Bivariate Weibull Distributions Derived From Copula Functions In The Presence Of Cure Fraction And Censored Data

Em’ilio A. Coelho-Barros¹*, Jorge Alberto Achcar², Josmar Mazucheli³

¹Universidade Tecnológica Federal do Paraná, DAMAT/UTFPR
²Universidade de São Paulo, FMRP/USP
³Universidade Estadual de Maringá, DEs/UEM

Abstract: In this paper we introduce bivariate Weibull distributions derived from copula functions in presence of cure fraction, censored data and covariates. Two copula functions are explored: the FGM (Farlie-Gumbel-Morgenstern) copula and the Gumbel copula. Inferences for the proposed models are obtained under the Bayesian approach, using standard MCMC (Markov Chain Monte Carlo) methods. An illustration of the proposed methodology is given considering a medical data set.

Key words: Bayesian methods, Bivariate Weibull, Censored observations, Copula functions, Cure fraction.

1. Introduction

A long-term survivor mixture model, also known as standard cure rate model, assumes that the studied population is a mixture of susceptible individuals, who experience the event of interest and non susceptible individuals that will never experience it. These individuals are not at risk with respect to the event of interest and are considered immune, non susceptible or cured Maller and Zhou (1996). Different approaches, parametric and non-parametric, have been considered to model the proportion of immunes and interested readers can refer, for example, to Boag (1949), Berkson (1952), Haybittle (1965), Farewell (1982, 1986), Meeker (1987), Dunsmuir et al., (1989), Taylor, (1995), Gamel et al. (1990, 1999), Ghitany and Maller (1992), Yamaguchi, (1992), Copas and Heydary (1997), Ng and McLachlan (1998), De Angelis et al. (1999), Peng and Dear (2000), Sy and Taylor (2000), Cancho and Bolfarine, (2001), Yu et al. (2004), Kannan et al., (2010).

Cure models for paired and clustered survival data have been considered by a number of authors, Wienke et al. (2003,2006) introduced a model for a cure fraction in bivariate time-to-event data motivated by the article of Chatterjee and Shih (2001). Chatterjee and Shih (2003) used a correlated gamma frailty model by a copula function to analyse bivariate survival data.

* Corresponding author.
Considering univariate lifetimes and following Maller and Zhou (1996), the standard cure rate model assumes that a certain fraction $p$ in the population is cured or never fail with respect to the specific cause of death or failure, while the remaining $(1 - p)$ fraction of the individuals is not cured, leading to the survival function for the entire population written as,

$$ S(t) = p + (1 - p)S_0(t), $$

where $p \in (0,1)$ is the mixing parameter and $S_0(t)$ denotes a proper survival function for the non cured group in the population.

From the mixture survival function, (1), the probability density function is obtained from $f(t) = -dS(t)/dt$ and given by,

$$ f(t) = (1 - p)f_0(t), $$

where $f_0(t)$ is the probability density function for a susceptible individual. From $h(t) = f(t)/S(t)$, the hazard function is given by,

$$ h(t) = \frac{(1 - p)f_0(t)}{p + (1 - p)S_0(t)}. $$

An alternative to a standard long-term survivor mixture model is the long-term survivor non-mixture model suggested by Yakovlev et al. (1996); Tsodikov et al. (2003); Lambert et al. (2010) which defines an asymptote for the cumulative hazard and hence for the cure fraction. The non-mixture model or the promotion time cure fraction has been used by Lambert et al. (2007, 2010) to estimate the probability of cure fraction in cancer lifetime data.

In this paper we consider a generalization of the mixture cure fraction model (1) for the bivariate case considering different copula functions to capture the possible existing dependence between two lifetimes $T_1$ and $T_2$ for the susceptible individuals, censored data and the presence of covariates. Inferences for the proposed models are obtained using Bayesian methods where samples of the joint posterior distribution of interest is simulated using MCMC (Markov Chain Monte Carlo) methods, as the popular Gibbs sampling algorithm Gelfand and Smith (1990); Casella and George (1992) and the Metropolis-Hastings algorithm Chib and Greenberg (1995).

The paper is organized as follows: in Section 2 we introduce bivariate lifetime data assuming mixture cure fraction models; in Section 3 we present the likelihood function considering bivariate lifetime data in presence of cure fraction and censored data considering two special copula functions: Farlie-GumbelMorgenstern (FGM) copula and Gumbel copula; in Section 4 we introduce the Weibull marginal distributions for the lifetimes $T_1$ and $T_2$; in Section 5 we introduce an illustrative example. Finally, in Section 6 we present some concluding remarks.

2. Bivariate Lifetime Data Assuming Mixture Cure Fraction Models

In some areas of application, especially in medical and engineering studies, we could have two lifetimes $T_1$ and $T_2$ associated with each individual or unit. Usually, these data are assumed to be independent, but in many cases, the lifetime of one component could affect the lifetime of
the other component. This is the case as an example in the medical area of paired organs like kidneys, lungs, eyes, ears, dental implants among many others. In the literature we observe many papers related to bivariate lifetime parametric models Gumbel (1960); Freund (1961); Marshall and Olkin (1967); Downton (1970); Hawkes (1972); Block and Basu (1974); Hougaard (1986); Sarkar (1987); Arnold and Strauss (1988). An alternative is the use of copula functions Nelsen (2006); Trivedi and Zimmer (2005, 2006) to model bivariate lifetime data assuming different marginal lifetime distributions.

We also could have the presence of cure fraction or immunes for both lifetimes $T_1$ and $T_2$. In this way, let us assume mixture models for $T_1$ and $T_2$, given, respectively, by,

$$S_1 (t_1) = p_1 + (1 - p_1) S_{10} (t_1),$$

$$S_2 (t_2) = p_2 + (1 - p_2) S_{01} (t_2),$$

where $S_{10} (t_1)$ and $S_{01} (t_2)$ are the survival functions for the susceptible individuals in the lifetimes $T_1$ and $T_2$, respectively; $p_1$ and $p_2$ are, respectively, the cure probabilities in $T_1$ and $T_2$; $(1 - p_1)$ and $(1 - p_2)$ are, respectively, the probabilities of susceptible in lifetimes $T_1$ and $T_2$.

