On Absolutely Continuous Bivariate Generalized Exponential Power Series Distribution

Rasool Roozeger*& Debasis Kundu †

Abstract

Recently Mahmoudi and Jafari (2012) introduced generalized exponential power series distributions by compounding generalized exponential with the power series distributions. This is a very flexible distribution with some interesting physical interpretation. Kundu and Gupta (2011) introduced an absolute continuous bivariate generalized exponential distribution, whose marginals are generalized exponential distributions. The main aim of this paper is to introduce bivariate generalized exponential power series distributions. Two special cases namely bivariate generalized exponential geometric and bivariate generalized exponential Poisson distributions are discussed in details. It is observed that both the special cases are very flexible and their joint probability density functions can take variety of shapes. They have interesting copula structures and these can be used to study their different dependence structures and to compute different dependence measures. It is observed that both the models have six unknown parameters each, and the maximum likelihood estimators cannot be obtained in closed form. We have proposed to use EM algorithm to compute the maximum likelihood estimators of the unknown parameters. Some simulation experiments have been performed to see the effectiveness of the proposed EM algorithm. The analyses of two data sets have been performed for illustrative purposes, and it is observed that the proposed models and the EM algorithm work quite satisfactorily. Finally we provide the multivariate generalization of the proposed model.

Key Words and Phrases: Generalized exponential distribution; copula; Fisher information matrix; maximum likelihood estimators; EM algorithm.

AMS Subject Classifications: 62F10, 62F03, 62H12.

*Department of Statistics, Yazd University, Yazd 89195 - 741, IRAN.

†Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, INDIA. E-mail: kundu@iitk.ac.in, Phone no. 91-512-2597141, Fax no. 91-512-2597500.

1 Introduction

Marshall and Olkin (1997) provided a general method to introduce a parameter into a family of distributions by compounding it with the geometric distribution. They discussed in details two special cases namely exponential and Weibull distributions. The proposed method introduced an extra parameter to the family of distributions, hence brings more flexibility. Moreover, it has some interesting physical interpretation also. Since then, extensive work has been done on this method exploring different properties, and extending to some other distributions, see for example Ghitany et al. (2005, 2007), Pham and Lai (2007), Barreto-Souza et al. (2011, 2013), Louzada et al. (2014), Ristic and Kundu (2015, 2016) and the references cited therein.

In the same paper Marshall and Olkin (1997) briefly introduced a bivariate extension of the proposed model. They did not provide any properties and did not discuss any inferential issues also. Recently, Kundu and Gupta (2014) and Kundu (2015) considered bivariate Weibull geometric and bivariate generalized exponential geometric distributions, respectively, and discuss their properties and various inferential issues. In both the cases, the joint distribution has a singular component, and due to this reason, it is difficult to extend it to the multivariate case. Although, Marshall and Olkin (1997) had indicated about compounding the geometric distribution with an absolute continuous bivariate distribution, no details have been provided. This is an attempt towards that direction.

The main idea of Marshall and Olkin (1997) is to compound a family of continuous distributions with a geometric distribution. In recent times extensive work has been done to compound a family of continuous distribution with different discrete distributions. The main advantage of this procedure is that it produces a very flexible class of continuous distribution functions, and moreover often it can have some interesting physical interpretations also.

Attempt has been made mainly to compound different discrete distributions like Poisson, logarithmic or power series distributions with other continuous distributions. Mahmoudi and Jafari (2012) introduced a class of generalized exponential power series (GEP) distributions by compounding generalized exponential distribution with a power series distributions. The three parameter GEP distribution is a very flexible distribution. The probability density function (PDF) of the GEP is an unimodal function and it can take variety of shapes depending on the two shape parameters. The generalized exponential (GE) of Gupta and Kundu (1999) can be obtained as a special case of the GEP distribution. It can be used quite effectively to analyze skewed as well as heavy tail data. The hazard function of the GEP distribution can take all four different shapes; increasing, decreasing, bath-tub and unimodal. It may be mentioned that not too many three-parameter distributions have this flexibility. For some other development along this direction, see for example Adamidis and Loukas (1998), Kus (2007), Tahmasbi and Razaei (2008), Chahkandi and Ganjali (2009), Barreto-Souza et al. (2011) and Harandi and Alamatsaz (2016).

Recently Kundu and Gupta (2011) introduced an absolute continuous bivariate generalized exponential (BGE) distribution, whose marginals are generalized exponential (GE) distributions. The proposed distribution is very flexible and the joint probability density function can take different shapes. Further, it is observed that the proposed absolutely continuous BGE distribution can be obtained using Clayton copula, hence many dependence properties and dependence measures can be easily established from the copula properties. It can be obtained by some other methods also, and it has some interesting physical interpretations. Therefore, it provides a practitioner another option of a class of distributions which can be used to analyze a bivariate data set.

The main aim of this paper is to combine the BGE with the power series distributions. We name it as the absolutely continuous bivariate generalized exponential power series (BGEP) distributions. The proposed BGEP distribution is a very flexible distribution. The joint probability density function of the proposed BGEP can have variety of shapes. Therefore, it can be used quite effectively to analyze bivariate skewed data and bivariate heavy tail data. The BGE distribution of Kundu and Gupta (2011) can be obtained as a special case of the BGEP distribution. The marginals of the BGEP distribution are GEP distributions. We derive several properties of the proposed distribution. Several well known univariate distributions can be obtained as its marginals. Generating samples from a BGEP distribution is quite straight forward, hence simulation experiments can be performed quite conveniently. We consider two special cases namely: (a) Bivariate generalized exponential geometric (BGEG) and (b) Bivariate generalized exponential Poisson (BGEP_o) distributions in details and explore their specific properties.

