Journal of Data Science,17(3). P. 445 - 480,2019
DOI:10.6339/JDS.201907_17(3).0001

USE OF GRAPHICAL METHODS IN THE DIAGNOSTIC OF PARAMETRIC PROBABILITY DISTRIBUTIONS FOR BIVARIATE LIFETIME DATA IN PRESENCE OF CENSORED DATA

Jorge Alberto Achcar[1], Jose Rafael Tovar Cuevas[2], Fernando A. Moala[3*]

[1]*Department of Social Medicine, University of Sao Paulo, Ribeirao Preto, Brazil.*

[2]*Department of Statistics, Universidad del Valle, Calli, Colombia.*

[3]*Department of Statistics, UNESP - S~ao Paulo State University, Brazil.*

Abstract

The choice of an appropriate bivariate parametrical probability distribution for pairs of lifetime data in presence of censored observations usually is not a simple task in many applications. Each existing bivariate lifetime probability distribution proposed in the literature has different dependence structure. Commonly existing classical or Bayesian discrimination methods could be used to discriminate the best among different proposed distributions, but these techniques could not be appropriate to say that we have good fit of some particular model to the data set. In this paper, we explore a recent dependence measure for bivariate data introduced in the literature to propose a graphical and simple criterion to choose an appropriate bivariate lifetime distribution for data in presence of censored data.

Keywords: bivariate lifetime, Bayesian approach, censoring data, copula functions, diagnostic discrimination methods.

* Corresponding author
Email: f.moala@unesp.br

1. Introduction

In many different studies such as in medicine, engineering, economy, ecology, among several others, we have bivariate data that usually have some dependence structure (see for example, Davarzani et al. (2015) and Icuma et al (2016). To model this dependence there are many statistical models introduced in the literature, as for example, frailty models and copula function models (see for example, Goethals et al., 2008). In lifetime data applications, we usually have the presence of censored data (see for example, Wang (2011), Martinez et al. (2016) and Karvanen et al (2010). In the analysis of right censored univariate lifetime data, the non-parametric product-limit estimator proposed by Kaplan and Meier (1958) for the survival function $S(t) = P(T > t)$, where T is a random variable denoting a lifetime and t is a fixed value, usually is the first step in any lifetime statistical analysis, as in medical or engineering applications (see for example, Lawless,1982; or Klein and Moeschberger,1995). In this situation, it is common the use of graphical methods to choose an appropriate parametrical distribution by comparing the fitted survival function with plots of the Kaplan-Meier product-limit estimates for the survival function. For the situation of bivariate lifetimes T_1 and T_2 in presence of censored observations, different nonparametric estimators for the joint survival function $S(t_1, t_2) = P(T_1 > t_1, T_2 > t_2)$ were proposed in the literature (see for example, Campbell, 1981; Campbell and Foldes, 1982; Langberg and Shaked, 1982; Hanley and Parnes, 1983; Tsai, Leurgans and Crowley, 1986; Dabrowska, 1988; Prentice and Cai, 1992 or Puritt, 1993).

In general, the nonparametric estimates for the bivariate survival function introduced in the literature have great complexities to be used in practical situations.

A special case that occurs in some applications is given when the two lifetimes are subject to independent censorship by a single censoring variable (univariate censorship).

In this situation, Lin and Ying (1993) introduced a nonparametric estimator of the bivariate survival function in presence of censored data with nice statistical properties (strong consistency, weakly convergence to zero-mean Gaussian process) that is simple to be applied for bivariate lifetime data analysis in presence of univariate censoring.

In this paper, it is proposed the use of the Lin-Ying nonparametric estimator for the joint survival function in presence of univariate censorship in a new graphical diagnostic approch to compare different parametrical bivariate lifetime models exploring a recent dependence measure introduced by Ledwina (2015) for two random variables.

In this new graphical methodology, the LY nonparametric estimator of the joint survival function $S(t_1, t_2)$ in presence of univariate censoring and the Kaplan and Meier nonparametric estimators for the marginal survival functions $S_1(t_1)$ and $S_2(t_2)$ are used in the Ledwina dependence measure.

Three bivariate lifetime parametric models are considered in this study in presence of censored data: the Block and Basu (1974) exponential distribution and two bivariate Weibull distributions derived from copula functions (see for example, Flores, 2016). The main goal of this paper is to introduce a simple graphical method to choose an appropriate bivariate lifetime parametrical distribution based on plots of the empirical Ledwina measure of dependence.

The paper is organized as follows: in section 2, it is presented the Lin and Ying estimator; in section 3, it is introduced the Ledwina dependence measure; in section 4, it is introduced the Block and Basu bivariate distribution; in section 5, it is introduced some special copula functions; section 6 presents five applications; finally section 7 presents some concluding remark.

2. The Lin and Ying estimator

The Lin-Ying (1993) non-parametric estimator simply denoted as the LY estimator for the joint survival function is based on the idea that the bivariate survival function of paired survival times can be expressed as the ratio of the bivariate at-risk probability to the survival function of the censoring time.

The LY estimator of the joint survival function $S(t_1, t_2)$ in presence of univariate censoring is given as follows: let $(T_{1i}, T_{2i}), i = 1,2,\dots,n$ be bivariate pairs of lifetimes and C_i, i = 1,…, n be identically distributed censoring times with survival function $G(c) = P(C > c)$. Define the indicator variables,

$$d_{1i} = \begin{cases} 1 & \text{if } T_{1i} \le C_i \\ 0 & \text{if } T_{1i} > C_i \end{cases} \text{ and } d_{2i} = \begin{cases} 1 & \text{if } T_{2i} \le C_i \\ 0 & \text{if } T_{2i} > C_i \end{cases} \quad (2.1)$$

Also define, $U_i = min(T_{1i}, C_i)$ and $V_i = min(T_{2i}, C_i)$, $i = 1,2,\dots,n$. In this way, a nonparametric estimator of $S(t_1, t_2) = P(T_1 > t_1, T_2 > t_2) = \frac{P(U>t_1,V>t_2)}{G(z)}$ where $z = max(t_1, t_2)$ is obtained using the following results:

(i) $P_n(U > t_1, V > t_2) = \frac{\sum_{i=1}^{n} I(U>t_1,V>t_2)}{n}$(empirical survival function)

(ii) To get an estimator of $G(\cdot)$ it is observed that in this case the data are given by $(R_i, \delta c_i)$, $i = 1,2,\dots,n$ where, $R_i = min(C_i, Z_i)$, $Z_i = max(T_{1i}, T_{2i})$ or $R_i = max(U_i, V_i)$, and $\delta c_i = I[C_i \le Z_i] = 1 - (\delta_{1i})(\delta_{2i})$.

In this case it is used the Kaplan-Meier product-limit estimator. Let $\{c_k; k \ge 1\}$ be the ordered sequence of distinct time points where the censoring occurs, and let d_k be the number of censored observations, at c_k. Also, let $n_k = \sum_{i=1}^{n} I(R_i \ge c_k)$ Thus,

$$\hat{G}(t) = \prod_{k:c_k<t} \frac{n_k - d_k}{n_k} \quad (2.2)$$

(iii) In this way, the LY estimator of the joint survival function is given by,

$$\hat{S}(t_1, t_2) = P(T_1 > t_1, T_2 > t_2) = \frac{\sum_{i=1}^{n} I(U_i > t_1, V_i > t_2)}{n\hat{G}(z)} \quad (2.3)$$

where $z = max(t_1, t_2)$.

Some remarks:

(i) The LY estimator is simpler than any other nonparametric estimator for the bivariate survival function introduced in the literature and it is reduced to the usual empirical survival function in the absence of censoring.

(ii) The LY estimator (2.3)is a natural generalization of the univariate product-limit (Kaplan-Meier) estimator.

(iii) From (2.3) it is possible to get a nonparametric estimator for the joint cumulative distribution function from the result, $F(t_1, t_2) = P(T_1 \leq t_1, T_2 \leq t_2) = 1 - \left(P(T_1 > t_1) + P(T_2 > t_2)\right) + P(T_1 > t_1, T_2 > t_2)$.

3. The Ledwina dependence measure

It is important to point out that in applications usually there are great difficulties to decide by an appropriate bivariate lifetime model which adequately models the strength of dependence between two random variables. Dependence scalar measure indexes or global measures of dependence for two random variables were studied by many authors (see for example, Jogdeo, 1982; Lancaster, 1982; Drouet and Kotz, 2001; Balakrishnan and Lai, 2009). Usually the use of scalar dependence bivariate measures could be not the best way to represent complex dependence structures and other dependence measures were introduced in the literature, as local indexes (see, for example, Drouet and Kotz, 2001; Kowalczyk and Pleszczynska, 1997; Bjerve and Doksum, 1993 and Bairamov et al., 2003).

In a recent paper, Ledwina (2015) introduced a new function valued measure of dependence of two random variables T_1 and T_2 and presented the main properties of this measure. The proposed measure has simple form and its definition explores only cumulative distribution functions.

The measure takes values in the interval $[-1,1]$ and treats both variables in a symmetrical way. The measure preserves the correlation order, or equivalently the concordance order, which is the quadrant order restricted to the class of distributions with fixed marginals.

The Ledwina dependence measure in the general case assuming two random variables T_1 and T_2 is given by,

$$q(t_1, t_2) = \frac{F(t_1, t_2) - F_1(t_1)F_2(t_2)}{w(t_1, t_2)}, \text{ for } (t_1, t_2) \in D \tag{3.1}$$

where $\omega(t_1, t_2) = \left[F_1(t_1)F_2(t_2)\left(1 - F_1(t_1)\right)\left(1 - F_2(t_2)\right)\right]^{\frac{1}{2}}$, $D = \left\{(t_1, \ t_2), 0 < F_1(t_1) < 1, 0 < F_2(t_2) < 1\right\}$, $F_1(t_1)$ and $F_2(t_2)$ are the marginal distribution functions for T_1 and T_2, that is , $F_1(t_1) = P(T_1 \leq t_1)$, $F_2(t_2) = P(T_2 \leq t_2)$; $\mathrm{F}(t_1, t_2)$ is the joint distribution function for T_1 and T_2 given by $\mathrm{F}(t_1, t_2) = P(T_1 \leq t_1, T_2 \leq t_2)$.

From (3.1)it is seen that q treats both variables T_1 and T_2 symmetrically and a knowledge of q and the marginal distributions allows one to recover the joint distribution function for T_1 and T_2 (for properties of q and more details, see Ledwina, 2015). Observe that $-1 < q(t_1, t_2) < 1$.

An empirical simple estimator for the Ledwina dependence (3.1) considering the bivariate data set $(t_{1\mathrm{i}}, t_{2\mathrm{i}})$, $i = 1,2, \ldots, n$ not in the presence of censored data is obtained replacing $\mathrm{F}(t_1, t_2)$,$F_1(t_1)$ and $F_2(t_2)$ by their respective empirical estimators,

$$F_n(t_1,t_2)=\frac{\sum_{i=1}^{n} I(T_1\le t_1, T_2\le t_2)}{n}$$

$$F_n(t_1)=\frac{\sum_{i=1}^{n} I(T_1\le t_1)}{n} \qquad (3.2)$$

$$F_n(t_2)=\frac{\sum_{i=1}^{n} I(T_2\le t_2)}{n}$$

for fixed values t_1 and t_1 , where $I(\cdot)$ is the indicator function.

Similarly, we could consider the survival functions $S(t_1,t_2)$, $S_1(t_1)$ and $S_2(t_2)$ in place of $F(t_1,t_2)$, $F_1(t_1)$ and $F_2(t_2)$ in (3.1) obtaining a dependence measure given by,

$$r(t_1,t_2)=\frac{S(t_1,t_2)-S_1(t_1)S_2(t_2)}{w(t_1,t_2)},\ \text{for}\ (t_1,t_2)\in D \qquad (3.3)$$

where $\omega(t_1,t_2)=\left[S_1(t_1)S_2(t_2)\big(1-S_1(t_1)\big)\big(1-S_2(t_2)\big)\right]^{\frac{1}{2}}$,$D=\{(t_1,\ t_2), 0<S_1(t_1)<1, 0<S_2(t_2)<1\}$.

An empirical estimator of $r(t_1,t_2)$ is given replacing $S(t_1,t_2)$, $S_1(t_1)$ and $S_2(t_2)$ by the empirical estimators,

$$S_n(t_1,t_2)=\frac{\sum_{i=1}^{n} I(T_1> t_1, T_2> t_2)}{n}$$

$$S_n(t_1)=\frac{\sum_{i=1}^{n} I(T_1> t_1)}{n} \qquad (3.4)$$

$$S_n(t_2)=\frac{\sum_{i=1}^{n} I(T_2> t_2)}{n}$$

4. The Block and Basu distribution

To analyse bivariate lifetimes it is possible to assume different parametric distributions introduced in the literature (see for example, Freund, 1961; Marshall and Olkin, 1967; Sarkar, 1987; Downton, 1970; Gumbel, 1960; Hawkes, 1972; Hougaard, 1986; Arnold and Strauss, 1988). Among all these bivariate lifetime distributions, one model has been very well explored in the literature: the Block and Basu (1974) bivariate exponential distribution, denoted as the BB distribution. The BB distribution with parameters , λ_1 , λ_1 and λ_3 for the lifetimes T_1 and T_2 has a joint density function given by,

$$f(t_1,t_2)=\begin{cases}\dfrac{\lambda\lambda_1\lambda_{23}}{\lambda_{12}}\exp\{-\lambda_1 t_1-\lambda_{23}t_2\} & \text{if } t_1<t_2\\[2ex] \dfrac{\lambda\lambda_2\lambda_{13}}{\lambda_{12}}\exp\{-\lambda_{13} t_1-\lambda_{2}t_2\} & \text{if } t_1\ge t_2\end{cases} \qquad (4.1)$$

where $\lambda_{12} = \lambda_1 + \lambda_2$, $\lambda_{13} = \lambda_1 + \lambda_3$, $\lambda_{23} = \lambda_2 + \lambda_3$ and $\lambda = \lambda_1 + \lambda_2 + \lambda_3$.

