

The Long-Term Bivariate Survival FGM Copula Model: An Application to a Brazilian HIV Data

Francisco Louzada^{1*}, Adriano K. Suzuki², Vicente G. Cancho¹,
Fátima L. Prince³ and Gilberto A. Pereira³

¹*Universidade de São Paulo*, ²*Universidade Federal de São Carlos* and
³*Universidade Federal do Triângulo Mineiro*

Abstract: In this paper we propose a new bivariate long-term distribution based on the Farlie-Gumbel-Morgenstern copula model. The proposed model allows for the presence of censored data and covariates in the cure parameter. For inferential purpose a Bayesian approach via Markov Chain Monte Carlo (MCMC) is considered. Further, some discussions on the model selection criteria are given. In order to examine outlying and influential observations, we develop a Bayesian case deletion influence diagnostics based on the Kullback-Leibler divergence. The newly developed procedures are illustrated on artificial and real HIV data.

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

1. Introduction

Multivariate survival data are present in various areas. Considering particularly bivariate survival data, we can observe two lifetimes for a same client, patient or equipment. For instance, in the medical area we can have interest in studying the lifetimes of paired 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 damages of dual generators in a power plant or the lifetimes of motors in a twin-engine airplane. In financial applications we may have bivariate lifetimes until default.

In general, multivariate survival data are correlated and the study of that dependence has been focus of many researches. In literature, an extensive list of papers on modeling multivariate survival data can be found, for example Aslanidou

*Corresponding author.

et al. (1998), Gao *et al.* (2006), Hougaard (2000) and Hanagal (2011). Clayton (1978), Vaupel *et al.* (1979), Hougaard (1986), Oakes (1989), Wienke (2011) and Hanagal (2011) consider frailty models in that one or more random effects are included in the model in order to model the dependence between the observations.

More recently, the copula models (see for instance, Embrechts *et al.* (2003), Trivedi and Zimmer (2005), Chiou and Tsay (2008), Nelsen (2006), Zhang *et al.* (2010) and Jaworski (2010)) have become a popular tool to model the dependence of multivariate data, especially in biological areas, actuarial sciences and finances. A copula is a function that connects the marginal distributions to restore the joint distribution. Different copula functions represent different dependence structures between variables (Nelsen, 2006). 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 Hougaard (1989), Oakes (1989), Shih and Louis (1995) and Gustafson *et al.* (2003). Particularly, from a Bayesian perspective, Romeo, Tanaka and Pedroso de Lima (2006) considered an application of the Archimedian copula family for modeling the dependence of bivariate survival. In their analysis, the authors 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. Univariately, models which consider that 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 Boag (1949) and Berkson and Gage (1952). 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 interested readers can refer to Maller and Zhou (1996), Peng *et al.* (1998), Rodrigues *et al.* (2009), Mazucheli *et al.* (2009) and Perdona and Louzada-Neto (2011) amongst others.

In this paper, considering a parallel path to the work of Romeo, Tanaka and Pedroso de Lima (2006), but differently from them, for the first time, we consider

a bivariate distribution based on the Farlie-Gumbel-Morgenstern (FGM) copula (see, Conway (1983)) with long term survival models as marginal distributions.

The main objective of the paper is to present the FGM long-term bivariate survival copula model and to develop diagnostic measures from a Bayesian perspective based on the Kullback-Leibler (K-L) divergence as proposed by Peng and Dey (1995). For each individual lifetime variable we consider long-term Weibull mixture models with covariates included in the cure proportion. For inferential purpose joint estimation was performed and a sampling based approach via Markov Chain Monte Carlo (MCMC) was considered. Applications on artificial and on a real HIV data illustrate our approach.

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 illustrates our approach on an artificial dataset generated according to the proposed bivariate model and on a real HIV dataset, where we also compare the fitting of our model with usual copula models, namely, the Positive Stable, Frank and Clayton ones, as well as with the independence case. Some final remarks in Section 6 ends the paper.

2. The Bivariate Long-Term Survival Model Based on the FGM Copula

In order to define a copula we first suppose that C_ϕ is a joint survival 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} denote, 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 by

$$S_{pop}(t_1, t_2) = C_\phi(S_{pop1}(t_1), S_{pop2}(t_2)), \quad t_1, t_2 > 0, \quad (1)$$

and

$$f_{pop}(t_1, t_2) = c_\phi(S_{pop1}(t_1), S_{pop2}(t_2))f_{pop1}(t_1)f_{pop2}(t_2), \quad t_1, t_2 > 0, \quad (2)$$

respectively. Notice that the marginal distributions and the dependence structure can be visualized separately and this dependence structure is represented by a copula.

The marginal distributions of T_j , $j = 1, 2$, were assumed to be driven by a mixture model (Maller and Zhou, 1996). The idea is to enable modeling a proportion of long-term survivors. Mixture models for long-term survivors have been widely used for fitting data where some individuals may never suffer the

cause of failure under study (Maller and Zhou, 1996). In this type of modeling, it is assumed that, due to some unobserved prognostic factors, a certain fraction p_j 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

$$S_{popj}(t_j) = p_j + (1 - p_j) S_0(t_j), \quad (3)$$

$j = 1, 2$, where $S_0(t_j)$ denotes the survival function for the non-cured group in the population. 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 $S_0(t_j)$ are the Gompertz, Exponential and Weibull distributions. Yamaguchi (1992) considered the generalized log-gamma distribution for the cure rate in the context of accelerated failure-time regression models. Peng, Dear and Denham (1998) proposed a generalized F mixture model for the cure rate, which includes the most popular survival models as particular cases. Perdoná and Louzada-Neto (2011) proposed a general failure rate model which accommodates comprehensive families of cure rate models as particular cases.

Following Conway (1983), considering the Farlie-Gumbel-Morgenstern distribution, hereafter FGM copula, we have the copula distribution function given as

$$C_\phi(u, v) = uv[1 + \phi(1 - u)(1 - v)], \quad (4)$$

where $0 \leq u, v \leq 1$ and $-1 \leq \phi \leq 1$. For $\phi > 0$ we have evidence for positive dependence between u and v . For $\phi < 0$ we have evidence for negative dependence between u and v .

Consider (T_1, T_2) comes from the FMG copula (4) then the joint long-term survival of (T_1, T_2) is given by

$$S_{pop}(t_1, t_2) = S_{pop_1}(t_1)S_{pop_2}(t_2)(1 + \phi(1 - S_{pop_1}(t_1))(1 - S_{pop_2}(t_2))), \quad (5)$$

where ϕ parameter measures the intensity of the dependence between the lifetimes. Observe that when $\phi = 0$, $S_{pop}(t_1, t_2) = S_{pop_1}(t_1)S_{pop_2}(t_2)$, leading to the conclusion that the random variables T_1 and T_2 are independent.

