A New Class of Bivariate Sushila Distributions in Presence of Right-Censored and Cure Fraction

Ricardo Puziol de Oliveira\textsuperscript{a}, Marcos Vinicius de Oliveira Peres\textsuperscript{a}, Jorge Alberto Achcar\textsuperscript{a} and Edson Z. Martinez\textsuperscript{a}

\textsuperscript{a}Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil

Abstract. The present study introduces a new bivariate distribution based on the Sushila distribution to model bivariate lifetime data in presence of a cure fraction, right-censored data and covariates. The new bivariate probability distribution was obtained using a methodology used in the reliability theory based on fatal shocks, usually used to build new bivariate models. Additionally, the cure rate was introduced in the model based on a generalization of standard mixture models extensively used for the univariate lifetime case. The inferences of interest for the model parameters are obtained under a Bayesian approach using MCMC (Markov Chain Monte Carlo) simulation methods to generate samples of the joint posterior distribution for all parameters of the model. A simulation study was developed to study the inferential properties of the new methodology. The proposed methodology also was applied to analyze a set of real medical data obtained from a retrospective cohort study that aimed to assess specific clinical conditions that affect the lives of patients with diabetic retinopathy. For the discrimination of the proposed model with other usual models used in the analysis of bivariate survival data, some Bayesian techniques of model discrimination were used and the model validation was verified from usual Cox-Snell residuals, which allowed us to identify the adequacy of the proposed bivariate cure rate model.

1 Introduction

In medical research, many studies are developed to evaluate how some subject-specific factors impact the probability of an exposed patient experiencing the event of interest (such as death, for example). Conversely, researchers may also be interested in understanding which conditions explain the existence of an underlying cluster of right-censored patients. In this case, it would be essential to access the mechanism responsible for characterizing “cured” patients.

Standard survival analysis techniques, as the Cox proportional hazards model (Cox, 1972), provide no straightforward estimation for cure rates, which we acknowledge as the primary motivation to assume mixture or non-mixture cure fraction models. In this way, several approaches have been proposed in the literature to model cure rates on univariate lifetime data (Farewell, 1982; De Angelis et al., 1999; Cancho and Bolfarine, 2001; Price and Manatunga, 2001; Yu et al., 2004; Yin and Ibrahim, 2005; Lambert et al., 2006; Lu, 2010; Othus et al., 2012; Achcar, Coelho-Barros and Mazucheli, 2012). The bivariate case was treated by Wienke, Lichtenstein and Yashin (2003) and Wienke, Locatelli and Yashin (2006); in each case, the authors have introduced model-based structures for the analysis of bivariate time-to-event data cure rates. Besides, a general multivariate cure rate model was derived and well-discussed by Cancho, Dey and Louzada (2016).

Lambert (2007) describes a cure rate model that incorporates the expected (or historical) value for the mortality of each individual, and thus it is possible to estimate cure fractions when some patients die of other causes than the investigated one. Such an approach may be advantageous in practice since the information of cure at the individual level will rarely be
available. Lambert also points out that the relative survival curve usually stays at a plateau level after several years in oncology studies. This plateau effect occurs when the mortality rate of diseased individuals is the same as the expected mortality rate in the general population, or equivalently, the excess mortality rate is equal to zero; that is, there is a cure fraction in the target population.

According to Vahidpour (2016), there are two most usual approaches for univariate cure models: the mixture cure rate models, also known as standard cure rate models (De Angelis et al., 1999; Tsokikov, Ibrahim and Yakovlev, 2003; Lambert et al., 2006), which have been widely used to model survival data in presence of cure rate; and the non-mixture cure rate models which are not so popular (Achcar, Coelho-Barros and Mazucheli, 2012; Vahidpour, 2016). However, the main goal of this study is to work with bivariate lifetime data. In this way, Wienke, Locatelli and Yashin (2006) proposed a mixture approach for bivariate cure rate models where the two lifetimes associated to the same individual are denoted by $T_1$ and $T_2$ with marginal survival functions given, respectively, by,

$$ S_1(t) = \rho_1 + (1 - \rho_1)S_{01}(t) \quad \text{and} \quad S_2(t) = \rho_2 + (1 - \rho_2)S_{02}(t) $$

where $\rho_k$ (proportion of cured individuals) and $S_{0k}(t)$ (survival function for the non-cured individuals) are associated to each lifetime $T_k, k = 1, 2$. Defining the indicators variables,

$$ V_k = \begin{cases} 1 & \text{if the individual in the } k\text{-th event is susceptible} \\ 0 & \text{if the individual in the } k\text{-th event is cured,} \end{cases} $$

where $k = 1, 2$ and $\psi = \text{cov}(V_1, V_2)$ such that $0 \leq \psi \leq \min(\rho_1, \rho_2) - \rho_1 \rho_2$, we have,

i.) $\phi_{11} = P(V_1 = 1, V_2 = 1) = P(V_1 = 1)P(V_2 = 1) + \psi$ which indicates the probability that an individual is susceptible to both events;

ii.) $\phi_{10} = P(V_1 = 1, V_2 = 0) = P(V_1 = 1)P(V_2 = 0) - \psi$ which indicates the probability that an individual is susceptible for the first event but not for the second event;

iii.) $\phi_{01} = P(V_1 = 0, V_2 = 1) = P(V_1 = 0)P(V_2 = 1) - \psi$ which indicates the probability that an individual is susceptible for the second event but not for the first event;

iv.) $\phi_{00} = P(V_1 = 0, V_2 = 0) = P(V_1 = 0)P(V_2 = 0) + \psi$ which indicates the probability that an individual is not susceptible to both events;

Thus, the joint long-term survival function for the bivariate lifetimes $T_1$ and $T_2$ is given by,

$$ S(t_1, t_2) = \phi_{11}S_0(t_1, t_2) + \phi_{10}S_{10}(t_1) + \phi_{01}S_{20}(t_2) + \phi_{00} $$

where $S_0(t_1, t_2)$ is the joint survival function for the susceptible individuals; $S_{10}(t_1)$ is the marginal survival function for $T_1$; $S_{20}(t_2)$ is the marginal survival function for $T_2$, $\phi_{11} + \phi_{10} + \phi_{01} + \phi_{00} = 1$, $\rho_1 = \phi_{00} + \phi_{01}$ and $\rho_2 = \phi_{00} + \phi_{10}$.

The primary goal of this paper is to explore the structure of mixture cure fraction models to analyze long-term lifetimes depending on potential risk factors assuming bivariate lifetime data. Specifically, we aim to introduce a mixture cure fraction model based on the bivariate Sushila distribution derived from fatal shock method introduced by Marshall and Olkin (1967a), whose continuous counterpart is widely used in medical applications nowadays, mainly due to the great flexibility of its hazard rate curves. The Sushila probability density function (pdf) is given by,

$$ f_X(x | \alpha, \theta) = \frac{\theta^2}{\alpha(\theta + 1)} \left(1 + \frac{x}{\alpha}\right)^{\theta - 1} \exp\left\{-\frac{\theta}{\alpha} x\right\}, \quad x \in \mathbb{R}_+, $$

\[\text{(4)}\]
where \( \theta \in \mathbb{R}_+ \) and \( \alpha \in \mathbb{R}_+ \). Shanker et al. (2013) have shown that this model is a two component mixture of an exponential distribution with scale parameter \( \theta/\alpha \) and a gamma distribution having shape parameter equal to 2 and a scale parameter \( \theta/\alpha \) with mixing proportion given by \( \theta/(\theta + 1) \). A comprehensive discussion about the mathematical properties of the Sushila distribution such as moments, hazard function, stochastic orderings, parameter estimation, among others is also presented by Shanker et al. (2013). The corresponding survival function is given by

\[
S_X(x | \alpha, \theta) = \frac{\alpha(\theta + 1) + \theta x}{\alpha(\theta + 1)} \exp\left\{ -\frac{\theta}{\alpha} x \right\}, \quad x \in \mathbb{R}_+.
\]  

(5)

Also, as pointed out by Shanker et al. (2013), when \( \alpha = 1 \) the univariate Sushila distribution has the Lindley distribution (Ghitany, Atieh and Nadarajah, 2008) as a special case which is widely used to analyze lifetime data. This fact leads us to believe that this probability distribution can be very useful in the analysis of lifetime data, especially in the bivariate case.

