Statistical Methods and Applications manuscript No.
(will be inserted by the editor)

On a new absolutely continuous bivariate generalized exponential distribution

S. M. Mirhosseini · M. Amini · D. Kundu · A. Dolati

Received: date / Accepted: date

Abstract In this paper we studied a three-parameter absolutely continuous bivariate distribution whose marginals are generalized exponential distributions. The proposed three-parameter bivariate distribution can be used quite effectively as an alternative to the Block and Basu bivariate exponential distribution. The joint probability density function, the joint cumulative distribution function and its associated copula have simple forms. We derive different properties of this new distribution. The maximum likelihood estimators of the unknown parameters can be obtained by solving simultaneously three non-linear equations. We propose to use EM algorithm to compute the maximum likelihood estimators, which can be implemented quite conveniently. One data set has been analyzed for illustrative purposes. Finally we propose some generalization of the proposed model.

Keywords Generalized exponential · Bivariate exponential distribution · Dependence · Measure of association · EM algorithm

Mathematics Subject Classification (2000) MSC Primary 62E15 · MSC Secondary 62H10

1 INTRODUCTION

The two-parameter generalized exponential (GE) distribution proposed by Gupta and Kundu [10] has been used quite successfully to analyze lifetime data and it has received some attention in recent years; see, e.g.[3, 15,

S. M. Mirhosseini
Department of Statistics, Faculty of Mathematical Science, Ferdowsi University of Mashhad,P.O. Box 91775-1159, Mashhad, Iran

M. Amini
Department of Statistics, Faculty of Mathematical Science, Ferdowsi University of Mashhad,P.O. Box 91775-1159, Mashhad, Iran

D. Kundu
Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India

A. Dolati
Department of Statistics, Yazd University, Yazd, 89195-741, Iran
Tel.: +983518122222
Fax: +983518210695
E-mail: adolati@yazd.ac.ir

16]. A two-parameter GE distribution with the shape parameter α and the scale parameter λ has the following cumulative distribution function (CDF)

$$F_{GE}(x;\alpha,\lambda) = \left(1-e^{-\lambda x}\right)^{\alpha}. \tag{1}$$

From now on a GE with the CDF (1) will be denoted by GE(α, λ). When α = 1, it becomes an exponential distribution with parameter λ, and we will denote it by Exp(λ). It is well known that the probability density function (PDF) and the hazard function (HF) of a GE distribution can take different shapes. If the shape parameter is less than or equal to one, the PDF and the HF are decreasing functions. When the shape parameter is greater than one, then PDF is a unimodal, and the corresponding HF is an increasing function. For some recent developments on the GE distribution, and for its different applications, the readers are referred to [11, 18] and the references cited therein.

In many reliability and life testing experiments, bivariate data arise quite naturally. Among the different bivariate lifetime model, Marshall-Olkin bivariate exponential model [22] is the most popular one. It is a singular distribution, that is a pair (X,Y) distributed as Marshall-Olkin model has this property that $P(X=Y)>0$. Block and Basu [5] introduced a three-parameter absolutely continuous bivariate exponential distribution by removing the singular part of the Marshall-Olkin model. Although Block and Basu bivariate exponential distribution does not have exponential marginals, but its marginals have decreasing PDFs and HFs. Several authors have been studied applications of the Block and Basu model; see, e.g [2]. A wide survey on different bivariate distributions can be found in [4, 17].

Recently, different bivariate exponential and bivariate generalized exponential (BGE) distributions have been proposed and studied in the literature; see, e.g. [19–21, 24, 28, 32]. Most of these models are singular distributions, and they can be used if there are ties in the data. It is not trivial to extend the univariate GE distribution to the bivariate or the multivariate case. According to Joe [13], Section 4.1, a parametric family of distributions should satisfy four desirable properties: (a) There should exist an interpretation like a mixture or other stochastic representation. (b) The margins should belong to the same parametric family and numerical evaluation should be possible. (c) The bivariate dependence between the margins should be described by a parameter and cover a wide range of dependence. (d) The distribution and density functions should preferably have a closed-form representation; at least numerical evaluation should be possible. The aim of this paper is to study a new absolutely continuous bivariate generalized exponential distribution, whose marginals are univariate GE distributions and fulfill all of the properties (a)–(d). This new three-parameter BGE distribution is obtained using the distribution of minimum order statistics of two independent samples of ordinary exponential distribution when the sample size is random; see, e.g, [7]. All joint or marginal distributions and density functions, the corresponding moments, the copula and its density have simple analytic representations that can be easily employed in applications. Moreover, the proposed distribution has positive quadrant dependent property which implies that the measures of association such as Pearson's moment correlation, Kendall's tau and Spearman's rho for this distribution vary between 0 and 1, which is a suitable range of dependence for applications.

The maximum likelihood estimators (MLEs) can be obtained by simultaneously solving three non-linear equations. The MLEs cannot be obtained in explicit forms. For known shape parameter the MLEs of the scale parameters can be obtained by using the EM algorithm. In the proposed EM algorithm, at each E-step the maximization can be performed explicitly. Hence the implementation of the EM algorithm is quite simple in this case. Finally, the shape parameter can be estimated by maximizing the profile log-likelihood function.

The rest of the paper is organized as follows. We discuss the derivation of the BGE distribution in Section 2. In Section 3 we study its different properties. Statistical inference is provided in Section 4. The analysis of a real data set is presented in Section 5. Finally we propose some generalizations and conclude the paper in Section 6.

2 BGE distribution: PDF and CDF

Following the idea given in [7], let $\{X_1, X_2, \ldots\}$ and $\{Y_1, Y_2, \ldots\}$ be two sequences of mutually independent and identically distributed (i.i.d.) random variables. It is assumed that for $k \in \{1,2,3,\ldots\}$, $X_k \sim \text{Exp}(\lambda_1)$ and $Y_k \sim \text{Exp}(\lambda_2)$. Consider a sequence of independent Bernoulli trials in which the kth trial has probability of $\frac{\alpha}{k}$, $0 < \alpha \leq 1$, $k \in \{1,2,3,...\}$ and let N be the trial number on which the first success occurs. The discrete random variable N has the probability mass function

$$P(N = n) = (1-\alpha)\left(1 - \frac{\alpha}{2}\right)\ldots\left(1 - \frac{\alpha}{n-1}\right)\frac{\alpha}{n}; \;\; n = 1,2,\ldots,$$

and its probability generating function is given by, see, e.g., [26]

$$g(t) = E(t^N) = 1 - (1-t)^{\alpha}; \quad t \in [0,1]. \tag{2}$$

Let

$$U = \min(X_1, ..., X_N) \quad \text{and} \quad V = \min(Y_1, ..., Y_N).$$

The joint survival function of (U,V) is then

$$\begin{aligned}
\overline{F}(u,v) &= P\{U \geq u, V \geq v\} \\
&= \sum_{n=1}^{\infty} [P(X_i \geq u)P(Y_i \geq v)]^n P(N = n) \\
&= g\left(e^{-(\lambda_1 u + \lambda_2 v)}\right) \\
&= 1 - \{1 - e^{-(\lambda_1 u + \lambda_2 v)}\}^{\alpha}.
\end{aligned} \tag{3}$$

Since $\overline{F}(u,v) = 1 - F_1(u) - F_2(v) + F(u,v)$, the associated joint distribution function is given by

$$F(u,v) = (1-e^{-\lambda_1 u})^{\alpha} + (1-e^{-\lambda_2 v})^{\alpha} - \{1 - e^{-(\lambda_1 u + \lambda_2 v)}\}^{\alpha}, \; u,v \geq 0, \lambda_1, \lambda_2 > 0, 0 < \alpha \leq 1. \tag{4}$$

From (4), by taking $v \to \infty$ or $u \to \infty$, it follows that

$$U \sim \text{GE}(\alpha, \lambda_1) \quad \text{and} \quad V \sim \text{GE}(\alpha, \lambda_2). \tag{5}$$

Therefore, it follows that the marginals of (4) are GE distributions. A pair (U,V) distributed as (4) is said to have BGE with parameters α, λ_1 and λ_2, and it will be denoted by BGE$(\alpha, \lambda_1, \lambda_2)$. The joint density function of BGE is given by

$$f(u,v) = \lambda_1 \lambda_2 \alpha e^{-(\lambda_1 u + \lambda_2 v)} \{1 - \alpha e^{-(\lambda_1 u + \lambda_2 v)}\} \{1 - e^{-(\lambda_1 u + \lambda_2 v)}\}^{\alpha - 2}. \tag{6}$$

It can be easily seen that the joint PDF of BGE is a decreasing function in both arguments, and it resembles the joint PDF of the Block and Basu bivariate exponential distribution. Plots of the CDF and PDF of the BGE distribution are provided in Fig. 1.