3. Inference

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

Let (T_{i1}, T_{i2}) denote the i -th bivariate lifetime and (C_{i1}, C_{i2}) the censored bivariate lifetime, 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 $S_{pop}(t_1|\gamma_1)$ and $S_{pop}(t_2|\gamma_2)$ be the survival functions of T_{i1} and T_{i2} , respectively, where γ_1 and γ_2 are parameter vectors of q_1 and q_2 elements associated to each one of the marginal distributions.

Considering the bivariate survival function $S_{pop}(t_1, t_2|\phi, \gamma_1, \gamma_2)$ given in (5), the contribution of the i -th individual for the log-likelihood of $\theta = (\phi, \gamma_1, \gamma_2)$ is given by

$$\begin{aligned} \ell_i(\theta) = & \delta_{i1}\delta_{i2} \log \left(\frac{\partial^2 S_{pop}(t_1, t_2|\theta)}{\partial t_{i1} \partial t_{i2}} \right) + \delta_{i1}(1 - \delta_{i2}) \log \left(\frac{-\partial S_{pop}(t_1, t_2|\theta)}{\partial t_{i1}} \right) \\ & + \delta_{i2}(1 - \delta_{i1}) \log \left(\frac{-\partial S_{pop}(t_1, t_2|\theta)}{\partial t_{i2}} \right) + (1 - \delta_{i1})(1 - \delta_{i2}) \log S_{pop}(t_1, t_2|\theta). \end{aligned} \quad (6)$$

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 (5), 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, γ_1 and γ_2 , respectively. In order to guarantee proper posteriors, we adopt proper priors with known hyper-parameters. Thus, we consider a prior distribution of $\theta = (\phi, \gamma_1, \gamma_2)$ given by

$$\pi(\theta) \propto (1 - \phi)^{r_1 - 1} (1 + \phi)^{r_2 - 1} \prod_{j=1}^2 \pi(\gamma_j). \quad (7)$$

(7) implies that $(1 - \phi)/2$ follows a $Beta(r_1, r_2)$ distribution, $\pi(\gamma_j)$ is the prior distribution of γ_j and that γ_1 , γ_2 and ϕ are mutually independent. Besides, $\phi \in (-1, 1)$.

Combining (7) with the likelihood function, $L(\theta) = \exp(\sum_{i=1}^n \ell_i(\theta))$, where $\ell_i(\theta)$ is given in (6), we straightforwardly obtain the joint posterior distribution of θ , $\pi(\theta|\mathcal{D})$, where \mathcal{D} is the observed dataset.

For parametric specification of the marginals it is possible to introduce covariates in each one of the components γ_j , $j = 1, 2$. For simplicity and to avoid restrictions in the parametric space, we consider $p_j = \exp(\mathbf{x}_j^\top \beta_j) / (1 + \exp(\mathbf{x}_j^\top \beta_j))$, where β_j is corresponding to the vector of unknown coefficients of order $m_j \times 1$, associated to the covariates \mathbf{x}_j , $j = 1, 2$.

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 (Gilks, Richardson and Spiegelhater, 1996). 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 Gelfand and Smith (1990) and Gamerman and Lopes (2006). 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.

We assume T_j with a Weibull mixture distribution which survival distribution is given by (3) with parameters α_j and λ_j , and $p_j = \exp(\beta_{0j} + \beta_{1j}\mathbf{x}) / (1 + \exp(\beta_{0j} + \beta_{1j}\mathbf{x}))$, $j = 1, 2$. We choose the following independent prior distributions: $\beta_{kj} \sim N(\mu_{kj}, \sigma_{kj}^2)$, $\lambda_j \sim \text{Gamma}(a_j, b_j)$ and $\alpha_j \sim \text{Gamma}(c_j, d_j)$, where $k = 0, 1$ and $j = 1, 2$.

Combining the likelihood function (6) and the prior distribution (7), 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. The full conditional posterior densities for each parameter is given by

$$\begin{aligned} \pi(\phi | \mathcal{D}, \theta_{(-\phi)}) &\propto (1 - \phi)^{r_1 - 1} (1 + \phi)^{r_2 - 1} \left(\prod_{i=1}^n \Delta_i \right), \\ \pi(\beta_{0j} | \mathcal{D}, \theta_{(-\beta_{0j})}) &\propto \left(\prod_{i=1}^n f_{pop_j}^{\delta_i} \left(1 + e^{\beta_{0j} + \beta_{1j}x_i} \right)^{-(1-\delta_{ji})} \Delta_i \right) e^{\frac{-1}{2\sigma_{0j}^2} (\beta_{0j}^2 - 2\mu_{0j}\beta_{0j})}, \\ \pi(\beta_{1j} | \mathcal{D}, \theta_{(-\beta_{1j})}) &\propto \left(\prod_{i=1}^n f_{pop_j}^{\delta_i} \left(1 + e^{\beta_{0j} + \beta_{1j}x_i} \right)^{-(1-\delta_{ji})} \Delta_i \right) e^{\frac{-1}{2\sigma_{1j}^2} (\beta_{1j}^2 - 2\mu_{1j}\beta_{1j})}, \\ \pi(\lambda_j | \mathcal{D}, \theta_{(-\lambda_j)}) &\propto \left(\prod_{i=1}^n f_{pop_j}^{\delta_i} \left(1 - p_{ji} \right)^{(1-\delta_{ji})} \Delta_i \right) \lambda_j^{n+a_j - (1 + \sum_{i=1}^n \delta_{ji})} \\ &\quad \times e^{-\lambda_j \left[b_j + \sum_{i=1}^n t_{ji}^{\alpha_j} (1 - \delta_{ji}) \right]}, \end{aligned}$$

and

$$\pi(\alpha_j | \mathcal{D}, \theta_{(-\alpha_j)}) \propto \left(\prod_{i=1}^n f_{popj}^{\delta_i} (1 - p_{ji})^{(1-\delta_{ji})} \Delta_i t_{ji}^{(\alpha_j-1)(1-\delta_{ji})} \right) \alpha_j^{c_j-1} \times e^{-\left[d_j \alpha_j + \sum_{i=1}^n \lambda_j t_{ji}^{\alpha_j} (1-\delta_{ji}) \right]}$$

for $j = 1, 2$ with $\mathcal{D} = (\mathbf{t}_1, \mathbf{t}_2, \mathbf{d}_1, \mathbf{d}_2, \mathbf{x})$, $f_{popj} = -S'_{popj}(t_{ji})$,

$$\Delta_i = \nu_{1i}^{(1-\delta_1)(1-\delta_2)} \nu_{2i}^{\delta_1 \delta_2} \nu_{3i}^{\delta_1(1-\delta_2)} \nu_{4i}^{\delta_2(1-\delta_1)},$$

