(t_1)S_{pop2}(t_2)$, leading to the conclusion that the random variables T_1 and T_2 are independent.

At least in principle, other choices of C_ϕ in (2.3) can be made straightforwardly, leading to different copula survival models with promotion time models as marginal distributions. For instance, according to the dependence structure, other copula models can be equally considered such as the Gaussian copula as well as the Frank, Clayton and positive stable frailty families (Archimedean copulas) amongst others. The construction of the inferential procedure will be analogous to what shall be presented below, differing in the choice of the *a priori* distribution for the parameter dependence of the copula chosen. We however, in terms of comparison, consider in the numerical sections the positive stable frailty family.

3 INFERENCE

For inference, we adopt a full Bayesian approach and assume lifetimes in presence of random censoring. The likelihood function, prior distributions for the parameters in the model, details of the MCMC algorithm and the model comparison are described below.

Let (T_{i1}, T_{i2}) the i th bivariate lifetime and (C_{i1}, C_{i2}) the random censored bivariate lifetimes, for $i = 1, \dots, n$. Suppose that (T_{i1}, T_{i2}) and (C_{i1}, C_{i2}) are independent. For each individual i observed quantities are represented by the random variables $t_{ij} = \min(T_{ij}, C_{ij})$ and $\delta_{ij} = I(t_{ij} = T_{ij})$, which denotes a censorship indicator, $j = 1, 2$. Let $S_{pop}(t_1|\boldsymbol{\gamma}_1)$ and $S_{pop}(t_2|\boldsymbol{\gamma}_1)$ be the survival functions of T_{i1} and T_{i2} , respectively, where $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ are parameter vectors of q_1 and q_2 elements associated to each one of the marginal distributions.

Considering the bivariate survival function $S_{pop}(t_1, t_2|\phi, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2)$ given in (2.9), the contribution of the i th individual for the log-likelihood of $\boldsymbol{\theta} = (\phi, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2)$ is given by [30]

$$\begin{aligned} \ell_i(\boldsymbol{\theta}) = & \delta_{i1}\delta_{i2} \log \left(\frac{\partial^2 S_{pop}(t_1, t_2|\boldsymbol{\theta})}{\partial t_{i1} \partial t_{i2}} \right) + \delta_{i1}(1 - \delta_{i2}) \log \left(\frac{-\partial S_{pop}(t_1, t_2|\boldsymbol{\theta})}{\partial t_{i1}} \right) \\ & + \delta_{i2}(1 - \delta_{i1}) \log \left(\frac{-\partial S_{pop}(t_1, t_2|\boldsymbol{\theta})}{\partial t_{i2}} \right) + (1 - \delta_{i1})(1 - \delta_{i2}) \log S_{pop}(t_1, t_2|\boldsymbol{\theta}). \end{aligned} \quad (3.1)$$

3.1 Prior and posterior densities

The use of the Bayesian method besides being an alternative statistical approach, it allows the incorporation of previous knowledge of the parameters through an informative prior distribution. When there is not such previous knowledge one may consider a non-informative prior structure. For instance, in order to carry out a Bayesian inference procedure for the bivariate survival model of giving in (2.9), consider as in Section 2, that the parametric distribution family of the marginal lifetimes T_1 and T_2 are known and indexed by the parameter vectors, $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$, respectively. In order to guarantee proper posterior distributions, we adopt proper prior distributions according to the variation of the parametric space, but ensuring non-informativeness according to the fixed hyper-parameters, which lead to such situation. Thus, we consider a prior distribution of $\boldsymbol{\theta} = (\phi, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2)$ given by

$$\pi(\boldsymbol{\theta}) \propto (1 - \phi)^{r_1 - 1} (1 + \phi)^{r_2 - 1} \prod_{j=1}^2 \pi(\boldsymbol{\gamma}_j). \quad (3.2)$$

Equation (3.2) implies that $(1 - \phi)/2$ follows a $Beta(r_1, r_2)$ distribution, $\pi(\boldsymbol{\gamma}_j)$ is the prior distribution of $\boldsymbol{\gamma}_j$ and that $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2$ and ϕ are mutually independent, also $\phi \in (-1, 1)$.

Combining (3.2) with the likelihood function, $L(\boldsymbol{\theta}) = \exp(\sum_{i=1}^n \ell_i(\boldsymbol{\theta}))$, where $\ell_i(\boldsymbol{\theta})$ is given in (3.1), we straightforwardly obtain the joint posterior distribution of $\boldsymbol{\theta}$, $\pi(\boldsymbol{\theta}|D)$, where D is the observed dataset.

For cure rate model (2.1), the parameter vectors $\boldsymbol{\gamma}_j$, consists of the parameters of the distribution chosen for $S_0(t)$ and the parameter θ_j , $j = 1, 2$. Following [7] we incorporate covariates for the

parametric cure rate model (2.1) through the cure rate parameter θ as follows: For each subject i , $i = 1, \dots, n$, and for each time j , $j = 1, 2$, we consider $\theta_{ji} \equiv \theta(\mathbf{x}'_i \beta_j) = \exp(\mathbf{x}'_i \beta_j)$, where $\mathbf{x}'_i = (x_{i1}, \dots, x_{iq})$ denote the $q \times 1$ vector of covariates for the i th subject and $\beta_j = (\beta_{j1}, \dots, \beta_{jq})'$ denote the corresponding vector of regression coefficients. We choose the following independent prior distributions: $\beta_{0j} \sim Normal(\mu_{0j}, \sigma_{0j}^2)$, $\beta_{1j} \sim Normal(\mu_{1j}, \sigma_{1j}^2)$ $j = 1, 2$.

3.2 Computation

In the Bayesian approach, the target distribution for inference is the posterior of the parameters of interest. For this, we need to obtain the marginal posterior densities of each parameter, which are obtained by integrating the joint posterior density with respect to each parameter.

We point out that for any marginal distribution of T_1 and T_2 the joint posterior distribution is not tractable analytically but Markov chain Monte Carlo (MCMC) methods such as the Gibbs sampler, can be used to draw samples, from which features of the marginal posterior distributions of interest can be inferred [21]. For the estimation procedure we consider joint estimation where all the model parameters are estimated simultaneously in the MCMC algorithm. The Gibbs sampler is an iterative procedure of a broad class of methods generically named Markov Chain Monte Carlo (MCMC). Many practical aspects of the MCMC methodology are described in [18, 17]. This method is applicable in situations where one is not able to generate samples directly from the joint posterior density. It however requires the full conditional densities for generating samples.

For each time T_j we assume that $S_0(t_j)$ given in (2.1) has the Weibull distribution given in (2.7) with parameters α_j and λ_j , $j = 1, 2$. We choose the following independent prior distributions, $\alpha_j \sim Gamma(a_j, b_j)$, $\lambda_j \sim Gamma(c_j, d_j)$ $j = 1, 2$.