The joint survival function for the BB distribution is given by,

$$S(t_1,t_2) = P\{T_1 > t_1, T_2 > t_2\} = \begin{cases} S_1(t_1,t_2) & \text{if } t_1 < t_2 \\ S_2(t_1,t_2) & \text{if } t_1 > t_2 \end{cases} \tag{4.2}$$

where

$$S_1(t_1,t_2) = \frac{\lambda}{\lambda_{12}} \exp\{-\lambda_1 t_1 - \lambda_{23} t_2\} - \frac{\lambda_3}{\lambda_{12}} \exp\{-\lambda_2 t_2\}$$

and

$$S_2(t_1,t_2) = \frac{\lambda}{\lambda_{12}} \exp\{-\lambda_{13} t_1 - \lambda_2 t_2\} - \frac{\lambda_3}{\lambda_{12}} \exp\{-\lambda_1 t_1\}$$

Considering the BB bivariate exponential distribution, the marginal distributions for T_1 and T_2 are mixtures of exponentials having an absolutely continuous joint distribution with positive correlation coefficient. Bayesian analysis for bivariate lifetime distributions has been considered as a good alternative to get the inferences of interest when compared to the use of standard classical inference methods, in terms of better accuracy for the obtained inferences of interest. Bayesian inference is based on the Bayes formula, where the posterior distribution is obtained combining the likelihood function with the prior distributions for the parameters of the model. With the use of existing simulation methods, it is possible to get accurate Monte Carlo estimates for all parameters of interest. In this way, the use of MCMC (Markov Chain Monte Carlo) methods (see for example, Gelfand and Smith, 1990 or Chib and Greenberg, 1995) gives a good alternative to obtain accurate posterior summaries of interest.

Bayesian approaches for the BB lifetime distribution using MCMC methods in presence or not of covariates and censored data are presented by many authors (see for example, Achcar and Leandro, 1998; or Achcar and Santos, 2011).

4.1 The likelihood function

Suppose either T_1 or T_2 can be censored and that censoring is independent of the lifetimes. Let us subdivide the n observations into four cases:

C_1: both t_{1i} and t_{2i} are observed lifetimes, $i = 1,2,\ldots,n$;

C_2 t_{1i} is a lifetime and t_{2i} is a censoring time (that is, we only know that $T_{2\mathrm{i}} \geq t_{2i}$);

C_3: t_{1i} is a censoring time and t_{2i} is a lifetime;

C_4: both t_{1i} and t_{2i} are censoring times.

The likelihood function for a continuous model (see for example, Lawless, 1982, page 479) is given by,

$$L = \prod_{i\in C_1} f(t_{1i},t_{2i}) \prod_{i\in C_2}\left(-\frac{\partial S(t_{1i},t_{2i})}{\partial t_{1i}}\right) \prod_{i\in C_3}\left(-\frac{\partial S(t_{1i},t_{2i})}{\partial t_{2i}}\right) \prod_{i\in C_4} S(t_{1i},t_{2i}) \tag{4.3}$$

Assuming the BB bivariate exponential distribution with density function (4.1) and survival function (4.2), we have,

$$-\frac{\partial S\left(t_{1i},t_{2i}\right)}{\partial t_{1i}}=\begin{cases} S'_{1t_1}\left(t_{1i},t_{2i}\right) & \text{if } t_{1i}<t_{2i} \\ S'_{2t_1}\left(t_{1i},t_{2i}\right) & \text{if } t_{1i}\geq t_{2i} \end{cases} \tag{4.4}$$

where

$$S'_{1t_1}\left(t_{1i},t_{2i}\right)=\frac{\lambda\lambda_1}{\lambda_{12}}\exp\left\{-\lambda_1 t_{1i}-\lambda_{23}t_{2i}\right\}$$

and

$$S'_{2t_1}\left(t_{1i},t_{2i}\right)=\frac{\lambda\lambda_{13}}{\lambda_{12}}\exp\left\{-\lambda_{13}t_{1i}-\lambda_2 t_{2i}\right\}-\frac{\lambda\lambda_3}{\lambda_{12}}\exp\left\{-\lambda_1 t_{1i}\right\}$$

Also,

$$-\frac{\partial S\left(t_{1i},t_{2i}\right)}{\partial t_{2i}}=\begin{cases} S'_{1t_2}\left(t_{1i},t_{2i}\right) & \text{if } t_{1i}<t_{2i} \\ S'_{2t_2}\left(t_{1i},t_{2i}\right) & \text{if } t_{1i}\geq t_{2i} \end{cases} \tag{4.5}$$

where

$$S'_{1t_2}\left(t_{1i},t_{2i}\right)=\frac{\lambda\lambda_{23}}{\lambda_{12}}\exp\left\{-\lambda_1 t_{1i}-\lambda_{23}t_{2i}\right\}-\frac{\lambda\lambda_3}{\lambda_{12}}\exp\left\{-\lambda_2 t_{2i}\right\}$$

and

$$S'_{2t_2}\left(t_{1i},t_{2i}\right)=\frac{\lambda\lambda_2}{\lambda_{12}}\exp\left\{-\lambda_{13}t_{1i}-\lambda_2 t_{2i}\right\}$$

Let us define the indicator variables δ_{1i}, $\delta_{2i}$and v_i, by,

$$d_{ji}=\begin{cases} 1 & \text{if } t_{ji} \text{ is an observed lifetime} \\ 0 & \text{if } t_{ji} \text{ is a censored observation} \end{cases} \tag{4.6}$$

for $j=1,2;\ i=1,\ldots,n$ and

$$v_i=\begin{cases} 1 & \text{if } t_{1i}<t_{2i} \\ 0 & \text{if } t_{1i}\geq t_{2i} \end{cases} \tag{4.7}$$

In this way, the log-likelihood function for λ_1, λ_2 and λ_3 is given by,

$$l\left(\lambda_1,\lambda_2,\lambda_3\right)=\log L\left(\lambda_1,\lambda_2,\lambda_3\right)=\sum_{i=1}^{n}v_i\delta_{1i}\delta_{2i}\log f_1\left(t_{1i},t_{2i}\right)+\sum_{i=1}^{n}\left(1-v_i\right)\delta_{1i}\delta_{2i}iog$$

$$f_2\left(t_{1i},t_{2i}\right)+\sum_{i=1}^{n}v_i\delta_{1i}\left(1-\delta_{2i}\right)\log S'_{1t_1}\left(t_{1i},t_{2i}\right)+\sum_{i=1}^{n}\left(1-v_i\right)\delta_{1i}\left(1-\delta_{2i}\right)\log S'_{2t_1}\left(t_{1i},t_{2i}\right)+$$

$$\sum_{i=1}^{n}v_i\left(1-\delta_{1i}\right)\delta_{2i}\log S'_{1t_2}\left(t_{1i},t_{2i}\right)+\sum_{i=1}^{n}\left(1-v_i\right)\left(1-\delta_{1i}\right)\delta_{2i}\log S'_{2t_2}\left(t_{1i},t_{2i}\right)+\sum_{i=1}^{n}v_i\left(1-\delta_{1i}\right)$$

$$\left(1-\delta_{2i}\right)\log S_1\left(t_{1i},t_{2i}\right)+\sum_{i=1}^{n}\left(1-v_i\right)\left(1-\delta_{1i}\right)\left(1-\delta_{2i}\right)\log S_2\left(t_{1i},t_{2i}\right) \tag{4.8}$$

For a Bayesian analysis of the BB distribution in the presence of censored observations, let us assume independent gamma distributions for the parameters λ_k, that is

$$\lambda_k \sim \text{Gamma}\left[a_k, b_k\right] \tag{4.9}$$

for $k = 1,2,3$; a_k and b_k are known hyperparameters; $Gamma[a_k, b_k]$ denotes a gamma distribution with mean $\frac{a_k}{b_k}$ and variance $\frac{a_k}{b_k^2}$.

The joint posterior distribution for λ_1, λ_2 and λ_3 is given by,

$$\pi\left(\lambda_1, \lambda_2, \lambda_3 \mid t_1, t_2\right) \propto \left(\prod_{k=1}^{3} \lambda_k^{a_k-1} e^{-b_k \lambda_k}\right) L\left(\lambda_1, \lambda_2, \lambda_3\right) \tag{4.10}$$

where $L(\lambda_1, \lambda_2, \lambda_3)$ is the likelihood function and $t_1$and t_2 are the vectors of observed data.

To simulate samples for joint posterior distribution (4.10), it is used MCMC methods.

In this way, it is simulated samples for the joint posterior distribution (4.10), from the full posterior conditional distributions $\pi(\lambda_1|\lambda_2, \lambda_3, t_1, t_2)$, $\pi(\lambda_2|\lambda_1, \lambda_3, t_1, t_2)$ and $\pi(\lambda_3|\lambda_1, \lambda_2, t_1, t_2)$using the Gibbs sampling algorithm or the Metropolis-Hastings algorithm when the full posterior conditional distributions have unknown forms.

5. Use of copula functions

The use of copula functions (Sklar, 1959) is becoming very popular in the construction of new parametrical multivariate distributions (see for example, Durante and Sempi, 2015 or Nelsen, 2006). Copula functions can be used to link marginal distributions with a joint distribution.

For specified univariate marginal distribution functions $F_1(t_1), F_2(t_2), \ldots, F_{\mathrm{m}}(t_{\mathrm{m}})$, the function,

$$C\left(F_1\left(t_1\right), F_2\left(t_2\right), \ldots, F_m\left(t_m\right)\right) = F\left(t_1, t_2, \ldots, t_m\right) \tag{5.1}$$

which is defined using a copula function C, results in a multivariate distribution function with univariate marginal distributions specified as $F_1(t_1)$,$F_2(t_2), \ldots, F_{\mathrm{m}}(t_{\mathrm{m}})$.

It is important to point out that any multivariate distribution function F can be written in the form of a copula function (Sklar, 1959); that is, if $F(t_1, t_2, \ldots, t_m)$ is a joint multivariate distribution function with univariate marginal distribution functions $F_1(t_1)$,$F_2(t_2), \ldots, F_{\mathrm{m}}(t_{\mathrm{m}})$, thus there exists a copula function $C(U_1, U_2, \ldots, U_m)$ such that,

$$F\left(t_1, t_2, \ldots, t_m\right) = C\left(F_1\left(t_1\right), F_2\left(t_2\right), \ldots, F_m\left(t_m\right)\right) \tag{5.2}$$

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 an 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 terms of the dependence structures.

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 $U_1 = F_1(t_1)$ and $U_2 = F_2(t_2)$ where U_1 and U_2 have uniform

$(0,1)$ distributions, but are usually dependent if T_1 and T_2 are dependent (T_1 and T_2 independent, implies that U_1 and $U_2$are independent).

Specifying dependence between T_1 and $T_2$is the same as specifying dependence between U_1 and U_2. Assuming U_1 and U_2 uniform random variables, the problem reduces to specifying a bivariate distribution between two uniforms, that is, a copula.

In this section, we introduce two special copula functions which are explored in our study: the Gumbel-Barnett and the Farlie-Gumbel-Morgenstern copulas. In all cases, $u = F_1(t_1) = P\ (T_1 \leq t_1),\ v = F_2(t_2) = P\ (T_2 \leq t_2)$; $C(u,\ v) = F\ (t_1, t_2)\ =\ P\ (T_1 \leq t_1, T_2 \leq t_2)$ and θ is the dependence parameter.

Also, let us assume the marginal Weibull distributions given by,

$$u = F_1(t_1) = 1 - \exp\left(-\left(\frac{t_1}{\lambda_1}\right)^{p_1}\right) \tag{5.3}$$

and

$$v = F_2(t_2) = 1 - \exp\left(-\left(\frac{t_2}{\lambda_2}\right)^{p_2}\right) \tag{5.4}$$

where p_1 and p_2 are the shape parameters of the Weibull distributions.

A copula based measure of dependence is also introduced by Ledwina(2015) given by,

$$q(u,v) = \frac{C(u,\mathrm{v}) - u\mathrm{v}}{w(u,\mathrm{v})}$$

where $w(u,v)\ =\ [uv(1\ -\ u)(1\ -\ v)]^{0.5}$ and $B_1(u,v) < q(u,v) < B_2(u,v)$; $B_1(u,v) = w(u,v)[max(u+v-1,0) - uv]$; $B_2(u,v)\ =\ w(u,v)[min(u,v) - uv]$.