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

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

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

and

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

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 (Chib and Greenberg, 1995). The simulations were performed using the OpenBUGS software (Spiegelhalter *et al.*, 2007). 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 Spiegelhalter *et al.* (2002), the expected Akaike information criterion (EAIC) by Brooks (2002), and the expected Bayesian (or Schwarz) information criterion (EBIC) by Carlin and Louis (2001) 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 a likelihood function given by

$$g(t_{1i}, t_{2i}|\theta) = \begin{cases} f_{pop1}(t_{1i})f_{pop2}(t_{2i})(1 + \phi(1 - 2S_{pop1}(t_{1i}))(1 - 2S_{pop2}(t_{2i}))), & \text{if } \delta_{1i} = 1 \text{ and } \delta_{2i} = 1, \\ f_{pop1}(t_{1i})S_{pop2}(t_{2i})(1 + \phi(1 - 2S_{pop1}(t_{1i}))(1 - S_{pop2}(t_{2i}))), & \text{if } \delta_{1i} = 1 \text{ and } \delta_{2i} = 0, \\ S_{pop1}(t_{1i})f_{pop2}(t_{2i})(1 + \phi(1 - S_{pop1}(t_{1i}))(1 - 2S_{pop2}(t_{2i}))), & \text{if } \delta_{1i} = 0 \text{ and } \delta_{2i} = 1, \\ S_{pop1}(t_{1i})S_{pop2}(t_{2i})(1 + \phi(1 - S_{pop1}(t_{1i}))(1 - S_{pop2}(t_{2i}))), & \text{if } \delta_{1i} = 0 \text{ and } \delta_{2i} = 0, \end{cases}$$

for $i = 1, \dots, n$, where $f_{popj}(\cdot)$ and $S_{popj}(\cdot)$, $j = 1, 2$, are Weibull mixture density function and Weibull mixture survival function, respectively. The EAIC, EBIC and DIC criteria can be calculated using the MCMC output by means of $\widehat{\text{EAIC}} = Dbar + 2q$, $\widehat{\text{EBIC}} = Dbar + q \log(n)$ and $\widehat{\text{DIC}} = Dbar + \widehat{\rho}_D = 2Dbar - Dhat$, 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

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

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 Gelfand, Dey and Chang (1992). 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} (Chen, Ibrahim and Sinha, 2001) can be written as

$$\text{CPO}_i = \int_{\Theta} g(t_{1i}, t_{2i}|\theta) \pi(\theta|\mathcal{D}^{(-i)}) d\theta, \quad i = 1, \dots, n.$$

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 (Chen, Shao and Ibrahim, 2000) 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 $B = \sum_{i=1}^n \log(\widehat{CPO}_i)$. 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 (Cook and Weisberg, 1982), in which the effects are studied instead of completely removing cases from the analysis. This reasoning shall form the basis for our Bayesian global influence methodology. Then, 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,

$$K(P, P_{(-i)}) = \int \pi(\theta|\mathcal{D}) \log \left[\frac{\pi(\theta|\mathcal{D})}{\pi(\theta|\mathcal{D}^{(-i)})} \right] d\theta. \quad (8)$$

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 Peng and Dey (1995) and Weiss (1996), it may be difficult for a practitioner to judge the cutoff point of the divergence measure so as to determine whether a small subset of observations is influential or not. In this context, we will use the proposal given by Peng and Dey (1995) and Weiss (1996) by considering the following. Consider a biased coin, which has success probability p . Then the divergence between the biased and an unbiased coin is

$$K(P, P_{(-i)}) = K(f_0, f_1) = \int -\log \left(\frac{f_0(x)}{f_1(x)} \right) f_1(x) dx, \quad (9)$$

where $f_0(x) = p^x(1-p)^{1-x}$ and $f_1(x) = 0.5$, $x = 0, 1$. Now if $K(f_0, f_1) = d_{K-L}(p)$ then it can be easily checked that d_{K-L} , satisfies the following equation

$$d_{K-L}(p) = -\frac{\log[4p(1-p)]}{2}. \quad (10)$$

It is not difficult to see that d_{K-L} increases as p moves away from 0.5. In addition, $d_{K-L}(p)$ is symmetric about $p = 0.5$ and d_{K-L} achieves its minimum at $p = 0.5$. In this point, $d_{K-L}(0.5) = 0$ and $f_0 = f_1$. Therefore, if we consider $p > 0.80$ (or $p \leq 0.20$) as a strong bias in a coin, then $d_{K-L}(0.80) = 0.223$. Then, this equation implies that i -th case is considered influential when $K(P, P_{(-i)}) > 0.223$.

Similarly, the calibration value of $K(P, P_{(-i)})$ for the i -th case can be obtained by solving the equation $K(P, P_{(-i)}) = -\log[4p_i(1-p_i)]/2$ for p_i . After some algebraic manipulation it can be shown that $p_i = \{1 + \sqrt{1 - \exp[-2K(P, P_{(-i)})]}\}/2$. This equation implies that $0.5 \leq p_i \leq 1$. That is, if $p_i > 0.8$ then the i -th case is considered influential.

For our case it can be shown that (8) can be expressed as a posterior expectation

$$\begin{aligned} K(P, P_{(-i)}) &= \log E_{\theta|\mathcal{D}}\{[g(t_{1i}, t_{2i}|\theta)]^{-1}\} + E_{\theta|\mathcal{D}}\{\log[g(t_{1i}, t_{2i}|\theta)]\} \\ &= -\log(\text{CPO}_i) + E_{\theta|\mathcal{D}}\{\log[g(t_{1i}, t_{2i}|\theta)]\}, \end{aligned} \quad (11)$$

where $E_{\theta|\mathcal{D}}(\cdot)$ denotes the expectation with respect to the joint posterior $\pi(\theta|\mathcal{D})$. Thus (11) 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

$$K(\widehat{P}, \widehat{P}_{(-i)}) = -\log(\widehat{\text{CPO}}_i) + \frac{1}{Q} \sum_{q=1}^Q \log[g(t_{1i}, t_{2i}|\theta_q)]. \quad (12)$$

5. Application

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 this section we consider an artificial sample generated according to (5), assuming that the observed individual lifetime T_j has a Weibull distribution with parameters α_j and λ_j with probability density function given by

$$f(t_j) = \alpha_j \lambda_j t_j^{\alpha_j - 1} \exp\{-\lambda_j t_j^{\alpha_j}\}, \quad j = 1, 2. \quad (13)$$