Combining the likelihood function (3.1) and the prior distribution (3.2), we can obtain the joint posterior density of all unobservables which is not tractable analytically but MCMC methods such as the Gibbs samples, can be used to draw samples, from which features of marginal posterior distributions of interest can be inferred. Considering $\mathcal{D} = (\mathbf{t}_1, \mathbf{t}_2, \mathbf{d}_1, \mathbf{d}_2, \mathbf{x})$, where for each individual i , $i = 1, 2$, \mathbf{t}_i and \mathbf{d}_i , are the observed vectors of lifetimes and censoring t_{ij} and δ_{ij} for $j = 1, 2$, respectively.

The full conditional posterior densities for each parameter are given by

$$\begin{aligned} \pi(\phi|\mathcal{D}, \theta_{(-\phi)}) &\propto (1 - \phi)^{r_1 - 1} (1 + \phi)^{r_2 - 1} \left(\prod_{i=1}^n \Delta_i \right), \\ \pi(\beta_{0j}|\mathcal{D}, \theta_{(-\beta_{0j})}) &\propto \left(\prod_{i=1}^n \Delta_i S_{popj}(t_{ji}) \right) \exp \left(\left\{ \sum_{i=1}^n \delta_{ji} + \frac{\mu_{0j}}{n\sigma_{0j}^2} \right\} \beta_{0j} - \frac{\beta_{0j}^2}{2\sigma_{0j}^2} \right), \\ \pi(\beta_{1j}|\mathcal{D}, \theta_{(-\beta_{1j})}) &\propto \left(\prod_{i=1}^n \Delta_i S_{popj}(t_{ji}) \right) \exp \left(\left\{ \sum_{i=1}^n \delta_{ji} x_i + \frac{\mu_{1j}}{n\sigma_{1j}^2} \right\} \beta_{1j} - \frac{\beta_{1j}^2}{2\sigma_{1j}^2} \right), \end{aligned}$$

$$\pi(\lambda_j | \mathcal{D}, \theta_{(-\lambda_j)}) \propto \left(\prod_{i=1}^n \Delta_i S_j(t_{ji}) \right) \lambda_j^{n+c_j-1} \exp \left(\left[\sum_{i=1}^n \theta_{ji} S_j(t_{ji}) \right] - d_j \lambda_j \right) \quad \text{and}$$

$$\pi(\alpha_j | \mathcal{D}, \theta_{(-\alpha_j)}) \propto \left(\prod_{i=1}^n \Delta_i t_{ji}^{\alpha_j-1} S_j(t_{ji}) \right) \alpha_j^{n+a_j-1} \exp \left(\left[\sum_{i=1}^n \theta_{ji} S_j(t_{ji}) \right] - b_j \alpha_j \right),$$

where

$$\theta_{ji} = \exp(\beta_{0j} + \beta_{1j} x_i), \quad S_j(t_{ji}) = \exp(-\lambda_j t_{ji}^{\alpha_j})$$

and

$$S_{popj}(t_{ji}) = x \exp(-\theta_{ji} (1 - S_j(t_{ji}))), \quad \text{for } j = 1, 2,$$

with

$$\Delta_i = v_{1i}^{(1-\delta_{1i})(1-\delta_{2i})} v_{2i}^{\delta_{1i}\delta_{2i}} v_{3i}^{\delta_{1i}(1-\delta_{2i})} v_{4i}^{\delta_{2i}(1-\delta_{1i})},$$

$$v_{1i} = 1 + \phi \prod_{j=1}^2 (1 - S_{popj}(t_{ji})),$$

$$v_{2i} = 1 + \phi \left(1 + \prod_{j=1}^2 (1 - 2S_{popj}(t_{ji})) \right),$$

$$v_{3i} = 1 + \phi \left(1 - 2S_{pop1}(t_{ji}) - S_{pop2}(t_{ji}) + 2 \prod_{j=1}^2 S_{popj}(t_{ji}) \right)$$

and

$$v_{4i} = 1 + \phi \left(1 - S_{pop1}(t_{ji}) - 2S_{pop2}(t_{ji}) + 2 \prod_{j=1}^2 S_{popj}(t_{ji}) \right), \quad \text{for } j = 1, 2.$$

The conditional densities above do not belong to any known parametric density family. In order to generate our samples we then implement an Metropolis-Hasting algorithm within Gibbs iterations ([9]). The simulations were performed using the OpenBUGS software ([43]). OpenBUGS codes are available by mailing to one of the authors.

3.3 Model comparison criteria

In the literature, there are various methodologies which intend to analyze the suitability of a distribution, as well as selecting the best fit among a collection of distributions. In this paper we shall inspect some of the Bayesian model selection criteria; namely, the deviance information criterion (*DIC*) proposed by [42], the expected Akaike information criterion (*EAIC*) by [4], and the expected Bayesian (or Schwarz) information criterion (*EBIC*) by [5] was used. These

criteria are based on the posterior mean of the deviance, $E\{D(\theta)\}$, which is also a measure of fit and can be approximated from the MCMC output by

$$Dbar = \frac{1}{V} \sum_{v=1}^V D(\theta_v)$$

where the index v indicates the v -th realization of a total of V realizations and

$$D(\theta) = -2 \sum_{i=1}^n \log(g((t_{1i}, t_{2i}|\theta)),$$

where $g(\cdot)$ is pdf corresponding to our model. For observed data we have that $g(t_{1i}, t_{2i}) = f_1(t_{1i})f_2(t_{2i})\Lambda_i$, for $i = 1, \dots, n$, with $\Lambda_i = 1 + \alpha(1 - 2S_{pop1}(t_{1i}) - 2S_{pop2}(t_{2i}) + 4S_{pop1}(t_{1i})S_{pop2}(t_{2i}))$, $f_{popj}(\cdot)$ and $S_{popj}(\cdot)$, $j = 1, 2$, are Weibull promotion density function and Weibull promotion survival function, respectively. The EAIC, EBIC and DIC criteria can be calculated using the MCMC output by means of $\widehat{EAIC} = Dbar + 2q$, $\widehat{EBIC} = Dbar + q \log(n)$ and $\widehat{DIC} = Dbar + \widehat{\rho}_D = 2Dbar - Dhat$, respectively, where q is the number of parameters in the model and ρ_D is the effective number of parameters, defined as $E\{D(\theta)\} - D\{E(\theta)\}$, where $D\{E(\theta)\}$ is the deviance evaluated at the expected values of the posterior distributions, which can be estimated as

$$Dhat = D\left(\frac{1}{V} \sum_{v=1}^V \phi^{(v)}, \frac{1}{V} \sum_{v=1}^V \alpha_1^{(v)}, \frac{1}{V} \sum_{v=1}^V \alpha_2^{(v)}, \frac{1}{V} \sum_{v=1}^V \lambda_1^{(v)}, \frac{1}{V} \sum_{v=1}^V \lambda_2^{(v)}, \frac{1}{V} \sum_{v=1}^V \beta_1^{(v)}, \frac{1}{V} \sum_{v=1}^V \beta_2^{(v)}\right).$$

Comparing alternative models, the preferred model is the one with the smallest criteria values.