The joint survival function for $T_1$ and $T_2$ is given by,

$$S (t_1, t_2) = P (T_1 > t_1, T_2 > t_2) = 1 - F_1 (t_1) - F_2 (t_2) + F (t_1, t_2),$$

where $F_1 (t_1) = P (T_1 \leq t_1)$ is the marginal cumulative distribution function for $T_1$; $F_2 (t_2) = P (T_2 \leq t_2)$ is the marginal cumulative distribution function for $T_2$; $F (t_1, t_2) = P (T_1 \leq t_1, T_2 \leq t_2)$ is the joint cumulative distribution function for $T_1$ and $T_2$.

In this way, we get from (4),

$$F_1 (t_1) = 1 - S_1 (t_1) = 1 - p_1 - (1 - p_1) S_{10} (t_1);$$

$$F_2 (t_2) = 1 - S_2 (t_2) = 1 - p_2 - (1 - p_2) S_{01} (t_2).$$

Thus, from (5), we get,

$$S (t_1, t_2) = p_1 + p_2 - 1 + (1 - p_1) S_{10} (t_1) + (1 - p_2) S_{01} (t_2) + F (t_1, t_2).$$

Let us define the binary indicator random variables $V_1$ and $V_2$ where $V_1 = 1$ if the individual is susceptible for the lifetime $T_1$ and $V_1 = 0$ if the individual is immune; in the same way, $V_2 = 1$ if the individual is susceptible for the lifetime $T_2$ and $V_2 = 0$ if the individual is immune, thus $P (V_1 = 1) = 1 - p_1$, $P (V_1 = 0) = p_1$, $P (V_2 = 1) = 1 - p_2$ and $P (V_2 = 0) = p_2$.

Let us assume $V_1$ independent of $V_2$, that is,

$$P (V_1 = v_1, V_2 = v_2) = P (V_1 = v_1) P (V_2 = v_2),$$

where $v_j = 1$ or $0$ for $j = 1, 2$. That is,
\[ \varphi_{00} = P(V_1 = 0, V_2 = 0) = p_1 p_2; \]

\[ \varphi_{10} = P(V_1 = 1, V_2 = 0) = (1 - p_1) p_2; \]

\[ \varphi_{01} = P(V_1 = 0, V_2 = 1) = p_1 (1 - p_2); \]

\[ \varphi_{11} = P(V_1 = 1, V_2 = 1) = (1 - p_1)(1 - p_2); \]

where \( \varphi_{00} + \varphi_{10} + \varphi_{01} + \varphi_{11} = 1. \)

Assuming \( T_1 \) independent of \( T_2 \), we have, \( S(t_1, t_2) = S_1(t_1) S_2(t_2) \); from (4), we have,

\[ S(t_1, t_2) = p_1 p_2 + (1 - p_1) p_2 S_{10}(t_1) + p_1 (1 - p_2) S_{01}(t_2) + (1 - p_1)(1 - p_2) S_{10}(t_1) S_{01}(t_2), \]

assuming \( T_1 \) dependent of \( T_2 \), we get from (9),

\[ S(t_1, t_2) = \varphi_{00} + \varphi_{10} S_{10}(t_1) + \varphi_{01} S_{01}(t_2) + \varphi_{11} S_{110}(t_1, t_2), \]

where \( S_{110}(t_1, t_2) \) is the joint survival function for \( T_1 \) and \( T_2 \) for the susceptible individuals.

Different bivariate lifetime distributions could be used for \( S_{110}(t_1, t_2) \) Block and Basu (1974); Marshall and Olkin (1967); Sarkar (1987). Another possibility is the use of copula functions Nelsen (2006). Copula functions can be used to link marginal distributions with a joint distribution. For specified univariate marginal distribution functions \( F_1(t_1), F_2(t_2), ..., F_m(t_m) \), the function,

\[ C(F_1(t_1), F_2(t_2), ..., F_m(t_m)) = F(t_1, t_2, ..., t_m), \]

which is defined using a copula function \( C \), results in a multivariate distribution function with univariate distributions specified as \( F_1(t_1), F_2(t_2), ..., F_m(t_m) \).

It is important to point out that any multivariate distribution function \( F \) can be written in the form of a copula function Sklar (2010), that is, if \( F(t_1, t_2, ..., t_m) \) is a joint multivariate distribution function with univariate marginal distribution functions \( F_1(t_1), F_2(t_2), ..., F_m(t_m) \), thus there exist a copula function \( C(u_1, u_2, ..., u_m) \), such that,

\[ F(t_1, t_2, ..., t_m) = C(F_1(t_1), F_2(t_2), ..., F_m(t_m)). \]

If every \( F_i \) is continuous, then \( C \) is unique. For the special case of bivariate distributions, we have \( m = 2 \).

The approach to formulate a multivariate distribution using a copula is based on the idea that a simple transformation can be made of each marginal variable in such a way that each transformed marginal variable has a uniform distribution. Once this is done, the dependence structure can be expressed as a multivariate distribution on the obtained uniforms, and a copula
is precisely a multivariate distribution on marginally uniform random variables. In this way, there are many families of copulas which differ in the detail of the dependence they represent. In the bivariate case, let \( T_1 \) and \( T_2 \) be two random variables with continuous distribution functions \( F_1 \) and \( F_2 \).