This paper is organized as follows. In Section 2, we present the Marshall-Olkin method to generate bivariate distributions where the major advantage of this method is introduction a bivariate model with few parameters and a flexible dependence structure that can be given by any probability model. In addition, we present the bivariate Sushila model using the Marshall-Olkin method and its major properties and also introducing a regression structure. We also describe the adopted Bayesian approach, the associated numerical procedures considered for inferential purposes, some Bayesian discrimination criteria and the model validation approach based on the Cox-Snell residuals. In Section 3, we verify the empirical properties of the Bayesian estimators through an intensive Monte Carlo simulation study. In Section 4, we present an application of the proposed methodology to a medical data set. General comments and concluding remarks are addressed in Section 5.

### 2 Materials and Methods

#### 2.1 An Extended Marshall and Olkin Family of Distributions

Marshall and Olkin (1967a) introduced a method to obtain an extended family of distributions using a fatal shock model. Fatal shocks can refer, for example, to damage caused to biological organs by illness or environmental causes of damage acting on a technical system. In particular, this model assumes three independent sources of shocks presented in the environment of a system consisting of two components such that: a shock from source 1 destroys the component 1, it occurs at a random time \( W_1 \); a shock from source 2 destroys the component 2, it occurs at a random time \( W_2 \); a shock from source 3 destroys both components, it occurs at a random time \( W_3 \). In other words:

**Definition.** Suppose that the components of a two-component system fail after receiving an overall fatal shock. Independent Poisson processes \( W_1(t, \delta_1), W_2(t, \delta_2), W_3(t, \delta_3) \), with parameters \( \delta_i, i = 1, 2, 3 \), govern the occurrence of fatal shocks. Then, we have:

- Events in the process \( W_1(t, \delta_1) \) are fatal shocks transmitted to component 1.
- Events in the process \( W_2(t, \delta_2) \) are fatal shocks transmitted to component 2.
- Events in the process \( W_3(t, \delta_3) \) are fatal shocks transmitted equally and independently to both components.
Thus if \( X = \min(W_1, W_3) \) and \( Y = \min(W_2, W_3) \) denote, respectively, the lifetimes of the first and second components, we have that the probability of the system is working until an overall failure (in other words, the joint survival function (sf) function) is given by,

\[
P(X > x, Y > y) = P(\{W_1(x) = 0, W_2(y) = 0, W_3(\max(x, y)) = 0\})
= P(\{\min(W_1, W_3) > x\}, \{\min(W_2, W_3) > y\})
= P(\{W_1 > x, W_3 > x\}, \{W_2 > y, W_3 > y\})
= P(W_1 > x, W_2 > y, W_3 > \max(x, y)).
\tag{6}
\]

Since the random variables \( W_j, (j = 1, 2, 3) \) are mutually independent, we have,

\[
P(X > x, Y > y) = P(W_1 > x)P(W_2 > y)P(W_3 > z),
\tag{7}
\]

where \( z = \max(x, y) \). For this model, the dependence structure of the random variables \( X \) and \( Y \) is related to the common source of shock 3.

### 2.2 Bivariate Sushila Distribution

Let \( X_i, i = 1, 2 \) be two independent random variables having Sushila distributions with parameters \( \alpha_i, \theta_i > 0, i = 1, 2 \). For the latent variable, \( X_3 \), we may assume a one-parameter distribution to make our model more simply, especially to obtain the cross-factorial moment, for example, in an analytical form. Furthermore, assuming a one-parameter distribution implies a simpler likelihood function in terms of parsimony. Thus let us assume a latent random variable \( X_3 \) having a exponential distribution with parameter \( \theta_3 > 0 \) which is the simplest one-parameter continuous distribution. In this way, using the Marshall-Olkin fatal shock method described in section 2.1, the new proposed joint survival function is given by,

\[
S(x, y) = \left[ \frac{\alpha_1 \omega_1 + \theta_1 x}{\alpha_1 \omega_1} \right] \left[ \frac{\alpha_2 \omega_2 + \theta_2 y}{\alpha_2 \omega_2} \right] \exp \left\{-\frac{\theta_1}{\alpha_1} x - \frac{\theta_2}{\alpha_2} y - \theta_3 z \right\}.
\tag{8}
\]

where \( x = \min(x_1, x_3), y = \min(x_2, x_3), z = \max(x, y), x_1 > 0, x_2 > 0, x_3 > 0, w_1 = \theta_1 + 1 \) and \( w_2 = \theta_2 + 1 \). The joint survival function defined by (8) is denoted as a bivariate Sushila (BS) distribution and special cases of its contour plot are illustrated in Figure 1, assuming different values for the parameters of the proposed model.

**Figure 1** Contour plots of the joint survival function for the BS model assuming different parameter values by \( \alpha_1 = \alpha_2 = 1.50, \theta_1 = \theta_2 = 0.50, \) and \( \theta_3 = 0.5 \) (left panel), 1.0 (central panel) and 1.5 (right panel).
2.3 Bivariate Hazard Function

According to Vaidyanathan et al. (2016), a primary limitation of the definition for the traditional hazard function in bivariate case is that this function is defined from \( \mathbb{R}^2 \to \mathbb{R} \), that is, \( h(x, y) \) is not a vector quantity. To overcome this limitation, Johnson and Kotz (1975) defined the bivariate hazard rate function in vector form based on the derivatives of the logarithm of the joint survival function in relation to \( x \) and \( y \). In this way, assuming the BS distribution, the vector components of the joint hazard rate function are given by,

\[
- \frac{\partial \log S(x, y)}{\partial x} = \begin{cases} 
\left( \frac{\alpha_2 \omega_2 + \theta y}{\alpha_1 \omega_1 \omega_2} + \frac{\alpha_2 \omega_1 + \alpha_1 \theta_1 \alpha_2 x + \theta_2 \alpha_1 + \theta_2 x}{\alpha_1 \omega_1 \omega_2} \right) \\
\times \exp \left\{ -\frac{\theta_1}{\alpha_1} x - \frac{\theta_2}{\alpha_2} y - \theta_3 x \right\} & \text{if } x > y \\
\left( \frac{\alpha_2 \omega_2 + \theta y}{\alpha_1 \omega_1 \omega_2} \right) \exp \left\{ -\frac{\theta_1}{\alpha_1} x - \frac{\theta_2}{\alpha_2} y - \theta_3 x \right\} & \text{if } x < y \\
\end{cases}
\]

and

\[
- \frac{\partial \log S(x, y)}{\partial y} = \begin{cases} 
\left( \frac{\alpha_1 \omega_1 + \theta_1 x}{\alpha_2 \omega_2 \omega_1} + \frac{\alpha_2 \omega_1 + \alpha_2 \theta_2 \omega_2 y + \theta_2 \alpha_2 + \theta_2 y}{\alpha_2 \omega_2 \omega_2} \right) \\
\times \exp \left\{ -\frac{\theta_1}{\alpha_1} x - \frac{\theta_2}{\alpha_2} y - \theta_3 y \right\} & \text{if } x < y \\
\left( \frac{\alpha_1 \omega_1 + \theta_1 x}{\alpha_2 \omega_2 \omega_1} \right) \exp \left\{ -\frac{\theta_1}{\alpha_1} x - \frac{\theta_2}{\alpha_2} y - \theta_3 x \right\} & \text{if } x > y \\
0 & \text{if } x = y \\
\end{cases}
\]