Another criteria which is one of the most used in applied works is derived from the conditional predictive ordinate (CPO) statistics. For a detailed discussion on the CPO statistics and its applications to model selection, see [19]. Let \mathcal{D} the full data and $\mathcal{D}^{(-i)}$ denote the data with the i th observation deleted. We denote the posterior density of θ given $\mathcal{D}^{(-i)}$ by $\pi(\theta|\mathcal{D}^{(-i)})$, for $i = 1, \dots, n$. For the i th observation, the CPO_i can be written as

$$CPO_i = \int_{\Theta} g((t_{1i}, t_{2i})|\theta)\pi(\theta|\mathcal{D}^{(-i)}) = \left\{ \int_{\Theta} \frac{\pi(\theta|\mathcal{D}^{(-i)})}{g((t_{1i}, t_{2i})|\theta)} d\theta \right\}^{-1}, i = 1, \dots, n.$$

For the proposed model a closed form of the CPO_i is not available. However, a Monte Carlo estimate of CPO_i can be obtained by using a single MCMC sample from the posterior distribution $\pi(\theta|\mathcal{D})$. Let $\theta_1, \theta_2, \dots, \theta_Q$ be a sample of size Q of $\pi(\theta|\mathcal{D})$ after the burn-in. A Monte Carlo approximation of CPO_i ([8]) is given by

$$\widehat{CPO}_i = \left\{ \frac{1}{Q} \sum_{q=1}^Q \frac{1}{g((t_{1i}, t_{2i})|\theta_q)} \right\}^{-1}.$$

A summary statistics of the CPO_i 's is $B = \sum_{i=1}^n \log(\widehat{CPO})$. The larger is the value of B , the better is the fit of the model.

4 BAYESIAN CASE INFLUENCE DIAGNOSTICS

The best known perturbation schemes are based on case deletion ([13]), in which the effects are studied of completely removing cases from the analysis. This reasoning will form the basis for our Bayesian global influence methodology and in doing so it will be possible to determine which subjects might be influential for the analysis. In order to investigate if some of the observations are influential for the analysis, we considered a Bayesian case influence diagnostic procedure based on the Kullback-Leibler (K-L) divergence between P and $P_{(-i)}$, denoted by $K(P, P_{(-i)})$, where P denotes the posterior distribution of θ for the full data, and $P_{(-i)}$ denotes the posterior distribution of θ dropping the i -th observation. Specifically,

$$K(P, P_{(-i)}) = \int \pi(\theta|\mathcal{D}) \log \left[\frac{\pi(\theta|\mathcal{D})}{\pi(\theta|\mathcal{D}^{(-i)})} \right] d\theta. \quad (4.1)$$

The $K(P, P_{(-i)})$ measures the effect of deleting of i -th observation from the full data on the joint posterior distribution of θ . As pointed by [36], calibration of $K(P, P_{(-i)})$ can be done by solving for p_i the equation $K(P, P_{(-i)}) = K[B(0.5), B(p_i)] = -\log[4p_i(1-p_i)]/2$ where $B(p)$ denotes the Bernoulli distribution with success probability p . This implies that describing outcomes using $\pi(\theta|\mathcal{D})$ instead of $\pi(\theta|\mathcal{D}^{(-i)})$ is compatible with describing an unobserved event as having probability p_i when correct probability is 0.5. After some algebraic manipulation it can be shown that $p_i = \frac{1}{2}\{1 + \sqrt{1 - \exp[-2K(P, P_{(-i)})]}\}$. This equation implies that $0.5 \leq p_i \leq 1$.

That is, if $p_i \gg 0.5$ then the i th case is considered influential.

For our case it can be shown that (4.1) can be expressed as a posterior expectation

$$\begin{aligned} K(P, P_{(-i)}) &= \log E_{\theta|\mathcal{D}} \{ [g((t_{1i}, t_{2i})|\theta)]^{-1} \} + E_{\theta|\mathcal{D}} \{ \log [g((t_{1i}, t_{2i})|\theta)] \} \\ &= -\log(CPO_i) + E_{\theta|\mathcal{D}} \{ \log [g((t_{1i}, t_{2i})|\theta)] \}, \end{aligned} \quad (4.2)$$

where $E_{\theta|\mathcal{D}}(\cdot)$ denotes the expectation with respect to the joint posterior $\pi(\theta|\mathcal{D})$. Thus (4.2) can be estimated by sampling from the posterior distribution of θ via MCMC methods. Let $\theta_1, \theta_2, \dots, \theta_Q$ be a sample of size Q of $\pi(\theta|\mathcal{D})$. Then, a Monte Carlo estimate of $K(P, P_{(-i)})$ is given by

$$K(\widehat{P}, \widehat{P}_{(-i)}) = -\log(\widehat{CPO}_i) + \frac{1}{Q} \sum_{q=1}^Q \log [g((t_{1i}, t_{2i})|\theta_q)]. \quad (4.3)$$

5 APPLICATIONS

In this section, results from simulation studies and a real data example are presented in order to illustrate the performance of the proposed methodology.

5.1 Artificial Data

In the study we considered the survival model given in (2.9), assuming that $S_0(t_j)$ has the Weibull distribution given in (2.7) with parameters α_j and λ_j , $j = 1, 2$. The covariates x_i were generated from a Bernoulli distribution with parameter 0.5 and the artificial bivariate data (T_{i1}, T_{i2}) , $i = 1, \dots, n$, were generated assuming $n = 160$ according to the following steps: First we generated $T_{i1} = (-\log(1 - u_{i1})/\lambda_{i1})^{1/\alpha_1}$ where $u_{i1} \sim U(0, 1)$. Next T_{i2} was generated using a random variable $u_{i2} \sim U(0, 1)$ and the solution of the nonlinear equation, $w_i + \phi(2u_{i1} - 1)(w_i^2 - w_i) - u_{i2} = 0$, considering $T_{i2} = (-\log(1 - w_i)/\lambda_{i2})^{1/\alpha_2}$. For each covariate value k , $k = 0, 1$, we fixed the cure rate $\pi_{kj} = \exp \theta_{kj}$ value, $j = 1, 2$. For t_{kj} simulating a cutoff point t_{kj}^* which was chosen so that the higher times represent π_{kj} % of the lifetimes, and were assumed censored. The following cure rate values was fixed: $\pi_{01} = 0.4$, $\pi_{11} = 0.3$, $\pi_{02} = 0.3$ and $\pi_{12} = 0.2$. We considered the following values for the other parameters of the model $\alpha_1 = 2$, $\lambda_1 = 0.5$, $\alpha_2 = 1.5$, $\lambda_2 = 0.2$ and $\phi = 0.6$.

For sake of illustration we also compare our FGM bivariate copula model with the well known positive stable frailty (PSF) model ([25]). The joint survival function of the PSF model is given by

$$S_{pop}(t_1, t_2) = ((-\ln(S_{pop1}(t_1)))^{\frac{1}{\phi}} + (-\ln(S_{pop2}(t_2)))^{\frac{1}{\phi}})^{\phi}. \quad (5.1)$$

When $\phi \rightarrow 1$, we obtain $S_{pop}(t_1, t_2) = S_{pop}(t_1)S_{pop}(t_2)$. The Kendall's tau⁷ is given by $\tau_{\phi} = 1 - \phi$.