5.1 Gumbel-Barnett copula (GB copula)

The Gumbel-Barnett copula (Gumbel,1960; Barnett, 1980) is defined by,

$$C(u,v) = u + v - 1 + (1-u)(1-v)\exp\{-\theta \ln(1-u)\ln(1-v)\} \tag{5.5}$$

for $0 \leq \theta \leq 1$. Independence corresponds to $\theta = 0$.

The joint density function $f(t_1, t_2)$ for the random variables T_1 and T_2 is obtained from the second derivative of the joint distribution function with respect to $t_1\ and\ t_2$, that is,

$$f(t_1,t_2) = f_1(t_1)\, f_2(t_2) \exp\{-\theta \log(1-F_1(t_1)) \log(1-F_2(t_2))\}\{1-\theta-\theta\log(1-F_1(t_1)) - \theta\log(1-F_2(t_2)) + \theta^2 \log(1-F_1(t_1))\log(1-F_2(t_2))\} \tag{5.6}$$

(

Also,

$$F(t_1,t_2) = F_1(t_1) + F_2(t_2) - 1 + (1-F_1(t_1))(1-F_2(t_2))\exp\{-\theta \ln(1-F_1(t_1))\ln(1-F_2(t_2))\} \tag{5.7}$$

From (5.7), we have,

$$F_I\left(t_1,t_2 \mid \lambda_1,\lambda_2,p_1,p_2,\theta\right)=1-\exp\left(-\left(\frac{t_1}{\lambda_1}\right)^{p_1}\right)-\exp\left(-\left(\frac{t_2}{\lambda_2}\right)^{p_2}\right)+\exp\left\{-\left(\frac{t_1}{\lambda_1}\right)^{p_1}-\left(\frac{t_2}{\lambda_2}\right)^{p_2}-\theta\left(\frac{t_1}{\lambda_1}\right)^{p_1}\left(\frac{t_2}{\lambda_2}\right)^{p_2}\right\} \quad (5.8)$$

where $t_1 > 0$ and $t_2 > 0$.

Observe that if $\theta = 0$, we have independent random variables. Also, observe that the joint bivariate survival function is given by,

$$S\left(t_1,t_2\right)=1-F_1\left(t_1\right)-F_2\left(t_2\right)+F\left(t_1,t_2\right) \quad (5.9)$$

where $F_1(t_1)$and $F_2(t_2)$ are given by (5.4)and $F(t_1, t_2)$is given by (5.8), that is,

$$S_I\left(t_1,t_2\right)=\exp\left\{-\left(\frac{t_1}{\lambda_1}\right)^{p_1}-\left(\frac{t_2}{\lambda_2}\right)^{p_2}-\theta\left(\frac{t_1}{\lambda_1}\right)^{p_1}\left(\frac{t_2}{\lambda_2}\right)^{p_2}\right\} \quad (5.10)$$

The Gumbel-Barnett copula model is denoted as ”GB model”.

5.2 Farlie-Gumbel-Morgenstern copula (FGM copula)

The Farlie-Gumbel-Morgenstern copula (Morgenstern, 1956; Nelsen, 2006) is defined by,

$$C(u,v)=uv[1+\theta(1-u)(1-v)] \quad (5.11)$$

for $-1 \leq \theta \leq 1$. In other words, the copula C captures the dependence structure between T_1 and T_2. Independence corresponds to $\theta = 0$.

The joint density function $f(t_1, t_2)$for the random variables T_1 and T_2 is given by,

$$f(t_1,t_2) = f_1(t_1)f_2(t_2)[1 + \theta f_1(t_1)f_2(t_2)(1 - 2F_1(t_1))\,(1 - 2F_2(t_2))]$$

From(5.4) and(5.11) we have,

$$F_{II}\left(t_1,t_2 \mid \lambda_1,\lambda_2,p_1,p_2,\theta\right)=\left[1-\exp\left(-\left(\frac{t_1}{\lambda_1}\right)^{p_1}\right)\right]\left[1-\exp\left(-\left(\frac{t_2}{\lambda_2}\right)^{p_2}\right)\right][1+\theta\exp\{-\left(\frac{t_1}{\lambda_1}\right)^{p_1}-\left(\frac{t_2}{\lambda_2}\right)^{p_2}\}] \quad (5.13)$$

The joint survival function is given by,

$$S_{II}\left(t_1,t_2 \mid \lambda_1,\lambda_2,p_1,p_2,\theta\right)=\exp\left\{-\left(\frac{t_1}{\lambda_1}\right)^{p_1}-\left(\frac{t_2}{\lambda_2}\right)^{p_2}\right\}\left\{1+\theta\left[1-\exp\left(-\left(\frac{t_1}{\lambda_1}\right)^{p_1}\right)\right]\right.$$

$$\left[1-\exp\left(-\left(\frac{t_2}{\lambda_2}\right)^{p_2}\right)\right]\} \tag{5.14}$$

The Farlie-Gumbel-Morgenstern copula model is denoted as "FGM model".

The likelihood function for λ_1, λ_2, p_1, p_1, and θ assuming GB or FGM copula models with marginal Weibull distributions is obtained from(4.3).

6. Applications

In the estimation of the Ledwina dependence measure (3.3) considering each proposed parametrical distribution function, we use Bayesian methods. As we have posterior distributions with no closed form, we use the OpenBugs software (Spiegelhalter et al, 2003) to obtain the Bayesian estimates of the parameters (use of Markov Chain Monte Carlo or MCMC methods, see for example Gelfand and Smith, 1990 or Chib and Greenberg, 1995). For all models, 15,000 Gibbs samples were simulated from the conditional distributions for each parameter. From these generated samples, we discarded the first 5,000 samples to eliminate the effect of the initial values considering a spacing of size 100 to get the final simulated sample. Convergence of the algorithm was verified graphically.

6.1 Application 1: Holt and Prentice (1974) data set

In Table 1, it is presented a bivariate lifetime data set introduced by Holt and Prentice (1974) that corresponds to the survival times in days of closely (T_1) and poorly (T_2) matched skin grafts on the same burn patient (n = 11). Table 1 also presents the censoring indicators for the two lifetimes where δ = 1 for a complete observation and δ = 0 for a censored observation.

Table 1: Holt and Prentice (1974) data set

Obs	t_1	t_2	δ_1	δ_2
1	37	29	1	1
2	19	13	1	1
3	57	15	0	1
4	93	26	1	1
5	16	11	1	1
6	22	17	1	1
7	20	26	1	1
8	18	21	1	1
9	63	43	1	1
10	29	15	1	1

In Table 2, it is presented the posterior summaries of interest using the three parametrical models presented in sections 4 and 5. It is assumed $\text{Gamma}(1,100)$ prior distributions for the parameters λ_k, $k = 1, 2$ of the BB distribution, uniform $U(0,500)$ prior distributions for the parameters λ_k, $k = 1, 2$ of the Weibull marginal distributions and $U(0,5)$prior distributions for the shape parameters of the marginal Weibull distributions assuming the GB and FGM copula models. Furthermore, let us assume the prior distributions $U[0, 1]$ and $U[0, 0.2]$ for the parameter θ in the GB and FGM models, respectively.

In Table 3, it is presented the LY non-parametrical estimator of the joint survival function and the Bayesian estimators of the joint survival function assuming the BB, GB and FGM models using the Bayesian estimators given in Table 2. In Table 3, it is also presented the Kaplan-Meier (KM) estimates for the marginal survival functions of the lifetimes T_1 and T_2.

Table 2: Posterior means and posterior standard deviations (in parenthesis) for the parameters of the three models

model	λ_1		λ_2		λ_3
BB	0.006 (0.004)		0.018 (0.010)		0.025 (0.010)
Copula	λ_1	λ_2	p_1	p_2	θ
GB	63.08 (23.51)	27.68 (4.201)	1.427 (0.424)	2.304 (0.548)	0.191 (0.174)
FGM	57.14 (18.48)	27.41 (3.757)	1.535 (0.396)	2.427 (0.536)	0.114 (0.055)

Table 3: Estimators for the joint survival function $S(t_1, t_2)$

pair	t_1	t_2	LY	KM $S(t_1)$	KM $S(t_2)$	BB	GB	FGM
1	16	11	0.909	0.909	0.909	0.544	0.768	0.779
2	19	13	0.727	0.727	0.818	0.486	0.696	0.708
3	18	21	0.454	0.818	0.454	0.368	0.490	0.503
4	22	17	0.363	0.545	0.545	0.405	0.570	0.583
5	20	26	0.272	0.636	0.272	0.298	0.335	0.343
6	29	15	0.363	0.454	0.636	0.383	0.554	0.561
7	37	29	0.181	0.363	0.181	0.214	0.186	0.196
8	57	15	0.363	0.363	0.636	0.203	0.316	0.297
9	60	40	0.181	0.363	0.090	0.099	0.025	0.029
10	63	43	0.000	0.181	0.000	0.085	0.013	0.017
11	93	26	0.000	0.000	0.272	0.060	0.055	0.053

In Table 4, it is presented the Bayesian estimates of the Ledwina dependence measure (3.3) assuming the BB, GB and FGM and the LY non-parametric estimator for $S(t_1,\ t_2)$. For the estimation of the empirical Ledwina dependence measure (use of LY estimator) we used the KM product-limit estimators for the univariate lifetimes T_1 and T_2, for the estimation of the BB Ledwina dependence measure we used Bayesian estimators for the parameters of the marginal exponential distributions of the univariate lifetimes T_1 and T_2 , and for the estimation of the GB and FGM Ledwina dependence measures we used Bayesian estimators for the parameters of the marginal Weibull distributions of the univariate lifetimes T_1 and T_2

Table 4: Ledwina dependence measure(3.3) for each model

pair	t_1	t_2	LED EMP	LED BB	LED GB	LED FGM
1	16	11	1.0000	0.1604	-0.0148	0.0117
2	19	13	0.7698	0.1528	-0.0217	0.0152
3	18	21	0.43033	0.0712	-0.0424	0.0203

4	22	17	0.26666	0.1335	-0.0360	0.0204
5	20	26	0.46291	0.0491	-0.0572	0.0216
6	29	15	0.31053	0.1354	-0.0352	0.0211
7	37	29	0.62361	0.1057	-0.0861	0.0260
8	57	15	0.57143	-0.0190	-0.0217	0.0222
9	60	40	1.07571	0.0716	-0.0685	0.0148
10	63	43	*	0.0683	-0.0503	0.0115
11	93	26	*	-0.0665	-0.0975	0.0183

In Figure 1 it is presented the plots of the Ledwina dependence measures considering each fitted model. From these plots, it is possible to see that the BB bivariate exponential distribution is better in comparison with the other copula models, since the plot of the estimated Ledwina BB measure is more close to the empirical Ledwina measure using the LY non-parametric estimator, although the data set has a very small sample size. The two copula models (GB and FGM models) give similar fit for the data set.

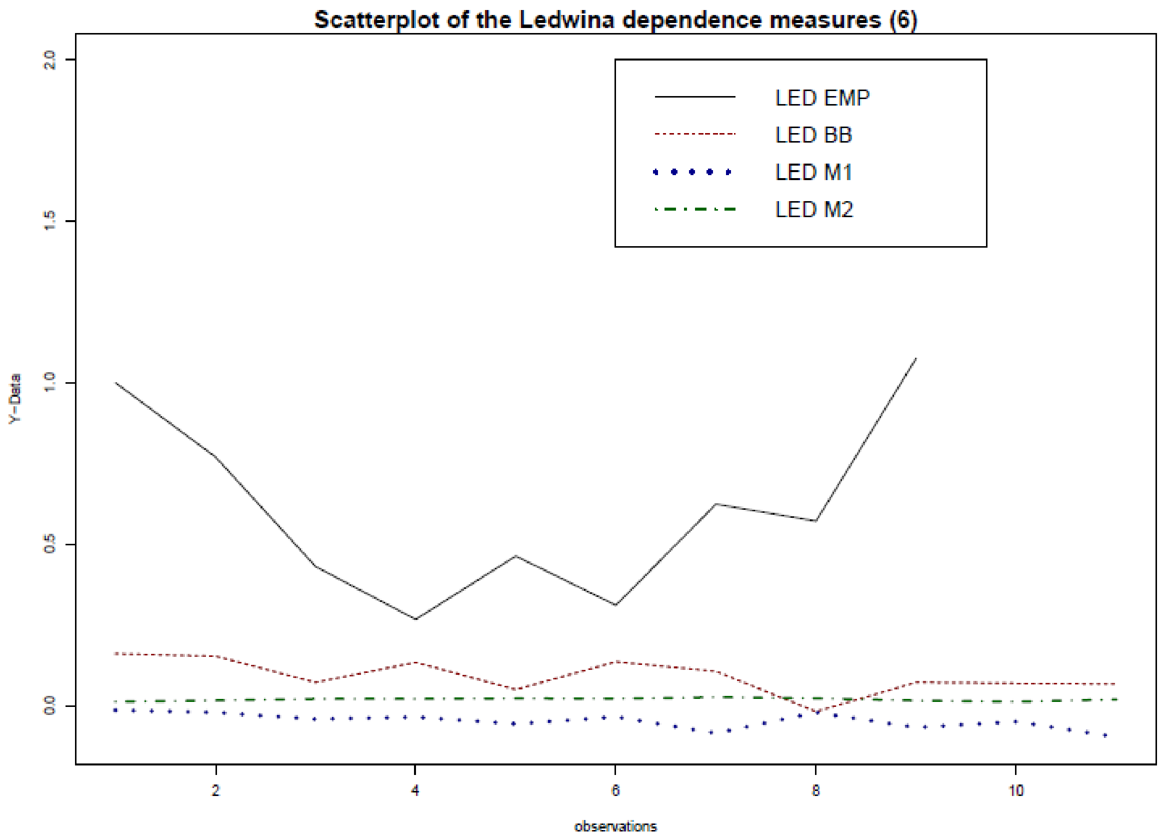

Figure 1: Plots of the Ledwina dependence measures for each fitted model

6.2 Application 2: bone marrow transplant data set

Bone marrow transplants are a standard treatment for acute leukemia. Prognosis for recovery may depend on risk factors known at the time of transplantation, such as patient and/or donor age and sex, the stage of initial disease, the time from prognosis to transplantation among many other factors. The final prognosis may change as the patient posttransplantation history develops with occurrence of events at random times during the recovery process, such as development of accurate or chronic graft-versus-host disease (GVHD), return of the platet count to normal levels, or development of infections. In this study, 137 patients with leukemia received a combination of 16 mg/kg of oral Busulfan (BU) and 120 mg/kg of intravenous cyclophosphamide (Cy) and were treated at one of four hospitals: 76 at Ohio State University hospitals (OSU) in Columbus; 21 at Hahnemann University (HU) in Philadelphia; 23 at St. Vincent's Hospital (SVH) in Sidney, Australia and 17 at Alfred Hospital (AH) in Melbourne, Australia (data set introduced by Klein and Moeschberger, 1997, page 464).