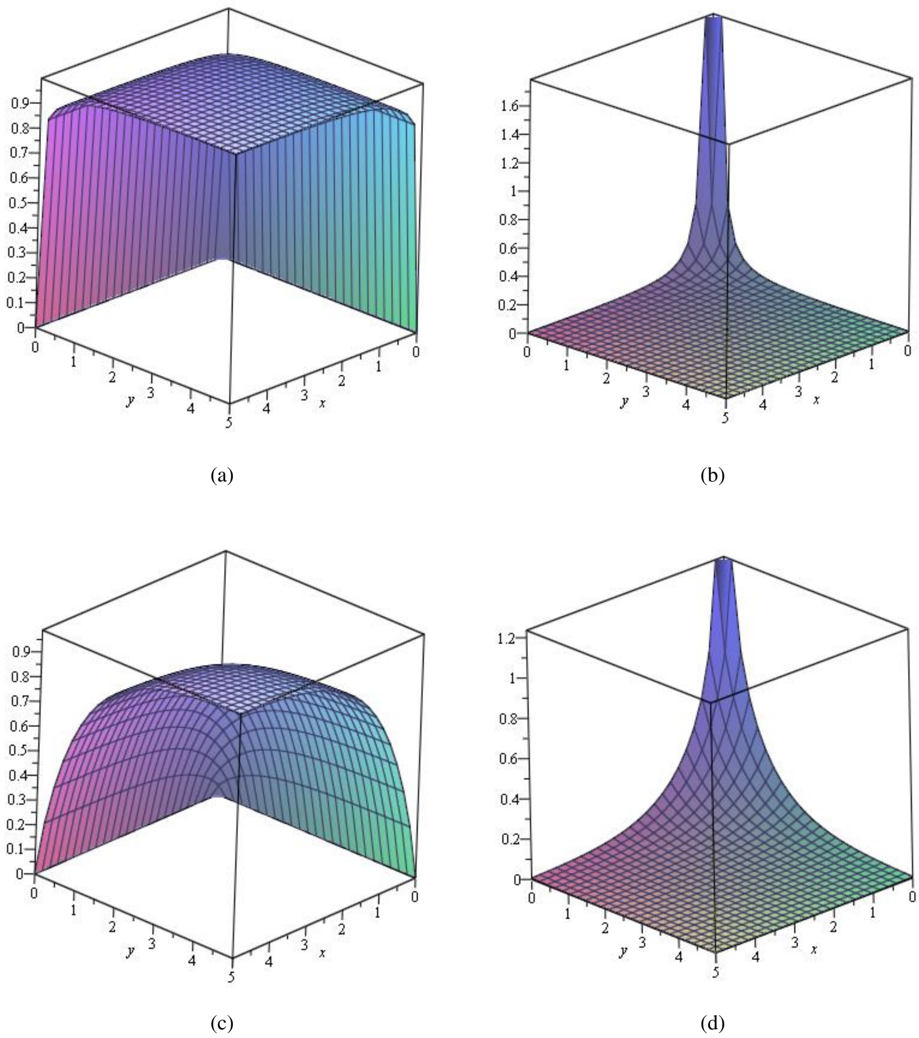

Fig. 1 Plots of the joint CDF (left panels) and PDF (right panels) of BGE distribution. In (a) , (b) $(\alpha, \lambda_1, \lambda_2) = (0.1, 1, 1)$ and in (c), (d) $(\alpha, \lambda_1, \lambda_2) = (0.9, 1, 1)$.

Since $0 < 1 - e^{-(\lambda_1 u + \lambda_2 v)} < 1$, for $u, v > 0$, by using the binomial series expansion we have

$$(1 - e^{-(\lambda_1 u + \lambda_2 v)})^\alpha = \sum_{j=0}^{\infty} \binom{\alpha}{j} (-1)^j e^{-j(\lambda_1 u + \lambda_2 v)},$$

and the joint survival function (3) can be rewritten as

$$\overline{F}(u, v) = \sum_{j=1}^{\infty} \binom{\alpha}{j} (-1)^{j+1} e^{-j(\lambda_1 u + \lambda_2 v)}.$$

We observe that the infinite series is summable, differentiable and hence by differentiating with respect to u and v we get

$$f(u,v)=\lambda_1\lambda_2\sum_{j=1}^{\infty}\binom{\alpha}{j}(-1)^{j+1}j^2e^{-j(\lambda_1u+\lambda_2v)}. \tag{7}$$

3 PROPERTIES

3.1 MOMENTS & CONDITIONAL MOMENTS

First we provide expressions for the joint moment generating function (mgf) and Pearson's correlation coefficient of the BGE model.

Proposition 1 *If the random vector* $(U,V)\sim \mathrm{BGE}(\lambda_1,\lambda_2,\alpha)$ *then*

(a) the joint m.g.f. of (U,V)*, for* $|t|<\lambda_1$ *and* $|s|<\lambda_2$*, is given by*

$$M_{U,V}(t,s)=\sum_{j=1}^{\infty}\binom{\alpha}{j}(-1)^{j+1}\frac{\lambda_1\lambda_2j^2}{(j\lambda_1-t)(j\lambda_2-s)}, \tag{8}$$

(b) the Pearson's correlation coefficient of (U,V) *is given by*

$$\mathrm{Corr}(U,V)=\frac{1}{\kappa(\alpha,1)}\left(\sum_{j=1}^{\infty}\frac{\binom{\alpha}{j}(-1)^{j+1}}{j^2}-\gamma(\alpha,1)^2\right), \tag{9}$$

where

$$\gamma(\alpha,\beta)=\Psi(\alpha+\beta)-\Psi(\beta),\quad \text{and } \kappa(\alpha,\beta)=\Psi'(\beta)-\Psi'(\alpha+\beta),$$
$$\Psi(t)=\frac{d}{dt}\ln\Gamma(t),$$

is the digamma function, $\Psi'(.)$ *is its derivative and*

$$\Gamma(t)=\int_0^{\infty}x^{t-1}e^{-x}dx,\quad t>0,$$

is the gamma function [9].

Proof Starting from $M_{U,V}(t,s)=E(e^{tU+sV})$, using (7) we have

$$M_{U,V}(t,s)=\lambda_1\lambda_2\int_0^{\infty}\int_0^{\infty}\sum_{j=1}^{\infty}\binom{\alpha}{j}(-1)^{j+1}j^2e^{-(j\lambda_1-t)x-(j\lambda_2-s)y}dxdy.$$

Since the quantity inside the summation is absolutely integrable, interchanging the summation and integration we have the required result. For part (b) using (7), direct calculation shows that the product moment could be obtained as

$$E(UV)=\sum_{j=1}^{\infty}\frac{\binom{\alpha}{j}(-1)^{j+1}}{\lambda_1\lambda_2j^2}.$$

Since $U\sim \mathrm{GE}(\alpha,\lambda_1)$ and $V\sim \mathrm{GE}(\alpha,\lambda_2)$, from [10], using

$$E(U)=\frac{\Psi(\alpha+1)-\Psi(1)}{\lambda_1},\quad E(V)=\frac{\Psi(\alpha+1)-\Psi(1)}{\lambda_2}$$

and

$$\mathrm{Var}(U)=\frac{\Psi'(1)-\Psi'(\alpha+1)}{\lambda_1^2},\quad \mathrm{Var}(V)=\frac{\Psi'(1)-\Psi'(\alpha+1)}{\lambda_2^2}$$

the expression for correlation coefficient is immediate.

The following result gives the conditional moments of BGE model.

Proposition 2 *If the random vector* $(U,V)\sim \mathrm{BGE}(\lambda_1,\lambda_2,\alpha)$*, then the conditional expectation of* V *given* U*, is given by*

$$E(V|U=u)=\frac{\lambda_1}{\lambda_2}\frac{1}{h_U(u)},$$

where $h_U(u)=f_U(u)/\bar{F}_U(u)$ *is hazard rate function of* U.

Proof The conditional density function of $(V|U=u)$ is given by

$$f_{V|U=u}(v)=\frac{\lambda_2}{\alpha e^{-\lambda_1 u}(1-e^{-\lambda_1 u})^{\alpha-1}}\sum_{j=1}^{\infty}\binom{\alpha}{j}(-1)^{j+1}j^2 e^{-(j\lambda_1 u+\lambda_2 v)}.$$

Thus,

$$\begin{aligned}
E(V|U=u) &= \int_0^\infty x f_{X_2|X_1=u}(x)dx \\
&= \frac{1}{\alpha(1-e^{-\lambda_1 u})^{\alpha-1}}\sum_{j=1}^{\infty}\binom{\alpha}{j}(-1)^{j+1}je^{-(j-1)\lambda_1 u}\left(\int_0^\infty j\lambda_2 x e^{-j\lambda_2 x}dx\right) \\
&= \frac{1}{\alpha\lambda_2(1-e^{-\lambda_1 u})^{\alpha-1}}\sum_{j=1}^{\infty}\binom{\alpha}{j}(-1)^{j+1}e^{-\lambda_1(j-1)u} \\
&= \frac{1-(1-e^{-\lambda_1 u})^{\alpha}}{\alpha\lambda_2 e^{-\lambda_1 u}(1-e^{-\lambda_1 u})^{\alpha-1}} \\
&= \frac{\lambda_1}{\lambda_2}\frac{\overline{F}_U(u)}{f_U(u)}.
\end{aligned}$$

It is known that (see, e.g., [10]) the hazard rate function of univariate GE distribution with the shape parameter α, has a decreasing hazard function if $\alpha\le 1$. Therefore, we have the following result.