For more details on frailty models where the random effects have a positive stable distribution interested readers can refer to [28, 32].

The following independent priors were considered to perform the Gibbs sampler:

$$\beta_{ij} \sim N(0, 10^3), \quad \lambda_j \sim \text{Gamma}(1, 0.001) \quad \text{and} \quad \alpha_j \sim \text{Gamma}(1, 0.001),$$

where $i = 0, 1$ and $j = 1, 2$. For parameter of PSF model we assume $\phi \sim \text{Beta}(1, 1)$ and for FGM copula that $1 - 2\pi(\phi) \sim \text{Beta}(1, 1)$, such choices guarantee that $\phi \in (-1, 1)$ and ensure non-informativeness.

The simulations were performed using the OpenBUGS software ([43]). For each generated data set we simulate two chain of size 50,000 for each parameter, disregarding the first 10,000 iterations to eliminate the effect of the initial values and to avoid autocorrelation problems, we consider a spacing of size 20, obtaining a effective sample of size 4,000 upon which the posterior inference is based on. For each sample the posterior mean of the parameter and the EAIC, EBIC, DIC and B are recorded. The MCMC convergence was monitored according to the methods recommended by [14] (CODA package). The number of iterations is considered sufficient for the approximate convergence since in all cases the Gelman-Rubin diagnostic is very close to 1.

Table 1 presents the summary for the FGM bivariate long-term survival model parameters. Table 2 presents the Bayesian criteria for the FGM and PSF models. The FGM model outperforms the PSF one in all considered criteria.

Table 1: Simulated data. Posterior mean, standard deviation (SD) and HPD (95%) interval for the FGM bivariate long-term survival model parameters.

	Parameter	Mean	SD	HPD (95%)
Time 1	α_1	2.362	0.210	(1.972; 2.793)
	λ_1	0.365	0.162	(0.060; 0.696)
	β_{01}	0.01	0.149	(-0.281; 0.292)
	β_{11}	0.135	0.191	(-0.233; 0.499)
Time 2	α_2	1.678	0.149	(1.382; 1.974)
	λ_2	0.163	0.029	(0.112; 0.226)
	β_{02}	0.369	0.162	(0.054; 0.703)
	β_{12}	0.103	0.178	(-0.236; 0.457)
Copula	ϕ	0.587	0.212	(0.151; 0.942)

Table 2: Simulated data. Bayesian criteria.

Model	Criterion			
	EAIC	EBIC	DIC	B
FGM	839.750	867.426	830.3	-415.100
PSF	843.506	871.183	834.3	-417.128

5.2 Influence of outlying observations

One of our main goals in this study is to show the need for robust models to deal with the presence of outliers in the data. In order to do so, we generated a sample of length 160 with fixed parameters $\alpha_1 = 2$, $\lambda_1 = 0.5$, $\alpha_2 = 1.5$, $\lambda_2 = 0.2$ and $\phi = 0.6$. The percentage of censor was controlled as follows: for each covariate value k , $k = 0, 1$, we fixed the cure rate $\pi_{kj} = \exp\{\theta_{kj}\}$ value, $j = 1, 2$. For t_{kj} simulating a cutoff point t_{kj}^* which was chosen so that the higher times represent $\pi_{kj}\%$ of the lifetimes, and were assumed censored.

We selected cases 90, 107 and 120 for perturbation. The perturbation scheme were structured as following. To create influential observation artificially in the dataset, we choose one, two or three of these selected cases. For each case we perturbed one or both lifetimes as follows $\tilde{t}_i = t_i + 7S_t$, $i = 1, 2$, where S_t is the standard deviations of the t_i 's. For case 107 we perturbed only the lifetime t_1 and for case 90, the lifetime t_2 , and for case 120, both lifetimes were perturbed.

The MCMC computations were made in a similar fashion to those in the last section and further to monitor the convergence of the Gibbs samples we also used the methods recommended by [14]. Table 3 shows that the posterior inferences are sensitive to the perturbation of the selected case(s). In Table 3, Dataset a denotes the original simulated data set with no perturbation, and Datasets b to g denote datasets with perturbed cases.

Table 3: Simulated data. Posterior means and standard deviations (SD) for the FGM bivariate survival model parameters according to different perturbation schemes.

Dataset	Perturbed case	β_{01}	β_{11}	α_1	λ_1	β_{02}	β_{12}	α_2	λ_2	ϕ
		Mean [SD]								
a	Original Sample	0.010	0.135	2.362	0.365	0.369	0.103	1.678	0.163	0.587
		0.149	0.191	0.210	0.162	0.162	0.178	0.149	0.029	0.212
b	107	0.362	0.137	1.641	0.426	0.366	0.098	1.672	0.163	0.579
		0.231	0.197	0.136	0.108	0.162	0.186	0.152	0.030	0.218
c	90	0.0112	0.124	2.385	0.810	0.703	0.082	1.342	0.129	0.611
		0.1500	0.195	0.207	0.113	0.245	0.173	0.112	0.032	0.212
d	120	0.3737	0.143	1.627	0.419	0.698	0.081	1.341	0.130	0.630
		0.2520	0.195	0.129	0.110	0.247	0.179	0.109	0.033	0.207
e	{90, 107}	0.3595	0.137	1.641	0.427	0.681	0.086	1.343	0.133	0.599
		0.2300	0.198	0.132	0.108	0.239	0.183	0.109	0.032	0.221
f	{107, 120}	0.6657	0.097	1.486	0.294	0.705	0.077	1.333	0.130	0.621
		0.3340	0.197	0.123	0.095	0.266	0.186	0.111	0.033	0.213
g	{90, 107, 120}	0.652	0.099	1.487	0.295	0.886	0.048	1.245	0.112	0.619
		0.305	0.198	0.119	0.093	0.279	0.184	0.099	0.032	0.217
h	42	0.0023	0.138	2.398	0.809	0.666	0.132	1.355	0.130	0.626
		0.1490	0.194	0.212	0.112	0.243	0.181	0.109	0.033	0.209
i	{42, 120}	0.3491	0.145	1.639	0.429	0.858	0.108	1.250	0.110	0.659
		0.2150	0.200	0.129	0.103	0.296	0.187	0.102	0.031	0.204

Table 4 displays the fit of different cases of the perturbed data set. We can observe that the original simulated data (Dataset a) had the best fit.

Table 4: Simulated data. Bayesian criteria for each perturbed version fitting bivariate model based on the FGM bivariate survival model parameters according to different perturbation schemes.