In the analysis considered in this section, it is assumed as lifetimes, the time (in days) to accurate graft-versus-host disease (TA) with 111 censored observa- tions and 26 not-censored observations and the time (in days) to chronic graft- versus-host disease (TC) with 76 censored observations and 61 not-censored observations, that is, we have a great number of censored observations as it is common in medical studies. In this situation, usually it is very difficult to decide by an appropriate parametrical model to be fitted by the data set.

In Table 5, it is presented the posterior summaries of interest using the three parametrical models presented in sections 4 and 5 assuming the same prior distributions given in application 1.

Table 5: Posterior means and posterior standard deviation (in parenthesis) for the parameters of the three models

model	λ_1		λ_2		λ_3
BB	0.0003 (0.0000)		0.0009 (0.0001)		0.0000 (0.0000)
Copula	λ_1	λ_2	p_1	p_2	θ
GB	260700 (2827000)	1277 (267.20)	0.326 (0.058)	0.697 (0.069)	0.094 (0.086)
FGM	258700 (1089000)	1232 (249.30)	0.327 (0.061)	0.710 (0.071)	0.295 (0.128)

In Table 6 (see at the end of this paper) it is presented the LY non-parametrical estimator of the joint survival function and the Bayesian estimators for the parameters of the joint survival function assuming the BB, GB and FGM models and using the Bayesian estimators given in Table 5.

In Table 7 (placed at the end of this paper), it is presented the Bayesian estimators of the Ledwina dependence measure (3.3) assuming the BB, GB and FGM and the LY non-parametric estimator of $S(t_1, t_2)$. For the estimation of the empirical Ledwina dependence measure (use of LY estimator) we have used the KM product-limit estimators for the univariate lifetimes T_1 and T_2 ; for the estimation of the BB Ledwina dependence measure we used Bayesian estimators for the parameters of the marginal exponential distributions for the univariate lifetimes T_1 and T_2 and for the estimation of the GB and FGM Ledwina dependence measures we used Bayesian estimators for the parameters of the marginal Weibull distributions of the univariate lifetimes T_1 and T_2 .

In Figure 2 it is presented the plots of the Ledwina dependence measures considering each fitted model. From these plots, it is possible to see that the BB bivariate exponential distribution is better fitted by the data in comparison with the other copula models, since the plot of the estimated Ledwina BB measure is more close to the empirical Ledwina measure using the LY non-parametric estimator. The good fit of the BB bivariate exponential distribution is also observed in Figure 3.

It is important to point out that other bivariate lifetime distributions could be assumed to this data set with possible better fit. In this way, we have a powerful graphical approach to find appropriate bivariate distributions for bivariate lifetime data in presence of censoring.

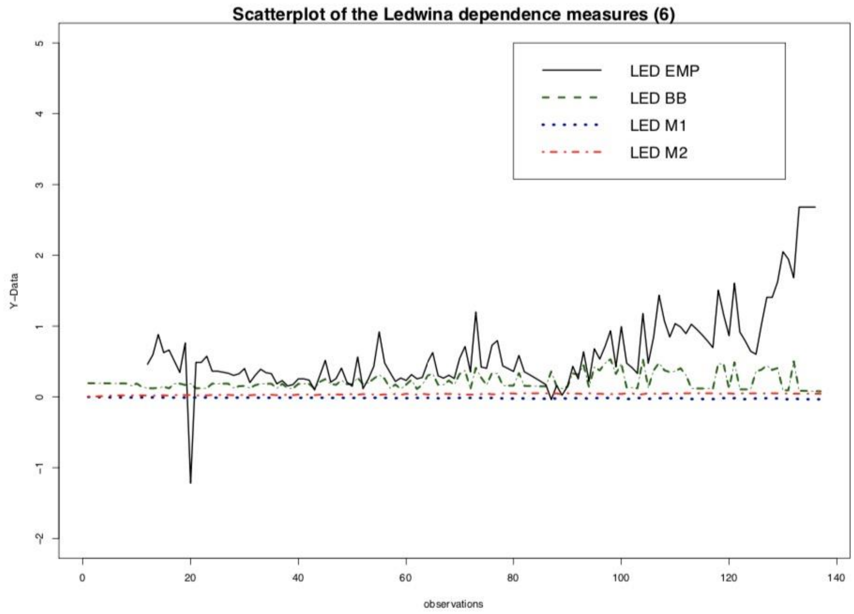

Figure 2: Plots of the Ledwina dependence measures for each fitted model

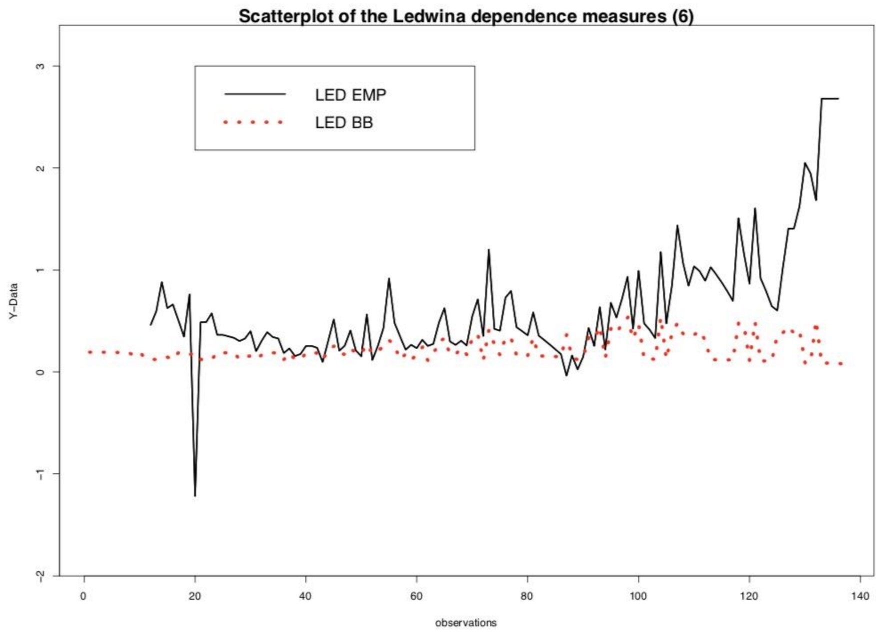

Figure 3: Plots of the Ledwina dependence measures considering the empirical LY nonparametric estimator and the BB bivariate lifetime distribution

6.3 Application 3: UEFA Champion League data

In this application, let us consider a data set obtained from Meintanis (2007) and introduced in Table 8. The data represent 37 football matches from the UEFA Champion's League 2004-

2005 and 2005-2006 seasons where at least one goal scored by the home team and at least one goal scored directly from a penalty kick, foul kick or any other direct kick by any team have been considered. In this way T_1 represents the time in minutes of the first kick goal scored by any team and T_2 represents the first goal of any type scored by the home team. In this case, all possibilities are open, for example, $T_1 < T_2$, $T_1 > T_2$ or $T_1 = T_2 = Y$.

Let us assume marginal exponential distributions with distribution functions given by,

$$F_i(t_i) = P\{T_i \leq t_i\} = 1 - \exp(-\lambda_i t_i), i = 1, 2 \quad (6.1)$$

and gamma prior distributions with hyperparameters $a = 0.1$ and $b = 0.1$ for the parameters λ_1 and λ_2 assuming the GB and FGM copula models (denoted as M1 and M2 models).

Table 8: UEFA Champion's League data (2004-2006)

match	T_1	T_2	match	T_1	T_2
Lyon×Real Madrid	26	20	Internazionale×Bremen	34	34
Milan×Fenerbahce	63	18	Real Madrid×Roma	53	39
Chelsea×Anderlecht	19	19	Man. United×Fenerbahce	54	7
Club Brugge×Juventus	66	85	Bayern×Ajax	51	28
Fenerbahce×PSV	40	40	Moscow×PSG	76	64
Internazionale×Rangers	49	49	Barcelona×Shakhtar	64	15
Panathinaikis×Bremen	8	8	Leverkusen×Roma	26	48
Ajax-Arsenal	69	71	Arsenal×Panathinaikos	16	16
Man. United×Benfica	39	39	Dynamo Kyiv×Real Madrid	44	13
Real Madrid×Rosenborg	82	48	Man. United×Sparta	25	14
Vilarreal×Benfica	72	72	Bayern×M. TelAviv	55	11
Juventus×Bayern	66	62	Bremen×Internazionale	49	49
Club Brugge×Rapid	25	9	Anderlecht×Valencia	24	24
Olympiacos×Lyon	41	3	Panathinaikos×PSV	44	30
Internazionale×Porto	16	75	Arsenal×Rosenborg	42	3
Schalke×PSV	18	18	Liverpool×Olympiacos	27	47
Barcelona×Bremen	22	14	M. TelAviv×Juventus	28	28
Milan×Schalke	42	42	Bremen×Panathinaikos	2	2
Rapid×Juventus	36	52			

Table 9 presents the posterior summaries of interest using the three parametrical models presented in sections 4 and 5.

Table 9: Posterior means and posterior standard deviation (in parenthesis) for the parameters of the three models

model	λ_1	λ_2	λ_3
BB	0.0001 (0.0004)	0.0007 (0.0015)	0.0444 (0.0053)
Copula	λ_1	λ_2	θ
GB	0.0244 (0.0040)	0.0297 (0.0049)	0.1695 (0.1575)
FGM	0.0254 (0.0040)	0.0323 (0.0053)	0.6748 (0.2850)

In Table 10, it is presented the KM non-parametrical estimator of the joint survival function and the Bayesian estimators of the joint survival function assuming the GB and FGM models using the Bayesian estimators given in Table 9.

Table 10: Estimators for the joint survival function $S(t_1, t_2)$

pair	t_1	t_2	KM	BB	GB	FGM
1	26	20	0.4864	0.1205	0.2838	0.3096
2	63	18	0.1621	0.1073	0.1157	0.1378
3	19	19	0.5675	0.0713	0.3517	0.3702
4	66	85	0.000002	0.0243	0.0090	0.0174
5	40	40	0.2432	0.0275	0.1001	0.1279
6	49	49	0.1351	0.0183	0.0565	0.0797
7	8	8	0.8648	0.1173	0.6510	0.6455
8	69	71	0.0270	0.0407	0.0137	0.0253
9	39	39	0.2702	0.0288	0.1065	0.1347
10	82	48	0.0008	0.0352	0.0220	0.0376
11	72	72	0.00009	0.0065	0.0120	0.0228
12	66	62	0.0810	0.0159	0.0211	0.0360
13	25	9	0.5405	0.2464	0.3007	0.3254
14	41	3	0.4594	0.2491	0.3406	0.3315
15	16	75	0.0270	0.0486	0.0676	0.0689
16	18	18	0.6216	0.0746	0.3724	0.3891
17	22	14	0.6486	0.1719	0.3809	0.3991
18	42	42	0.2162	0.0251	0.0883	0.1152
19	36	52	0.1351	0.1064	0.0753	0.0968
20	34	34	0.3513	0.0361	0.1449	0.1741
21	53	39	0.1621	0.0625	0.0714	0.0981
22	54	7	0.2433	0.1680	0.2159	0.2212
23	51	28	0.1891	0.1020	0.1112	0.1408
24	76	64	0.000005	0.0197	0.0143	0.0266
25	64	15	0.1621	0.1110	0.1258	0.1447
26	26	48	0.2162	0.1327	0.1157	0.1347
27	16	16	0.6756	0.0816	0.4173	0.4300
28	44	13	0.3243	0.1812	0.2249	0.2461
29	25	14	0.5945	0.1861	0.3528	0.3731
30	55	11	0.2162	0.1511	0.1827	0.1978
31	49	49	0.1351	0.0183	0.0565	0.0797
32	24	24	0.5135	0.0568	0.2633	0.2885
33	44	30	0.2433	0.0940	0.1257	0.1565
34	42	3	0.4324	0.2438	0.3325	0.3234
35	27	47	0.2432	0.1371	0.1159	0.1362
36	28	28	0.4054	0.0474	0.2079	0.2361
37	2	2	0.9729	0.1539	0.8996	0.8919

In Table 11, it is presented the Bayesian estimates of the Ledwina dependence measure (3.3) assuming the BB, GB and FGM and the LY non-parametric estimator for $S(t_1, t_2)$: for the estimation of the empirical Ledwina dependence measure (use of LY estimator) we used the KM product-limit estimators for the univariate lifetimes T_1 and T_2 for the estimation of the BB Ledwina dependence measure we used Bayesian estimators for the parameters of the marginal exponential dis- tributions of the univariate lifetimes T_1 and T_2, and for the estimation of the GB and FGM Ledwina dependence measures we used Bayesian estimators for the parameters of the marginal Weibull distributions of the univariate lifetimes T_1 and T_2.