The artificial bivariate data (T_{i1}, T_{i2}) , $i = 1, \dots, n$, was generated assuming $n = 160$ according to the following steps: First we generated

$$T_{i1} = (-\log(1 - u_{i1})/\lambda_1)^{1/\alpha_1},$$

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 $T_{i2} = (-\log(1 - w_i)/\lambda_2)^{1/\alpha_2}$. The covariates x_i were generated from a Bernoulli distribution with parameter 0.5. For each lifetime t_j , for each covariate level, \mathbf{x} (0 or 1), we have a proportion of long-term survivals. We fixed the p_{kj} values, representing the cure rate for the level k for the lifetime j , for $k = 0, 1$ and $j = 1, 2$. For each covariate level from each generated lifetime t_j we fixed a cutoff point t_{kj}^* , which was chosen so that the higher times represent the amount of $p_{kj}\%$ of long-term survivals. The following values were considered: $p_{01} = 0.5$, $p_{11} = 0.1$, $p_{02} = 0.3$, $p_{12} = 0.2$, $\alpha_1 = 2$, $\alpha_2 = 1.5$, $\lambda_1 = 0.5$, $\lambda_2 = 0.1$ and $\phi = 0.6$.

For sake of illustration we also compare our FGM bivariate copula model with some bivariate survival models induced by some well known copula models, namely, Positive Stable Frailty (PSF), Clayton and Frank copulas (Hougaard (1986), Clayton (1978) and Frank (1979)). The joint survival function of the PSF model is given by (Hougaard, 1986),

$$S_{pop}(t_1, t_2) = ((-\ln(S_{pop1}(t_1)))^{\frac{1}{\phi}} + (-\ln(S_{pop2}(t_2)))^{\frac{1}{\phi}})^{\phi}, \quad \phi \in (0, 1]. \quad (14)$$

When $\phi \rightarrow 1$, we obtain $S_{pop}(t_1, t_2) = S_{pop1}(t_1)S_{pop2}(t_2)$.

The joint survival function induced by the Clayton copula is given by,

$$S_{pop}(t_1, t_2) = \left(S_{pop1}(t_1)^{-\phi} + S_{pop2}(t_2)^{-\phi} - 1 \right)^{-1/\phi}, \quad \phi \in \mathfrak{R}^+. \quad (15)$$

When $\phi \rightarrow 0$, we obtain $S_{pop}(t_1, t_2) = S_{pop1}(t_1)S_{pop2}(t_2)$. While, the joint survival function induced by the Frank copula is given by,

$$S_{pop}(t_1, t_2) = \log_{\phi} \left(1 + \frac{(\phi^{S_{pop1}(t_1)} - 1)(\phi^{S_{pop2}(t_2)} - 1)}{\phi - 1} \right), \quad \phi \in (0, 1). \quad (16)$$

When $\phi \rightarrow 1$, we obtain $S_{pop}(t_1, t_2) = S_{pop1}(t_1)S_{pop2}(t_2)$. For more details on the PSF, Clayton and Frank copulas interested readers can refer to Nelsen (2006).

The following independent priors were considered to perform the Gibbs sampler: $\alpha_j \sim \text{Gamma}(1, 0.001)$, $\lambda_j \sim \text{Gamma}(1, 0.001)$, $\beta_{kj} \sim N(0, 10^3)$, $k = 0, 1$, $j = 1, 2$. For parameter of PSF model we assume $\phi \sim \text{Beta}(1, 1)$, for Clayton copula $\phi \sim \text{Gamma}(1, 0.001)$, for Frank copula $\phi \sim \text{Beta}(1, 1)$ and for FGM copula that $\frac{1}{2}(1 - \phi) \sim \text{Beta}(1, 1)$, such choices guarantee that $\phi \in (-1, 1)$ and ensure non-informativeness. 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 simulations were performed using the OpenBUGS software (Spiegelhalter *et al.*, 2007). The Bayes estimates were based on posterior samples recorded every 20-th iteration from 50,000 Gibbs samples after a burn-in of 10,000 samples. The MCMC convergence was monitored according to the methods recommended by Cowless and Carlin (1996) (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, PSF, Frank and Clayton models. The FGM model outperforms is better when compared to the other models 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	2.087	0.178	(1.745; 2.457)
	λ_1	0.674	0.078	(0.528; 0.837)
	β_{01}	-0.427	0.301	(-1.055; 0.115)
	β_{11}	-2.056	0.550	(-3.161; -1.031)
Time 2	α_2	1.716	0.163	(1.406; 2.044)
	λ_2	0.108	0.024	(0.067; 0.159)
	β_{02}	-1.133	0.347	(-1.86; -0.538)
	β_{12}	-0.563	0.484	(-1.541; 0.362)
Copula	ϕ	0.668	0.204	(0.230; 0.977)

Table 2: Simulated data. Bayesian criteria

Model	Bayesian criteria			
	EAIC	EBIC	DIC	B
FGM	939.650	967.326	930.400	-465.243
PSF	947.184	974.860	937.300	-469.120
Frank	941.447	969.124	932.400	-466.336
Clayton	940.140	967.8161	931.100	-465.642

In Table 1, the 95% HPD intervals for ϕ are large. In order to verify what conditions are needed to get narrower intervals we performed a simulation study in which we vary the ϕ parameter values and the sample size n , with $\phi = 0.2, 0.6, 0.8$ and $n = 100, 200, 300, 400, 500$. The other parameters were fixed the same values as in Subsection 5.1: $p_{01} = 0.5$, $p_{11} = 0.1$, $p_{02} = 0.3$, $p_{12} = 0.2$, $\alpha_1 = 2$, $\alpha_2 = 1.5$,

$\lambda_1 = 0.5$ and $\lambda_2 = 0.1$. For each set up 1,000 generated samples were considered, from which we fitted our modeling and retrieve the amplitude of the 95% HPD interval for ϕ . The results of this study were condensed in Table 3, where we can observe that the amplitude 95% HPD interval becomes narrower with the sample size. This fact is more evident for large values of ϕ .

Table 3: Amplitude of the 95% HPD interval for parameter ϕ

n	ϕ		
	0.2	0.6	0.8
100	0.789	0.725	0.739
200	0.717	0.663	0.675
300	0.637	0.567	0.583
400	0.521	0.465	0.449
500	0.482	0.438	0.416
600	0.363	0.383	0.278

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 use the same sample previously simulated. We selected cases 5, 35, 95 (both observed lifetimes) and 142 (lifetime 1 is observed and lifetime 2 is censored) 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 $\tilde{t}_i = t_i + 5S_t, i = 1, 2$, where S_t is the standard deviations of the t_i 's. For case 5 we perturbed only the lifetime t_1 and for case 35, the lifetime t_2 , and for cases 95 and 142, 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 Cowless and Carlim (1996). Table 4 shows that the posterior inferences are sensitive to the perturbation of the selected case(s). In Table 4, Dataset (a) denotes the original simulated data set with no perturbation, and Datasets (b) to (g) denote datasets with perturbed cases.