The probability integral transform can be applied separately to the two random variables to define \( U = F_1(t_1) \) and \( V = F_2(t_2) \), where \( U \) and \( V \) have uniform \((0,1)\) distributions, but are usually dependent if \( T_1 \) and \( T_2 \) are dependent \((T_1 \text{ and } T_2 \text{ independent, implies that } U \text{ and } V \text{ are independent})\). Specifying dependence between \( T_1 \) and \( T_2 \) is the same as specifying dependence between \( U \) and \( V \). With \( U \) and \( V \) uniform random variables, the problem reduces to specifying a bivariate distribution between two uniforms, that is, a copula.

As a first copula function appropriated to model weak dependences, we consider a Farlie-Gumbel-Morgenstern (FGM) copula Morgenstern (1956), given by,

\[
F_{110}(t_1,t_2) = F_{10}(t_1)F_{01}(t_2)[1 + \theta_1S_{10}(t_1)S_{01}(t_2)],
\]

where the dependence parameter \( \theta_1 \) is defined in the interval \([-1,1]\).

From (5), the joint survival function \( S_{110}(t_1,t_2) \) for \( T_1 \) and \( T_2 \) considering the susceptible individuals is given by,

\[
S_{110}(t_1,t_2) = 1 - F_{10}(t_1) - F_{01}(t_2) + F_{110}(t_1,t_2),
\]

that is, for the susceptible individuals,

\[
S_{110}(t_1,t_2) = S_{10}(t_1)S_{01}(t_2)[1 + \theta_1F_{10}(t_1)F_{01}(t_2)].
\]

From (11) we get,

\[
S(t_1,t_2) = \varphi_{00} + \varphi_{10}S_{10}(t_1) + \varphi_{01}S_{01}(t_2) + \varphi_{11}A(t_1,t_2),
\]

where \( A(t_1,t_2) = S_{10}(t_1)S_{01}(t_2)[1 + \theta_1F_{10}(t_1)F_{01}(t_2)] \).

As a second copula function, we assume the Gumbel copula Gumbel (1960) given by,

\[
F_{110}(t_1,t_2) = \frac{F_{10}(t_1) + F_{01}(t_2) - 1 + S_{10}(t_1)S_{01}(t_2)\times\exp\{-\theta_2\log[S_{10}(t_1)]\log[S_{01}(t_2)]\}},
\]

where \( 0 \leq \theta_2 \leq 1 \).

That is, from (15), we get,

\[
S_{110}(t_1,t_2) = S_{10}(t_1)S_{01}(t_2)\exp\{-\theta_2\log[S_{10}(t_1)]\log[S_{01}(t_2)]\}.
\]

From (11) we get,
\[
S(t_1,t_2) = \phi_{00} + \phi_{10}S_{10}(t_1) + \phi_{01}S_{01}(t_2) + \phi_{11}S_{10}(t_1)S_{01}(t_2)B(t_1,t_2),
\]
where \( B(t_1,t_2) = \exp\{-\theta_2 \log[S_{10}(t_1)]\log[S_{01}(t_2)]\} \).

3. Bivariate Lifetime Data In Presence Of Cure Fraction And Censored Data

Let us assume \( n \) pairs of a random sample of lifetimes \( t_{1i} \) and \( t_{2i}, i = 1,\ldots,n \), that can be right censored where censoring is independent of the lifetimes divided into four classes:

- \( C_1 \): both lifetimes \( t_{1i} \) and \( t_{2i} \) are observed;
- \( C_2 \): \( t_{1i} \) is a lifetime and \( t_{2i} \) is a censored time (that is, we only know that \( T_{2i} \geq t_{2i} \));
- \( C_3 \): \( t_{1i} \) is a censoring time and \( t_{2i} \) is a lifetime (that is, we only know that \( T_{1i} \geq t_{1i} \));
- \( C_4 \): both \( t_{1i} \) and \( t_{2i} \) are censoring times.

It is important to note that we could have other censoring schemes but this will be not considered in this paper.

The contribution of the \( i^{th} \) individual for the likelihood function Lawless (1982) is given by,

\[
L = \prod_{i \in C_1} f(t_{1i},t_{2i}) \prod_{i \in C_2} \left[-\frac{\partial S(t_{1i},t_{2i})}{\partial t_{1i}}\right] \prod_{i \in C_3} \left[-\frac{\partial S(t_{1i},t_{2i})}{\partial t_{2i}}\right] \prod_{i \in C_4} S(t_{1i},t_{2i}),
\]
where \( f(t_{1i},t_{2i}) \) is the joint probability density function for \( T_1 \) and \( T_2 \); \( S(t_{1i},t_{2i}) \) is the joint survival function; \( \frac{\partial S(t_{1i},t_{2i})}{\partial t_{1i}} \) and \( \frac{\partial S(t_{1i},t_{2i})}{\partial t_{2i}} \) are the partial derivatives of \( S(t_{1i},t_{2i}) \), with respect to \( t_{1i} \) and \( t_{2i} \), respectively.

Let us define the indicator variables \( \delta_{1i} \) and \( \delta_{2i} \), by,

\[
\begin{align*}
1 &\text{ if } t_{ji} \text{ is an observed lifetime } \quad \text{(22)} \\
0 &\text{ if } t_{ji} \text{ is a censored lifetime }
\end{align*}
\]
for \( j = 1,2; i = 1,2,\ldots,n \).