Table 11: Ledwina dependence measure (3.3) for each model

pair	T_1	T_2	LED EMP	LED BB	LED GB	LED FGM
1	26	20	0.4441	-0.6780	-0,0746	0.1669
2	63	18	0.2109	0.0400	-0,0846	0.1330
3	19	19	0.3834	-1.2631	-0,0660	0.1621
4	66	85	0.0090	0.2564	-0,0814	0.0627
5	40	40	0.2632	-0.3003	-0,0964	0.1426
6	49	49	0.4023	-0.1719	-0,0987	0.1212
7	8	8	0.6047	-4.3109	-0,0326	0.1091
8	69	71	0.2745	0.3470	-0,0882	0.0750
9	39	39	0.2720	-0.3194	-0,0958	0.1448
10	82	48	0.0055	0.1764	-0,0926	0.0893
11	72	72	-0.0571	-0.0260	-0,0866	0.0717
12	66	62	0.6261	0.0007	-0,0932	0.0879
13	25	9	0.4891	-0.9140	-0,0725	0.1664
14	41	3	0.0910	-0.7171	-0,0381	0.0929
15	16	75	0.0580	0.0815	-0,0733	0.0885
16	18	18	0.3714	-1.3758	-0,0635	0.1602
17	22	14	0.2086	-0.9808	-0,0623	0.1595
18	42	42	0.3275	-0.2655	-0,0974	0.1380
19	36	52	0.2360	0.2647	-0,0954	0.1265
20	34	34	0.3874	-0.4361	-0,0917	0.1550
21	53	39	0.2781	0.0080	-0,0987	0.1314
22	54	7	0.1973	-0.1940	-0,0588	0.1171
23	51	28	0.1950	0.0083	-0,0941	0.1458
24	76	64	-0.0580	0.1022	-0,0891	0.0771
25	64	15	0.2677	0.0164	-0,0799	0.1292
26	26	48	0.2582	0.2128	-0,0889	0.1356
27	16	16	0.3448	-1.6488	-0,0582	0.1550
28	44	13	0.0922	-0.1781	-0,0733	0.1490
29	25	14	0.2223	-0.7680	-0,0650	0.1609
30	55	11	0.2310	-0.0899	-0,0709	0.1322
31	49	49	0.4023	-0.1719	-0,0987	0.1212
32	24	24	0.4379	-0.8548	-0,0769	0.1659
33	44	30	0.2871	-0.0961	-0,0939	0.1516
34	42	3	0.2592	-0.6856	-0,0383	0.0923
35	27	47	0.2309	0.2248	-0,0898	0.1373
36	28	28	0.4248	-0.6452	-0,0839	0.1637
37	2	2	1.0000	-22.0708	-0,0088	0.0353

Figure 4 shows the plots of the Ledwina dependence measures considering each fitted model. From these plots, it is observed better fit of the M2 (FGM) model given that, plot of estimated Ledwina model very close to the plot of the empirical Ledwina measure.

6.4 Application 4: Recurrence Times of Infection for Kidney Patients

In this application, let us consider a survival data set introduced by McGilchrist and Aisbett (1991) related to kidney infection where the recurrence of infection of 38 kidney patients,

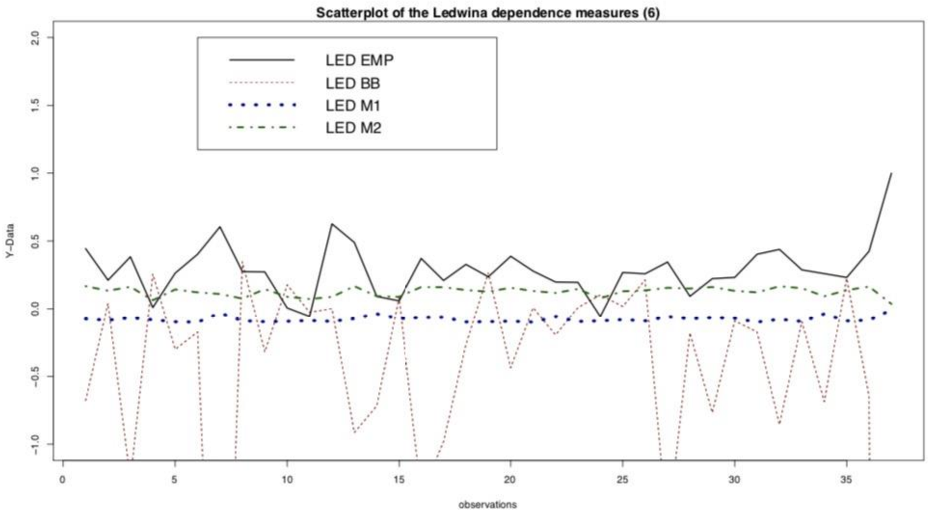

Figure 4: Plots of the Ledwina dependence measures for each fitted model

using portable dialysis machines, are recorded. Infections may occur at the location of insertion of the catheter. The time recorded, called infection time, is either the survival time (in days) of the patient until an infection occurred and the catheter had to be removed, or the censored time, where the catheter was removed by others reasons. The catheter is reinserted after some time and the second infection time is again observed or censored (data set in Table 12). To analyse this application, let us assume the three bivariate distributions obtained from a BB distribution, a GB and a FGM copula function models under marginal Weibull distributions, introduced in sections 3 and 5. As a first analysis of the recurrence times of infections, let us assume the BB exponential distribution with density (4.1). For a Bayesian analysis of the BB distribution, it is assumed gamma prior distributions (4.9) for the parameters λ_1,λ_2 and λ_3, with hyperparameters values $a_1 = a_2 = a_3 = 1$ and $a_1 = a_2 = a_3 = 100$.

Table 12: Recurrence times of infections in 38 kidney patients

patient	t_1	t_2	δ_1	δ_2
1	8	16	1	1
2	23	13	1	0
3	22	28	1	1
4	447	318	1	1
5	30	12	1	1

6	24	245	1	1
7	7	9	1	1
8	511	30	1	1
9	53	196	1	1
10	15	154	1	1
11	7	333	1	1
12	141	8	1	0
13	96	38	1	1
14	149	70	0	0
15	536	25	1	0
16	17	4	1	0
17	185	117	1	1
18	292	114	1	1
19	22	159	0	0
20	15	108	1	0
21	152	562	1	1
22	402	24	1	0
23	13	66	1	1
24	39	46	1	0
25	12	40	1	1
26	113	201	0	1
27	132	156	1	1
28	34	30	1	1
29	2	25	1	1
30	130	26	1	1
31	27	58	1	1
32	5	43	0	1
33	152	30	1	1
34	190	5	1	0
35	119	8	1	1
36	54	16	0	0
37	6	78	0	1
38	63	8	1	0

For a Bayesian analysis of bivariate Weibull distribution under GB and FGM copula functions with joint distribution functions given by (5.8) and (5.13), respectively, let us assume uniform prior distributions with hyperparameter values $a_1 = a_2 = 0$; $b_1 = b_2 = 500$, $c_1 = c_2 = 0$; $d_1 = d_2 = 2$; $e = 0$ and $f = 0.2$ for λ_1, λ_2, p_1, p_2 and θ. To simulate samples of the joint posterior distribution for the parameters of the three models, we have used the WinBUGS software which requires only the specification of the joint distribution for the data and the prior distributions for the parameters of the model. In the simulation procedure, we discarded the first 5000 simulated Gibbs samples ("burnin-samples ") to eliminate the effect of the initial values for the parameters of the model. Choosing every 20th simulated Gibbs sample, we obtained a final sample of size 5000 to get the posterior summaries of interest. Convergence of the Gibbs sampling algorithm was monitored using standard existing methods as time series for the simulated samples.

Table 13 presents the posterior means and the posterior standard-deviations for the parameters for the three parametrical models presented in sections 3 and 5.

Table 13: Posterior means and posterior standard deviation (in parenthesis) for the parameters of the three models

model	λ_1		λ_2		λ_3
BB	0.006 (0.002)		0.005 (0.002)		0.003 (0.002)
Copula	λ_1	p_1	λ_2	p_2	θ
GB	135.3 (32.96)	0.805 (0.110)	144.6 (31.66)	0.999 (0.141)	0.077 (0.055)
FGM	136.1 (32.89)	0.827 (0.011)	143.2 (30.22)	0.997 (0.014)	0.099 (0.057)

In Table 14, it is presented the LY non-parametrical estimator of the joint survival function and the Bayesian estimators of the joint survival function assuming the BB, GB and FGM models using the Bayesian estimators given in Table 13. In Table 14, it is also presented the Kaplan-Meier (KM) estimates for the marginal survival functions of the lifetimes T_1 and T_2.

In Table 15, it is presented the Bayesian estimates of the Ledwina dependence measure (3.3) assuming the BB, GB and FGM and the LY non-parametric estimator for $S(t_1\ t_2)$. For the estimation of the empirical Ledwina dependence measure (use of LY estimator) we used the KM product-limit estimators for the univariate lifetimes T_1 and T_2, for the estimation of the BB Ledwina dependence measure we used Bayesian estimators for the parameters of the marginal exponential dis- tributions of the univariate lifetimes T_1 and T_2, and for the estimation of the GB and FGM Ledwina dependence measures we used Bayesian estimators for the parameters of the marginal Weibull distributions of the univariate lifetimes T_1 and T_2.

Table 14: Estimators for the joint survival function $S(t_1\ t_2)$

patient	t1	t2	LY	BB	GB	FGM
1	8	16	0.632	0.8464	0.8053	0.8127
2	23	13	0.530	0.4038	0.7162	0.7266
3	22	28	0.418	0.7054	0.6493	0.6607
4	447	318	0.000	0.0025	0.0053	0.0081
5	30	12	0.474	0.3901	0.6810	0.6917
6	24	245	0.107	0.1342	0.1400	0.1453
7	7	9	0.737	0.8944	0.8551	0.8618
8	511	30	0.000	0.0078	0.0418	0.0415
9	53	196	0.143	0.1657	0.1541	0.1655
10	15	154	0.280	0.2910	0.2870	0.2932
11	7	333	0.054	0.0726	0.0915	0.0909
12	141	8	0.339	0.1848	0.3344	0.3387
13	96	38	0.237	0.2089	0.3533	0.3666
14	149	70	0.160	0.1153	0.2001	0.2138
15	536	25	0.000	0.0064	0.0389	0.0380
16	17	4	0.731	0.4498	0.8048	0.8133
17	185	117	0.043	0.0645	0.1133	0.1265
18	292	114	0.054	0.0296	0.0632	0.0719
19	22	159	0.216	0.2679	0.2593	0.2676

20	15	108	0.320	0.4140	0.3950	0.4032
21	152	562	0.000	0.0048	0.0052	0.0071
22	402	24	0.201	0.0203	0.0740	0.0739
23	13	66	0.388	0.5731	0.5396	0.5486
24	39	46	0.240	0.5522	0.4976	0.5118
25	12	40	0.418	0.6958	0.6535	0.6627
26	113	201	0.107	0.1090	0.0961	0.1089
27	132	156	0.080	0.1334	0.1177	0.1321
28	34	30	0.279	0.3298	0.5799	0.5928
29	2	25	0.697	0.8291	0.8106	0.8142
30	130	26	0.284	0.1783	0.3120	0.3214
31	27	58	0.249	0.5521	0.5034	0.5164
32	5	43	0.491	0.7145	0.6888	0.6944
33	152	30	0.120	0.1476	0.2653	0.2742
34	190	5	0.238	0.1289	0.2582	0.2591
35	119	8	0.391	0.2179	0.3819	0.3875
36	54	16	0.404	0.3248	0.5519	0.5633
37	6	78	0.406	0.5496	0.5339	0.5389
38	63	8	0.421	0.3245	0.5490	0.5583