2.4 Marginal Distributions

An important highlight of the proposed model is that, even the marginal distributions of the random variables \( X_1 \) and \( X_2 \) are given by generalized Sushila distributions, and they have closed forms for the expected values and variances. Also, the proposed bivariate model inherits most of the properties of the Sushila model in the univariate case, especially for the hazard function that has increasing shape, that is, this bivariate distribution could be very useful, especially to model positive data. The marginal survival functions can be expressed, respectively, by,

\[ S(x) = S_S(x, \theta, \alpha_1)S_E(x, \theta_3), \ i = 1, 2 \]

where \( S_S(\cdot), S_E(\cdot) \) are the survival functions of the Sushila and Exponential distributions, respectively.
The main advantage of the proposed model is that the means and variances for the marginal probability distributions have closed forms. The marginal means and variances, for \( i = 1, 2 \), are given respectively, by,
\[
\mathbb{E}[X_i] = \frac{(\alpha_i \theta_i \theta_3 + \alpha_i \theta_3 + \theta_4^2 + 2\theta_i)\alpha_i}{(\theta_i + 1)(\alpha_i \theta_3 + \theta_4)^2}
\]
and,
\[
\text{Var}[X_i] = \frac{[\theta_i^4 + (2\alpha_i \theta_3 + 4)\theta_i^3 + (\theta_3^2 \alpha_i^2 + 6\alpha_i \theta_3 + 2)\theta_i^2 + (2\alpha_i^2 \theta_3^2 + 4\theta_3 \alpha_i)\theta_i + \theta_3^2 \alpha_i^2] \alpha_i^2}{(\theta_i + 1)^2 (\alpha_i \theta_3 + \theta_4)^4}
\]

According to Balakrishnan and Ristić, 2016; Ristić et al., 2018, to evaluate a bivariate model, it is necessary that the marginal distributions fit to the marginal data sets assuming the bivariate model. Since our model has closed forms for the expected values, we may reparametrize the marginal survival function in terms of the marginal mean to consider a bivariate model. Thus, given the vector \( \mathbf{q}_i \) of covariates (and risk factors) on the probability of a subject being a long-term survivor, we have to specify a monotonic, invertible, and twice differentiable link function, say \( g \), in which \( \theta_i = g^{-1}(\mathbf{x}_i^\top \mathbf{\beta}) \) \((i = 1, \ldots, n)\) is well defined into \( \mathbb{R}_+ \), where \( \mathbf{\beta} = (\beta_0, \beta_1, \ldots, \beta_q) \) is the vector of \( q + 1 \) regression coefficients and \( \mathbf{x}_i^\top = (1, x_{1i}, \ldots, x_{qi}) \) is a vector of covariates that may include, for example, dummy variables, cross-level interactions, and polynomials. For the requested purpose, one may choose any suitable mapping \( g : \mathbb{R} \to \mathbb{R}_+ \). The popular choice in this context is the \( \log \) link function
\[
\log (\theta_i) = \mathbf{x}_i^\top \mathbf{\beta},
\]
but the users can adopt other specifications (e.g., probit, log-log, complementary log-log, among others).

### 2.5 Inference Methods

In many applications related to lifetime data, it is common the presence of censored data, that could be right, left or interval censoring. In this section, let us assume the presence of right censored data, that is, associated to each lifetime \( X_j, j = 1, 2 \), there is a fixed censoring time \( C_j \) and the data are given by \( T_j = \min(X_1, C_1) \) and \( T_2 = \min(X_2, C_2) \). The likelihood function for the parameters of the BS distribution based on a sample of size \( n \) \((i = 1, \ldots, n)\) has the dataset classified in four regions:

- \( R_1 \): Both, \( X_{1i} \) and \( X_{2i} \), are complete observations;
- \( R_2 \): \( X_{1i} \) are complete and \( X_{2i} \) are censored;
- \( R_3 \): \( X_{1i} \) are censored and \( X_{2i} \) are complete;
- \( R_4 \): Both, \( X_{1i} \) and \( X_{2i} \), are censored observations.

Thus, given the vector \( \mathbf{t}_i = (t_{1i}, t_{2i}), i = 1, 2, \ldots, n \) of observed/censored lifetimes, the likelihood function of \( \zeta = (\theta_1, \theta_2, \theta_3, \alpha_1, \alpha_2) \) can be written as
\[
L(\zeta) \propto \prod_{i \in R_1} f(t_{1i}, t_{2i}) \prod_{i \in R_2} \left[ -\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}} \right] \prod_{i \in R_3} \left[ -\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}} \right] \prod_{i \in R_4} S(t_{1i}, t_{2i}). \tag{9}
\]
Another way to express the contribution of the \(i\)-th subject from a random sample \((t_{1i}, t_{2i}, \delta_{1i}, \delta_{2i}), i = 1, \ldots, n\), for the likelihood function is given by

\[
L_i = \left[ \frac{\partial^2 S(t_1, t_2)}{\partial t_{1i} \partial t_{2i}} \right]_{\delta_{1i}=1, \delta_{2i}=1}^{\delta_{1i}=1, \delta_{2i}=0} \times \left[ \frac{-\partial S(t_1, t_2)}{\partial t_{1i}} \right]_{\delta_{1i}=1, \delta_{2i}=1}^{\delta_{1i}=0, \delta_{2i}=1} [S(t_1, t_2)]^{-1}_{\delta_{1i}=1, \delta_{2i}=1}, \tag{10}
\]

where \(\delta_{ji}\) is a censoring indicator variable, that is, \(\delta_{ji} = 1\) for an observed lifetime and \(\delta_{ji} = 0\) for a censored lifetime.