Corollary 1 *If the random vector* $(U,V)\sim \mathrm{BGE}(\lambda_1,\lambda_2,\alpha)$*, then* $E(V|U=u)$*, is an increasing function in u.*

The following results will be useful for development of the EM algorithm. If (U,V) and N are same as defined before, then the joint PDF of U, V and N can be written as

$$f(u,v,n)=n^2\lambda_1\lambda_2 e^{-n(\lambda_1 u+\lambda_2 v)}P(N=n);\ \ u>0,v>0,n=1,2,\ldots. \tag{10}$$

Hence the conditional probability of $N=n$, given U and V can be written as

$$P(N=n|U=u,V=v)=Kn^2e^{-nA}P(N=n);\ \ n=1,2,\ldots, \tag{11}$$

where $A=\lambda_1 u+\lambda_2 v$, and

$$K^{-1}=\alpha e^{-A}\{1-\alpha e^{-A}\}\{1-e^{-A}\}^{\alpha-2}.$$

Proposition 3 *If* $(U,V) \sim \text{BGE}(\lambda_1, \lambda_2, \alpha)$*, and* N *is same as defined before, then*

$$E(N|U=u,V=v) = K\alpha e^{-A}(1-e^{-A})^{\alpha-3}\left\{e^{-2A}(\alpha^2-6\alpha+3)+e^{-A}(3\alpha+1)-2\right\}.$$

Proof If we denote

$$h(t) = E\left(te^{-A}\right)^N,$$

then it follows that

$$\sum_{n=1}^{\infty} n^3 e^{-A} P(N=n) = h^{'''}(1) + 3h^{''}(1) + h^{'}(1).$$

Now the result follows by using the fact that from (2)

$$h(t) = 1-\left(1-te^{-A}\right)^{\alpha}; \quad t\in[0,1].$$

3.2 Stress-strength parameter & Distribution of the Minimum

The stress-strength parameter, $R = P(U < V)$, is useful for data analysis purposes. The following result gives a convenient form for the stress-strength parameter of BGE model.

Proposition 4 *If* $(U,V) \sim \text{BGE}(\lambda_1, \lambda_2, \alpha)$*, then*

$$P(U<V) = \frac{\lambda_2}{\lambda_1+\lambda_2}.$$

Proof From (7) we have

$$\begin{aligned}
P(U<V) &= \int_0^{\infty}\int_0^{y} f(x,y)dxdy \\
&= \sum_{j=1}^{\infty}\binom{\alpha}{j}(-1)^{j+1} j^2 \int_0^{\infty}\int_0^{y}\lambda_1\lambda_2 e^{-j(\lambda_1 x+\lambda_2 y)}dxdy \\
&= \sum_{j=1}^{\infty}\binom{\alpha}{j}(-1)^{j+1}\int_0^{\infty}\lambda_2 j e^{-j\lambda_2 y}(1-e^{-j\lambda_1 y})dy \\
&= \frac{\lambda_2}{\lambda_1+\lambda_2}\sum_{j=1}^{\infty}\binom{\alpha}{j}(-1)^{j+1} \\
&= \frac{\lambda_2}{\lambda_1+\lambda_2},
\end{aligned}$$

which completes the proof.

Proposition 5 *If* $(U,V) \sim \text{BGE}(\lambda_1, \lambda_2, \alpha)$*, then*

$$W = \min(U,V) \sim GE(\alpha, \lambda_1+\lambda_2).$$

Proof

$$\begin{aligned}
P(W \geq w) &= P(U\geq w, V\geq w) \\
&= \sum_{n=1}^{\infty}[P(X_i\geq w)P(Y_i\geq w)]^n P(N=n) \\
&= g\left(e^{-(\lambda_1+\lambda_2)w}\right) \\
&= 1-\left\{1-e^{-(\lambda_1+\lambda_2)w}\right\}^{\alpha}
\end{aligned}$$

which completes the proof.

3.3 Rényi Entropy

The entropy of a random vector (U,V) with the joint pdf $f(x,y)$ is a measure of variation of the uncertainty. Rényi entropy [27] is defined by

$$I_R(\beta) = \frac{1}{1-\beta}\log\{T(\beta)\},$$

for $\beta > 0$, where

$$T(\beta) = \int f^{\beta}(x,y)dxdy.$$

Proposition 6 *Suppose that* $(U,V) \sim \mathrm{BGE}(\lambda_1,\lambda_2,\alpha)$. *Then*

$$T(\beta) = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{\binom{\beta}{j}\binom{\beta(\alpha-2)}{k}\alpha^{j+\beta}(\lambda_1\lambda_2)^{\beta-1}}{(j+k+\beta-1)^2}. \tag{12}$$

Proof From (6) and substituting the transformations $x = e^{-\lambda_1 u}$ and $y = e^{-\lambda_2 v}$, we have

$$T(\beta) = \int_0^1\int_0^1 \alpha^{\beta}(\lambda_1\lambda_2)^{\beta-1}(xy)^{\beta-1}\Big(1-\alpha xy\Big)^{\beta}\Big(1-xy\Big)^{(\alpha-2)\beta}dxdy.$$

By applying the binomial series expansion to the integrand and integrating with respect to x and y, we have the required result.

3.4 Dependence properties

In what follows we discuss the dependence properties of the BGE distribution through its associated copula. In view of *Sklar's Theorem* [33], solving the equation

$$C_\alpha\{F_1(x),F_2(y)\} = F(x,y),$$

for the function $C_\alpha : [0,1]\times[0,1] \to [0,1]$ yields the underlying copula associated with the pair (U,V) having the BGE distribution defined by (4) as

$$\begin{aligned} C_\alpha(u,v) &= u+v-\{1-(1-u^{\frac{1}{\alpha}})(1-v^{\frac{1}{\alpha}})\}^{\alpha} \\ &= u+v-uv\{u^{-\frac{1}{\alpha}}+v^{-\frac{1}{\alpha}}-1\}^{\alpha}, \end{aligned} \tag{13}$$

for all $u,v \in (0,1)$ and $0 < \alpha \le 1$. The theory and applications of copulas are well documented in [25]. The survival copula $\widehat{C}_\alpha(u,v) = u+v-1+C_\alpha(1-u,1-v)$, associated to C_α is given by

$$\widehat{C}_\alpha(u,v) = 1-(1-u)(1-v)\{(1-u)^{-\frac{1}{\alpha}}+(1-v)^{-\frac{1}{\alpha}}-1\}^{\alpha}, \tag{14}$$

for all $u,v \in (0,1)$ and $0 < \alpha \le 1$. The copula (13) is not a member of the Archimedean family of copulas [25], but its associated survival copula (14) turns to be an Archimedean copula with the strict generator $\phi(t) = -\ln(1-(1-t)^{\frac{1}{\alpha}})$ (see, Example 3 in [7]).

Recall that for two copulas C_1 and C_2, we say that C_2 is *more concordant than* C_1 (written $C_1 \prec_c C_2$) if $C_1(u,v) \le C_2(u,v)$, or equivalently $\widehat{C}_1(u,v) \le \widehat{C}_2(u,v)$, for all $u,v \in (0,1)$. A pair (X_1,X_2) with the copula C is *positively quadrant dependent* (written PQD) if $\Pi \prec_c C$, where $\Pi(u,v) = uv$ is the product copula [25].

The dependence properties of the BGE distribution depend only on the parameter α. The following result provides the dependence ordering of the BGE family of distributions with respect to the parameter α.

Proposition 7 *The copula C_α defined by (13) is negatively ordered with respect to α; that is for $\alpha_1, \alpha_2 \in (0,1]$, such that $\alpha_1 \leq \alpha_2$, we have $C_{\alpha_2} \prec_c C_{\alpha_1}$.*

Proof Since $\widehat{C}_\alpha$ is an Archimedean copula with the generator $\phi_\alpha(t) = -\ln(1-(1-t)^{\frac{1}{\alpha}})$, and $\phi_{\alpha_2} \circ \phi_{\alpha_1}^{-1}(t) = -\ln(1-(1-e^{-t})^{\frac{\alpha_2}{\alpha_1}})$ is non-decreasing in t, for $\alpha_1 \leq \alpha_2$, in view of Theorem 4.4.2 in [25], we have $\widehat{C}_{\alpha_2} \prec_c \widehat{C}_{\alpha_1}$, or equivalently, $C_{\alpha_2} \prec_c C_{\alpha_1}$.