In this way, we rewrite the likelihood function (21) as,

\[
L = \prod_{i=1}^{n} \left[f(t_{1i},t_{2i})\right]^{\delta_{1i}\delta_{2i}} \prod_{i=1}^{n} \left[-\frac{\partial S(t_{1i},t_{2i})}{\partial t_{1i}}\right]^{\delta_{1i}(1-\delta_{2i})} \times \prod_{i=1}^{n} \left[-\frac{\partial S(t_{1i},t_{2i})}{\partial t_{2i}}\right]^{(1-\delta_{1i})\delta_{2i}} \prod_{i=1}^{n} S(t_{1i},t_{2i})^{(1-\delta_{1i})(1-\delta_{2i})}
\]

(23)
3.1 Farlie-Gumbel-Morgenstern Copula

From (17), we get expressions for $f(t_1, t_2)$, $-\frac{\partial S(t_1, t_2)}{\partial t_1}$, $-\frac{\partial S(t_1, t_2)}{\partial t_2}$ and $S(t_1, t_2)$ needed in the likelihood function (23), given by:

1. $f(t_1, t_2) = \frac{\partial^2 F(t_1, t_2)}{\partial t_1 \partial t_2} = -\frac{\partial^2 S(t_1, t_2)}{\partial t_1 \partial t_2}$

From (17), we get,

$$\frac{\partial A(t_1, t_2)}{\partial t_1} = -f_{10}(t_1) S_{01}(t_2) + \theta_1 S_{01}(t_2) F_{01}(t_2) [-f_{10}(t_1) F_{10}(t_1) + f_{10}(t_1) S_{10}(t_1)]$$

$$= -f_{10}(t_1) S_{01}(t_2) + \theta_1 S_{01}(t_2) F_{01}(t_2) f_{10}(t_1) [1 - 2F_{10}(t_1)].$$

(24)

and

$$\frac{\partial A(t_1, t_2)}{\partial t_2} = f_{10}(t_1) f_{01}(t_2) + \theta_1 f_{10}(t_1) f_{01}(t_2) [1 - 2F_{10}(t_1)] [1 - 2F_{01}(t_2)]$$

(25)

From the expression (25), we get,

$$f(t_1, t_2) = \phi_{11} f_{10}(t_1) f_{01}(t_2) [1 + \theta_1 [1 - 2F_{10}(t_1)][1 - 2F_{01}(t_2)]].$$  

(26)

2. From (17), we get,

$$-\frac{\partial S(t_1, t_2)}{\partial t_1} = f_{10}(t_1) \{\phi_{10} + \phi_{11} S_{01}(t_2) - \phi_{11} \theta_1 S_{01}(t_2) F_{01}(t_2) [1 - 2F_{10}(t_1)]\}$$

(27)

3. From (17), we get,

$$-\frac{\partial S(t_1, t_2)}{\partial t_2} = f_{01}(t_2) \{\phi_{01} + \phi_{11} S_{10}(t_1) - \phi_{11} \theta_1 S_{10}(t_1) F_{10}(t_1) [1 - 2F_{01}(t_2)]\}$$

(28)

4. Observe that $S(t_1, t_2)$ needed in the likelihood function (23) is given by (17).

3.2 Gumbel Copula

From (20), we get,

$$\frac{\partial B(t_1, t_2)}{\partial t_1} = \frac{\theta_2 f_{10}(t_1) \log [S_{01}(t_2)] B(t_1, t_2)}{S_{10}(t_1)}$$

(29)

and,

$$\frac{\partial B(t_1, t_2)}{\partial t_2} = \frac{\theta_2 f_{01}(t_2) \log [S_{10}(t_1)] B(t_1, t_2)}{S_{01}(t_2)}$$

(30)
In this way, we get,

\[
\frac{\partial S(t_1, t_2)}{\partial t_1} = -\phi_{10} f_{10}(t_1) + \phi_{11} S_{01}(t_2) \left[ -f_{10}(t_1) B(t_1, t_2) + S_{10}(t_1) \frac{\partial B(t_1, t_2)}{\partial t_1} \right];
\]

\[
\frac{\partial S(t_1, t_2)}{\partial t_2} = -\phi_{10} f_{01}(t_2) + \phi_{11} S_{01}(t_1) \left[ -f_{01}(t_2) B(t_1, t_2) + S_{01}(t_2) \frac{\partial B(t_1, t_2)}{\partial t_2} \right].
\]

Different marginal parametric lifetime distributions could be used for the lifetimes \(T_1\) and \(T_2\), as exponential, Weibull, log-normal, gamma or generalized gamma distributions.

### 4. Weibull Marginal Distributions For The Lifetimes \(T_1\) And \(T_2\)

Considering the bivariate lifetime distributions in presence of cure fractions derived from mixture models with a dependence structure given by FGM or Gumbel copula functions, we assume as a special model, a popular lifetime model extensively used in medical or engineering applications: the Weibull distribution. In this way, we consider marginal Weibull distributions for the lifetimes \(T_1\) and \(T_2\) of susceptible individuals with densities and survival functions given, respectively, by

\[
f_{10}(t_1) = \frac{\lambda_1}{\mu_1} \left( \frac{t_1}{\mu_1} \right)^{\lambda_1-1} \exp \left[ -\left( \frac{t_1}{\mu_1} \right)^{\lambda_1} \right];
\]

\[
f_{01}(t_2) = \frac{\lambda_2}{\mu_2} \left( \frac{t_2}{\mu_2} \right)^{\lambda_2-1} \exp \left[ -\left( \frac{t_2}{\mu_2} \right)^{\lambda_2} \right];
\]

\[
S_{10}(t_1) = \exp \left[ -\left( \frac{t_1}{\mu_1} \right)^{\lambda_1} \right];
\]

\[
S_{01}(t_2) = \exp \left[ -\left( \frac{t_2}{\mu_2} \right)^{\lambda_2} \right].
\]
The Weibull distribution is very popular to analyse lifetime data, since its hazard function \( h(t) = f(t)/S(t) \) is increasing if the shape parameter \( \lambda > 1 \); is decreasing if \( \lambda < 1 \) or is constant (an exponential distribution) if \( \lambda = 1 \) Lawless (1982).