Under a Bayesian framework, one may assume that no expert information is available to justify the choice of informative \textit{priors} for the model parameters. In this context, we specify \textit{prior} distributions such that, even for moderate sample sizes, the information provided by the data should dominate the \textit{prior} information. The non-informative \textit{prior} distributions assumed in this work are given by

\[
\begin{align*}
\phi_{ij} &\sim \text{Beta}(1,1)(i, j = 0, 1), \\
\theta_i &\sim \text{Gamma}(0.01, 0.01)(i = 1, 2, 3), \\
\alpha_i &\sim \text{Gamma}(0.01, 0.01)(i = 1, 2),
\end{align*}
\]

and, for the regression approach,

\[
\beta \sim \text{Normal}_{q+1}(0, 10^2 I_q),
\]

where \(I_{q+1}\) is a identity matrix of dimension \(q+1\). Thus, not assuming the regression structure, a Bayesian approach for the mixture cure rate model based on the BS distribution can be considered by writing the unnormalized joint \textit{posterior} distribution of the vector \(\zeta\) as

\[
\pi(\zeta; t, \delta) \propto \exp \{ \ell(\zeta; t, \delta) \} \prod_{i=1,2,3} \pi(\theta_i) \prod_{i=1,2} \pi(\alpha_i) \prod_{i=0,1} \prod_{j=0,1} \pi(\phi_{ij}). \tag{11}
\]

where \(\ell(\cdot)\) is the logarithm of the likelihood function. Assuming the regression structure, we have,

\[
\pi(\zeta; t, \delta) \propto \exp \{ \ell(\zeta; t, \delta) \} \prod_{i=1,2,3} \pi(\theta_i) \prod_{i=1,2} \pi(\alpha_i) \prod_{i=0,1} \prod_{j=0,1} \pi(\phi_{ij}) \pi(\beta). \tag{12}
\]

From this point of view, inferences for vector \(\zeta\) are entirely based on the marginal \textit{posterior} densities, which can be obtained by integrating Equation (12), apart from the normalizing constant, which we assume to be finite. However, deriving analytical expressions for these densities is infeasible, mainly due to the complexity of the associated log-likelihood function in presence of censored data and covariates. It is important to point out that the use of Bayesian methods for lifetime data analysis assuming censored data and different parametrical models has been considered by many authors using standard numerical methods, Laplace approximation methods and Monte Carlo simulation methods (see for example, Achcar, Broekmeyer and Hunter, 1985; Achcar and Bolfarine, 1986a,b, 1988, 1989; Cancho, Bolfarine and Achcar, 1999; Oliveira et al., 2021). In this study we use Markov Chain Monte Carlo (MCMC) methods (Gelfand and Smith, 1990; Chib and Greenberg, 1995; Achcar and Leandro, 1998) to generate pseudo-random samples from the marginal \textit{posterior} densities. In this work, we have adopted the Metropolis-within-Gibbs (MwG) sampler algorithm.

To apply the proposed Bayesian approach to fit the mixture BS cure rate model, we have adopted the MwG algorithm for MCMC sampling. For each generated sample, a chain with \(N = 200,000\) values was generated for each component of \(\zeta\), considering a burn-in period.
of 5% of the chain’s size. To obtain pseudo-independent samples from the joint posterior distribution (12), one out every 100 generated values was kept, resulting in chains of size 2,000 for each parameter. All computations were performed using the \texttt{R2jags} package (Su and Yajima, 2012), which is available in the \texttt{R} environment (R Core Team, 2015).

2.6 Model comparison criteria

Under a Bayesian approach, a criterion often used to compare among different Bayesian models is the deviance information criterion (DIC) proposed by Spiegelhalter et al. (2002). The DIC criteria is given by

\[ \text{DIC} = D(\hat{\theta}) + 2p_D = 2D(\bar{\theta}) - D(\hat{\theta}), \]

where \( D(\hat{\theta}) \) is the deviance calculated in the posterior mean of the parameter of interest obtained using MCMC simulation methods and \( p_D \) is the effective number of parameters in the model, with \( p_D = D(\hat{\theta}) - D(\bar{\theta}) \), where \( \bar{\theta} \) is the posterior mean of the deviance. Another criteria considered in this study, is the \textit{EAIC} (extended Akaike information criterion) proposed by Brooks (2002). This criteria is defined by

\[ \text{EAIC} = D(\overline{\theta}) + 2k \]

where \( k \) is the number of parameters estimated by the model. We also consider the use of the \textit{EBIC} (extended Bayesian Information criterion), introduced by Carlin and Louis (2000), expressed by

\[ \text{EBIC} = D(\overline{\theta}) + k \ln(n) \]

where \( n \) is the size sample. Smaller values of DIC, EAIC and EBIC indicate better model fit.

In the discrimination of the proposed models, we also use the logarithms of the pseudo marginal likelihood functions (LPML) to choose the best fitted Bayesian model. The LPML is obtained from the conditional predictive ordinates (CPO) (Gelfand, Dey and Chang, 1992). The \( \text{CPO}_i \) for the \( i \)-th observation, \( i = 1, \ldots, n \), is given by

\[ \text{CPO}_i = \int f(D|\theta) f(\theta|D_{[-i]}) \, d\theta, \]

where \( \theta \) is the complete vector of parameter, \( D_{[-i]} \) is the sample without the \( i \)-th observation and \( f(\theta|D_{[-i]}) \) is the posterior density of \( \theta \) given \( D_{[-i]} \). Frequently, \( \text{CPO}_i \) does not have a closed form and is very complicated to calculate. However, according to Dey, Chen and Chang (1997) it is possible to obtain an approximation based on MCMC methods for the \( \text{CPO}_i \), expressed by

\[ \overline{\text{CPO}}_i = \left[ \frac{1}{N} \sum_{n=1}^{N} \frac{1}{f(D_i|\theta_n)} \right]^{-1}, \quad i = 1, \ldots, n, \]

where \( N \) is the number of iterations applied in the implementation of the MCMC procedure after the burn-in period (the burn-in is used to eliminate the effect of the initial values of the iterative algorithm) and where \( \theta_n \) is the vector of the samples obtained at the \( n \)-th iteration (Chen, Shao and Ibrahim, 2012). In this way, a numerical approximation for the LPML criteria is obtained by

\[ \text{LPML} = \sum_{i=1}^{n} \ln \overline{\text{CPO}}_i. \]

According to Geisser and Eddy (1979), the larger values of LPML indicate better fit of the model.
2.7 Model validation

Model validation procedures play an essential role when evaluating the suitability of any fitted model. In general, residual metrics are widely used in such a context, being those measures responsible to indicate departures from the underlying model assumptions by quantifying the data variability that the fitted model did not accommodate. Assessing model suitability using residual metrics is common nowadays; however, deriving appropriate residuals is not always easy for non-normal models typically adopted to describe cure rates in survival studies. In this way, we consider here a popular residual metric proposed by Cox and Snell (Cox and Snell, 1968), which can be straightforwardly used in our context to assess the appropriateness of the proposed model when used in the analysis of real datasets. The Cox-Snell residuals are defined by

$$r_i = - \log[S(t_i)].$$

According to Cox and Snell (1968), if the obtained model fit is adequate, then the Cox-Snell residuals should follow an Exponential distribution with mean equal to one (if $T$ has survival distribution $S(t)$, then $-\log[S(T)] \sim \text{Exp}(1)$). However, if a survival time is right-censored, then the corresponding Cox-Snell residual, say $r_i^+$, is lower than $r_i$, which was obtained from an uncensored observation with the same lifetime. In this way, Crowley and Hu (1977) proposed the modified Cox-Snell residuals, based on the mean and the median of an Exp(1) distribution. These modified residuals were derived by assuming that the difference between the cumulative hazard functions, $H(t_i)$ and $H(t_i^+)$, also follow Exp(1) distributions. Thus, the modified Cox-Snell residuals for censored observations are defined by

$$r_i^+ = 1 - \log[S(t_i)] \quad \text{or} \quad r_i^+ = \log(2) - \log[S(t_i)].$$

In this work, we assume the following residuals for the marginal survival distributions of the proposed model:

$$\kappa_i = \begin{cases} - \log[S(t_i)], & \text{if } t_i \text{ was observed}, \\ 1 - \log[S(t_i)], & \text{if } t_i \text{ was right-censored}. \end{cases}$$

The overall procedure for obtaining to obtain residuals is summarized in the following.