Data names	Bayesian criteria			
	EAIC	EBIC	DIC	B
a	839.750	867.426	830.300	-415.100
b	870.308	897.985	861.100	-434.932
c	857.951	885.628	848.400	-426.561
d	888.006	915.682	878.500	-445.262
e	888.387	916.063	878.900	-446.388
f	898.624	926.301	889.100	-449.301
g	908.782	936.458	899.200	-451.618
h	856.841	884.518	847.400	-425.651
i	897.509	925.185	888.100	-448.408

Now we consider the sample from the posterior distributions of the parameters of the model based on the FGM bivariate survival model to compute the K-L divergence and calibration of this divergence, as describe in Section 4. The results in Table 5 show, before perturbation (Dataset a), that all the selected cases are not influential with small $K(P, P_{(-i)})$ and related calibration close to 0.596. However, after perturbations (Dataset b to i), the $K(P, P_{(-i)})$ increases and the corresponding calibrations become larger than 0.5, indicating those cases are influential.

Table 5: Case influence diagnostics for the simulated data.

Data names	Case number	$K(P, P_{(-i)})$	Calibration
a	42	0.019	0.596
	90	0.010	0.570
	107	0.010	0.570
	120	0.010	0.570
b	107	4.896	1.000
c	90	2.660	0.999
d	120	6.715	1.000
e	90	2.047	0.996
	107	5.602	1.000
f	107	0.789	0.945
	120	4.680	1.000
g	90	0.544	0.907
	107	0.737	0.939
	120	1.572	0.989
h	42	2.184	0.997
i	42	0.602	0.918
	120	4.593	1.000

In the Figure 1 shows the $K(P, P_{(-i)})$ for the FGM bivariate survival model. Clearly we can see that $K(P, P_{(-i)})$ performed well to identifying influential case(s), providing larger $K(P, P_{(-i)})$ when compared to the other cases.

We point out that, the way in which our copula survival model was constructed, we expect that $K(P, P_{(-i)})$ detects the perturbed observations if the data is generated from a model assuming a particular copula and is adjusted assuming the other copulas. This is due to the fact that the marginals are the same, independent of the assumed copula function, differing only in the structure of dependence between them. Nevertheless, we considered the data sets generated according to the of FGM copula according to Subsection 5.1, with their perturbed observations and fitted the model assuming the FGM copula, as well as the Archimedean copulas PS, Frank and Clayton. In all situations $K(P, P_{(-i)})$ detects the influential points in the same manner.

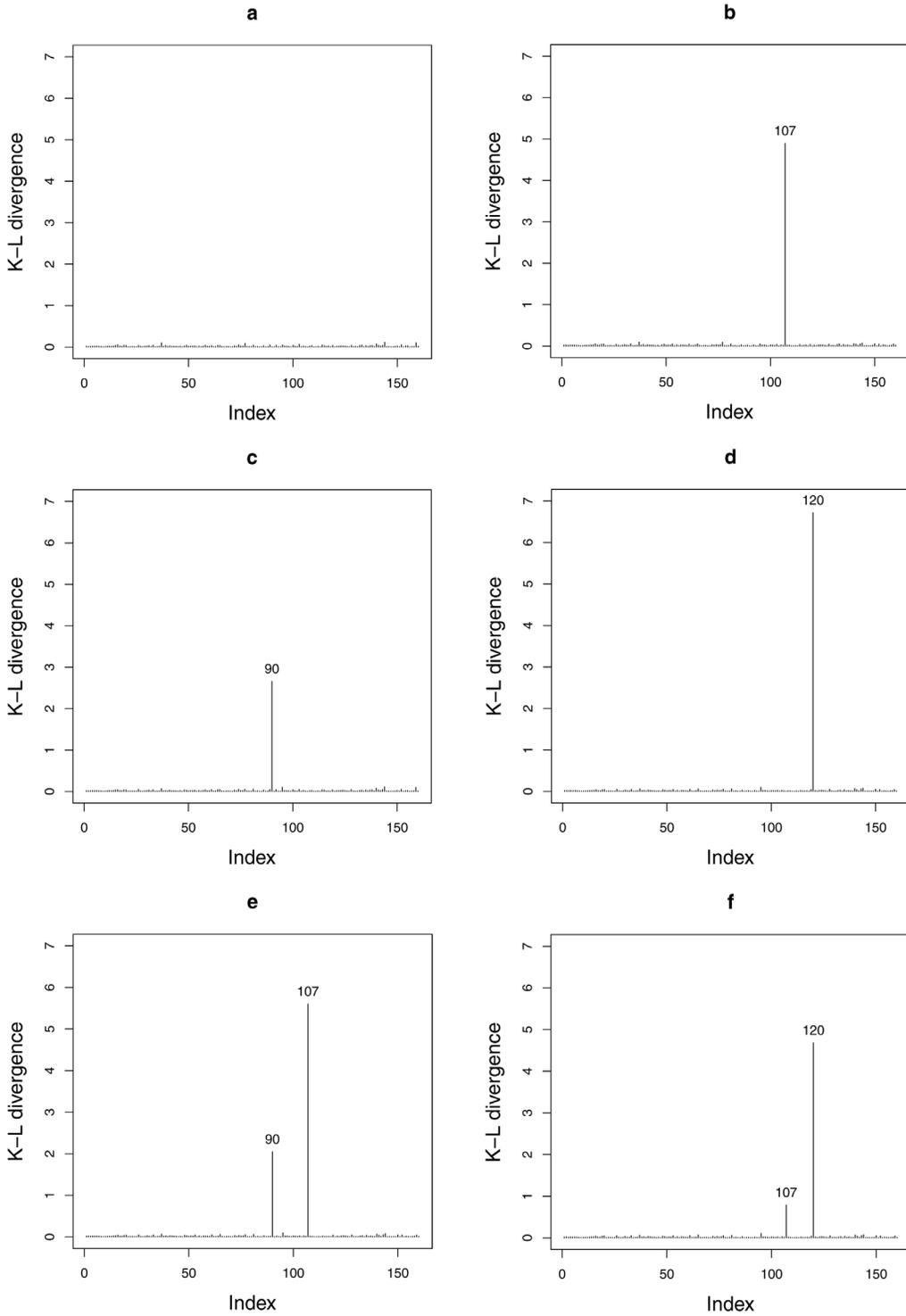


Figure 1: (to be continue next page).