From (2), the means for lifetimes \( T_1 \) and \( T_2 \) are given respectively, by,

\[
E(T_1) = (1 - p_1) \mu_1 \Gamma \left[ 1 + \frac{1}{\lambda_1} \right]
\]

\[
E(T_2) = (1 - p_2) \mu_2 \Gamma \left[ 1 + \frac{1}{\lambda_2} \right]
\]

where \( p_1 = \phi_{00} + \phi_{01} \) and \( p_2 = \phi_{00} + \phi_{10} \).

We get inferences for the vector parameter \( \phi_r = (\lambda_1, \lambda_2, \mu_1, \mu_2, \phi_{00}, \phi_{10}, \phi_{01}, \theta_r) \); where \( r = 1 \) (FGM copula), \( r = 2 \) (Gumbel copula), \( \phi_{00} + \phi_{10} + \phi_{01} + \phi_{11} = 1 \), \( \lambda_1 \) and \( \lambda_2 \) are shape parameters, \( \mu_1 \) and \( \mu_2 \) are scale parameters for the marginal Weibull distributions for \( T_1 \) and \( T_2 \) using Bayesian methods where samples of the joint posterior distribution of interest are simulated using MCMC (Markov Chain Monte Carlo) methods, as the popular Gibbs sampling algorithm Gelfand and Smith (1990); Casella and George (1992) and the Metropolis-Hastings algorithm. Chib and Greenberg (1995). Assuming prior independence among the parameters, we consider uniform or gamma prior distributions for the parameters \( \lambda_1, \lambda_2, \mu_1, \mu_2 \) and \( \theta_r, r = 1, 2 \).

A Dirichlet prior distribution is considered for the incidence parameters \( \phi_{11}, \phi_{10}, \phi_{01}, \) and \( \phi_{00}, \) where \( \phi_{11} + \phi_{10} + \phi_{01} + \phi_{00} = 1, \) with density generically given by,

\[
f(x_1, \ldots, x_{k-1}; w_1, \ldots, w_k) = \frac{1}{B(w)} \prod_{i=1}^{k} x_i^{w_i-1}
\]

for all \( x_1, \ldots, x_{k-1} > 0 \) satisfying \( x_1 + \cdots + x_{k-1} < 1 \), where \( x_k \) is given by \( 1 - x_1 - \cdots - x_{k-1} \). The normalizing constant is the multinomial beta function, which can be expressed in terms of the gamma function:

\[
B(w) = \frac{\prod_{i=1}^{k} \Gamma(w_i)}{\Gamma \left( \sum_{i=1}^{k} w_i \right)}.
\]

In the presence of covariates associated to each bivariate lifetime \( T_1 \) and \( T_2 \), we could assume the following regression model,

\[
\mu_{1i} = \alpha_{1} \exp \left( \beta_1' x_i \right)
\]

\[
\mu_{2i} = \alpha_{2} \exp \left( \beta_2' x_i \right)
\]

where \( x_i = (x_{1i}, \ldots, x_{si})' \) is a vector of covariates and \( \beta_l = (\beta_{1l}, \ldots, \beta_{sl}), \) \( l = 1, 2 \) denotes a vector regression parameters.

We also assume Normal distributions for the regression coefficients \( \beta_{sl}, \) and Gamma prior distributions for \( \alpha_l, l = 1, 2; s = 1, \ldots, k. \)

5. An Example
In this application, we consider a data set Huster et al. (1989) with 197 patients were a 50% random sample of the patients with “high-risk” diabetic retinopathy as defined by the Diabetic Retinopathy Study (DRS). 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 initiation of treatment to the time when visual acuity dropped below 5/200 two visits in a row (call it “blindness”). Thus there is a built-in lag time of approximately 6 months (visits were every 3 months). Survival times in this data set are therefore the actual time 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. This data set can be obtained in http://www.mayo.edu/research/documents/diabetes.html/DOC-10027460

In the analysis considered here, we assume as lifetimes, the times (in months) to blindness for the eye randomized to laser treatment ($T_1$), with 143 censored observations and 54 not-censored observations, and the time (in months) to blindness for the eye randomized that not received the treatment ($T_2$), with 96 censored observations and 101 not-censored observations. Two covariates, the age at diagnosis of diabetes (in years) and the type of diabetes (juvenile, adult) were considered. Observe that among the censored observations, some patients will never have the occurrence of the event of interest, that is, they are immunes.

As a preliminary analysis of the data, we present in Figure 1, the Kaplan-Meier nonparametric estimates for the survival functions considering the two lifetimes associated to each patient (treated and untreated eyes). From these plots, we have some indication of cure fraction (a large number of censoring times at the end of study). From the Kaplan-Meier plots of Figure 1, we also observe better results for the treated eyes, that is, larger times to blindness in comparison with the untreated eyes. From the Kaplan-Meier estimates we have estimates for the mean time to blindness: 57.83 months for the treated eyes and 43.53 months for the untreated eyes.

In medical research, it is common to assume independence between $T_1$ and $T_2$, which could be not appropriated.

![Figure 1: Kaplan-Meier nonparametric estimates for the survival function (0:censored; 1:blindness).](image-url)
As a first model, we assume a bivariate Weibull lifetime distribution derived from the FGM copula in presence of cure fraction presented in Section 3.1 with independent priors $\lambda_i \sim \text{Gamma}(0.001,0.001)$, $\mu_i \sim \text{Gamma}(0.001,0.001)$, $i = 1,2$, and $\theta_1 \sim U(-1,1)$, were $\text{Gamma}(a,b)$ denotes a gamma distribution with mean $a/b$ and variance $a/b^2$ and $U(c,d)$ denotes an uniform distribution defined in the interval $(c,d)$. We also assume a Dirichlet prior given by (38) for $\phi_{11}$, $\phi_{10}$, $\phi_{01}$ and $\phi_{00}$ with hyperparameter values $w_1 = w_2 = w_3 = w_4 = 1$ (noninformative priors for the incidence parameters).