Table 15: Ledwina dependence measure (3.3) for each model

patient	T_1	T_2	LED EMP	LED BB	LED GB	LED FGM
1	8	16	-1.5315	0.0164	-0.0078	0.0088
2	23	13	-0.8038	-3.6514	-0.0105	0.0113
3	22	28	-0.7997	0.0326	-0.0147	0.0151
4	447	318	-0.0930	0.0064	-0.0355	0.0078
5	30	12	-0.5889	-3.2455	-0.0111	0.0117
6	24	245	0.0484	0.0322	-0.0284	0.0155
7	7	9	-2.0167	0.0118	-0.0057	0.0065
8	511	30	-0.2838	-0.0905	-0.0228	0.0085
9	53	196	0.0652	0.0561	-0.0364	0.0208
10	15	154	0.0788	0.0324	-0.0233	0.0167
11	7	333	0.1438	0.0126	-0.0174	0.0081
12	141	8	0.2581	-1.2645	-0.0139	0.0108
13	96	38	-0.2513	-0.7625	-0.0262	0.0209
14	149	70	-0.0256	-0.3363	-0.0359	0.0228
15	536	25	*	-0.0909	-0.0206	0.0075
16	17	4	*	-7.8965	-0.0053	0.0060
17	185	117	-0.2020	-0.1701	-0.0426	0.0219
18	292	114	-0.0455	-0.0846	-0.0418	0.0177
19	22	159	0.0054	0.0390	-0.0268	0.0185
20	15	108	-0.2109	0.0349	-0.0214	0.0176
21	152	562	*	0.0262	-0.0291	0.0065
22	402	24	1.0403	-0.1716	-0.0226	0.0100
23	13	66	-0.1838	0.0322	-0.0173	0.0162
24	39	46	-0.3166	0.0493	-0.0220	0.0202

25	12	40	-0.6546	0.0278	-0.0137	0.0141
26	113	201	0.0804	0.0783	-0.0433	0.0210
27	132	156	-0.1242	0.0870	-0.0434	0.0226
28	34	30	-0.4972	-1.8611	-0.0176	0.0172
29	2	25	-2.2910	0.0099	-0.0056	0.0062
30	130	26	-0.1982	-0.7212	-0.0236	0.0179
31	27	58	-0.4005	0.0450	-0.0214	0.0196
32	5	43	-0.9097	0.0184	-0.0102	0.0105
33	152	30	-0.2637	-0.5684	-0.0257	0.0183
34	190	5	*	-1.1875	-0.0115	0.0080
35	119	8	-0.2932	-1.4790	-0.0134	0.0111
36	54	16	-0.3475	-1.9086	-0.0153	0.0147
37	6	78	-0.6090	0.0221	-0.0137	0.0127
38	63	8	-0.6750	-2.4530	-0.0115	0.0111

Figure 5 shows the plots of the Ledwina dependence measures considering each fitted model. From these plots, it is possible to see that the BB bivariate exponential distribution is better in comparison with the other copula models, since the plot of the estimated Ledwina BB measure is more close to the empirical Ledwina measure using the LY non-parametric estimator, although the data set has a very small sample size. The two copula models (GB and FGM models) give similar fit for the data set.

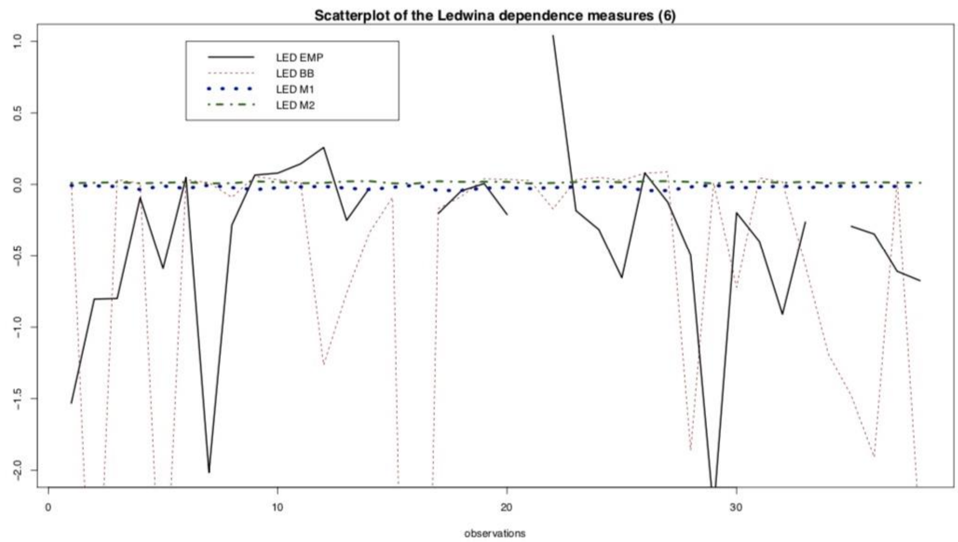

Figure 5: Plots of the Ledwina dependence measures for each fitted model

6.5 Application 5: Thyroid cancer data

As our last application, we used a medical data set related to differentiated cancer o thyroid. In this data set, there are the observations of 15 patients treated by the head and neck surgeon Gabriel Sanchez de Guzman who works in a level IV hospital located in Bogotá Colombia. Each patient received surgery to remove the thyroid gland and then was treated with redioiodine

I-131. Each individual was followed after finalized the therapy until to be observed a thyroglobulin level greater than or equal 2.0 *ng/ml* which is an indicator of the presence of recurrence/persistence of the cancer. In this way, the time recorded for each patient is the recurrence/persistence time (in months). Fourteen individuals presented two recurrence/persistence events and at the date of the last control, one individual did not present recurrence/persistence (censored time). The data were analysed in the same form as considered for the other data sets. For the Bayesian analysis, we considered the same prior distributions for the parameters that were assumed to analyse the infection for Kidney data and, to get the posterior summaries of interest the same simulation approach was used. Table 16 presents the posterior summaries of interest and the Kaplan-Meier (KM) estimates for the marginal survival functions of the lifetimes are reported in Table 17.

Table 16: Posterior means and posterior standard deviation (in parenthesis) for the parameters of the three models

model	λ_1		λ_2		λ_3
BB	0.000004 (1.0e−12)		0.00009 (1.0e−12)		0.0026 (1.0e−12)
Copula	λ_1	p_1	λ_2	p_2	θ
GB	850.7 (29.75)	0.7478 (0.1448)	849 (29.02)	1.47 (0.2765)	0.2002 (0.188)
FGM	846.7 (28.63)	0.7677 (0.1521)	855.8 (28.21)	1.504 (0.2876)	0.4288 (0.2824)

Table 17: Estimators for the joint survival function $S(t_1\ t_2)$

pair	t_1	t_2	LY	KM S(t_1)	KM S(t_2)	BB	GB	FGM
1	88	460	0.5333	0.7333	0.7143	0.3045	0.5391	0.5721
2	3634	645	0.0000	0.0000	*	0.1174	0.0214	0.0334
3	1634	2304	0.0000	0.1333	0.0000	0.0023	0.0017	0.0049
4	875	1353	0.0667	0.2667	0.0714	0.0282	0.0356	0.0617
5	322	50	0.4000	0.4000	0.9286	0.8475	0.6006	0.6096
6	925	455	0.1333	0.2000	0.7857	0.2800	0.2118	0.2528
7	2900	420	0.0667	0.0667	0.8571	0.2331	0.0511	0.0642
8	713	560	0.0667	0.3333	0.5714	0.2195	0.2203	0.2698
9	122	699	0.2667	0.6000	0.4286	0.1642	0.3572	0.3959
10	243	519	0.2667	0.4667	0.6429	0.2572	0.3976	0.4440
11	92	672	0.2667	0.6667	0.5000	0.1766	0.3923	0.4274
12	48	1238	0.1333	0.9333	0.2143	0.0414	0.1491	0.1628
13	136	974	0.1333	0.5333	0.3571	0.0809	0.2130	0.2459
14	81	1270	0.0667	0.8000	0.2143	0.0380	0.1296	0.1467
15	66	1138	*	0.8667	0.2857	0.0534	0.1756	0.1947

From the results showed in Tables 17 and 18 and the plots presented in Figure 6, the model with the best fit is the GB model since for this model we observed values of the Ledwima measure more close to the empirical Ledwina measure using the LY non-parametric estimator. It is important to point out that in this example there is a very small sample size, usually a great difficult to fit a parametrical model, but as illustrative purposes, it was possible to fit marginal Weibull distributions and a dependence structure type Gumbel Barnett.

Table 18: Ledwina dependence measure (3.3)for each model

Pair	t_1	t_2	LED EMP	LED BB	LED GB	LED FGM
1	88	460	-0.0227272	-1.1928977	0.0067614	0.1755114
2	3634	645	*	*	*	*
3	1634	2304	*	*	*	*
4	875	1353	0.4432026	0.0947547	0.1612955	0.3977240
5	322	50	0.2182179	3.8801868	1.8597619	1.9334105
6	925	455	-0.1666667	0.7500000	0.3237500	0.5800000
7	2900	420	0.1048284	2.0676107	-0.0782807	0.0758565
8	713	560	-0.5773503	0.0844375	0.0879016	0.3022429
9	122	699	0.1111111	-0.3158333	0.4883333	0.6495833
10	243	519	-0.1889823	-0.2292355	0.3677594	0.5650569
11	92	672	-0.1889823	-0.5719548	0.3452233	0.4944720
12	48	1238	-0.5345225	-1.4561060	-0.3765043	-0.2391988
13	136	974	-0.1889822	-0.4119765	0.1497684	0.2896625
14	81	1270	-0.2941742	-0.5051461	0.1686599	0.2944194
15	66	1138	*	-1.1819818	-0.3692746	-0.2422164

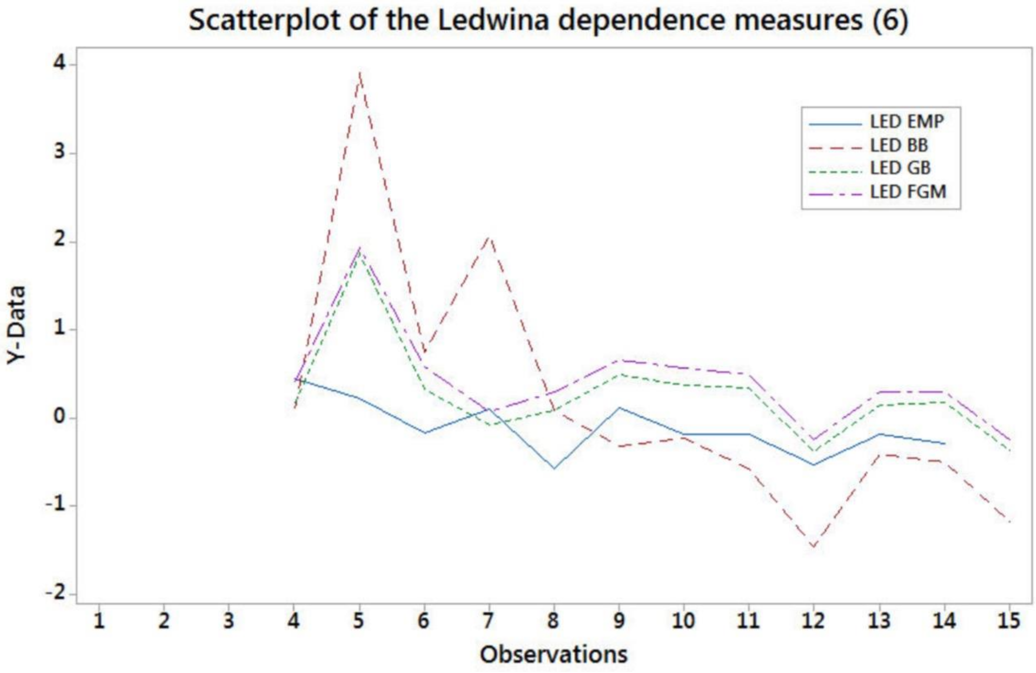

Figure 6: Plots of the Ledwina dependence measures for each fitted model

7. Concluding remarks

Usually in applications with bivariate lifetime data in presence of censoring, the choice of an appropriate parametrical distribution is not easy, since each assumed parametrical model has a different dependence structure. The proposed graphical method based on the Ledwina dependence measure introduced in this paper could be very useful to decide by an appropriate bivariate lifetime distribution in applications with bivariate data sets in presence of univariate censoring. In this way, the LY non-parametric estimator for the survival function that is a generalization of the popular Kaplan-Meier product limit estimator could be a starting point in each application. Plots of the Ledwina dependence measure considering the LY estimator could be compared to the plots of the Ledwina dependence measure assuming any parametrical lifetime distribution, implying in an useful graphical tool to be used in applications considering bivariate lifetime data. It is important to point out that we could use our proposed methodology in applications of bivariate lifetime analysis in presence of univariate censoring mechanism considering any parametrical model introduced in the literature to be fitted by bivariate lifetime data. Each proposed parametrical model have some physical interpretations as the degree of dependence associated to each distribution, but in applications usually it is difficult for a statistician to decide if the proposed model is good in terms of good fit. From the obtained results of this study it was observed that in practical applications standard copula based models should be used with caution, since every copula function implies in different dependence structures which could not capture the existing dependence structure for the dataset if the model is not appropriate. Finally, it is important to point out that our proposed graphical approach is useful in discrimination of proposed models but also in the indication of the good fit of a proposed model for bivariate lifetime data in presence of censored data. As alternative to our methodology, the popular existing model selection criteria AIC (Akaike information criterion), BIC (Bayesian information criterion) and DIC (deviance information criterion) only indicates the best among different models, but all proposed models could be wrong.