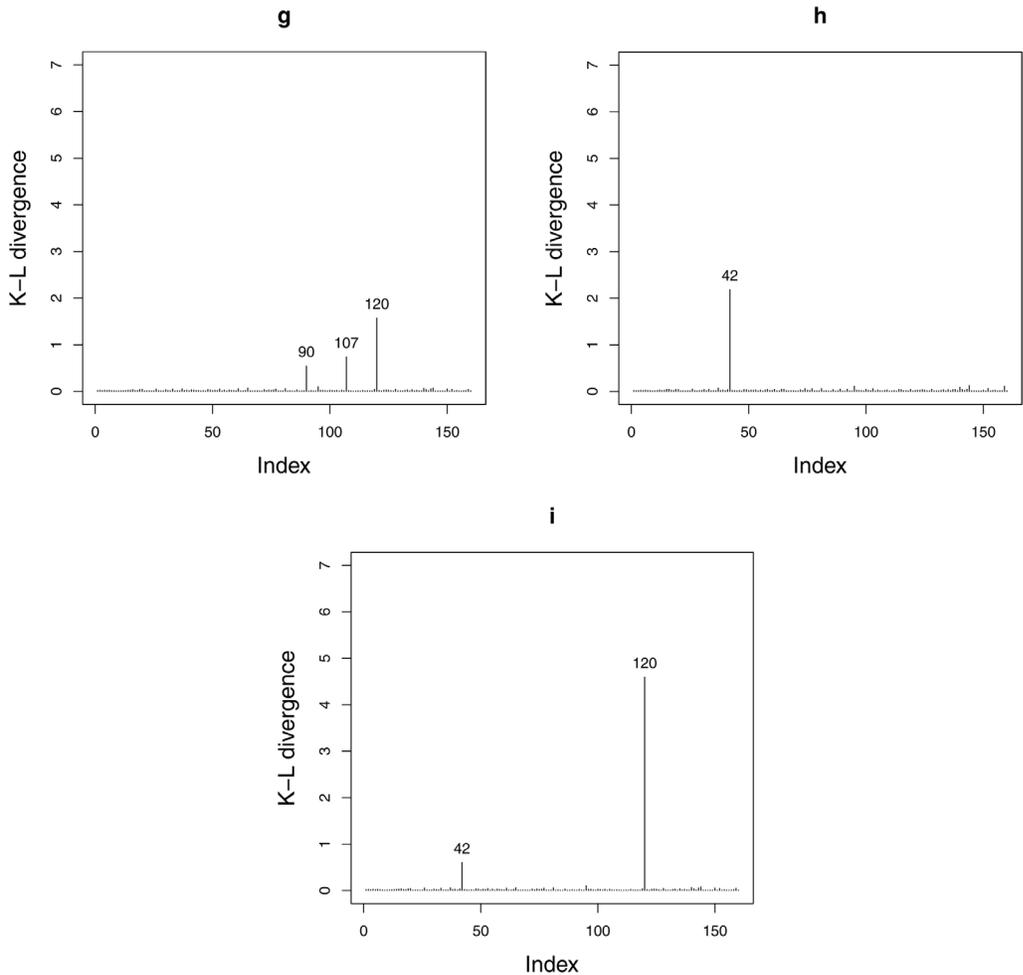


Figure 1: Simulation data. Index plot of $K(P, P_{(-i)})$ from fitting a bivariate survival model based on FGM copula.

5.3 Real data

In this section we consider a Bayesian analysis of the dataset presented in [44] that consist of follow up times for 197 diabetic patients under 60 years of age. The main purpose of the study is to assess the efficacy of photocoagulation treatment for proliferative retinopathy. The treatment was randomly assigned to one eye of each patient and the other, was considered as a control. The first component is a vector of lifetimes which represents the time up to visual loss for the treatment eye (T_1), while the second component (T_2) is the time up to visual loss for the control eye. The subjects could be censored, which happened for 73% of the treated eyes and 49% of the untreated eyes. As [39] we consider as covariate the dichotomized patient age (less than 20 years and older than 20 years). For sake of illustration the perturbation scheme were structured

as following. To create an influential observation artificially in the dataset we selected the case 100 and perturbed both lifetimes as follows: $\tilde{t}_i = t_i + 7S_t$, $i = 1, 2$, where S_t is the standard deviations of the t_i 's.

The data set was then reanalyzed though our approach by adopting the model in (2.1) with promotion time Weibull as marginal. The Bayes estimates were based on posterior samples recorded every 20th iteration from 50,000 Gibbs samples after a burn-in of 10,000 samples. Similarly to what was considered when the simulation study, we choose the following independent proper prior distributions: $\beta_{ij} \sim N(0, 10^3)$, $\lambda_j \sim \text{Gamma}(1, 0.001)$ and $\alpha_j \sim \text{Gamma}(1, 0.001)$, where $i = 0, 1$ and $j = 1, 2$. For parameter of PSF model we assume $\phi \sim \text{Beta}(1, 1)$ and for FGM copula that $1 - 2\pi(\phi) \sim \text{Beta}(1, 1)$, such choices guarantee that $\phi \in (-1, 1)$ and ensure non-informativeness. To monitor the convergence of the Gibbs samples we uses the methods recommended by [14].

In Table 6 we report posterior summaries for the parameters under the FGM long-term bivariate survival copula model considering Weibull and Exponential promotion marginals for $S_0(t)$ given in (2.1), considering the original sample as well as the perturbed one. Following [20], we have checked the sensitivity of the routine use for gamma prior on the variance components and found that results are fairly robust under different prior. We also checked the sensitivity analysis for the variance components parameters for various choices of prior parameters by changing only on parameter at a time and keeping all other parameters constant to their default values. The posterior summaries of the parameters do not present remarkable difference and not impair the results in Table 6. Considering both the original and perturbed samples, we observed a small change in the posterior estimates of all parameters, showing the importance of verifying the posterior inference sensitivity to perturbations. Also, for both samples, the HPD credibility intervals for the location parameter of the marginal Weibull distributions contains the 1, leading to evidence in favor of a Exponential distribution to the marginal ones. Then, considering the FGM long-term bivariate survival copula model with exponential promotion marginals for $S_0(t)$ given in (2.1), the Kendall's tau (2.6) are equals to 0.184 and 0.173, respectively, for the original sample and the perturbed one, leading to the the existence of a weak dependence between the marginal quantities.

Considering the FGM long-term bivariate survival copula model with Exponential promotion marginals for $S_0(t)$ given in (2.1), we compute the K-L divergences and related calibrations for the original sample, but do not find highly influential cases. The $K(P, P_{(-i)})$ are smaller that 0.096 and the corresponding calibrations are smaller than 1. For the perturbed sample however, for the case 100, the $K(P, P_{(-i)})$ is equals to 1.019 and the corresponding calibration is equal to 0.966 (detecting the influence point). For the remaining cases, $K(P, P_{(-i)})$ are smaller that 0.069 with corresponding calibrations smaller than 1. Figure 2 shows the index plot of $K(P, P_{(-i)})$ from fitting the bivariate survival model based on FGM copula considering both samples (original and perturbed).

Table 7 presents the model comparison criteria discussed in Section 3 for comparing the FGM long-term bivariate survival copula model with Weibull promotion marginal distributions with

Table 6: Diabetic retinopathy study data. Summary results from the posterior distribution, mean and standard deviation (SD) for parameter under model based on FGM copula.