Considering the rjags library Plummer (2011) in R software R Development Core Team (2011), a single chain has been used considering the simulation for each parameter for 200,000 times with a burn sample size of 50,000 to eliminate the possible effect of the initial values. We choose the simulated values from 100 to 100, to get approximated uncorrelated values which result in a final chain of size 2,000. We automatically generate initial values by sampling from the prior for the parameters, since the priors are not totally diffuse (vague). Convergence of the Gibbs Sampling algorithm was monitored using standard graphical methods, as the trace plots and the estimated posterior densities of the simulated samples Brooks and Gelman (1998). In Figure 2 we have an example considering the parameter $\lambda_1$ from the FGM copula (first fitted model), all the other parameters have similar behavior. Note we did not show the full conditional distributions since we consider the rjags library, making it unnecessary to use, however in appendix A we show the required codes to obtain the Bayesian estimates.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{trace_plot.png}
\caption{Trace plot and estimated posterior density of the simulated samples considering the parameter $\lambda_1$ from the FGM copula.}
\end{figure}

In Table 1, we have the posterior summaries of interest (Monte Carlo estimates of the posterior means for each parameter). We also have in Table 1, the Monte Carlo estimate of DIC (Deviance Information Criterion), introduced by Spiegelhalter et al. (2002), used as a discrimination criterion for different models. Smaller values of DIC indicates better models. The deviance can be expressed as,

$$D(\theta) = -2\log L(\theta | y) + c,$$

(41)
where $L(\theta \mid y)$ is the likelihood function for the unknown parameters in $\theta$ given the observed data $y$ and $c$ is a constant not required for comparing models. Spiegelhalter et al. (2002) defined the DIC criterion by,

$$\text{DIC} = D(\hat{\theta}) + 2n_D,$$

where $D(\theta)$ is the deviance evaluated at the posterior mean $\theta^*$ and $n_D$ is the effective number of parameters in the model, namely $n_D = D^* - D(\theta^*)$, where $D^* = E[D(\theta)]$ is the posterior deviance measuring the quality of the goodness-of-fit of the current model to the data.

In Table 1, we have the posterior summaries of interest assuming independent marginal Weibull distributions for the lifetimes $T_1$ and $T_2$ in presence of cure fraction, that is, with $\theta_1 = 0$, assuming the same priors for the parameters of interest and the same Gibbs sampling steps using the rjags library in R software used for the bivariate Weibull distribution derived from FGM copula in presence of cure fraction (results in Table 1).

From the obtained Monte Carlo estimates for DIC assuming the two models given in Tables 1 and 2 (presence or not of dependence between $T_1$ and $T_2$) we observe that the bivariate Weibull distribution derived from a FGM copula function gives better fit for the data, since the obtained DIC value is smaller for the dependent model than for the independent model.

Table 2: Posterior summaries (independent Weibull distributions in presence of cure fraction)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>S.D.</th>
<th>95% Credible Interval</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>0.98</td>
<td>0.17</td>
<td>(0.71;1.34)</td>
<td></td>
</tr>
</tbody>
</table>
Considering “FGM copula model” in the presence of cure fraction and the covariates age and type of diabetes (see (40)), that is,

$$
\mu_{1i} = \alpha_1 \exp(\beta_{11}x_{1i} + \beta_{12}x_{2i});
$$

$$
\mu_{2i} = \alpha_2 \exp(\beta_{21}x_{1i} + \beta_{22}x_{2i});
$$

where $i = 1,\ldots,n$ ($n = 197$) $x_{1i}$ denotes the age at diagnosis of diabetes (in years) and $x_{2i}$ is a binary variable denoting the type of diabetes ($1 =$ juvenile; $0 =$ adult).

We consider the same priors for the parameters, $\lambda_1$, $\lambda_2$, $\theta_{11}$, $\varphi_{10}$, $\varphi_{01}$ and $\varphi_{00}$, and the same Gibbs sampling steps using the rjags library in R software used for the bivariate Weibull distribution derived from FGM copula in presence of cure fraction. For the others parameters we consider independent priors $\alpha_l \sim \text{Gamma}(0.001,0.001)$, $\beta_{l1} \sim \text{N}(0,1000)$ and $\beta_{l2} \sim \text{N}(0,1000)$, $l = 1,2$, where $\text{N}(e,f)$ denotes a normal distribution with mean $e$ and variance $f$. In Table 3, we have the posterior summaries of interest assuming “FGM copula model” in the presence of cure fraction and the covariates age and type of diabetes.

From the results of Table 3, we observe that both covariates (age at diagnosis and type of diabetes) do not present significant effects on the response since the 95% credible intervals contain the zero value. In fact the Monte Carlo estimates of DIC are very similar considering the dependent model in presence or not of covariates. Considering the independent model, we observe a DIC value a little bit larger, an indication that the dependence model is better fitted by the data.

Table 3: Posterior summaries (“FGM copula model”, considering cure fraction and covariates)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>S.D.</th>
<th>95% Credible Interval</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1$</td>
<td>102.05</td>
<td>138.99</td>
<td>(6.20;514.62)</td>
<td>1669.76</td>
</tr>
</tbody>
</table>
In this case, we also could model the data considering other existing copula functions since the FGM copula is appropriate only to model weak dependence Nelsen (2006). Observe that the 95% credible interval of the dependence parameter $\theta_1$ includes the zero value, implying that $\theta_1$ is insignificant and it is enough to justify that $T_1$ and $T_2$ are not dependent at least through FGM copula function.

Now, let us assume a bivariate Weibull lifetime distribution derived from the Gumbel copula in presence of cure fraction presented in Section 3.2 with independent priors $\lambda_l \sim \text{Gamma}(0.001, 0.001)$, $\mu_l \sim \text{Gamma}(0.001, 0.001)$, $l = 1, 2$, and $\theta_2 \sim U(0, 1)$. We also assume a Dirichlet prior given by (38) for $\varphi_{11}$, $\varphi_{10}$, $\varphi_{01}$ and $\varphi_{00}$ with hyperparameter values $w_1 = w_2 = w_3 = w_4 = 1$ (noninformative priors for the incidence parameters).