- Step 1: Fit the assumed model;
- Step 2: Estimate the Cox-Snell residuals, $\hat{\kappa}_i$, for each observation;
- Step 3: Estimate the empirical survival function of the Cox-Snell residuals, $\hat{S}_e(\hat{\kappa}_i)$, via the Kaplan and Meier (1958) non-parametric estimator of the survival function;
- Step 4: Compute $-\log[\hat{S}_e(\hat{\kappa}_i)]$;
- Step 5: Plot $\hat{\kappa}_i$ vs. $-\log[\hat{S}_e(\hat{\kappa}_i)]$.

After following these steps, one should evaluate the results carefully. In general, if the displayed dots are close to a straight line with unit slope and intercept equal to zero, then the fitted model can be considered adequate.

3 Simulation Study

The empirical properties of an estimator can be accessed through Monte Carlo simulations. In this way, we have conducted an intensive simulation study aiming to evaluate the performance of the proposed Bayesian approach in some specific situations. The simulation process
was carried out by generating 500 pseudo-random samples of sizes $n = 25, 50, 75$ and 100 with 40% of right-censored observations from the proposed BS-Type I cure rate model derived from Equation (3), with $\alpha_i = 0.2, i = 1, 2, \theta_i = 0.5, i = 1, 2, 3$ and varying the cure rate proportions for each time in two scenarios: low cure rate: 20% and high cure rate: 80%.

The inverse-transform method was considered in the process of generating the pseudo-random samples for each component of the BS model; then we used the fatal shock structure to generate pseudo-random samples for the bivariate model. As previously described, the assumed Bayesian point estimator is the posterior expected value, and here we have evaluated the performance of such an estimator by assessing its bias (BIAS) and its mean squared error (MSE). Thus, using samples of the model parameters, one may obtain approximate Monte Carlo estimators for these measures as

$$\widehat{\text{BIAS}}(\hat{\beta}) \approx \frac{1}{500} \sum_{i=1}^{500} (\hat{\beta}_i - \beta)$$

and

$$\widehat{\text{MSE}}(\hat{\beta}) \approx \frac{1}{500} \sum_{i=1}^{500} (\hat{\beta}_i - \beta)^2,$$

where $\beta = \alpha_i (i = 1, 2), \theta_j (j = 1, 2, 3)$, or $\rho_k (k = 1, 2)$ are the model parameters, and $\{\hat{\beta}\}_{i=1}^{500}$ is the sequence of posterior means based on chains of size 2,000.

Remark. Note that the cure rate proportion denoted by $\rho_k (k = 1, 2)$ is obtained by $\rho_1 = \phi_{00} + \phi_{10}$ (Time 1) and $\rho_2 = \phi_{00} + \phi_{01}$ (Time 2).

The variance of $\hat{\beta}$ can be estimated as the difference between the MSE and the square of the bias. Besides, the coverage probability (CP) of the 95% highest posterior density (HPD) intervals was estimated by the approximation

$$\widehat{\text{CP}}(\hat{\beta}) \approx \frac{1}{500} \sum_{i=1}^{500} \delta_i(\hat{\beta}),$$

where $\delta_i(\hat{\beta})$ assumes 1 if the $i$th HPD interval contains the actual value $\theta$ and zero otherwise.

The obtained results are presented in Table 1. We have noticed in our study that the biases of $\alpha_i (i = 1, 2)$ and $\theta_j (j = 1, 2, 3)$ are positive, and $\rho_k (k = 1, 2)$ has a negative bias in most cases. However, the overall accuracy of the Bayesian estimators improved with increasing sample sizes since the estimated MSEs decreases as $n$ increases. Additionally, we observed that the estimated CPs of the HPD intervals for $\alpha_i (i = 1, 2)$, $\theta_j (j = 1, 2, 3)$ and $\rho_k (k = 1, 2)$ were converging to the nominal level (95%).

4 An Application to a Diabetic Retinopathy Data Set

Diabetic retinopathy is a chronic progressive, potentially sight-threatening disease of retinal microvasculature associated with the prolonged hyperglycaemia. It is caused by damage to the blood vessels of the light-sensitive tissue at the back of the eye called the retina, which processes light and vision for the brain. Over time, diabetes damages the blood vessels in the retina, which could cause the retinal tissue to swell, resulting in blurred vision. According to The National Eye Institute (The National Eye Institute, 2018), diabetic retinopathy disease may progress through four stages:

1. Mild nonproliferative retinopathy: Small areas of balloon-like swelling in the retina’s tiny blood vessels, called microaneurysms, occur at this stage.
2. Moderate nonproliferative retinopathy: As the disease progresses, blood vessels that nourish the retina may swell and distort.
3. Severe nonproliferative retinopathy: Many more blood vessels are blocked, depriving blood supply to areas of the retina. These areas secrete growth factors that signal the retina to grow new blood vessels.
Table 1 Results of the simulation study for the proposed BS-Type I cure rate model.