Table 5 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.

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 6 show, before perturbation (Dataset (a)), that all the selected cases

Table 4: Simulated data. Posterior means and standard deviations (SD) for the FGM bivariate survival model parameters according to different perturbation schemes

Data names	Perturbed case	Parameters								
		α_1	λ_1	β_{01}	β_{11}	α_2	λ_2	β_{02}	β_{21}	ϕ
		Mean SD	Mean SD	Mean SD	Mean SD	Mean SD	Mean SD	Mean SD	Mean SD	Mean SD
a	None	2.087 0.178	0.674 0.078	-0.427 0.301	-2.056 0.550	1.716 0.163	0.108 0.024	-1.133 0.347	-0.563 0.484	0.668 0.204
b	5	1.815 0.156	0.606 0.081	-0.684 0.397	-2.475 1.202	1.713 0.165	0.108 0.024	-1.139 0.383	-0.546 0.487	0.648 0.216
c	35	2.084 0.186	0.666 0.082	-0.451 0.298	-2.049 0.637	1.402 0.124	0.131 0.025	-1.567 0.579	-1.000 1.084	0.664 0.199
d	95	1.761 0.143	0.589 0.078	-0.755 0.413	-2.602 1.177	1.361 0.121	0.131 0.024	-1.876 0.899	-1.108 1.395	0.714 0.194
e	{5, 35}	1.827 0.154	0.604 0.077	-0.662 0.355	-2.287 0.819	1.384 0.128	0.130 0.025	-1.753 0.719	-1.030 1.303	0.624 0.219
f	{35,95}	1.753 0.143	0.586 0.080	-0.744 0.399	-2.728 1.238	1.257 0.098	0.134 0.025	-2.409 1.204	-1.611 1.666	0.667 0.202
g	{5,35,95}	1.657 0.127	0.545 0.075	-0.916 0.496	-3.237 1.514	1.258 0.102	0.132 0.024	-2.615 1.302	-1.549 1.777	0.628 0.219
h	142	1.710 0.135	0.600 0.078	-0.719 0.378	-2.622 1.250	1.738 0.155	0.105 0.023	-1.112 0.316	-0.503 0.441	0.690 0.206
i	{95,142}	1.599 0.117	0.540 0.072	-1.012 0.521	-3.453 1.495	1.408 0.115	0.132 0.025	-1.538 0.530	-0.566 0.688	0.722 0.192

Table 5: 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	939.650	967.326	930.400	-465.243
b	955.960	983.637	946.500	-474.896
c	958.307	985.983	947.800	-476.402
d	974.706	1002.382	962.600	-485.541
e	974.323	1002.000	963.000	-485.467
f	984.981	1012.658	973.000	-489.721
g	994.623	1022.299	982.600	-494.208
h	958.374	986.051	948.200	-475.905
i	986.177	1013.854	975.700	-491.553

are not influential with small $K(P, P_{(-i)})$ and related calibration close to 0.622. However, after perturbations (Datasets (b) to (g)), the $K(P, P_{(-i)})$ increases and the corresponding calibrations become larger than 0.80, indicating those cases are influential.

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.

Table 6: Simulated data. Case influence diagnostics

Data names	Case number	$K(P, P_{(-i)})$	Calibration
a	5	0.013	0.581
	35	0.031	0.622
	95	0.021	0.602
	142	0.030	0.621
b	5	1.629	0.990
c	35	2.324	0.998
d	95	3.759	1.000
e	5	1.514	0.988
	35	2.162	0.997
f	35	0.526	0.903
	95	2.551	0.998
g	5	0.439	0.882
	35	0.593	0.917
	95	1.799	0.993
h	142	2.006	0.995
i	95	3.465	1.000
	142	0.845	0.951

5.3 HIV Dataset

Opportunistic infections are important causes of morbidity and mortality in people with Human immunodeficiency virus (HIV), and deaths and hospitalizations are the major events resulting from these infections. In addition, opportunistic infections are important indicators of the impact of interventions in the HIV-infected population (Candiani *et al.*, 2007). Moreover, HIV-infected persons are subject to numerous infectious and neoplastic complications related to their disease. In general, apart from infections acquired in living outside the hospital when hospitalized, the HIV-infected individual is also more susceptible to hospital infections by his possible poor health condition. Some of these risk factors for nosocomial infections, particularly those of nosocomial bloodstream are the use of invasive devices, antibiotic exposure and length of stay in hospital (Tumbarello *et al.*, 1998).

Overall, the general health condition of the patient admitted to the service as well as the occurrence of an infection acquired during hospitalization, can further extend his length of stay in hospital. In addition, the HIV-infected patient may require one or more hospital admissions and length of stay in this environment can contribute significantly to serious complications, which may lead to death of the patient. Moreover, the timing of a given hospitalization may somehow depend on the time the patient been hospitalized previously. So it seems reasonable to