8. Acknowledgments

The authors are grateful to Gabriel Sánchez de Guzmán, a head and neck surgeon who works in the Hospital San Ignacio in Bogotá, and who kindly provided the data set used in our fifth application. We also are grateful to Jennyfer Portilla Yela by her valuable help to obtain the results of the last application. We would like to thank the Editor-in-Chief and the referees for their careful reading and for the comments which greatly improved the paper.

References

[1] Achcar, J.A.; Leandro, R.A. (1998). Use of Markov Chain Monte Carlo methods in a Bayesian analysis of the Block and Basu bivariate exponential distribution. Annals of the Institute of Statistical Mathematics, 50, 3, 403-416.

[2] Achcar, J.A.; Santos, C.A. (2011). A Bayesian Analysis for the Block-Basu Bivariate Exponential Distribution in the Presence of Covariates and Censored Data. Journal of Applied Statistics, v. 38, p. 2213-2223.

[3] Arnold, B.C.; Strauss, D. (1988). Bivariate distributions with exponential conditionals. Journal of the American Statistical Association, 83, 402, 522-527.

[4] Bairamov, I.; Kotz, S.; Kozubowski, T.J. (2003). A new measure of linear local dependence. Statistics, 37, 243-258.

[5] Balakrishnan, N.; Lai, C.D. (2009). Continuous Bivariate Distributions, Springer, Dordrecht.

[6] Barnett, V. (1980). Some bivariate uniform distributions. Communications in Statistics - Theory and Methods, 9, 453-461.

[7] Bjerve, S.; Doksum, K. (1993). Correlation curves: measures of association as function of covariate values. The Annals of Statistics, 21, 890-902.

[8] Block, H.W.; Basu, A.P. (1974). A continuous bivariate exponential extension. Journal of American Statistical Association, 69, 1031-1037.

[9] Campbell, G. (1981). Nonparametric bivariate estimation with randomly censored data. Biometrika, 68, 417-422.

[10] Campbell, G.; Foldes, A. (1982). Large sample properties of nonparametric bivariate estimators with censored data. In Nonparametric Statistical Inference, Colloquia Mathematica-Societatis Janos Bolyai, Ed. B.V. Gnedenko, M.L. Puri and I. Vincze, pp 103-122, Amsterdam: North-Holland.

[11] Chib,S.; Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. American Statistician, 49 , pp. 327-335.

[12] Dabrowska, D.M. (1988). Kaplan-Meier estimate on the plane. Annals of Statistics, 16, 1475-1489.

[13] Davarzani , N.; Achcar, J.A.; Smirnov, E.N.; Peeters, R. (2015). Bivariate lifetime geometric distribution in presence of cure fractions, Journal of Data Science 13, 755-770.

[14] Downton, F. (1970). Bivariate exponential distributions in reliability theory. Journal of the Royal Statistical Society, B, 32, 408-417.

[15] Drouet M.D.; Kotz, S. (2001). Correlation and Dependence. Imperial College Press, London.

[16] Durante, F.; Sempi, C. (2015). Principles of copula theory. CRC Press, 332 pages.

[17] Flores, Y.S. (2016) General multivariate dependence using associated copulas, Revstat Statistical Journal, 14 , No. 1, pp. 1-28.

[18] Freund, J.E. (1961). A bivariate extension of the exponential distribution. Journal of the American Statistical Association, 56, 971-977.

[19] Gelfand, A.E.; Smith, A.F.M. (1990). Sampling-based approaches to calculating marginal densities. J.Am. Statist. Assoc. 85, pp. 398-410.

[20] Goethals,K.; Paul Janssen, P.; Duchateau, L. (2008). Frailty models and copulas: similarities and differences. Journal of Applied Statistics, Vol. 35, No. 9, 1071-1079.

[21] Gumbel, E.J. (1960). Bivariate exponential distributions. Journal of the American Statistical Association, 55, 698-707.

[22] Hanley, J.A.; Parnes, M.N. (1983). Nonparametric estimation of a multivariate distribution in the presence of censoring. Biometrics, 39, 129-139 .

[23] Hawkes, A.G. (1972). A bivariate exponential distribution with applications to reliability. Journal of the Royal Statistical Society, B, 34, 129-131.

[24] Holt, J.D.; Prentice, R.L. (1974). Survival analysis in twin studies and matched pairs experiments. Biometrika, 61, 17-30.

[25] Hougaard, P. (1986). A class of multivariate failure time distributions. Biometrika, 3, 73, 671-678.

[26] Icuma , T.R.; Buzatto, I.P.C.; Tiezzi, D.G.; Achcar, J.A.; Davarzani, N. (2016). Use of Bivariate Lifetime Distributions Assuming Continuous or Discrete Data Applied to Patients with Breast Cancer, Journal of Data Science, 14, 657-680.

[27] Jogdeo, K. (1982). Dependence, concepts of, in: S. Kotz, N.L. Johnson (Eds), Encyclopedia of Statistical Sciences, Vol. 2. Wiley, New York, pp. 324- 334.

[28] Kaplan, E.L.; Meier, P. (1958). Nonparametric estimation from incomplete observations. Journal of the American Statistical Association, 53 (282): 457- 481。

[29] Karvanen, J.; Saarela, O; Kuulasmaa, K. (2010). Nonparametric Multiple Imputation of Left Censored Event Times in Analysis of Follow-up Data. Journal of Data Science, 8, 151-172.

[30] Klein J.P.; Moeschberger, M.L. (1997). Survival Analysis: Techniques for Censored and Truncated Data. Springer, Berlin.

[31] Kowalczyk, T.; Pleszczynska, E. (1997). Monotonic dependence functions of bivariate distributions. The Annals of Statistics, 5, 1221-1227.

[32] Lancaster, H.O. (1982). Dependence, measures and indices of, in: S. Kotz, N.L. Johnson (Eds), Encyclopedia of Statistical Sciences, Vol. 2. Wiley, New York, pp. 334-339.

[33] Langberg, N.A.; Shaked, M. (1982). On the identifiability of multivariate life distributions. Annals of Probability, 10,773-779.

[34] Lawless, J.F. (1982). Statistical Models and Methods for Lifetime Data. Wiley, New York, 1982.

[35] Ledwina, T. (2015). Visualizing association structure in bivariate copulas using new dependence function. In: Stochastic models, Statistics and their applications, Ed: Steland, A.; Rafajlowicz, E.; Szajowski, K;. Springer Proceedings in Mathematics and Statistics, 19-27.

[36] Lin, D.Y.; Ying, Z. (1993). A simple nonparametric estimator of the bivariate survival function under univariate censoring. Biometrika, 80, 3, 573-581.

[37] Marshall, A.W.; Olkin, I. (1967). A generalized bivariate exponential distribution. Journal of Applied Probability, 4, 291-302.

[38] Martinez, E.Z.; Achcar, J.A.; Peres, M.V.O.; Queiroz, J.A.M. (2016). A brief note on the simulation of survival data with a desired percentage of right-censored data. Journal of Data Science, 14, 701-712.

[39] McGilchrist, C.A. and Aisbett, C.W. (1991). Regression with frailty in survival analysis, Biometrics, 47, 461 466. pp. 461 466.

[40] Morgenstern, D. (1956). Einfachebeispielezweidimensionalerverteilungen. Mit- teilingsblatt fur MathematisheStatistik, 8, 234-235.

[41] Nelsen, R.B. (2006). An Introduction to Copulas. Springer, New York.

[42] Prentice, R.L.; Cai, J. (1992). Covariance and survival function estimation using censored multivariate failure time data. Biometrika, 79, 495-512.

[43] Pruitt, R.C. (1993). Small sample comparisons of five bivariate survival estimators. Journal of Statistical Computing and Simulation, 45,147-167.

[44] Sarkar, S.K. (1987). A continuous bivariate exponential distribution. Journal of the American Statistical Association, 82, 667-675.

[45] Sklar, A. (1959). Fonctions de r partition n dimensions et leursmarges. Publ. Inst. Stat. Univ. Paris, 8, 229-231.

[46] Spiegelhalter, D.J., Thomas, A., Best, N.G., Lunn, D. (2003). WinBugs user manual, version 1.4. MRC Biostatistics Unit Cambridge, U.K.

[47] Tsai, W.Y.; Leurgans, S.; Crowley, J. (1986). Nonparametric estimation of a bivariate survival function in the presence of censoring. Annals of Statistics, 13, 1351-1365.

[48] Wang., J. (2011). Estimation of Lifetime Distribution with Missing Censoring. Journal of Data Science, 9, 331-343.

Table 6: Estimators for the joint survival function $S(t_1\ t_2)$

pairs	t_1	t_2	LY	BB	GB	FGM
1	1	1	1.00000	0.998761	0.976321	0.977043
2	2	2	1.00000	0.997523	0.967970	0.968942
3	10	10	0.98507	0.987677	0.932058	0.934001
4	16	16	0.97744	0.980356	0.914288	0.916662
5	35	35	0.87121	0.957528	0.872530	0.875826
6	48	48	0.83969	0.942217	0.850027	0.853776
7	53	53	0.83077	0.936393	0.842154	0.846055
8	62	62	0.82946	0.926001	0.828827	0.832979
9	38	63	0.85156	0.931273	0.835869	0.839677
10	73	73	0.81102	0.913456	0.813748	0.818174

11	29	74	0.88889	0.923753	0.827687	0.831398
12	16	79	0.97600	0.922625	0.829971	0.833308
13	10	80	0.98374	0.923268	0.833989	0.837051
14	21	80	0.95122	0.920463	0.825625	0.829142
15	38	84	0.86179	0.912622	0.813298	0.817267
16	10	84	0.96748	0.919717	0.829805	0.832867
17	86	86	0.81978	0.898849	0.797298	0.802011
18	93	93	0.79361	0.891081	0.788953	0.793807
19	70	95	0.83787	0.895043	0.792377	0.796955
20	97	97	0.77538	0.886672	0.784325	0.789256
21	18	100	0.89210	0.903646	0.807512	0.810942
22	21	100	0.87542	0.902897	0.805654	0.809197
23	21	105	0.88175	0.898558	0.800929	0.804470
24	105	105	0.76190	0.877920	0.775352	0.780429
25	105	105	0.76190	0.877920	0.775352	0.780429
26	107	107	0.76013	0.875745	0.773163	0.778276
27	110	110	0.75832	0.872494	0.769920	0.775084
28	30	120	0.77575	0.883464	0.782743	0.786568
29	67	121	0.72136	0.873627	0.769570	0.774207
30	72	121	0.72136	0.872420	0.768315	0.773032
31	28	121	0.75832	0.883101	0.782787	0.786548
32	88	122	0.69497	0.867731	0.763810	0.768764
33	122	122	0.70833	0.859607	0.757383	0.762743
34	128	128	0.68805	0.853235	0.751356	0.756808
35	129	129	0.68565	0.852178	0.750366	0.755833
36	25	140	0.68565	0.867801	0.768123	0.771731
37	153	153	0.60017	0.827191	0.727735	0.733530
38	36	155	0.60941	0.852752	0.751312	0.755234
39	28	156	0.63428	0.853817	0.753868	0.757523
40	162	162	0.59036	0.818011	0.719749	0.725655
41	162	162	0.59036	0.818011	0.719749	0.725655
42	164	164	0.58696	0.815984	0.718007	0.723937
43	22	172	0.64182	0.842151	0.744574	0.747906
44	183	130	0.66838	0.838705	0.741867	0.747828
45	194	94	0.80482	0.865595	0.771873	0.777579
46	195	195	0.54244	0.785213	0.692403	0.698667
47	168	200	0.52815	0.787299	0.692272	0.698269
48	222	123	0.69842	0.835254	0.743022	0.749215
49	226	226	0.47783	0.755602	0.669036	0.675577
50	243	243	0.44130	0.739841	0.657018	0.663693
51	248	100	0.78656	0.847737	0.760154	0.766256
52	262	262	0.40866	0.722614	0.644173	0.650982
53	265	210	0.51147	0.759077	0.675467	0.682325