<table>
<thead>
<tr>
<th>n</th>
<th>Parameter</th>
<th>Bias</th>
<th>MSE</th>
<th>CP</th>
<th>n</th>
<th>Parameter</th>
<th>Bias</th>
<th>MSE</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>$\alpha_1$</td>
<td>0.3204</td>
<td>0.1026</td>
<td>0.9165</td>
<td>25</td>
<td>$\alpha_1$</td>
<td>0.4903</td>
<td>0.2362</td>
<td>0.9821</td>
</tr>
<tr>
<td></td>
<td>$\alpha_2$</td>
<td>0.2798</td>
<td>0.0783</td>
<td>0.9738</td>
<td></td>
<td>$\alpha_2$</td>
<td>0.8759</td>
<td>0.4761</td>
<td>0.9875</td>
</tr>
<tr>
<td></td>
<td>$\theta_1$</td>
<td>0.4155</td>
<td>0.1726</td>
<td>0.9980</td>
<td></td>
<td>$\theta_1$</td>
<td>0.3459</td>
<td>0.1605</td>
<td>0.9821</td>
</tr>
<tr>
<td></td>
<td>$\theta_2$</td>
<td>0.3420</td>
<td>0.1170</td>
<td>0.9945</td>
<td></td>
<td>$\theta_2$</td>
<td>0.6273</td>
<td>0.2517</td>
<td>0.9924</td>
</tr>
<tr>
<td></td>
<td>$\theta_3$</td>
<td>0.1697</td>
<td>0.0288</td>
<td>0.9844</td>
<td></td>
<td>$\theta_3$</td>
<td>0.2903</td>
<td>0.0843</td>
<td>0.9802</td>
</tr>
<tr>
<td></td>
<td>$\rho_1$</td>
<td>-0.0957</td>
<td>0.0226</td>
<td>0.9693</td>
<td></td>
<td>$\rho_1$</td>
<td>-0.1262</td>
<td>0.0119</td>
<td>0.9705</td>
</tr>
<tr>
<td></td>
<td>$\rho_2$</td>
<td>-0.1503</td>
<td>0.0226</td>
<td>0.9782</td>
<td></td>
<td>$\rho_2$</td>
<td>-0.1589</td>
<td>0.0171</td>
<td>0.9751</td>
</tr>
<tr>
<td>50</td>
<td>$\alpha_1$</td>
<td>0.2918</td>
<td>0.0851</td>
<td>0.9256</td>
<td>50</td>
<td>$\alpha_1$</td>
<td>0.3818</td>
<td>0.1458</td>
<td>0.9748</td>
</tr>
<tr>
<td></td>
<td>$\alpha_2$</td>
<td>0.2760</td>
<td>0.0762</td>
<td>0.9702</td>
<td></td>
<td>$\alpha_2$</td>
<td>0.6139</td>
<td>0.3769</td>
<td>0.9801</td>
</tr>
<tr>
<td></td>
<td>$\theta_1$</td>
<td>0.3498</td>
<td>0.1224</td>
<td>0.9909</td>
<td></td>
<td>$\theta_1$</td>
<td>0.3102</td>
<td>0.0962</td>
<td>0.9772</td>
</tr>
<tr>
<td></td>
<td>$\theta_2$</td>
<td>0.3091</td>
<td>0.0955</td>
<td>0.9939</td>
<td></td>
<td>$\theta_2$</td>
<td>0.4806</td>
<td>0.1787</td>
<td>0.9789</td>
</tr>
<tr>
<td></td>
<td>$\theta_3$</td>
<td>-0.0381</td>
<td>0.0015</td>
<td>0.9657</td>
<td></td>
<td>$\theta_3$</td>
<td>0.2619</td>
<td>0.0638</td>
<td>0.9757</td>
</tr>
<tr>
<td></td>
<td>$\rho_1$</td>
<td>-0.0950</td>
<td>0.0090</td>
<td>0.9671</td>
<td></td>
<td>$\rho_1$</td>
<td>-0.1106</td>
<td>0.0122</td>
<td>0.9507</td>
</tr>
<tr>
<td></td>
<td>$\rho_2$</td>
<td>-0.0950</td>
<td>0.0090</td>
<td>0.9671</td>
<td></td>
<td>$\rho_2$</td>
<td>-0.1192</td>
<td>0.0142</td>
<td>0.9534</td>
</tr>
<tr>
<td>75</td>
<td>$\alpha_1$</td>
<td>0.2342</td>
<td>0.0717</td>
<td>0.9634</td>
<td>75</td>
<td>$\alpha_1$</td>
<td>0.2383</td>
<td>0.0970</td>
<td>0.9622</td>
</tr>
<tr>
<td></td>
<td>$\alpha_2$</td>
<td>0.2447</td>
<td>0.0688</td>
<td>0.9681</td>
<td></td>
<td>$\alpha_2$</td>
<td>0.4333</td>
<td>0.1878</td>
<td>0.9778</td>
</tr>
<tr>
<td></td>
<td>$\theta_1$</td>
<td>0.2784</td>
<td>0.1032</td>
<td>0.9753</td>
<td></td>
<td>$\theta_1$</td>
<td>0.1203</td>
<td>0.0526</td>
<td>0.9636</td>
</tr>
<tr>
<td></td>
<td>$\theta_2$</td>
<td>0.2807</td>
<td>0.0849</td>
<td>0.9712</td>
<td></td>
<td>$\theta_2$</td>
<td>0.4125</td>
<td>0.1702</td>
<td>0.9728</td>
</tr>
<tr>
<td></td>
<td>$\theta_3$</td>
<td>0.0690</td>
<td>0.0048</td>
<td>0.9644</td>
<td></td>
<td>$\theta_3$</td>
<td>0.2124</td>
<td>0.0451</td>
<td>0.9744</td>
</tr>
<tr>
<td></td>
<td>$\rho_1$</td>
<td>-0.0220</td>
<td>0.0012</td>
<td>0.9544</td>
<td></td>
<td>$\rho_1$</td>
<td>-0.0886</td>
<td>0.0079</td>
<td>0.9488</td>
</tr>
<tr>
<td></td>
<td>$\rho_2$</td>
<td>0.0152</td>
<td>0.0009</td>
<td>0.9535</td>
<td></td>
<td>$\rho_2$</td>
<td>-0.0853</td>
<td>0.0073</td>
<td>0.9449</td>
</tr>
<tr>
<td>100</td>
<td>$\alpha_1$</td>
<td>0.1919</td>
<td>0.0536</td>
<td>0.9581</td>
<td>100</td>
<td>$\alpha_1$</td>
<td>0.1848</td>
<td>0.0341</td>
<td>0.9532</td>
</tr>
<tr>
<td></td>
<td>$\alpha_2$</td>
<td>0.1708</td>
<td>0.0783</td>
<td>0.9538</td>
<td></td>
<td>$\alpha_2$</td>
<td>0.3082</td>
<td>0.0950</td>
<td>0.9545</td>
</tr>
<tr>
<td></td>
<td>$\theta_1$</td>
<td>0.1146</td>
<td>0.0719</td>
<td>0.9615</td>
<td></td>
<td>$\theta_1$</td>
<td>0.0508</td>
<td>0.0026</td>
<td>0.9429</td>
</tr>
<tr>
<td></td>
<td>$\theta_2$</td>
<td>0.1259</td>
<td>0.0414</td>
<td>0.9598</td>
<td></td>
<td>$\theta_2$</td>
<td>0.3380</td>
<td>0.1142</td>
<td>0.9637</td>
</tr>
<tr>
<td></td>
<td>$\theta_3$</td>
<td>-0.0258</td>
<td>0.0007</td>
<td>0.9802</td>
<td></td>
<td>$\theta_3$</td>
<td>0.1039</td>
<td>0.0108</td>
<td>0.9544</td>
</tr>
<tr>
<td></td>
<td>$\rho_1$</td>
<td>-0.0124</td>
<td>0.0014</td>
<td>0.9594</td>
<td></td>
<td>$\rho_1$</td>
<td>0.0047</td>
<td>0.0010</td>
<td>0.9523</td>
</tr>
<tr>
<td></td>
<td>$\rho_2$</td>
<td>0.0061</td>
<td>0.0002</td>
<td>0.9502</td>
<td></td>
<td>$\rho_2$</td>
<td>0.0294</td>
<td>0.0009</td>
<td>0.9501</td>
</tr>
</tbody>
</table>

4. Proliferative diabetic retinopathy (PDR): At this advanced stage, growth factors secreted by the retina trigger the proliferation of new blood vessels, which grow along the inside surface of the retina and into the vitreous gel.

Vision loss by diabetic retinopathy is sometimes irreversible. However, early detection and treatment can reduce the risk of blindness by 95 percent. It is important to point out that diabetic retinopathy often lacks early symptoms, people diagnosticated with diabetes should get a comprehensive dilated eye exam at least once a year. The treatment for diabetic retinopathy is often delayed until it starts to progress to PDR, or when diabetic macular edema occurs (DME).