	Parameter	Weibull		Exponential	
		Mean	HPD (95%)	Mean	HPD (95%)
Original Sample	α_1	1.057	(0.800; 1.330)	—	—
	λ_1	0.024	(0.010; 0.048)	0.026	(0.020; 0.044)
	β_{01}	-0.570	(-0.961; 0.187)	-0.523	(-1.020; 0.212)
	β_{11}	-0.469	(-0.999; -0.130)	-0.433	(-1.044; 0.114)
	α_2	0.977	(0.804; 1.150)	—	—
	λ_2	0.019	(0.008; 0.033)	0.017	(0.006; 0.030)
	β_{02}	0.155	(-0.332; 1.035)	0.199	(-0.335; 1.023)
	β_{12}	0.363	(0.003; 0.669)	0.350	(0.031; 0.735)
	ϕ	0.830	(0.499; 0.994)	0.813	(0.4085; 0.993)
Perturbed Sample	α_1	0.843	(0.638; 1.073)	—	—
	λ_1	0.015	(0.002; 0.053)	0.015	(0.003; 0.031)
	β_{01}	0.427	(-0.784; 1.774)	-0.103	(-0.791; 1.034)
	β_{11}	0.239	(-0.212; 0.842)	-0.522	(-1.065; -0.011)
	α_2	0.943	(0.772; 1.143)	—	—
	λ_2	0.019	(0.009; 0.034)	0.014	(0.005; 0.025)
	β_{02}	-0.228	(-0.948; 0.174)	0.303	(-0.236; 1.072)
	β_{12}	0.308	(-0.206; 0.696)	0.348	(0.005; 0.740)
	ϕ	0.780	(0.385; 0.992)	0.809	(0.465; 0.993)

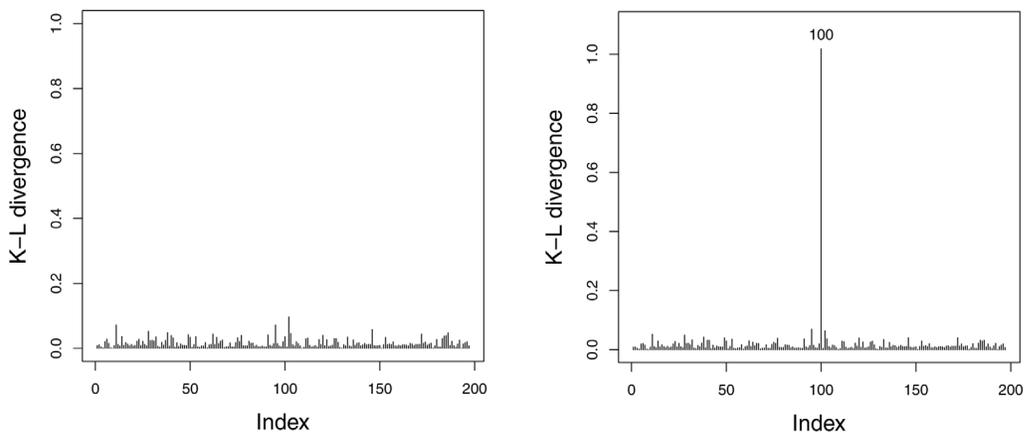


Figure 2: Diabetic Retinopathy study data. Index plot of $K(P, P_{(-i)})$ from fitting the bivariate survival model based on FGM copula.

its particular independence cases. There is positive evidence in favor The FGM model at the expense of the model which considers independence. Also, there is evidence in favor of the FGM long-term bivariate survival copula model with Exponential promotion marginal distributions.

Besides, we fitted the PSF model for the diabetic retinopathy study data considering both Weibull and Exponential marginal promotion distributions. Based on all Bayesian criteria, there is positive evidence in favor the FGM model at the expense of the PSF model, indicating that the FGM long-term bivariate survival copula model can be seen as a competitor to the PSF model, which is a well known model commonly used in literature for fitting bivariate lifetime data. As expected, all the criteria presented evidence in favor the original sample (without the influential point).

Table 7: Diabetic retinopathy study data. Bayesian criteria.

	Model	EAIC	EBIC	DIC	B
Original Data	FGM Weibull	1668.667	1698.216	1657	-828.157
	FGM Exponential	1665.295	1688.278	1657	-828.812
	PSF Weibull	1672.354	1701.903	1662	-831.180
	PSF Exponential	1671.071	1700.619	1659	-830.188
	Independence Weibull	1682.922	1709.187	1673	-837.187
	Independence Exponential	1677.419	1697.120	1671	-835.820
Perturbed Data	FGM Weibull	1685.183	1714.732	1670	-838.286
	FGM Exponential	1677.427	1700.409	1668	-835.423
	PSF Weibull	1682.526	1712.074	1668	-836.560
	PSF Exponential	1679.256	1702.239	1671	-836.787
	Independence Weibull	1694.613	1720.879	1684	-843.252
	Independence Exponential	1689.951	1709.650	1682	-842.850

6 SOME FINAL REMARKS

In this paper, we presented the FGM long-term bivariate survival copula model. Parameter estimation is based on a Bayesian approach via MCMC. We have used Bayesian case influence diagnostic based on the Kullback-Leibler divergence in order to study the sensitivity of the Bayesian estimates under perturbations in the model/data. Finally, we demonstrate our approach with an artificial and a real dataset.

In the two-step approach, the marginal parameters are estimated firstly and then, the copula parameter is estimated in a second step. This approach provides consistent, but not efficient estimators using a frequentist approach. However, from the Bayesian perspective, one can estimate all the model parameters simultaneously via the MCMC algorithm such that the assumption of independence in the first step is avoided. We however also considered a two-step estimation

method for all study. We omitted the results obtained since they are similar to those obtained by considering the joint estimation approach which we prefer.

Here we assume lifetimes in presence of random censoring. However, at least in principle, the proposed methodology can be developed straightforwardly for the type I and type II censoring particular cases. Interval and left censoring are also important and could be considered further.

ACKNOWLEDGMENTS

The researchers of Francisco Louzada and Vicente G. Cancho are supported by the Brazilian organization CNPq.

RESUMO. Neste artigo nós propomos um modelo bivariado de longa duração baseado na copula de Farlie-Gumbel-Morgenstern, onde assumimos marginais com estrutura de tempo de promoção. O modelo proposto permite a presença de dados censurados e de covariáveis. Para fins inferenciais foi considerada uma abordagem bayesiana usando métodos Monte Carlo em Cadeias de Markov. Além disso, algumas discussões sobre os critérios de seleção de modelos são apresentadas. A fim de detectar outliers e observações influentes, nós apresentamos um método bayesiano de análise de influência de deleção de caso baseado na divergência de Kullback-Leibler. Os procedimentos desenvolvidos são ilustrados em dados artificiais e reais.

Palavras-chave: Abordagem Bayesiana, Diagnóstico de influência de deleção de caso, modelagem copula, sobrevivência de longa duração.

REFERENCES

- [1] H. Aslanidou, D.K. Dey & D. Sinha. Bayesian analysis of multivariate survival data using Monte Carlo methods. *Canadian Journal of Statistics*, **1** (2008), 33–48.
- [2] J. Berkson & R.P. Gage. Survival cure for cancer patients following treatment. *Journal of the American Statistical Association*, **47**(259) (1952), 501–515.
- [3] J.W. Boag. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society B*, **11**(1) (1949), 15–53.
- [4] S.P. Brooks. Discussion on the paper by Spiegelhalter, Best, Carlin, and van der Linde. *Journal Royal Statistical Society B*, **64**(3) (2002), 616–618.
- [5] B.P. Carlin & T.A. Louis. Bayes and Empirical Bayes Methods for Data Analysis, 2nd ed., Chapman & Hall/CRC, Boca Raton (2001).
- [6] D.G. Clayton. A model for association in bivariate life-tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, **65** (1978), 141–151.
- [7] M.H. Chen, J.G. Ibrahim & D. Sinha. A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association*, **94**(447) (1999), 909–919.