Considering the rjags library Plummer (2011) in R software R Development Core Team (2011), a single chain has been used considering the simulation for each parameter for 210,000 times with a burn sample size of 50,000 to eliminate the possible effect of the initial values. We choose the simulated values from 100 to 100, to get approximated uncorrelated values which result in a final chain of size 2,000. Convergence of the Gibbs Sampling algorithm was monitored using standard graphical methods, as the trace plots of the simulated samples Brooks and Gelman (1998).

In Table 4, we have the posterior summaries of interest (Monte Carlo estimates of the posterior means for each parameter). We also have in Table 4, the Monte Carlo estimate of DIC (Deviance Information Criterion), introduced by Spiegelhalter et al. (2002), used as a discrimination criterion for different models. Smaller values of DIC indicates better models.
Table 4: Posterior summaries ("Gumbel copula model", considering cure fraction)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>S.D.</th>
<th>95% Credible Interval</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\lambda_1)</td>
<td>0.85</td>
<td>0.13</td>
<td>(0.63;1.13)</td>
<td></td>
</tr>
<tr>
<td>(\lambda_2)</td>
<td>0.94</td>
<td>0.09</td>
<td>(0.77;1.14)</td>
<td></td>
</tr>
<tr>
<td>(\mu_1)</td>
<td>62.14</td>
<td>23.27</td>
<td>(22.61;109.41)</td>
<td></td>
</tr>
<tr>
<td>(\mu_2)</td>
<td>28.99</td>
<td>5.36</td>
<td>(20.92;41.89)</td>
<td></td>
</tr>
<tr>
<td>(\theta_2)</td>
<td>0.55</td>
<td>0.20</td>
<td>(0.16;0.91)</td>
<td>1662.86</td>
</tr>
<tr>
<td>(\varphi_{11})</td>
<td>0.50</td>
<td>0.10</td>
<td>(0.29;0.67)</td>
<td></td>
</tr>
<tr>
<td>(\varphi_{10})</td>
<td>0.04</td>
<td>0.04</td>
<td>(0.001;0.14)</td>
<td></td>
</tr>
<tr>
<td>(\varphi_{01})</td>
<td>0.14</td>
<td>0.12</td>
<td>(0.005;0.41)</td>
<td></td>
</tr>
<tr>
<td>(\varphi_{00})</td>
<td>0.31</td>
<td>0.05</td>
<td>(0.21;0.40)</td>
<td></td>
</tr>
</tbody>
</table>

Finally, assuming “Gumbel copula model” in presence of cure fraction and covariates, we assumed the same regression model (43) and consider the same priors for the parameters, \(\lambda_1, \lambda_2, \theta_1, \varphi_{11}, \varphi_{10}, \varphi_{01}\) and \(\varphi_{00}\), and the same Gibbs sampling steps using the rjags library in R software used for the bivariate Weibull distribution derived from Gumbel copula in presence of cure fraction. For the others parameters we consider independent priors \(\alpha_l \sim \text{Gamma}(0.001,0.001)\), \(\beta_{1l} \sim \text{N}(0,1000)\) and \(\beta_{2l} \sim \text{N}(0,1000), l = 1,2\). In Table 5, we have the posterior summaries of interest assuming “Gumbel copula model” in the presence of cure fraction and the covariates age and type of diabetes. From the results of Table 5, we also conclude that both the covariates do not have significant effects on the response, but the dependent model derived from Gumbel copula gives smaller values of DIC, possibly an indication of better fit of the model for the data.

Table 5: Posterior summaries ("Gumbel copula model", considering cure fraction and covariates)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>S.D.</th>
<th>95% Credible Interval</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_1)</td>
<td>130.17</td>
<td>136.89</td>
<td>(11.42;505.52)</td>
<td></td>
</tr>
<tr>
<td>(\beta_{11})</td>
<td>-0.0001</td>
<td>0.02</td>
<td>(-0.05;0.05)</td>
<td></td>
</tr>
<tr>
<td>(\beta_{12})</td>
<td>-0.85</td>
<td>0.70</td>
<td>(-2.24;0.53)</td>
<td></td>
</tr>
<tr>
<td>(\alpha_2)</td>
<td>60.81</td>
<td>37.19</td>
<td>(19.91;156.39)</td>
<td></td>
</tr>
<tr>
<td>(\beta_{21})</td>
<td>-0.02</td>
<td>0.01</td>
<td>(-0.05;0.005)</td>
<td></td>
</tr>
<tr>
<td>(\beta_{22})</td>
<td>-0.15</td>
<td>0.43</td>
<td>(-1.01;0.67)</td>
<td></td>
</tr>
<tr>
<td>(\lambda_1)</td>
<td>0.88</td>
<td>0.14</td>
<td>(0.64;1.19)</td>
<td>1661.68</td>
</tr>
</tbody>
</table>
Bivariate Weibull Distributions Derived From Copula Functions In The Presence Of Cure Fraction And Censored Data

\[
\begin{array}{ccc}
\lambda_2 & 0.93 & 0.08 \quad (0.77;1.10) \\
\theta_2 & 0.53 & 0.20 \quad (0.13;0.90) \\
\phi_{11} & 0.49 & 0.10 \quad (0.31;0.66) \\
\phi_{10} & 0.03 & 0.03 \quad (0.0005;0.09) \\
\phi_{01} & 0.18 & 0.11 \quad (0.01;0.40) \\
\phi_{00} & 0.31 & 0.05 \quad (0.21;0.40)
\end{array}
\]