Corollary 2 *Suppose that $(U,V) \sim \mathrm{BGE}(\lambda_1, \lambda_2, \alpha)$. Then (U,V) is* PQD.

Proof As a consequence of Proposition 7, for $\alpha \leq 1$ we have that $\Pi(u,v) = C_1(u,v) \leq C_\alpha(u,v)$ for all $u, v \in (0,1)$.

Remark 1 Since the BGE distribution defined by (4) has the PQD property, it is suitable to describe the positive dependence of a random pair (U,V). However, it is very simple to consider a distribution to describe a negative dependence. It suffices to consider the copula C^* given by $C^*_\alpha(u,v) = u - C_\alpha(u, 1-v)$. It is obvious that the properties of a copula C^*_α can be obtained in a simple way from the corresponding properties of a copula C_α; see, [25] for detail.

Let (U,V) and (U',V') be two continuous random vectors with the same univariate marginals and the respective joint density functions f and g. The pair (U,V) is said to be more positive likelihood ratio dependent (PLRD) than the pair (U',V'), denoted by $(U,V) \prec_{\mathrm{PLRD}} (U',V')$, if

$$f(x_1,y_1)f(x_2,y_2)g(x_1,y_2)g(x_2,y_1) \geq f(x_1,y_2)f(x_2,y_1)g(x_1,y_1)g(x_2,y_2), \tag{15}$$

whenever $x_1 \leq x_2$ and $x_2 \leq y_2$ [31]. When U' and V' are independent, then the pair (U,V) is said to be PLRD and the condition (15) reduces to $f(x_1,y_1)f(x_2,y_2) \geq f(x_1,y_2)f(x_2,y_1)$. Holland and Wang [12] showed that a sufficient condition for PLRD in the case of continuous random variables is that $\frac{\partial^2}{\partial x \partial y} \ln f(x,y) \geq 0$.

Proposition 8 *Suppose that $(U,V) \sim \mathrm{BGE}(\lambda_1, \lambda_2, \alpha)$. Then (U,V) is PLRD.*

Proof Let (U,V) be a random vector with the joint distribution function $G^u(x,y) = x^\alpha + y^\alpha - (x+y-xy)^\alpha$, $x, y \in [0,1]$, $0 < \alpha \leq 1$. Let $g^u(.,.)$ be the density function associated with $G^u(.,.)$. Since $\frac{\partial^2}{\partial x \partial y} \ln g^u(x,y) \geq 0$, for all $x, y \in [0,1]$, $0 < \alpha \leq 1$, by the result of Holland and Wang [12] the pair (U,V) is PLRD, i.e., $(U',V') \prec_{\mathrm{PLRD}} (U,V)$, where U' and V' are independent with the same univariate marginal distribution as U and V. For $i = 1,2$, let $\phi_i(t)$ be the increasing mapping $t \to -\frac{1}{\lambda_i}\ln(1-t)$. Since the pair (U,V) has the same joint distribution as the pair $(\phi_1(U), \phi_2(V)$, in view of Theorem 9.D.2 in [31], we have the required result.

In the following we discuss the tail dependence properties of the BGE distribution. For a pair (U,V) with the copula D, the lower (resp, upper) tail dependence coefficient, λ_L (resp, λ_U) is defined by [13,25]

$$\lambda_L(U,V) = \lim_{u \to 0^+} \frac{D(u,u)}{u}, \tag{16}$$

and

$$\lambda_U(U,V) = 2 - \lim_{u \to 1^-} \frac{1 - D(u,u)}{1-u}. \tag{17}$$

Proposition 9 *Suppose that* $(U,V) \sim \mathrm{BGE}(\lambda_1, \lambda_2, \alpha)$. *Then*

$$\lambda_L(U,V) = 0, \quad \lambda_U(U,V) = 0.$$

Proof By taking into account (16), from (13), the lower tail dependence coefficient of (U,V) can be expressed as

$$\lambda_L(U,V) = \lim_{u \to 0^+} \frac{u(2-(2-u^{\frac{1}{\alpha}})^{\alpha})}{u} = 0.$$

By taking into account (17), the upper tail dependence coefficient of (U,V) can be expressed as

$$\lambda_U(U,V) = 2 - \lim_{u \to 1^-} \frac{1-u(2-(2-u^{\frac{1}{\alpha}})^{\alpha})}{1-u} = 0.$$

In the rest of this section we provide expressions for some well-known measures of association for a vector (U,V) having BGE distribution. The population version of two of the most common nonparametric measures of association between the components of a continuous random pair (U,V) are *Kendall's tau* (τ) and *Spearman's rho* (ρ) which depend only on the copula C of the pair (U,V), and are given by

$$\tau(C) = 4 \int_0^1 \int_0^1 C(u,v)\, dC(u,v) - 1, \tag{18}$$

and

$$\rho(C) = 12 \int_0^1 \int_0^1 C(u,v) du dv - 3. \tag{19}$$

See [25] for detail.

Proposition 10 *Suppose that* $(U,V) \sim BGE(\lambda_1, \lambda_2, \alpha)$. *Then*

$$\tau(U,V) = 1 + 4\alpha B(2, 2\alpha - 1)(\Psi(2) - \Psi(2\alpha + 1)),$$

and

$$\rho(U,V) = 9 - 12\alpha^2 \sum_{j=0}^{\infty} (-1)^j \binom{\alpha}{j} [B(j,\alpha)]^2,$$

where $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1} dx$, *denotes the beta function.*

Proof First recall that for a copula–based measure of association κ, satisfying Scarsini's axioms [30] and any pair of continuous random variables (T,S) with associated copula C, one has $\kappa(T,S) = \kappa(-T,-S)$, or equivalently, $\kappa(C) = \kappa(\widehat{C})$. Now the result follows from Proposition 9 in [7].

Remark 2 Note that as a consequence of the PQD property of the BGE distribution, for $(U,V) \sim \mathrm{BGE}(\lambda_1, \lambda_2, \alpha)$, we have $\mathrm{Corr}(U,V) \geq 0$, $\tau(U,V) \geq 0$ and $\rho(U,V) \geq 0$.

We provide the values of the Kendall's tau and the Spearman's rho for BGE distribution in Table 1. It is immediate that as α increases, both Kendall's tau and Spearman's rho decrease to 0, as it should be. Moreover, it is observed that $\dfrac{\rho_s(\alpha)}{\tau(\alpha)} = \dfrac{3}{2}$ as $\alpha \longrightarrow 1$, although we could not proved it theoretically.

α	$\rho(\alpha)$	$\tau(\alpha)$	$\frac{\rho_s(\alpha)}{\tau(\alpha)}$
0.0001	0.999999935	0.9998000258	1.000200
0.1001	0.952423077	0.8218848820	1.158828
0.2001	0.854503656	0.6770884892	1.262027
0.3001	0.739507866	0.5547812933	1.332972
0.4001	0.620781827	0.4487290130	1.383423
0.5001	0.504091758	0.3549780000	1.420065
0.6001	0.391946075	0.2708654200	1.447014
0.7001	0.285344232	0.1945181220	1.466929
0.8001	0.184556871	0.1245645610	1.481616
0.9001	0.089498325	0.0599718330	1.492339

Table 1 Kendall's tau and Spearman's rho associated with the family of distributions (4) for different values of $\alpha \in (0,1]$.