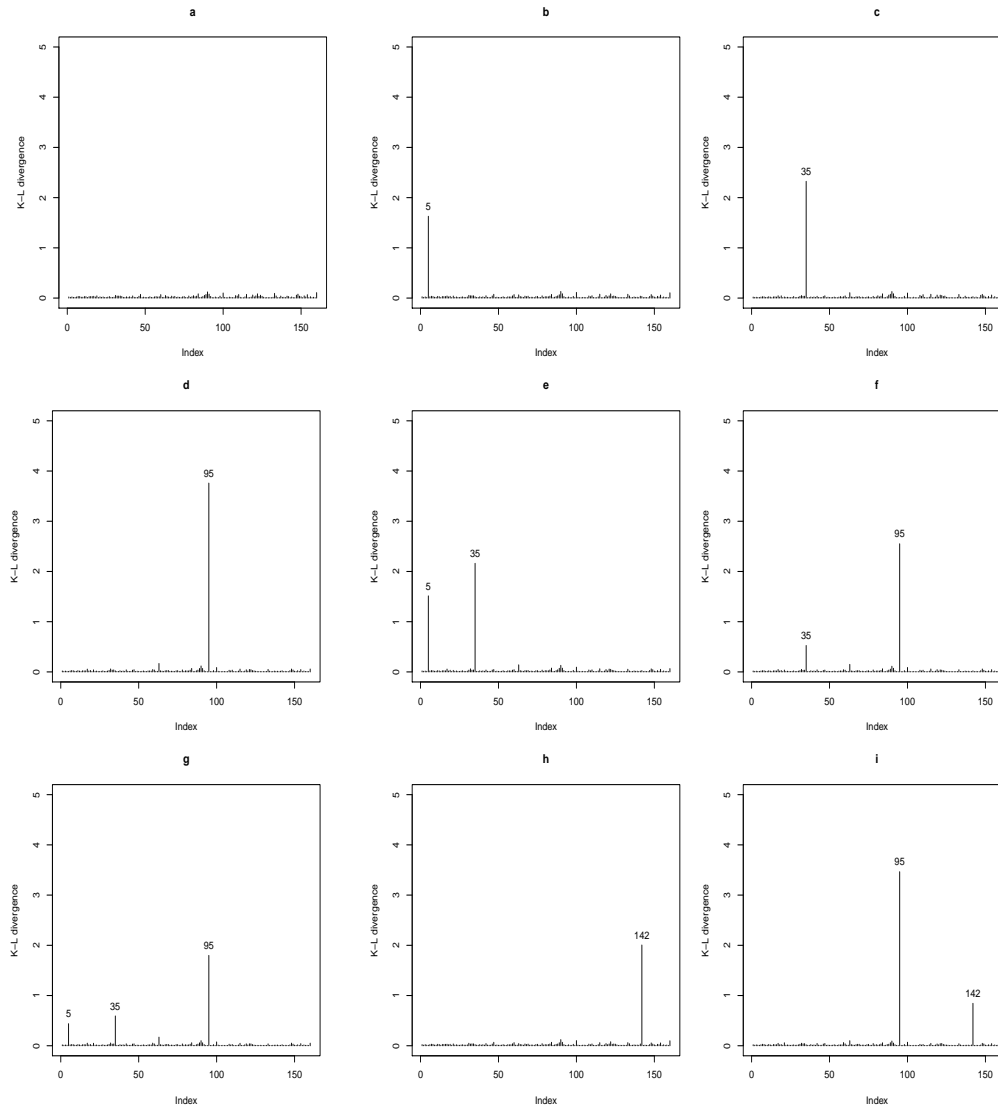


Figure 1: Artificial data. Index plot of $K(P, P_{(-i)})$ from fitting a bivariate survival model based on FGM copula

consider a long-term modeling to assess two times of hospitalization with a copulating dependence structure according to the variable gender, in order to see whether gender impacts the times of hospitalization.

In this context, we consider here 135 patients older than 18 years of age with HIV seen at the Serviço de Doenças Infecciosas e Parasitárias (DIP), Universidade Federal do Triângulo Mineiro (UFTM), Brazil, diagnosed with HIV between January 1996 and December 1998. The times of first (T_1) and second (T_2) hospitalization (in days) were modeled according to the gender of the patient. The

HIV data is an unpublished dataset and it was not statistically analyzed by other works in the best of our knowledge.

We then applied the proposed methodology to the HIV dataset. Figure 2 shows the scatter plot between (T_1) and (T_2) , the index plot of $K(P, P_{(-i)})$ from fitting the bivariate survival model based on FGM copula and the Kaplan-Meier curves for the variables T_1 and T_2 , dichotomized by the gender, along with the fit of the FGM Copula model with Weibull mixture marginals to real data, which will be presented later. Focusing only on the empirical curves, we clearly observe the presence of long-term survivals, since some of the curves stabilize above zero, leading to an indicative of long-term survivals.

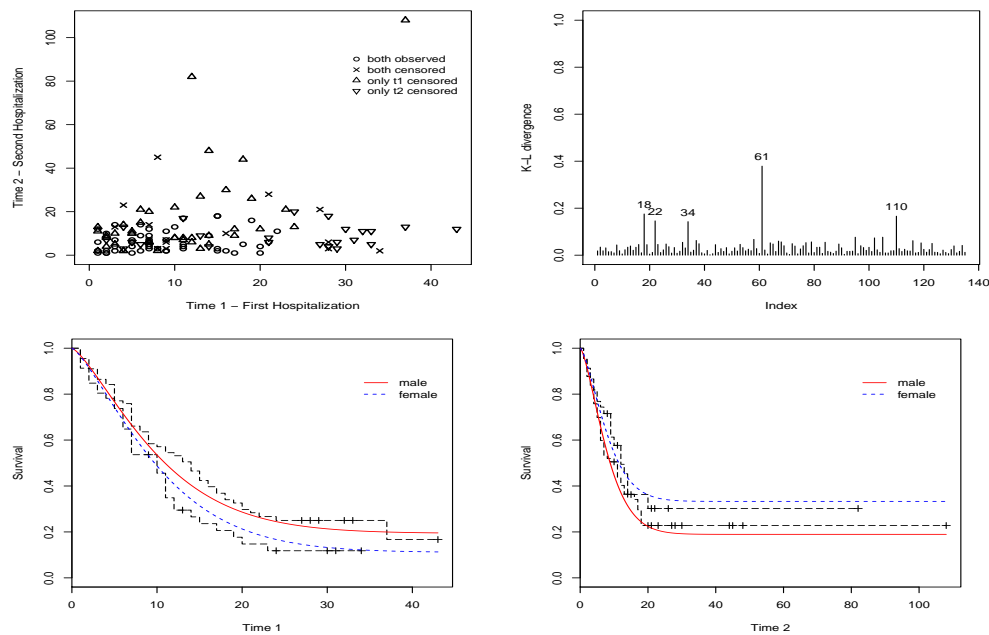


Figure 2: Real data. Scatter plot of T_1 vs T_2 (upper left panel). Index plot of $K(P, P_{(-i)})$ from fitting the bivariate survival model based on FGM copula (upper right panel). Kaplan-Meier curves for the variables T_1 (lower left panel) and T_2 (lower right panel) along with the fit of the FGM Copula model with Weibull mixture marginal

The data was analyzed through our approach by adopting the model in (5) with Weibull mixture marginals as in (13). The Bayes estimates were based on posterior samples recorded every 20-th iteration from 50,000 Gibbs samples after a burn-in of 10,000 samples. We choose the following independent prior distributions: $\alpha_j \sim \text{Gamma}(1, 0.001)$, $\lambda_j \sim \text{Gamma}(1, 0.001)$, $\beta_{kj} \sim N(0, 10^3)$, $k = 0, 1$, $j = 1, 2$ and $(1 - \phi)/2 \sim \text{Beta}(1, 1)$, such choices guarantee that $\phi \in (-1, 1)$ and ensure non-informativeness. To monitor the convergence of the Gibbs

samples we use the methods recommended by Cowless and Carlin (1996).