54	269	120	0.73790	0.826740	0.740650	0.747102
55	276	81	0.83608	0.856590	0.775218	0.781151
56	318	140	0.66584	0.799965	0.720123	0.727029
57	20	348	0.47398	0.711158	0.637026	0.638858
58	350	350	0.41142	0.647912	0.591254	0.598527
59	32	360	0.43427	0.700654	0.625950	0.628329
60	363	363	0.40518	0.637552	0.584202	0.591525
61	371	184	0.56299	0.755579	0.684093	0.691559
62	52	380	0.41051	0.683492	0.610223	0.613218
63	390	390	0.40417	0.616560	0.570082	0.577499
64	392	122	0.72302	0.797130	0.728613	0.735682
65	393	100	0.79415	0.813875	0.747320	0.754072
66	414	414	0.40886	0.598481	0.558085	0.565572
67	417	417	0.40213	0.596259	0.556620	0.564115
68	418	220	0.49720	0.720364	0.657562	0.665402
69	431	431	0.40103	0.585997	0.549879	0.557409
70	466	119	0.77464	0.782874	0.725914	0.733231
71	469	90	0.82636	0.804209	0.750977	0.757794
72	39	487	0.41953	0.618722	0.565697	0.567351
73	487	76	0.83907	0.810887	0.763165	0.769702
74	491	180	0.59117	0.733275	0.678928	0.686947
75	515	515	0.41910	0.528031	0.512445	0.520122
76	526	121	0.76674	0.768188	0.720510	0.728061
77	547	130	0.75065	0.757076	0.712079	0.719827
78	583	583	0.42573	0.485333	0.485372	0.493101
79	641	641	0.41813	0.451653	0.464179	0.471915
80	653	653	0.41028	0.444983	0.459989	0.467723
81	677	150	0.66507	0.715724	0.690040	0.698451
82	716	716	0.40945	0.411544	0.438990	0.446695
83	732	732	0.40099	0.403459	0.433907	0.441600
84	781	781	0.39221	0.379675	0.418916	0.426564
85	845	845	0.38310	0.350708	0.400543	0.408114
86	847	847	0.37363	0.349840	0.399990	0.407558
87	848	155	0.64094	0.678208	0.678789	0.687729
88	860	860	0.37120	0.344246	0.396420	0.403970
89	29	932	0.37893	0.403927	0.423999	0.421952
90	957	957	0.36857	0.305233	0.371268	0.378663
91	1030	210	0.54604	0.610690	0.635842	0.645797
92	1063	240	0.46194	0.587880	0.617116	0.627355
93	1074	120	0.76771	0.656767	0.696319	0.705151
94	1111	1111	0.38271	0.252173	0.336008	0.343104
95	1136	140	0.70116	0.632901	0.678927	0.688257
96	1156	180	0.61263	0.605790	0.650586	0.660535
97	1182	112	0.80118	0.641288	0.698957	0.707789
98	1199	91	0.84393	0.650949	0.716128	0.724417
99	1238	250	0.47865	0.553489	0.606185	0.616906
100	1258	120	0.81932	0.622480	0.690088	0.699224

101	1279	1279	0.43307	0.204751	0.302861	0.309584
102	1298	1298	0.42069	0.199984	0.299410	0.306089
103	25	1324	0.43307	0.277004	0.339842	0.335593
104	1330	96	0.86614	0.623527	0.707473	0.716087
105	1356	1356	0.43307	0.186106	0.289212	0.295755
106	1363	200	0.64262	0.559590	0.631719	0.642302
107	1377	123	0.86614	0.599475	0.684079	0.693462
108	1384	200	0.68694	0.556165	0.631135	0.641758
109	1433	236	0.58774	0.529865	0.608648	0.619690
110	1447	220	0.64158	0.535758	0.617459	0.628397
111	1470	180	0.69957	0.552705	0.641371	0.651906
112	1496	307	0.55433	0.486347	0.569905	0.581443
113	1499	1499	0.54134	0.155866	0.266052	0.272254
114	1527	1527	0.52721	0.150547	0.261819	0.267954
115	1535	1535	0.51181	0.149061	0.260627	0.266743
116	1562	1562	0.49494	0.144153	0.256657	0.262707
117	1568	1568	0.47637	0.143084	0.255785	0.261822
118	1602	139	0.86614	0.552720	0.665831	0.675876
119	1631	150	0.76990	0.542352	0.657257	0.667559
120	1674	1674	0.50949	0.125461	0.241036	0.246818
121	1799	140	0.86614	0.521050	0.660156	0.670475
122	1829	1829	0.51968	0.103523	0.221453	0.226870
123	1843	1843	0.49494	0.101741	0.219790	0.225174
124	1850	1850	0.46638	0.100862	0.218965	0.224333
125	1857	260	0.50524	0.457300	0.585622	0.597599
126	1870	230	0.62992	0.468649	0.601678	0.613426
127	2024	180	0.77952	0.469450	0.628166	0.639505
128	2133	250	0.67366	0.425480	0.585365	0.597683
129	2140	220	0.75787	0.436793	0.601781	0.613810
130	2204	2204	0.74240	0.065027	0.182018	0.186597
131	2218	2218	0.72179	0.063907	0.180729	0.185279
132	2252	150	0.86614	0.451370	0.643174	0.654218
133	2409	2409	0.86613	0.050430	0.164254	0.168414
134	2430	2430	0.86614	0.049134	0.162561	0.166679
135	2506	2506	0.86613	0.044715	0.156615	0.160587
136	2569	2569	0.86618	0.041355	0.151892	0.155745
137	2640	2640	*	0.037870	0.146782	0.150504

Table 7: Ledwina dependence measure (3.3)for each model

Pair	t_1	t_2	LED EMP	LED BB	LED GB	LED FGM
1	1	1	*	0.193353	-0.001011	0.003020
2	2	2	*	0.193299	-0.001439	0.004299
3	10	10	*	0.192865	-0.003247	0.009632
4	16	16	*	0.192540	-0.004110	0.012117
5	35	35	*	0.191512	-0.006062	0.017565
6	48	48	*	0.190811	-0.007077	0.020294

7	53	53	*	0.190541	-0.007427	0.021216
8	62	62	*	0.190056	-0.008015	0.022743
9	38	63	*	0.158984	-0.007458	0.021245
10	73	73	*	0.189464	-0.008672	0.024417
11	29	74	*	0.140135	-0.007529	0.021363
12	16	79	0.46131	0.124191	-0.006996	0.019877
13	10	80	0.59515	0.121273	-0.006516	0.018555
14	21	80	0.88159	0.128670	-0.007335	0.020793
15	38	84	0.62626	0.145115	-0.008186	0.023053
16	10	84	0.66211	0.121242	-0.006619	0.018808
17	86	86	0.50509	0.188766	-0.009381	0.026182
18	93	93	0.34641	0.188391	-0.009738	0.027054
19	70	95	0.76194	0.168492	-0.009377	0.026104
20	97	97	-1.21710	0.188176	-0.009935	0.027530
21	18	100	0.48801	0.122142	-0.007687	0.021586
22	21	100	0.48869	0.123999	-0.007878	0.022099
23	21	105	0.57499	0.123168	-0.008001	0.022386
24	105	105	0.36370	0.187748	-0.010315	0.028440
25	105	105	0.36370	0.187748	-0.010315	0.028440
26	107	107	0.34927	0.187640	-0.010407	0.028660
27	110	110	0.33457	0.187480	-0.010544	0.028983
28	30	120	0.30363	0.126312	-0.008832	0.024461
29	67	121	0.32780	0.152389	-0.010053	0.027631
30	72	121	0.40042	0.155847	-0.010167	0.027925
31	28	121	0.20316	0.124845	-0.008758	0.024258
32	88	122	0.30590	0.166027	-0.010519	0.028816
33	122	122	0.39079	0.186838	-0.011070	0.030210
34	128	128	0.34271	0.186518	-0.011322	0.030788
35	129	129	0.32798	0.186464	-0.011363	0.030882
36	25	140	0.18465	0.120759	-0.009002	0.024725
37	153	153	0.23087	0.185185	-0.012301	0.032978
38	36	155	0.15527	0.123779	-0.009841	0.026765
39	28	156	0.17473	0.120287	-0.009476	0.025810
40	162	162	0.25546	0.184707	-0.012630	0.033692
41	162	162	0.25546	0.184707	-0.012630	0.033692
42	164	164	0.23715	0.184601	-0.012701	0.033846

43	22	172	0.09767	0.118236	-0.009395	0.025447
44	183	130	0.30409	0.212919	-0.012027	0.032516
45	194	94	0.51533	0.251975	-0.010957	0.030137
46	195	195	0.20854	0.182957	-0.013746	0.036039
47	168	200	0.25894	0.171375	-0.013534	0.035480
48	222	123	0.40646	0.235889	-0.012180	0.032948
49	226	226	0.20903	0.181320	-0.014690	0.037923
50	243	243	0.15116	0.180425	-0.015172	0.038849
51	248	100	0.56437	0.271405	-0.011607	0.031708

52	262	262	0.11632	0.179428	-0.015685	0.039807
53	265	210	0.25600	0.198997	-0.014733	0.038198
54	269	120	0.43193	0.258080	-0.012449	0.033625
55	276	81	0.91799	0.313273	-0.011031	0.030383
56	318	140	0.48086	0.257357	-0.013401	0.035748
57	20	348	0.34759	0.121786	-0.011347	0.028620
58	350	350	0.21901	0.174843	-0.017778	0.043405
59	32	360	0.26658	0.114101	-0.012339	0.030880
60	363	363	0.23332	0.174171	-0.018054	0.043841
61	371	184	0.31530	0.241115	-0.014913	0.038884
62	52	380	0.25512	0.111622	-0.013517	0.033446
63	390	390	0.27556	0.172778	-0.018606	0.044685
64	392	122	0.48407	0.298986	-0.013257	0.035541
65	393	100	0.62609	0.328974	-0.012457	0.033763
66	414	414	0.29934	0.171545	-0.019073	0.045370
67	417	417	0.26522	0.171391	-0.019130	0.045452
68	418	220	0.30683	0.232394	-0.016015	0.041023
69	431	431	0.25960	0.170675	-0.019392	0.045822
70	466	119	0.54467	0.324662	-0.013505	0.036142
71	469	90	0.71187	0.372329	-0.012374	0.033602
72	39	487	0.35246	0.111143	-0.013777	0.032920
73	487	76	1.20037	0.410450	-0.011787	0.032223
74	491	180	0.42200	0.272072	-0.015460	0.040172
75	515	515	0.40492	0.166406	-0.020837	0.047702
76	526	121	0.72952	0.338212	-0.013828	0.036881
77	547	130	0.79477	0.331862	-0.014225	0.037744
78	583	583	0.43867	0.162991	-0.021873	0.048873
79	641	641	0.40002	0.160108	-0.022678	0.049677
80	653	653	0.36001	0.159515	-0.022837	0.049823
81	677	150	0.58521	0.336584	-0.015354	0.040176
82	716	716	0.35579	0.156421	-0.023629	0.050496
83	732	732	0.31274	0.155641	-0.023820	0.050643
84	781	781	0.26807	0.153265	-0.024381	0.051041
85	845	845	0.22168	0.150192	-0.025064	0.051450
86	847	847	0.17346	0.150096	-0.025085	0.051461
87	848	155	-0.03727	0.360602	-0.016042	0.041666
88	860	860	0.16111	0.149476	-0.025217	0.051530
89	29	932	0.02372	0.118366	-0.015190	0.031854
90	957	957	0.14771	0.144900	-0.026142	0.051921
91	1030	210	0.43184	0.332757	-0.018093	0.045647
92	1063	240	0.25501	0.314741	-0.018901	0.047057
93	1074	120	0.63670	0.445927	-0.015352	0.040307
94	1111	1111	0.21970	0.137805	-0.027418	0.052168
95	1136	140	0.67924	0.421245	-0.016240	0.042171
96	1156	180	0.53414	0.373881	-0.017579	0.044800
97	1182	112	0.71546	0.477327	-0.015239	0.040067

98	1199	91	0.93416	0.532028	-0.014298	0.037983
99	1238	250	0.42522	0.324452	-0.019562	0.048312
100	1258	120	0.99116	0.471168	-0.015716	0.041093
101	1279	1279	0.47604	0.130312	-0.028588	0.052051
102	1298	1298	0.41305	0.129481	-0.028708	0.052018
103	25	1324	0.33206	0.122792	-0.015721	0.029908
104	1330	96	1.17644	0.536794	-0.014769	0.039023
105	1356	1356	0.47604	0.126965	-0.029060	0.051896
106	1363	200	0.84522	0.374547	-0.018587	0.046753
107	1377	123	1.43709	0.479705	-0.016050	0.041802
108	1384	200	1.07223	0.376389	-0.018629	0.046839
109	1433	236	0.84625	0.350012	-0.019657	0.048634
110	1447	220	1.03570	0.363788	-0.019286	0.048021
111	1470	180	0.98839	0.404611	-0.018212	0.046115
112	1496	307	0.89459	0.310142	-0.021316	0.051217
113	1499	1499	1.02715	0.120897	-0.029843	0.051477
114	1527	1527	0.95526	0.119732	-0.029984	0.051378
115	1535	1535	0.87684	0.119400	-0.030023	0.051348
116	1562	1562	0.79096	0.118286	-0.030154	0.051247
117	1568	1568	0.69648	0.118039	-0.030182	0.051224
118	1602	139	1.50734	0.473532	-0.017046	0.043869
119	1631	150	1.15761	0.458208	-0.017496	0.044773
120	1674	1674	0.86505	0.113737	-0.030659	0.050784
121	1799	140	1.60573	0.488687	-0.017376	0.044544
122	1829	1829	0.91692	0.107640	-0.031268	0.050049
123	1843	1843	0.79097	0.107100	-0.031319	0.049978
124	1850	1850	0.64561	0.106831	-0.031343	0.049942
125	1857	260	0.60245	0.359609	-0.020992	0.051053
126	1870	230	1.01756	0.383764	-0.020283	0.049923
127	2024	180	1.40613	0.444725	-0.019082	0.047864
128	2133	250	1.40611	0.380850	-0.021178	0.051493
129	2140	220	1.62277	0.407215	-0.020417	0.050252
130	2204	2204	2.05062	0.093847	-0.032388	0.047971