4 Statistical Inference

4.1 Maximum Likelihood Estimation

In this section we discuss the MLEs of the parameters of BGE distribution, based on a random sample of size m, namely $\{(u_1,v_1),\ldots,(u_m,v_m)\}$ from BGE($\lambda_1,\lambda_2,\alpha$). The log-likelihood function becomes

$$l(\theta) = m\ln(\alpha\lambda_1\lambda_2) - \left(\lambda_1\sum_{i=1}^{m}u_i + \lambda_2\sum_{i=1}^{m}v_i\right) + (\alpha-2)\sum_{i=1}^{m}\ln\left(1-e^{-(\lambda_1 u_i+\lambda_2 v_i)}\right) + \sum_{i=1}^{m}\ln\left(1-\alpha e^{-(\lambda_1 u_i+\lambda_2 v_i)}\right) \quad (20)$$

where $\theta = (\alpha,\lambda_1,\lambda_2)$. The maximum likelihood estimates can be obtained by maximizing (20) with respect to the unknown parameters. The three normal equations become;

$$\frac{\partial l(\theta)}{\partial\alpha} = \frac{m}{\alpha} + \sum_{i=1}^{m}\ln\left(1-e^{-(\lambda_1 u_i+\lambda_2 v_i)}\right) - \sum_{i=1}^{m}\frac{e^{-(\lambda_1 x_i+\lambda_2 y_i)}}{1-\alpha e^{-(\lambda_1 u_i+\lambda_2 v_i)}} = 0,$$

$$\frac{\partial l(\theta)}{\partial\lambda_1} = \frac{m}{\lambda_1} - \sum_{i=1}^{m}u_i + (\alpha-2)\sum_{i=1}^{m}\frac{u_i e^{-(\lambda_1 u_i+\lambda_2 v_i)}}{1-e^{-(\lambda_1 u_i+\lambda_2 v_i)}} + \alpha\sum_{i=1}^{m}\frac{u_i e^{-(\lambda_1 u_i+\lambda_2 v_i)}}{1-\alpha e^{-(\lambda_1 u_i+\lambda_2 v_i)}} = 0,$$

$$\frac{\partial l(\theta)}{\partial\lambda_2} = \frac{m}{\lambda_2} - \sum_{i=1}^{m}v_i + (\alpha-2)\sum_{i=1}^{m}\frac{v_i e^{-(\lambda_1 u_i+\lambda_2 v_i)}}{1-e^{-(\lambda_1 u_i+\lambda_2 v_i)}} + \alpha\sum_{i=1}^{m}\frac{v_i e^{-(\lambda_1 u_i+\lambda_2 v_i)}}{1-\alpha e^{-(\lambda_1 u_i+\lambda_2 v_i)}} = 0.$$

Note that the Newton-Raphson method or other optimization routine may be used to maximize (20).

To avoid that we propose to use the profile likelihood method to compute the MLEs of the unknown parameters. For fixed α, the MLEs of λ_1 and λ_2 can be obtained by maximizing the profile log-likelihood function with respect to λ_1 and λ_2. We use EM algorithm to compute the MLEs of λ_1 and λ_2 for a given α, and finally we maximize the profile log-likelihood function of α, to compute the MLE of α. For implementing the EM algorithm, we treat the problem as a missing value problem. Suppose along with (u,v), we observe the associated N value also. Therefore, the complete observations are as follows: $\{(u_1,v_1,n_1),\ldots,(u_m,v_m,n_m)\}$. Based on the complete sample, the log-likelihood function without the additive constant becomes

$$l_c(\lambda_1,\lambda_2) = m\ln\lambda_1 + m\ln\lambda_2 - \lambda_1\sum_{i=1}^{m}n_i u_i - \lambda_2\sum_{i=1}^{m}n_i v_i. \quad (21)$$

For fixed α, the MLEs of λ_1 and λ_2 become

$$\widehat{\lambda}_1(\alpha) = \frac{m}{\sum_{i=1}^m n_i u_i} \quad \text{and} \quad \widehat{\lambda}_2(\alpha) = \frac{m}{\sum_{i=1}^m n_i v_i}. \tag{22}$$

Now we are in a position to provide the EM algorithm. For fixed α, at the E-step of the EM algorithm, we construct the pseudo log-likelihood function at the k-th iterate by replacing the true value of N by its expected value. It takes the following form;

$$\begin{aligned} l_s(\lambda_1,\lambda_2|\lambda_1^{(k)}(\alpha),\lambda_2^{(k)}(\alpha)) = m\ln\lambda_1 + m\ln\lambda_2 - \lambda_1\sum_{i=1}^m u_i E(N|u_i,v_i,\lambda_1^{(k)}(\alpha),\lambda_2^{(k)}(\alpha),\alpha) \\ -\lambda_2\sum_{i=1}^m v_i E(N|u_i,v_i,\lambda_1^{(k)}(\alpha),\lambda_2^{(k)}(\alpha),\alpha). \end{aligned} \tag{23}$$

Here $\lambda_1^{(k)}(\alpha)$ and $\lambda_2^{(k)}(\alpha)$, are the values of λ_1 and λ_2, respectively at the k-th iterate and $E(N|u_i,v_i,\lambda_1^{(k)}(\alpha),\lambda_2^{(k)}(\alpha),\alpha)$ can be obtained using Proposition 4. Hence at the M-step, the maximization can be easily performed to obtain $\lambda_1^{(k+1)}(\alpha)$ and $\lambda_2^{(k+1)}(\alpha)$ as

$$\lambda_1^{(k+1)}(\alpha) = \frac{m}{\sum_{i=1}^m C_i u_i} \quad \text{and} \quad \lambda_2^{(k+1)}(\alpha) = \frac{m}{\sum_{i=1}^m C_i v_i}, \tag{24}$$

where $C_i = E(N|u_i,v_i,\lambda_1^{(k)}(\alpha),\lambda_2^{(k)}(\alpha),\alpha)$. Continue the iteration until the convergence is met. Let us denote these estimates as $\widehat{\lambda}_1(\alpha)$ and $\widehat{\lambda}_2(\alpha)$. Finally maximize the profile log-likelihood function of α, to obtain the MLE of α, say $\widehat{\alpha}$. Therefore, the MLEs of α, λ_1 and λ_2 become $\widehat{\alpha}$, $\widehat{\lambda}_1(\widehat{\alpha})$ and $\widehat{\lambda}_2(\widehat{\alpha})$, respectively.

It can be easily seen that the density function of the BGE distribution satisfies all required conditions for the MLEs to be consistent and asymptotically normally distributed. We have the following result.

Proposition 11 *If $\widehat{\theta}$ is the MLE of θ, then*

$$\sqrt{n}(\theta - \hat{\theta}) \longrightarrow^d N_3(0, I^{-1}). \tag{25}$$

Here $\longrightarrow^d$ means convergence in distribution and $N_3(0,I^{-1})$, denotes the 3-variate normal distribution with mean vector 0 and the covariance matrix I^{-1}, and the matrix I is the Fisher information matrix; the elements of the matrix I are presented in the Appendix.

4.2 Testing of Hypotheses

We perform the following two testing of hypotheses problems.

PROBLEM 1: Testing whether the two marginals have the same distributions or not, can be carried out as follows:

$$H_0 : \lambda_1 = \lambda_2 \quad \text{versus} \quad H_1 : \lambda_1 \neq \lambda_2. \tag{26}$$

In this case the MLE of $\lambda = \lambda_1 = \lambda_2$ can be obtained along the same line as before. The pseudo log-likelihood function becomes;

$$l_s(\lambda|\lambda^{(k)}(\alpha)) = 2m\ln\lambda - \lambda\sum_{i=1}^m (u_i+v_i)E(N|u_i,v_i,\lambda^{(k)}(\alpha),\alpha), \tag{27}$$

where $\lambda^{(k)}$ denotes the estimate of λ at the k-th iteration. If $D_i = E(N|u_i, v_i, \lambda^{(k)}(\alpha)$, then

$$\lambda^{(k+1)}(\alpha) = \frac{2m}{\sum_{i=1}^{m} D_i(u_i + v_i)}. \tag{28}$$

Similarly, as before the estimate of α can be obtained by maximizing the profile log-likelihood function of α. Now if we denote the estimates of α and λ as $\widetilde{\alpha}$ and $\widetilde{\lambda}$, respectively, then using the standard likelihood ratio, which has the asymptotic distribution as follows:

$$2\left(l(\widehat{\alpha}, \widehat{\lambda}_1, \widehat{\lambda}_2) - l(\widetilde{\alpha}, \widetilde{\lambda}, \widetilde{\lambda})\right) \longrightarrow \chi_1^2.$$

PROBLEM 2: If we want to test whether the two components are independent or not, the following test can be performed:

$$H_0 : \alpha = 1 \quad \text{versus} \quad H_1 : \alpha \neq 1. \tag{29}$$

Under the null hypothesis the MLEs of λ_1 and λ_2 can be obtained as

$$\widetilde{\lambda}_1 = \frac{\sum_{i=1}^{m} u_i}{m} \quad \text{and} \quad \widetilde{\lambda}_2 = \frac{\sum_{i=1}^{m} v_i}{m}. \tag{30}$$

In this case, since α is in the boundary under H_0, the standard results do not apply. But using Theorem 3 of Self and Liang [29], it follows that

$$2\left(l(\widehat{\alpha}, \widehat{\lambda}_1, \widehat{\lambda}_2) - l(1, \widetilde{\lambda}_1, \widetilde{\lambda}_2)\right) \longrightarrow \frac{1}{2} + \frac{1}{2}\chi_1^2. \tag{31}$$

5 DATA ANALYSIS

In this section we provide the analysis of a real data set from McGilchrist and Aisbett [23]. The data has been obtained from an experiment, where M individuals are observed and times between recurrence of a particular type of event are recorded. In this study the recurrence time of infection in kidney patients who are using a portable dialysis machine are recorded. The infection occurs at the point of infection of the catheter, and when it occurs, the catheter has to be removed, the infection cleared up and then the catheter reinserted. Recurrence times are times from infection until next infection. For each patient two such recurrence times are given; namely first recurrence time (FRT) and second recurrence time (SRT). The data for 23 patients are reported in Table 2

Before, progressing further, we obtain some basic statistics of the first recurrence time and second recurrence time, and they are reported in Table 3. It is clear from Table 3 that both FRT and SRT have very long right tail. To get an idea about the shape of the empirical hazard function of the marginals, we provide the scaled TTT plots of FRT and SRT in Figure 2, as suggested by Aarset [1]. This plot provides an idea of the shape of the hazard function of a distribution. It has been shown in [1] that the scaled TTT transform is convex (concave) if the hazard rate is decreasing (increasing). For this data set, it indicates that both of variables have decreasing empirical hazard functions. Note that from Table 3 the Spearman's rho and Kendall's tau for FRT and SRT data are given by 0.266 and 0.184, respectively. This observation demonstrates an obvious positive dependence between involved data. This observation shows that the proposed BGE may be used for analyzing this bivariate data set.