To illustrate the usefulness of the proposed model, we considered a data set introduced by Huster, Brookmeyer and Self (1989) with 197 patients where 50% of the patients were classified by the authors as “high-risk” for diabetic retinopathy. Each patient had one eye randomized to laser treatment, and the other eye received no treatment. For each eye, the event of interest was the time from the beginning of the treatment to the time when visual acuity dropped below 2/200 (call it “blindness”). There was a built-in lag time of approximately six months (visits were every three months). Survival times in this data set are, therefore, the actual times to blindness in months, minus the minimum possible time to event (6.5 months). Censoring was caused by death, dropout or end of the study. It was considered for the bivariate analyzes that $T_1$ is the time up to visual loss for the treated eye, while $T_2$ is the time up to visual loss for the not treated or control eye. Table (2) shows the percentage of censoring...
data in the diabetic retinopathy data. It is observed that only 20% are not censored in both $T_1$ and $T_2$, the most of the data are censored in both times.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Description of the presence of censoring in diabetic retinopathy data set.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_1$</td>
<td>$T_2$</td>
</tr>
<tr>
<td>Censored data ($\delta_1 = 0$)</td>
<td>Censored data ($\delta_2 = 0$)</td>
</tr>
<tr>
<td>Completed data ($\delta_1 = 1$)</td>
<td>Censored data ($\delta_2 = 0$)</td>
</tr>
<tr>
<td>Censored data ($\delta_1 = 0$)</td>
<td>Completed data ($\delta_2 = 1$)</td>
</tr>
<tr>
<td>Completed data ($\delta_1 = 1$)</td>
<td>Completed data ($\delta_2 = 1$)</td>
</tr>
</tbody>
</table>

The fit of BS cure rate distribution was compared to the fit of traditional Marshall-Olkin Bivariate Lindley Distribution (BL) and Marshall-Olkin Bivariate Exponential Distribution (BE), for more details, see Oliveira (2019). Moreover, the parameters of the models were estimated under a Bayesian approach assuming independent non-informative (large variances) gamma prior distributions for each model parameter. The DIC, EAIC, EBIC, and LPML criteria were considered to discriminate the best model among all proposed models. The results are illustrated in Table 3. The estimated percentages for parameters $\phi_{00}$, $\phi_{01}$, $\phi_{10}$ and $\phi_{11}$ by the proposed bivariate BS cure rate model were consistent with the values observed in the real data in Table (2).

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Bayesian estimates assuming BS, BL and BE cure rate models for the diabetic retinopathy dataset.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>Parameter</td>
</tr>
<tr>
<td>BS</td>
<td>$\alpha_1$</td>
</tr>
<tr>
<td>BS</td>
<td>$\alpha_2$</td>
</tr>
<tr>
<td>BS</td>
<td>$\theta_1$</td>
</tr>
<tr>
<td>BS</td>
<td>$\theta_2$</td>
</tr>
<tr>
<td>BS</td>
<td>$\theta_3$</td>
</tr>
<tr>
<td>BS</td>
<td>$\phi_{00}$</td>
</tr>
<tr>
<td>BS</td>
<td>$\phi_{01}$</td>
</tr>
<tr>
<td>BS</td>
<td>$\phi_{10}$</td>
</tr>
<tr>
<td>BS</td>
<td>$\phi_{11}$</td>
</tr>
<tr>
<td>BS</td>
<td>$\rho_1$</td>
</tr>
<tr>
<td>BS</td>
<td>$\rho_2$</td>
</tr>
<tr>
<td>BS</td>
<td>$\theta_1$</td>
</tr>
<tr>
<td>BS</td>
<td>$\theta_2$</td>
</tr>
<tr>
<td>BS</td>
<td>$\theta_3$</td>
</tr>
<tr>
<td>BS</td>
<td>$\phi_{00}$</td>
</tr>
<tr>
<td>BS</td>
<td>$\phi_{01}$</td>
</tr>
<tr>
<td>BS</td>
<td>$\phi_{10}$</td>
</tr>
<tr>
<td>BS</td>
<td>$\phi_{11}$</td>
</tr>
<tr>
<td>BS</td>
<td>$\rho_1$</td>
</tr>
<tr>
<td>BS</td>
<td>$\rho_2$</td>
</tr>
<tr>
<td>BL</td>
<td>$\lambda_1$</td>
</tr>
<tr>
<td>BL</td>
<td>$\lambda_2$</td>
</tr>
<tr>
<td>BL</td>
<td>$\lambda_3$</td>
</tr>
<tr>
<td>BL</td>
<td>$\phi_{00}$</td>
</tr>
<tr>
<td>BL</td>
<td>$\phi_{01}$</td>
</tr>
<tr>
<td>BL</td>
<td>$\phi_{10}$</td>
</tr>
<tr>
<td>BL</td>
<td>$\phi_{11}$</td>
</tr>
<tr>
<td>BL</td>
<td>$\rho_1$</td>
</tr>
<tr>
<td>BL</td>
<td>$\rho_2$</td>
</tr>
</tbody>
</table>
Figure 2 presents the Kaplan-Meier marginal survival plots for both lifetimes and also the estimated survival BS, BL and BS functions considering the diabetic retinopathy data set. It is possible to note that BS and BE models display fitted curves suitable to the data. The estimated curves properly follow the empirical values related to the cure fraction plateau. It is worth noting that the BL model presents unsatisfactory curves, collaborating with the results of the selection criteria in Table 3, in which it presented the worst indicators. The BE and BE models showed a satisfactory fit, however, we can see in Figure 3 that the Cox-Snell residuals plots suggest a better fit of BE models considering the diabetic retinopathy data set.

Figure 2 Kaplan-Meier estimators versus Bayesian estimated survival functions for the marginal survival functions $T_1$ (panel (a)) and $T_2$ (panel (b)) assuming BS, BL and BE models.

Figure 3 Cox-Snell residual plots for the BS in $T_1$ (a), BS in $T_2$ (d), BL in $T_1$ (b), BL in $T_2$ (e), BE in $T_1$ (f) and BE in $T_2$ (f).
In order to illustrate an application of the regression model based on the BS cure rate model, in this analysis we consider the covariate type of diabetes: juvenile (diagnosis before age 20) and adult (diagnosis after age 20). This variable is represented in the regression model by using an indicator or dummy variable, where, where \( x_i = 0 \) if juvenile and \( x_i = 1 \) if adult. Thus, the proposed regression model is given by

\[
\log(\theta_1) = \beta_{10} + \beta_{11}x_i \quad \text{and} \quad \log(\theta_2) = \beta_{20} + \beta_{21}x_i.
\]

For all regression coefficients, we assumed normal prior distributions with mean 0 and large variance. The Bayesian estimates are shown in Table 4. It is possible to verify the significant effect of the covariate type of diabetes on the parameter \( \theta_1 \), since the value zero is not inside the 95% credible interval for the regression parameter \( \beta_{11} \).