Since, in this case, it is not possible to compare the Kaplan-Meier curves with the estimated survival functions, we could use other discrimination methods to choose the best model for the data set. A possibility to discriminate the proposed models in terms of better fit for the data not considering the presence of covariates is to compare the Monte Carlo estimates of the means for the lifetimes \(T_1\) and \(T_2\) using the bivariate Weibull distributions derived from copula functions. Standard non-parametric estimates for the means based on Kaplan-Meier estimates are given, respectively, by \(E(T_1) = 57.83\) and \(E(T_2) = 43.53\) (use of the R software). Using the bivariate Weibull distributions derived from copula functions in presence of cure fraction, we get estimates for the means of \(T_1\) and \(T_2\) from (37), given by:

- FGM copula in presence of cure fraction not considering the presence of covariates: \(E(T_1) = 36.08\) and \(E(T_2) = 25.62\).
- Independent lifetimes with Weibull distributions in presence of cure fraction not considering the presence of covariates: \(E(T_1) = 22.42\) and \(E(T_2) = 23.87\).
- Gumbel copula in presence of cure fraction not considering the presence of covariates: \(E(T_1) = 36.87\) and \(E(T_2) = 19.28\).

From these results, we also observe that the dependent models based on the copula functions are better fitted by the data, as compared to the independent Weibull model, since the obtained Bayesian estimates using copulas for the means of the marginal lifetimes are more close to the Kaplan-Meier estimates.

6. Concluding Remarks

The use of copula functions could be a good alternative to analyse bivariate lifetime data in presence of censored data, cure fraction and covariates. Observe that in many applications of lifetime modeling we could have the presence of a cure fraction for individuals that are “long term survivors” or “cured individuals”.

In this paper, we have used two copula functions, but we also could consider any other existing copula function to build new bivariate lifetime models. We also have used marginal Weibull distributions, since this distribution has a good flexibility of fit for lifetime data, as
constant, increasing or decreasing hazard functions. Other usual existing univariate lifetime distributions could be used (Gamma, log-normal, exponential, exponentiated Weibull distributions, among many others) to model the marginal distributions. This approach could be of great interest in many areas of interest, especially in medical or engineering studies.

The use of standard existing MCMC methods, especially with the rjags library in R software, could give a great simplification to obtain inferences for the proposed models. Using a real data set example, we observed that using both proposed models (FGM or Gumbel copulas) give similar results when observing the Monte Carlo estimates for the posterior means and credible intervals for all parameters, except for the association parameters \( \theta_1 \) and \( \theta_2 \), which have different interpretations Nelsen (2006). A small improvement for the fit of the models for the data is observed considering “Gumbel copula model” where the DIC values are smaller than considering “FGM copula model” (see Tables 1–5). Observe that prior information on the association parameter of copula functions usually it is not available from preliminary studies.

It is interesting to observe that we have great difficulties to get standard classical inferences for the parameters of the proposed models. In our case, we tried to find maximum likelihood estimators for the parameters of the models in our application using SAS software, but we did not get good results. Other problem with the use of classical approach: confidence intervals and hypothesis tests are based on asymptotic results. In our case, these results could be not accurate even using large sample sizes.

It is important to point out that any existing parametric lifetime distribution could be considered for the lifetimes of the susceptible individuals by choosing different parametric forms for \( f_{10}(t_1) \), \( f_{01}(t_2) \), \( S_{10}(t_1) \) and \( S_{01}(t_2) \) used in Sections 3.1 and 3.2 in the analysis of lifetimes in presence of cure fraction and censored data.

References


Received May 15, 2015; accepted November 1, 2015.
A rjags Codes

Listing 1: Main code of the rjags library.

```r
model.jags <- function ()
{
  for (i in 1:n)
  {
    phi[i] <- -1[i]
    zeros[i] ~ dpols(phi[i])
    if1[i] <- log(lambda1) - log(mu1) + (lambda1 - 1) * (log(t1[i]) - log(mu1)) -
               (t1[i]/mu1) * lambda1
    if2[i] <- log(lambda2) - log(mu2) + (lambda2 - 1) * (log(t2[i]) - log(mu2)) -
               (t2[i]/mu2) * lambda2
    s1[i] <- exp(-(t1[i]/mu1) * lambda1)
    s2[i] <- exp(-(t2[i]/mu2) * lambda2)
    f1[i] <- 1 - s1[i]
    f2[i] <- 1 - s2[i]
    if12[i] <- log(phi11) + if1[i] + if2[i] + log(1 + theta1 * (1 - 2 * f1[i]) *
                                                                    (1 - 2 * f2[i]))
    ldS1[i] <- if1[i] + log(phi10 + phi11 * s2[i] - phi11 * theta1 * s1[i] * f2[i] *
                              (1 - 2 * f1[i]))
    ldS2[i] <- if2[i] + log(phi01 + phi11 * s1[i] - phi11 * theta1 * s1[i] * f1[i] *
                              (1 - 2 * f2[i]))
    A[i] <- s1[i] * s2[i] * (1 + theta1 * f1[i] * f2[i])
    lS12[i] <- log(phi00 + phi10 * s1[i] + phi01 * s2[i] + phi11 * A[i])
    l[i] <- delta1[i] * delta2[i] * if12[i] + delta1[i] * (1 - delta2[i]) * ldS1[i] +
           (1 - delta1[i]) * delta2[i] * ldS2[i] + (1 - delta1[i]) * (1 - delta2[i]) *
           lS12[i]
  }

lambda ~ dgamma(1.0E-3,1.0E-3)
mu1 ~ dgamma(1.0E-3,1.0E-3)
lambda2 ~ dgamma(1.0E-3,1.0E-3)
mu2 ~ dgamma(1.0E-3,1.0E-3)
theta1 ~ dunif(-1,1)
phi11 <- p[1]
phi10 <- p[2]
phi01 <- p[3]
phi00 <- p[4]
p[1:4] ~ ddirich(aux[])
}
```