In Table 7 we report posterior summaries for the parameters of the FGM long-term bivariate survival copula model considering Weibull marginal mixture distribution. Following Gelman, Carlin and Rubin (2006), 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 priors. We also checked the sensitivity of the analysis for the variance component parameters for various choices of prior parameters by changing only on a parameter at a time and keeping all other parameters constant at their default values. The posterior summaries of the parameters do not present remarkable difference and not impair the results in Table 7. The HPD confidence intervals for β_{11} and β_{12} containing zero speaks against the gender impact in the times of hospitalization. These results are corroborated by the similarity between fitted bivariate survival FGM copula model curves under the Kaplan-Meier curves for the variables T_1 and T_2 , dichotomized by the gender presented in Figure 2.

Table 7: HIV data. Summary results from the posterior distribution, mean and standard deviation (SD) for parameter under model based on FGM copula

	Parameter	Mean	SD	HPD (95%)
Time 1	α_1	1.332	0.127	(1.088; 1.569)
	λ_1	0.040	0.012	(0.020; 0.069)
	β_{01}	-2.084	0.972	(-4.615; -0.697)
	β_{11}	0.658	0.970	(-1.138; 2.999)
Time 2	α_2	1.457	0.140	(1.194; 1.747)
	λ_2	0.040	0.013	(0.019; 0.070)
	β_{02}	-0.695	0.448	(-1.692; 0.081)
	β_{12}	-0.759	0.514	(-1.717; 0.301)
Copula	ϕ	0.559	0.259	(0.021; 0.966)

Although we have reanalyzed the data of HIV considering only the parameters that were significant (in Table 7), the amplitude of 95% HPD interval for ϕ showed no significant change. This fact was already expected, since the parameter ϕ is estimated regardless of the choice of the marginal. As pointed out by Romeo, Tanaka and Pedroso de Lima (2006), since the choice of the marginal distributions does not depend on a particular choice of a copula, it makes sense to consider the estimation of the marginals and the dependence parameters separately.

Table 8 presents the K-L divergences and related calibrations for the five observations which present the largest calibration values. Clearly we observe the procedure identifies the lifetime 61 as a possible influential case (corresponding calibration greater than 0.8). The result is corroborated by the $K(P, P_{(-i)})$ for the FGM bivariate survival model shown in upper right panel of Figure 2.

Table 8: Real data. Bayesian case influence diagnostic

	$K(P, P_{(-i)})$	Calibration
18	0.175	0.772
22	0.146	0.752
34	0.143	0.749
61	0.379	0.864
110	0.166	0.766

Besides the proposed FGM bivariate survival model, we also fitted to the data the bivariate survival models induced by a PSF, Clayton and Frank copulas, as presented in Section 5.1. Table 9 presents the model comparison criteria discussed in Section 3 for comparing the FGM long-term bivariate survival copula model with Weibull mixture marginal distribution with its particular independence case, as well as the long-term bivariate survival models induced by the PSF, Frank and Clayton copulas considering Weibull mixture distributions as marginals. Based on all Bayesian criteria, there is positive evidence in favor The FGM modeling, indicating that the FGM long-term bivariate survival copula model can be seen as a competitor to the well known bivariate survival models induced by the PSF, Frank and Clayton copulas commonly used in literature for fitting bivariate life time data.

Table 9: Real data. Bayesian criteria

	EAIC	EBIC	DIC	B
FGM	1319.258	1345.406	1311.000	-655.054
PSF	1319.812	1345.959	1311.000	-655.541
Frank	1322.804	1348.951	1316.000	-657.745
Clayton	1325.070	1351.217	1317.000	-659.217
Independence	1328.210	1351.450	1324.000	-662.220

6. Some Final Remarks

In this paper, we present the FGM long-term bivariate survival copula model. Parameter estimation is based on a Bayesian approach via MCMC. We propose a 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 illustrate our approach with an artificial and a real dataset.

In the two-step approach, the marginal parameters are estimated first 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 in the MCMC algorithm such that the assumption of independence in the first step is avoided. We however considered a two-step estimation method. We omitted the results since they are similar to those obtained by considering the joint estimation approach.

Acknowledgements

The research of Francisco Louzada and Vicente G. Cancho are supported by the Brazilian organization CNPq. The research of Adriano K. Suzuki is supported by the Brazilian organization CAPES. The research of Fatima L. Prince and Gilberto A. Pereira is supported by the Brazilian organization FUNEP-UFMT. We are grateful to the DIP-UFMT, especially to Dalmo Correia Filho and Mario L. S. Vergara, for their support in gathering and reading the records of HIV-infected patients considered in the present study.

References

- Aslanidou, H., Dey, D. K. and Sinha, D. (2008). Bayesian analysis of multivariate survival data using Monte Carlo methods. *Canadian Journal of Statistics* **1**, 33-48.
- Berkson, J. and Gage, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association* **47**, 501-515.
- Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society, Series B* **11**, 15-53.
- Brooks, S. P. (2002). Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde. *Journal of the Royal Statistical Society, Series B* **64**, 616-618.
- Candiani, T. M. S., Pinto, J., Araújo Cardoso, C. A., Carvalho, I. R., Dias, A. C. M., Carneiro, M. and Goulart, E. A. (2007). Impact of highly active antiretroviral therapy (HAART) on the incidence of opportunistic infections, hospitalizations and mortality among children and adolescents living with HIV/AIDS in Belo Horizonte, Minas Gerais State, Brazil. *Caderno de Saúde Pública* **23**, S414-S423.
- Carlin, B. P. and Louis, T. A. (2001). *Bayes and Empirical Bayes Methods for Data Analysis*, 2nd edition. Chapman and Hall, Boca Raton.

- Chen, M. H., Shao, Q. M. and Ibrahim, J. G. (2000). *Monte Carlo Methods in Bayesian Computation*. Springer, New York.
- Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *American Statistician* **49**, 327-335.
- Chiou, S. C. and Tsay, R. S. (2008). A copula-based approach to option pricing and risk assessment. *Journal of Data Science* **6**, 273-301.
- Clayton, D. G. (1978). A model for association in bivariate life-tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141-151.
- Conway, D. A. (1983). Farlie-Gumbel-Morgenstern distributions. In *Encyclopedia of Statistical Sciences* (Edited by Kotz and N. L. Johnson), Volume 3, 28-31. Wiley, New York.
- Cook, R. D. and Weisberg, S. (1982). *Residuals and Influence in Regression*. Chapman and Hall, Boca Raton.
- Cowless, M. K. and Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association* **91**, 883-904.
- Embrechts, P., Linskog, F. and McNeil, A. (2003). Modelling dependence with copulas and applications to risk management. In *Handbook of Heavy Tailed Distributions in Finance* (Edited by S. T. Rachev), 329-384. Elsevier, North-Holland. <http://www.math.ethz.ch/baltes/ftp/papers.html>.
- Frank, M. J. (1979). On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes Mathematicae* **19**, 194-226.
- Gamerman, D. and Lopes, H. F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, 2nd edition. Chapman and Hall, Boca Raton.
- Gao, F., Manatunga, A. K. and Chen, S. (2006). Developing multivariate survival trees with a proportional hazards structure. *Journal of Data Science* **4**, 343-356.
- Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**, 398-409.