<table>
<thead>
<tr>
<th>Param.</th>
<th>Mean</th>
<th>Standard Deviation (S.D.)</th>
<th>95% HPD Interval</th>
<th>DIC</th>
<th>EAIC</th>
<th>EBIC</th>
<th>LMPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_1 )</td>
<td>0.8050</td>
<td>0.3030</td>
<td>(0.3893, 1.5704)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>0.8753</td>
<td>0.3133</td>
<td>(0.3302, 1.3635)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_{10} )</td>
<td>-2.4646</td>
<td>0.2287</td>
<td>(-2.9427, -2.0524)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_{11} )</td>
<td>-0.8909</td>
<td>0.4123</td>
<td>(-1.1590, -0.1283)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_{20} )</td>
<td>-2.6651</td>
<td>0.2401</td>
<td>(-3.1519, -2.2282)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_{21} )</td>
<td>0.2072</td>
<td>0.2384</td>
<td>(-0.2583, 0.6704)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta_{10} )</td>
<td>0.0872</td>
<td>0.0192</td>
<td>(0.0527, 0.1284)</td>
<td>1562.6</td>
<td>1574.8</td>
<td>1579.1</td>
<td>790.6</td>
</tr>
<tr>
<td>( \theta_{11} )</td>
<td>0.0380</td>
<td>0.0150</td>
<td>(0.0131, 0.0715)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta_{20} )</td>
<td>0.0716</td>
<td>0.0169</td>
<td>(0.0428, 0.1077)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta_{21} )</td>
<td>0.0888</td>
<td>0.0237</td>
<td>(0.0481, 0.1413)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta_3 )</td>
<td>0.0104</td>
<td>0.0039</td>
<td>(0.0004, 0.0193)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \phi_{00} )</td>
<td>0.3436</td>
<td>0.0441</td>
<td>(0.2555, 0.4307)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \phi_{01} )</td>
<td>0.2975</td>
<td>0.0477</td>
<td>(0.2071, 0.3971)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \phi_{10} )</td>
<td>0.0811</td>
<td>0.0283</td>
<td>(0.0305, 0.1405)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \phi_{11} )</td>
<td>0.2778</td>
<td>0.0512</td>
<td>(0.1934, 0.3854)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho_1 )</td>
<td>0.6411</td>
<td>0.0538</td>
<td>(0.5339, 0.7305)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho_2 )</td>
<td>0.4248</td>
<td>0.0496</td>
<td>(0.3208, 0.5162)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4 shows the survival curves estimated by the Kaplan-Meier method and estimated by BS cure rate models for the diabetic retinopathy dataset in presence of a covariate. We can notice a satisfactory fit in both times.

![Figure 4](image-url)
5 Concluding Remarks

The use of existing parametric bivariate distributions could be a good alternative to analyze bivariate lifetime data in the presence of censored data and covariates. However, there are few flexible bivariate lifetime distributions introduced in the literature where, in general, bivariate lifetime datasets usually are analyzed using some existing exponential bivariate lifetime distributions as the popular Block and Basu bivariate exponential distribution (Block and Basu, 1974) or the Marshall-Olkin bivariate exponential distribution (Marshall and Olkin, 1967a,b) that could be not well fitted by the data in many applications.

In this study, we introduced a new bivariate distribution, denoted as bivariate Sushila distribution, obtained using the Marshall and Olkin (1967a) method to add a new parameter to the survival function of the Sushila distribution in order to propose a more flexible joint survival function based on the fatal shock model as an alternative for the existing bivariate models. Some properties of this new distribution were also discussed in this study and an extension to the class of distributions of Sushila type was provided.

In an application with real lifetime data presented in this study, we observed that, with the use the new BS distribution, it was possible to obtain in a simple way the Bayesian inferences of interest for bivariate lifetime datasets in presence of right-censored data and a covariate, with small computational costs (use of MCMC methods) as compared to other bivariate lifetime distributions obtained using other approaches, for example, using copula functions to construct new bivariate lifetime distributions assuming specified parametric marginal lifetimes distributions. We also observed in the real data application considering the analysis of the diabetic retinopathy data, that the proposed model was well fitted by the data as compared to other existing bivariate lifetime models. Additionally, we observed in this application, that under a non-parametric approach, 67.1% of the patients where the eyes were treated are non-susceptible to blindness (a value close to the Bayesian estimate of $\rho_1 = \phi_{00} + \phi_{01}$ given by 66.97%) and 39.4% of the patients with untreated eyes are non-susceptible for blindness (a value close to the Bayesian estimate of $\rho_2 = \phi_{00} + \phi_{10}$ given by 44.06%). That is, we have another indication that the proposed model has a good accuracy for the estimation of the long-term survivors. From the obtained results of this study, we can conclude that the use of the new proposed model could be of great interest in the analysis of bivariate lifetime data, especially in engineering, or medical studies.

Acknowledgments

The authors would like to thank the anonymous referees for the careful reading and thoughtful suggestions for improving this work’s content.
Appendix - Cross-Factorial Moment

For the proposed model given the fatal shock structure, the cross factorial moment has a closed form and is a monotonic increasing function given by,

\[ E[X_1X_2] = \frac{\alpha_1 \alpha_2}{[(\alpha_2 \theta_3 + \theta_2)\alpha_1 + \alpha_2 \theta_1]^3(\alpha_2 \theta_3 + \theta_2)^2(\theta_1 + 1)(\alpha_1 \theta_3 + \theta_1)^2(\theta_2 + 1)} \]

\[ \times \left\{ (\alpha_2 \theta_3 + \theta_2)^2(2 \theta_3^2(\theta_2 + 1)\alpha_2^2 + 3 \theta_2 \theta_3(\theta_2 + 2)\alpha_2 + \theta_2^2(\theta_2 + 2))(\theta_1 + 1)\theta_3 \alpha_1^4 \right. \]

\[ + (\alpha_2 \theta_3 + \theta_2)\theta_1 \left[ 7\left(\theta_1 + \frac{10}{7}\right)\theta_3^3(\theta_2 + 1)\alpha_2^3 + 13 \theta_2 \left(\theta_2 + \frac{22}{13}\right)\theta_1 + \frac{19}{13} \theta_2 + \frac{34}{13} \right] \]

\[ \times \theta_3^3\alpha_2^2 + 7 \theta_2^2(\theta_2 + 2) \left(\theta_1 + \frac{11}{7}\right)\theta_3^2 + \theta_2^3(\theta_2 + 2)(\theta_1 + 2) \left[ \alpha_1^3 \right. \]

\[ + 3 \alpha_2 \left[ 3\left(\theta_1 + \frac{16}{9}\right)\theta_3^3(\theta_2 + 1)\alpha_2^3 + \frac{20 \theta_3^2 \alpha_2^2 \theta_2}{3} \left(\theta_2 + \frac{3}{2}\right)\theta_1 + \frac{9}{5} \theta_2 + \frac{14}{5} \right] \]

\[ + \frac{14}{3} \theta_2^2 \left(\theta_2 + \frac{13}{7}\right)\theta_1 + \frac{13 \theta_2}{7} + \frac{25}{7} \theta_3 \alpha_2 + \theta_2^3(\theta_2 + 2)(\theta_1 + 2) \left[ \theta_1^2 \alpha_1^2 \right. \]

\[ + 3 \left(\frac{5}{3} \alpha_2 \theta_3 + \theta_2 \right)\theta_2^2(\theta_1 + 2)\theta_1^3(\theta_3(\theta_2 + 1)\alpha_2 + \theta_2^2 + 2 \theta_2)\alpha_1 \]

\[ + \alpha_2^3(\theta_1 + 2)\theta_1^4(\theta_3(\theta_2 + 1)\alpha_2 + \theta_2^2 + 2 \theta_2) \left\} \right. \]

Since the cross factorial moment has a closed-form, the covariance and the correlation coefficients can be directly obtained using the relations

\[ \text{Cov}(X_1, X_2) = E[X_1X_2] - E[X_1]E[X_2] \quad \text{and} \quad \rho = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} \]

Here we have to take into account that \( E[X_1X_2] > E[X_1]E[X_2] \) and therefore, \( 0 < \rho < 1 \).

Appendix - Code Bayesian Model and Traceplots

The scripts used to perform the analyses for this study are available in a GitHub repository: https://github.com/edsonzmartinez/BivariateSushila.

Figure 5 shows the traceplots of the corresponding MCMC chains for BS cure rate model considering diabetic retinopathy dataset, and the horizontal red lines indicate the posterior means.
A New Class of Bivariate Sushila Distributions

References