No.	FRT	SRT	No.	FRT	SRT
1	8	16	2	22	28
3	447	318	4	30	12
5	24	245	6	7	9
7	511	30	8	53	196
9	15	154	10	7	333
11	96	38	12	185	177
13	292	114	14	152	562
15	13	66	16	12	40
17	132	156	18	34	30
19	2	25	20	130	26
21	27	58	22	152	30
23	119	8			

Table 2 Kidney infection data of 23 patients. The patient No., first recurrence time (FRT), second recurrence time (SRT) are reported.

Statistics	FRT	SRT
Minimum	2	8
1st Quartile	13	26
Median	34	40
Mean	107.391	116.130
3st Quartile	152	177
Maximum	511	562
Standard deviation	136.163	135.868
Pearson's corr.	0.191	
Kendall's tau	0.184	
Spearman's rho	0.266	
rho/tau	1.449	

Table 3 Descriptive statistics of the data vector.

We divide all the data points by 100 mainly for computational purposes, it is not going to affect in the statistical inference. To get initial estimates of λ_1 and λ_2, we fit exponential distribution to the marginals, and obtain the initial estimates of λ_1 and λ_2 as 0.9311 and 0.8611. For each α, we use these initial estimates to start the EM algorithm. The profile log-likelihood function of α becomes a unimodal function, and it is reported in Figure 3. Finally maximizing the profile log-likelihood function of α, we obtain the MLEs of α, λ_1 and λ_2, and they are $\widehat{\alpha} = 0.7099$, $\widehat{\lambda}_1 = 0.7559$ and $\widehat{\lambda}_2 = 0.6551$, and the corresponding log-likelihood value is -49.0371. The associated 95% confidence intervals are 0.7099 ∓ 0.1116, 0.7559 ∓ 0.1457 and 0.6551 ∓ 0.1123, respectively. To see whether the GE distribution fits the marginal data or not, we compute the Kolmogorov-Smirnov (KS) distances of GE(0.7099,0.7559) and GE(0.7099,0.6551) to the empirical CDF of FRT and SRT data, respectively. It is observed that the KS distance between GE(0.7099,0.7559) and empirical CDF of FRT is 0.1731 and the associated p value is 0.4957. Similarly, the KS distance between GE(0.7099,0.6551) and empirical CDF of SRT is 0.1689 and the associated p value is 0.5275. Therefore, the GE distribution can be used to fit the marginals reasonably well, for the above data set.

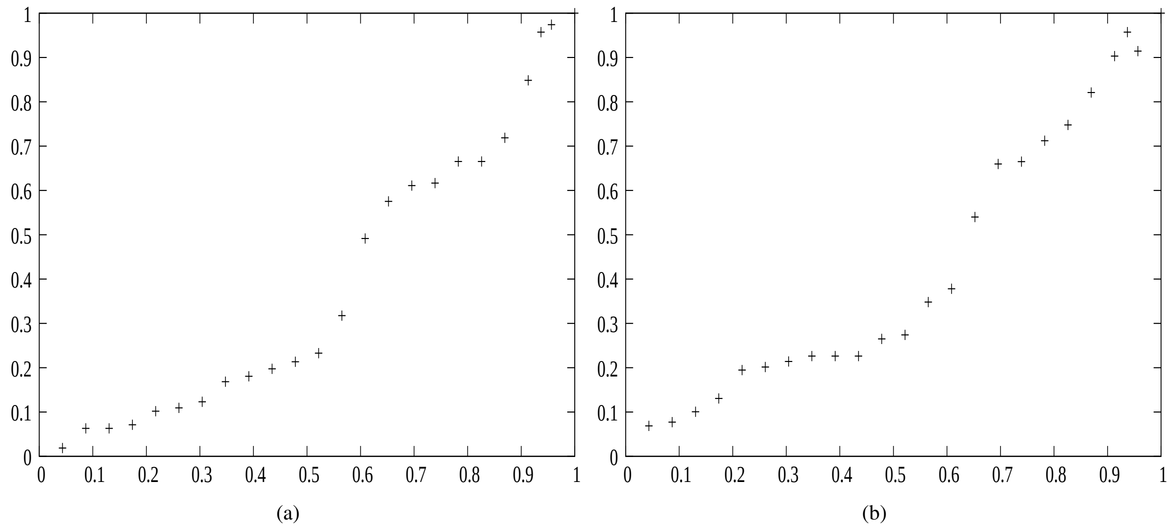

Fig. 2 Scaled TTT plots of (a) FRT and (b) SRT.

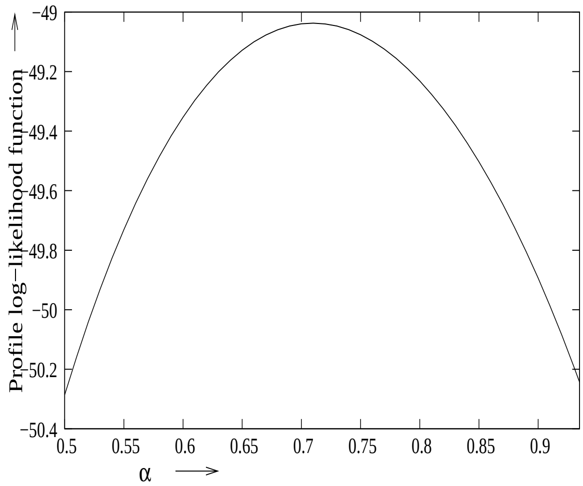

Fig. 3 Profile log-likelihood function of α.

We perform the following two testing of hypotheses problems; namely (26) and (29). In the first case (26), the value of the test statistic is 0.2102. Since the $p = 0.6466$, we do not reject the null hypothesis. In the second case the value of test statistic is 4.085. In this case $p = 0.0074$, hence we reject the null hypothesis. Note that from Table 1 we see that the population versions of Kendall's tau and Spearman's rho for BGE distribution with estimated parameter $\alpha = 0.7$ are given by 0.285 and 0.194, respectively. These values are close to the Spearman's rho and Kendall's tau between FRT and SRT, given in Table 3. This observation also shows that the proposed BGE may be used for analyzing this bivariate data set. The natural question that arises here is whether the BGE model fits these bivariate data or not. In the next section we provide a copula goodness-of-fit test.

For comparison purposes we have fitted (i) the three parameters Block and Basu [5] bivariate exponential model with the pdf

$$f(x,y)=\begin{cases}\frac{\lambda_1\lambda(\lambda_2+\lambda_3)}{\lambda_1+\lambda_2}e^{-\lambda_1 x-(\lambda_2+\lambda_3)y}, & x<y\\ \frac{\lambda_2\lambda(\lambda_1+\lambda_3)}{\lambda_1+\lambda_2}e^{-(\lambda_1+\lambda_3)x-\lambda_2 y}, & x>y\end{cases} \tag{32}$$

where $\lambda_1,\lambda_2,\lambda_3>0$ and $\lambda=\lambda_1+\lambda_2+\lambda_3$ and (ii) the five-parameter absolutely continuous bivariate generalized exponential distribution of the form

$$F(x,y)=[(1-e^{-\lambda_1 x})^{-\alpha_1}+(1-e^{-\lambda_2 x})^{-\alpha_2}-1]^{-\alpha},$$

which is constructed based on the Clayton's copula [25] proposed in Kundu and Gupta [21], to this data set.