- Gelfand, A. E., Dey, D. K. and Chang, H. (1992). Model determination using predictive distributions with implementation via sampling-based methods (with discussion). In *Bayesian Statistics* (Edited by Kotz and N. L. Johnson), Volume 4, 147-167. Oxford University Press, New York.
- Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (2006). *Bayesian Data Analysis*. Chapman and Hall, New York.
- Gilks, W. R., Richardson, S. and Spiegelhater, D. J. (1996). *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London.
- Gustafson, P., Aeschliman, D. and Levy, A. R. (2003). A simple approach to fitting Bayesian survival models. *Lifetime Data Analysis* **9**, 5-19.
- Hanagal, D. (2011). *Modeling Survival Data Using Frailty Models*. Chapman and Hall, Boca Raton.
- Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika* **73**, 671-678.
- Hougaard, P. (1989). Fitting a multivariate failure time distribution. *IEEE Transactions on Reliability* **38**, 444-448.
- Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. Springer, Heidelberg.
- Ibrahim, J. G., Chen, M. H. and Sinha, D. (2001). *Bayesian Survival Analysis*. Springer, New York.
- Jaworski, P. (2010). Copula theory and its applications, in *Proceedings of the Workshop Held in Warsaw, 25-26 September 2009 Volume 198 de Lecture Notes in Statistics* (Edited by P. Jaworski, F. Durante, W. Härdle and T. Rychlik). Springer, Dordrecht.
- Maller R. A. and Zhou X. (1996). *Survival Analysis with Long-Term Survivors*. Wiley, New York.
- Manatunga, A. K. and Oakes, D. (1999). Parametric analysis of matched pair survival data. *Lifetime Data Analysis* **5**, 371-387.
- Mazucheli, J., Achcar, J. A., Barros, E. A. C. and Louzada, F. (2009). Infant mortality model for lifetime data. *Journal of Applied Statistics* **36**, 1029-1036.
- Nelsen, R. B. (2006). *An Introduction to Copulas*, 2nd edition. Springer, New York.

- Romeo, J. S., Tanaka, N. I. and Pedroso de Lima, A. C. (2006). Bivariate survival modeling: a Bayesian approach based on copulas. *Lifetime Data Analysis* **12**, 205-222.
- Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association* **84**, 487-493.
- Paul, S., Gilbert, H. M., Ziecheck, W., Jacobs, J. and Sepkowitz, K. A. (1999). The impact of potent antiretroviral therapy on the characteristics of hospitalized patients with HIV infection. *AIDS* **13**, 415-418.
- Peng, F. and Dey, D. K. (1995). Bayesian analysis of outlier problems using divergence measures. *Canadian Journal of Statistics* **23**, 199-213.
- Peng, Y., Dear, K. B. G. and Denham, J. W. (1998). A generalized F mixture model for cure rate estimation. *Statistics in Medicine* **17**, 813-830.
- Perdoná, G. C. and Louzada-Neto, F. (2011). A general hazard model for lifetime data in the presence of cure rate. *Journal of Applied Statistics* **38**, 1395-1405.
- Rodrigues, J., Cancho, V. G., de Castro, M. and Louzada-Neto, F. (2009). On the unification of long-term survival models. *Statistics and Probabilities Letters* **79**, 753-759.
- Sahu, S. K. and Dey, D. K. (2000). A comparison of frailty and other models for bivariate survival data. *Lifetime Data Analysis* **6**, 207-228.
- Shih, J. H. and Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* **51**, 1384-1399.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B* **64**, 583-639.
- Spiegelhalter, D., Thomas, A., Best, N. and Lunn, D. (2007). *OpenBUGS: User Manual, version 3.0.2*. MRC Biostatistics Unit, Cambridge. <http://mathstat.helsinki.fi/openbugs/>.
- Tumbarello, M., Tacconelli, E., de Gaetano-Donati, K., Leone, F., Morace, G., Cauda, R. and Ortona, L. (1998). Nosocomial bloodstream infections in HIV-infected patients: attributable mortality and extension of hospital stay. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology* **19**, 490-497.

- The Diabetic Retinopathy Study Research Group. (1976). Preliminary report on effects of photocoagulation therapy. *American Journal of Ophthalmology* **81**, 383-396.
- Trivedi, P. K. and Zimmer, D. M. (2005). Copula modelling: an introduction for practitioners. *Foundations and Trends in Econometrics* **1**, 1-111.
- Vaupel, J. W., Manton, K. G. and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* **16**, 439-454.
- Weiss, R. (1996). An approach to Bayesian sensitivity analysis. *Journal of the Royal Statistical Society, Series B* **58**, 739-750.
- Wienke, A. (2011). *Frailty Models in Survival Analysis*. Chapman and Hall, Boca Raton.
- Yamaguchi, K. (1992). Accelerated failure-time regression models with a regression model of surviving fraction: an application to the analysis of “permanent employment” in Japan. *Journal of the American Statistical Association* **87**, 284-292.
- Zhang, S., Zhang, Y., Chaloner, K. and Stapleton, J. T. (2010). A copula model for bivariate hybrid censored survival data with application to the MACS study. *Lifetime Data Analysis* **16**, 231-249.

Received November 9, 2011; accepted February 29, 2012.

Francisco Louzada
Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo
Avenida Trabalhador são-carlense 400 - Centro
CEP: 13566-590 - São Carlos-SP, Brazil
louzada@icmc.usp.br

Adriano Kamimura Suzuki
Departamento de Estatística
Universidade Federal de São Carlos
Rodovia Washington Luiz, km 235
CEP: 13565-905 - São Carlos-SP, Brazil
adrianokamimura@gmail.com

Vicente Garibay Cancho
Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo
Avenida Trabalhador são-carlense 400 - Centro
CEP: 13566-590 - São Carlos-SP, Brazil
garibay@icmc.usp.br

Fátima L. Prince
Departamento de Enfermagem em Educação e Saúde Comunitária
Universidade Federal do Triângulo Mineiro
Praça Manoel Terra 330 - Centro
CEP: 38015-050 - Uberaba-MG, Brazil
fatima.prince@yahoo.com.br

Gilberto A. Pereira
Departamento de Enfermagem em Educação e Saúde Comunitária
Universidade Federal do Triângulo Mineiro
Praça Manoel Terra 330 - Centro
CEP: 38015-050 - Uberaba-MG, Brazil
pereira.gilberto@yahoo.com.br