- [8] M.H. Chen, Q.M. Shao & J. Ibrahim. Monte Carlo Methods in Bayesian Computation. Springer-Verlag, New York (2000).
- [9] S. Chib & E. Greenberg. Understanding the metropolis-hastings algorithm. *The American Statistician*, **49** (1995), 327–335.
- [10] S.C. Chiou & R.S. Tsay. A Copula-based Approach to Option Pricing and Risk Assessment. *Journal of Data Science*, **6** (2008), 273–301.
- [11] G. Colby & P. Rilstone. Simplified estimation of multivariate duration models with unobserved heterogeneity. *Computational Statistics*, **22** (2007), 17–29.
- [12] D.A. Conway. Farlie-Gumbel-Morgenstern distributions, in *Encyclopedia of Statistical Sciences 3*, S. Kotz & N.L. Johnson, eds., John Wiley & Sons, New York, (1983), 28–31.
- [13] R.D. Cook & S. Weisberg. Residuals and Influence in Regression. Chapman & Hall/CRC, Boca Raton (1982).
- [14] M.K. Cowless & B.P. Carlim. Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, **91** (1996), 883–904.
- [15] G. De Luca & G. Riviaccio. Archimedean copulae for risk measurement. *Journal of Applied Statistics*, **36**(8) (2009), 907–924.
- [16] P. Embrechts, F. Linskog & A. McNeil. Modelling dependence with copulas and applications to risks management, available at <http://www.math.ethz.ch/baltes/ftp/papers.html> (2003).
- [17] D. Gamerman & H.F. Lopes. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, 2nd ed., Chapman & Hall, Boca Raton (2006).
- [18] A.E. Gelfand & A.F.M. Smith. Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, **85**(410) (1990), 398–409.
- [19] A.E. Gelfand, D.K. Dey & H. Chang. Model determination using predictive distributions with implementation via sampling-based methods. Bayesian Statistics 4, Peñíscola, Oxford Univ. Press, New York, (1991), 147–167.
- [20] A. Gelman, J. Carlin & D. Rubin. Bayesian Data Analysis. Chapman & Hall/CRC, New York (2006).
- [21] W.R. Gilks, S. Richardson & D.J. Spiegelhater. Markov Chain Monte Carlo in Practice. Chapman & Hall, London (1996).
- [22] K. Goethals, P. Janssen & L. Duchateau. Frailty models and copulas: similarities and differences. *Journal of Applied Statistics*, **35**(9) (2008), 1071–1079.
- [23] P. Gustafson, D. Aeschliman & A.R. Levy. A simple approach to fitting Bayesian survival models. *Lifetime Data Analysis*, **9** (2003), 5–19.
- [24] D. Hanagal. Modeling Survival Data Using Frailty Models. Chapman & Hall/CRC, Boca Raton (2011).
- [25] P. Hougaard. A class of multivariate failure time distributions. *Biometrika*, **73** (1986), 671–678.
- [26] P. Hougaard. Fitting a multivariate failure time distribution. *IEEE Transactions on Reliability*, **38** (1989), 444–448.
- [27] P. Hougaard. Analysis of Multivariate Survival Data. Springer, Heidelberg (2000).
- [28] J.G. Ibrahim, M-H. Chen & D. Sinha. Bayesian Survival Analysis. Springer Verlag, New York (2001).

- [29] P. Jaworski. Copula Theory and Its Applications: Proceedings of the Workshop Held in Warsaw, 25-26 September 2009 Volume 198 de Lecture Notes in Statistics, eds. P. Jaworski, F. Durante, W. Härdle & T. Rychlik, Springer (2010).
- [30] J.F. Lawless. Statistical Models and Methods for Lifetime Data. New York: Wiley & Sons (2003).
- [31] R.A. Maller & X. Zhou. Survival Analysis with Long-Term Survivors. Wiley, New York (1996).
- [32] A.K. Manatunga & D. Oakes. Parametric Analysis of Matched Pair Survival Data. *Lifetime Data Analysis*, **5** (1999), 371–387.
- [33] R. Nelsen. An Introduction to Copulas, 2nd ed., Springer, New York (2006).
- [34] D. Oakes. Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, **84** (1989), 487–493.
- [35] D. Oakes. On Frailty Models and Copulas Recent Advances in Statistical Methods. Proceedings of Statistics 2001 Canada: The 4th Conference in Applied Statistics, Montreal, (2001), 218–224.
- [36] F. Peng & D. Dey. Bayesian analysis of outlier problems using divergence measures. The Canadian Journal of Statistics. *La Revue Canadienne de Statistique*, **23**(2) (1995), 199–213.
- [37] Y. Peng, K.B.G. Dear & J.W. Denham. A generalized F mixture model for cure rate estimation. *Statistics in Medicine*, **17** (1998), 813–830.
- [38] J. Rodrigues, M. de Castro, V.G. Cancho & F. Louzada Neto. On the unification of long-survival models. *Statistics and Probabilities Letters*, **79**(1) (2009), 753–759.
- [39] J.S. Romeo, N.I. Tanaka & A.C. Pedroso de Lima. Bivariate survival modeling: a Bayesian approach based on Copulas. *Lifetime Data Analysis*, **12** (2006), 205–222.
- [40] S.K. Sahu & D.K. Dey. A comparison of frailty and other models for bivariate survival data. *Lifetime Data Analysis*, **6** (2000), 207–228.
- [41] J.H. Shih & T.A. Louis. Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, **51** (1995), 1384–1399.
- [42] D.J. Spiegelhalter, N.G. Best, B.P. Carlin & A. van der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B*, **64** (2002), 583–639.
- [43] D. Spiegelhalter, A. Thomas, N. Best & D. Lunn. OpenBUGS User Manual, version 3.0.2, MRC Biostatistics Unit, Cambridge; software available at <http://mathstat.helsinki.fi/openbugs>.
- [44] The Diabetic Retinopathy Study Research Group. Preliminary report on the effect of photo-coagulation therapy. *American Journal of Ophthalmology*, **81** (1976), 383–396.
- [45] P.K. Trivedi & D.M. Zimmer. Copula modelling: an introduction for practitioners. *Foundations and Trends in Econometrics*, **1** (2005), 1–111.
- [46] J.W. Vaupel, K.G. Manton & E. Stallard. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, **16** (1979), 439–454.
- [47] A. Wienke. Frailty Models in Survival Analysis. Chapman & Hall/CRC, Boca Raton (2011).
- [48] A.Y. Yakovlev & A.D. Tsodikov. Stochastic Models of Tumor Latency and Their Biostatistical Applications. World Scientific, Singapore (1996).
- [49] S. Zhang, Y. Zhang, K. Chaloner & J.T. Stapleton. A copula model for bivariate hybrid censored survival data with application to the MACS study. *Lifetime Data Analysis*, **16** (2010), 231–249.