Model	Estimated parameters					Log-likelihood
BEG	$\hat{\alpha}=0.7099$	$\hat{\lambda}_1=0.7559$	$\hat{\lambda}_2=0.6551$			-49.0371
Block and Basu model	$\hat{\lambda}_1=0.9297$	$\hat{\lambda}_2=0.8586$	$\hat{\lambda}_3=0.0017$			-51.0796
Kundu and Gupta model	$\hat{\alpha}=1.5621$	$\hat{\alpha}_1=0.4160$	$\hat{\lambda}_1=0.6811$	$\hat{\alpha}_2=0.5392$	$\hat{\lambda}_2=0.7613$	-48.7047

Table 4 The MLEs and the values of Loglikelihood

The MLEs and corresponding log-likelihood values are given in Table 4. Therefore, based on the log-likelihood values, we can say that the proposed BGE model provides a better fit than Block and Basu [5] bivariate exponential model and is comparable with the Kundu and Gupta [21] model for this data set.

5.1 A COPULA GOODNESS-OF-FIT TEST

Once a model has been stated and estimated the natural question is to check whether the initial model assumptions are realistic. In other words we are faced with the so-called goodness-of-fit problem. As an advantage of the Sklar's Theorem [25] the marginal distributions and the copula can be chosen independently of one another. The univariate GE distribution provide adequate descriptions of the FRT and SRT data, individually. Since the copula of the BGE model has a closed and simple form, one can also try to perform a copula goodness-of-fit test. A review and comparison of goodness-of-fit procedures is given in [8]. Let $(x_1,y_1),\ldots,(x_n,y_n)$ be observations from a random vector (X,Y). When dealing with bivariate data, the most natural way of checking the adequacy of a copula model would be to compare the fitted copula and the empirical copula (see, e.g, [6]) of data defined by

$$C_n(\frac{i}{n},\frac{j}{n})=\frac{\text{number of pairs }(x,y)\text{ in the sample with }x\le x_{(i)},y\le y_{(j)}}{n},$$

where $x_{(i)}$ and $y_{(j)}$, $1\le i,j\le n$, denote order statistics from the sample. Figure 4 shows graph of the pairs (x_i,y_i) and (u_i,v_i), $i=1,\ldots,23$, for FRT and SRT data.

From simulations provided in [8] a good combination of power and conceptual simplicity is provided by the Cramer-von Mises statistic:

$$S_n=\sum_{i=1}^{n}\left(C_{\widehat{\alpha}}(u_i,v_i)-C_n(u_i,v_i)\right)^2,$$

where $C_{\widehat{\alpha}}$ is the fitted copula and

$$u_i = \frac{\text{rank of } x_i \text{ among } x_1,...,x_n}{n+1} \quad \text{and} \quad v_i = \frac{\text{rank of} y_i \text{ among } y_1,...,y_n}{n+1}.$$

This statistic measures how close the fitted copula is from the empirical copula of data. The P-value of the test is computed using a parametric bootstrap procedure described in Appendix A of [8]. To this end, we applied this procedure to both the FRT and SRT data. The bootstrap values $S_1^*,\ldots,S_{1000}^*$ of the Cramer-von Mises test statistic are generated and we found the proportion of these values that are larger than S_n = 0.0394 as P-value $\approx$ 0.3. Thus we may conclude that the copula C_α defined by (13) with the association parameter α = 0.7099 performs a good fit for this data set.

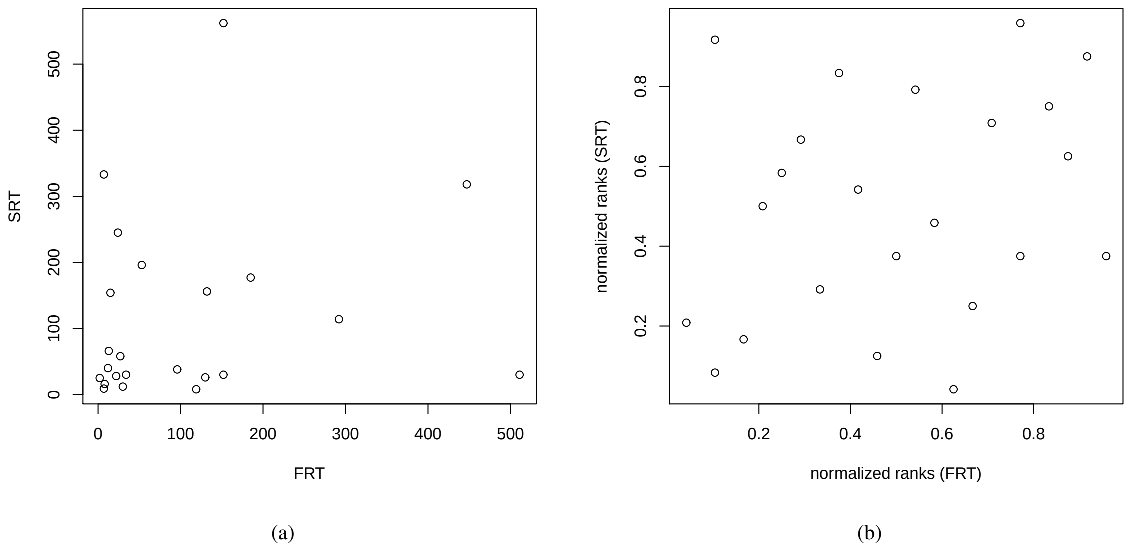

Fig. 4 Pairs of observations (a) and pairs of normalised ranks (b) for FRT and SRT data.

6 CONCLUSIONS

In this paper we studied a bivariate absolutely continuous generalized exponential distribution, whose marginals are generalized exponential distributions. It has three parameters and the marginals have decreasing hazard functions. Therefore, the proposed model can be used as an alternative to the quite popular Block and Basu bivariate exponential model. We derive different properties of the proposed model, and also provide the EM algorithm for computation of the MLEs of the unknown parameters. We have analyzed one real data set, and it is observed that the proposed model provides a good fit to the data set.

Now we briefly discuss different generalizations of the proposed model. (1) Although we have developed the methodology for the bivariate case, along the same line multivariate generalized exponential distribution also can be defined. Several properties can be obtained even for the multivariate case also. (2) Instead of taking exponential distribution for X and Y, it is possible to develop the methodology when they follow Weibull distribution. In this case we can obtain bivariate/ multivariate exponentiated Weibull distribution as an alternative to the already-existent bivariate Weibull models; see, e.g [14]. It will be interesting to develop different properties of this more flexible distributions. More work is needed along these directions.

ACKNOWLEDGEMENTS:

The authors would like to thank the referees for their helpful comments which improved significantly the earlier draft of the paper.

APPENDIX. FISHER INFORMATION MATRIX

From (20) we have

$$\begin{aligned}
\frac{\partial^2 l(\theta)}{\partial\alpha^2} &= -\frac{n}{\alpha^2} - \sum_{i=1}^{n}\left[\frac{e^{-(\lambda_1 x_i+\lambda_2 y_i)}}{1-\alpha e^{-(\lambda_1 x_i+\lambda_2 y_i)}}\right]^2,\\
\frac{\partial^2 l(\theta)}{\partial\lambda_1^2} &= -\frac{n}{\lambda_1^2} - (\alpha-2)\sum_{i=1}^{n}\frac{x_i^2 e^{-(\lambda_1 x_i+\lambda_2 y_i)}}{(1-e^{-(\lambda_1 x_i+\lambda_2 y_i)})^2} - \alpha\sum_{i=1}^{n}\frac{x_i^2 e^{-(\lambda_1 x_i+\lambda_2 y_i)}}{(1-\alpha e^{-(\lambda_1 x_i+\lambda_2 y_i)})^2},\\
\frac{\partial^2 l(\theta)}{\partial\lambda_2^2} &= -\frac{n}{\lambda_2^2} - (\alpha-2)\sum_{i=1}^{n}\frac{y_i^2 e^{-(\lambda_1 x_i+\lambda_2 y_i)}}{(1-e^{-(\lambda_1 x_i+\lambda_2 y_i)})^2} - \alpha\sum_{i=1}^{n}\frac{y_i^2 e^{-(\lambda_1 x_i+\lambda_2 y_i)}}{(1-\alpha e^{-(\lambda_1 x_i+\lambda_2 y_i)})^2},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l(\theta)}{\partial\alpha\partial\lambda_1} &= \sum_{i=1}^{n}\frac{x_i e^{-(\lambda_1 x_i+\lambda_2 y_i)}}{(1-e^{-(\lambda_1 x_i+\lambda_2 y_i)})^2} + \sum_{i=1}^{n}\frac{x_i e^{-(\lambda_1 x_i+\lambda_2 y_i)}}{(1-\alpha e^{-(\lambda_1 x_i+\lambda_2 y_i)})^2},\\
\frac{\partial^2 l(\theta)}{\partial\alpha\partial\lambda_2} &= \sum_{i=1}^{n}\frac{y_i e^{-(\lambda_1 x_i+\lambda_2 y_i)}}{(1-e^{-(\lambda_1 x_i+\lambda_2 y_i)})^2} + \sum_{i=1}^{n}\frac{y_i e^{-(\lambda_1 x_i+\lambda_2 y_i)}}{(1-\alpha e^{-(\lambda_1 x_i+\lambda_2 y_i)})^2},\\
\frac{\partial^2 l(\theta)}{\partial\lambda_1\partial\lambda_2} &= -(\alpha-2)\sum_{i=1}^{n}\frac{x_i y_i e^{-(\lambda_1 x_i+\lambda_2 y_i)}}{(1-e^{-(\lambda_1 x_i+\lambda_2 y_i)})^2} - \alpha\sum_{i=1}^{n}\frac{x_i y_i e^{-(\lambda_1 x_i+\lambda_2 y_i)}}{(1-\alpha e^{-(\lambda_1 x_i+\lambda_2 y_i)})^2}.
\end{aligned}$$

The Fisher information is $I(\theta) = [I(\theta_{ij})]$, where $I_{ij}(\theta) = -E\frac{\partial^2 l(\theta)}{\partial\theta_i\theta_j}$, and $\theta = (\theta_1,\theta_2,\theta_3) = (\alpha,\lambda_1,\lambda_2)$. We shall now present the exact expressions of $I_{ij}(\theta)$, for $i = 1,2,3$. Direct calculations show that

$$\begin{aligned}
I_{11} &= -E\left(\frac{\partial^2 l(\theta)}{\partial\alpha^2}\right) = \frac{n}{\alpha^2}\left(1+\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^k\alpha^{j+3}\binom{\alpha-2}{k}}{(j+k+3)^2}\right),\\
I_{22} &= -E\left(\frac{\partial^2 l(\theta)}{\partial\lambda_1^2}\right) = \frac{n}{\lambda_1^2}\left(1+2\alpha(\alpha-2)A_1+\alpha A_2\right),\\
I_{33} &= -E\left(\frac{\partial^2 l(\theta)}{\partial\lambda_2^2}\right) = \frac{n}{\lambda_2^2}\left(1+2\alpha(\alpha-2)A_1+\alpha A_2\right)
\end{aligned}$$

$$\begin{aligned}
I_{12} &= -E\left(\frac{\partial^2 l(\theta)}{\partial\alpha\partial\lambda_1}\right) = \frac{n\alpha}{\lambda_1}(B_1+B_2),\\
I_{13} &= -E\left(\frac{\partial^2 l(\theta)}{\partial\alpha\partial\lambda_2}\right) = \frac{n\alpha}{\lambda_2}(B_1+B_2),\\
I_{23} &= -E\left(\frac{\partial^2 l(\theta)}{\partial\lambda_1\partial\lambda_2}\right) = \frac{n\alpha}{\lambda_1\lambda_2}((\alpha-2)C_1+\alpha C_2),
\end{aligned}$$

where

$$A_1 = \sum_{j=0}^{\infty}(-1)^j\binom{\alpha-4}{j}\left(\frac{1}{(j+2)^4}-\frac{\alpha}{(j+3)^4}\right),$$
$$A_2 = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(-1)^j\binom{\alpha-2}{j}\alpha^k\frac{1}{(k+j+2)^2},$$

$$B_1 = \sum_{j=0}^{\infty}(-1)^{j+1}\binom{\alpha-3}{j}\left(\frac{1}{(j+2)^3}-\frac{\alpha}{(j+3)^3}\right),$$
$$B_2 = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^{j+1}\binom{\alpha-2}{j}\alpha^k}{(j+k+2)^3},$$
$$C_1 = \sum_{j=0}^{\infty}(-1)^j\binom{\alpha-4}{j}\left(\frac{1}{(j+2)^4)}-\frac{\alpha}{(j+3)^4}\right),$$
$$C_2 = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^{j}\binom{\alpha-2}{j}\alpha^k}{(j+k+3)^4}.$$

References

1. Aarset, M.V. (1987). How to identify a bathtub hazard rate?, *IEEE Transactions on Reliability*, **36**, 106–108.
2. Achcar, J. A., Coelho-Barros, E. A. and Mazucheli, J. (2013). Block and Basu bivariate lifetime distribution in the presence of cure fraction, *Journal of Applied Statistics*, **40**, 1864–1874.
3. Asgharzadeh, A. (2008). Approximate MLE for the scaled generalized exponential distribution under progressive type-II censoring, *Journal of the Korean Statistical Society*, **38**, 223–229.
4. Balakrishnan, N. and Lai, C. D. (2009). *Continuous Bivariate Distributions*, second edition. Springer, New York.
5. Block, H. and Basu, A. P. (1974). A continuous bivariate exponential extension, *Journal of the American Statistical Association*, **69**, 1031–1037.
6. Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés Un test non paramétrique díndépendance, *Académie Royale de Belgique Bulletin de la Classe des Sciences 5e Série*, **65**, 274–292.
7. Dolati, A., Amini, M. and Mirhosseini, S. M. (2014). Dependence properties of bivariate distributions with proportional (reversed) hazards marginals, *Metrika*, **77**, 333–347.
8. Genest, C., Remillard, B. and Beaudoin, D. (2008). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, **42**, 199–213.
9. Gradshteyn, I. S. and Ryzhik, I. M. (2000). *Table of Integrals, Series and Products*, sixth ed., Academic Press, San Diego.
10. Gupta, R. D. and Kundu, D. (1999). Generalized exponential distributions, *Austral. New Zealand J. Statist*, **41**, 173–188.
11. Gupta, R. D. and Kundu, D. (2007). Generalized exponential distribution; existing methods and some recent developments, *J. Statistical Planning and Inference*, **137**, 3537–3547.
12. Holland, P. W. and Wang, Y. J. (1987). Dependence functions for continous bivariates densities, *Commun. Statist. Theory and Methods*, **16:3**, 863–876.
13. Joe, H. (1997). *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.
14. Jose, K. K., M., Ristíc, M. M. and Joseph, A. (2011). Marshall–Olkin bivariate Weibull distributions and processes, *Statistical Papers*, **52**, 789–798.
15. Kannan, N., Kundu, D, Nair, P. and Tripathi, R. C. (2010). The generalized exponential cure rate model with covariates, *Journal of Applied Statistics*, **37:10**, 1625–1636.
16. Kim, N. (2014). Approximate MLE for the scale parameter of the generalized exponential distribution under random censoring, *Journal of the Korean Statistical Society*, **43:1**, 119–131.

17. Kotz, S., Balakrishnan, N. and Johnson, N.L. (2000). *Continuous multivariate distributions*, Vol. 1, second ed., Wiley: New York.
18. Kundu, D. and Gupta, R. D. (2008). Generalized exponential distribution: Bayesian estimations, *Computational Statistics and Data Analysis*, **52**, 1873–1883.
19. Kundu, D. and Gupta, R. D. (2009). Bivariate generalized exponential distribution, *J. Multivariate Analysis*, **100**, 581–593.
20. Kundu, D. and Gupta, R. D. (2010). Modified Sarhan–Balakrishnan singular bivariate distribution, *J. Statistical Planning and Inference*, **140**, 525–538.
21. Kundu, D. and Gupta, R. D. (2011). Absolute continuous bivariate generalized exponential distribution, *AStA Adv. Stat. Anal.*, **95**, 169–185.
22. Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution, *Journal of the American Statistical Association*, **62**, 30–44.
23. McGilchrist, C. A. and Aisbett, C. W. (1991). Regression with frailty in survival analysis, *Biometrics*, **47**, 461–466.
24. Mohsin, M., Kazianka, H., Pilz, J. and Gebhardt, A. (2014). A new bivariate exponential distribution for modeling moderately negative dependence, *Stat. Methods Appl*, **23**, 123–148.
25. Nelsen, R. B. (2006). *An Introduction to Copulas*. Second Edition, Springer, New York.
26. Pillai, R. N. and Jayakumar, K. (1995). Discrete Mittag–Leffler distributions, *Statistics and Probability Letters*, **23**, 271–274.
27. Rényi, A. (1961). On measures of entropy and information, in: Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, vol. I, University of California Press, Berkeley, pp. 547–561.
28. Sarhan, A. and Balakrishnan, N. (2007). A new class of bivariate distributions and its mixture, *Journal of Multivariate Analysis*, **98**, 1508–1527.
29. Self, S.G. and Linag, K. L. (1987). Asymptotic properties of the maximum likelihood estimators, and likelihood ratio test under non-standard conditions, *Journal of the American Statistical Association*, **82**, 605–610.
30. Scarsini, S. (1984). On measures of concordance. Stochastica, **8**, 201–218.
31. Shaked, M. and Shanthikumar, G. (2007). *Stochastic Orders*, Springer, New York.
32. Shoaee, S. and Khorram, E. (2012). A new absolute continuous bivariate generalized exponential distribution, *J. Statistical Planning and Inference*, **142**, 2203–2220.
33. Sklar, A. (1959). Fonctions de répartition à *n* dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, 229–231.