It is observed that BGEG and BGEP_o both have six unknown parameters each. The maximum likelihood estimators (MLEs) of the unknown parameters cannot be obtained in explicit form. They can be obtained to solve six normal equations simultaneously. To avoid that we have proposed an EM algorithm to compute the MLEs in both these cases. It is observed that at each 'M'-step of the EM algorithm, the maximization can be performed by solving three one-dimensional non-linear equations. Hence, the implementation of the EM algorithm is quite simple in practice. We have performed some simulation experiments and it is observed that the proposed EM algorithm is working quite satisfactorily. We have suggested a method to choose some good initial estimates also. Two data sets have been analyzed for illustrative purposes, and it is observed that the proposed models and the EM algorithm perform quite well.

Very recently Kundu et al. (2015) introduced an absolute continuous multivariate generalized exponential (MGE) distribution. It is observed that our proposed BGEP model can be extended along the same line to the multivariate case also. We call this new multivariate distribution as multivariate generalized exponential power series (MGEP) distribution. It is observed many of our results can be extended for the MGEP distribution. Finally we propose some open problems and conclude the paper.

Rest of the paper is organized as follows. In Section 2 we provide the BGEP distribution and discuss its different properties. Two special cases are discussed in Section 3. The estimation of the unknown parameters are provided in Section 4. The results of the simulation experiments and the analyses of two data sets are presented in Section 5 and Section 6, respectively. The MGEP distribution is introduced in Section 7, and finally conclusions appear in Section 8.

2 Model Description and Properties

Let us recall the following three definitions.

Definition 1: A random variable X is said to have a GE distribution with the shape parameter $\beta > 0$ and scale parameter $\theta > 0$, if the cumulative distribution function (CDF) of X for $x > 0$ is

$$F_X(x; \beta, \theta) = (1 - e^{-x\theta})^{\beta}; \tag{1}$$

and zero otherwise, see Gupta and Kundu (1999). A random variable X with the CDF (1) will be denoted by GE(β, θ).

Definition 2: The bivariate random vector (X_1, X_2) is said to have a BGE distribution with parameters $\alpha_1 > 0$, $\lambda_1 > 0$, $\alpha_2 > 0$, $\lambda_2 > 0$, $\alpha > 0$, if the joint cumulative distribution function (CDF) of (X_1, X_2), can be written as

$$F_{X_1,X_2}(x_1, x_2) = P(X_1 \le x_1, X_2 \le x_2) = \left[(1 - e^{-\lambda_1 x_1})^{-\alpha_1} + (1 - e^{-\lambda_2 x_2})^{-\alpha_2} - 1\right]^{-\alpha}, \tag{2}$$

for $x_1 > 0, x_2 > 0$, and zero otherwise, see Kundu and Gupta (2011). Form now on a bivariate

Table 1: Some useful quantities for different power series distributions

Distribution	a_n	$A(\theta)$	$A'(\theta)$	$A''(\theta)$	$A'''(\theta)$	s
Geometric	1	$\theta(1-\theta)^{-1}$	$(1-\theta)^{-2}$	$2(1-\theta)^{-3}$	$6(1-\theta)^{-4}$	1
Poisson	$(n!)^{-1}$	$e^{\theta}-1$	e^{θ}	e^{θ}	e^{θ}	∞
Logarithmic	n^{-1}	$-\ln(1-\theta)$	$(1-\theta)^{-1}$	$(1-\theta)^{-2}$	$2(1-\theta)^{-3}$	1
Binomial	$\binom{k}{n}$	$(1-\theta)^{k}-1$	$\frac{k}{(1+\theta)^{1-k}}$	$\frac{k(k-1)}{(1+\theta)^{2-k}}$	$\frac{k(k-1)(k-2)}{(1+\theta)^{3-k}}$	∞
Negative Binomial	$\binom{n-1}{k-1}$	$\frac{\theta^k}{(1-\theta)^k}$	$\frac{k\theta^{k-1}}{(1-\theta)^{k+1}}$	$\frac{k(k+2\theta-1)}{\theta^{2-k}(1-\theta)^{k+2}}$	$\frac{k(k^2+6k\theta+6\theta^2-3k-6\theta+2)}{\theta^{3-k}(1-\theta)^{k+3}}$	1

random variable (X_1, X_2) with the joint CDF (2) will be denoted by BGE$(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha)$.

Definition 3: A discrete random variable N is said to have a power series distributions if the probability mass function (PMF) of N can be written as

$$P(N = n) = \frac{a_n \theta^n}{A(\theta)}; \quad n = 1, 2, 3, \ldots. \tag{3}$$

Here $a_n \geq 0$, depends only on n, $A(\theta) = \sum_{n=1}^{\infty} a_n \theta^n$, and $\theta \in (0, s)$, where s can be ∞ also, is such that $A(\theta) < \infty$, see Noack (1950) or Johnson et al. (2005).

The following Table 1 summarizes some particular cases of the different truncated (at zero) power series (PS) distributions, namely for geometric, Poisson, logarithmic, binomial and negative binomial distributions. Here $\frac{dA(\theta)}{d\theta} = A'(\theta)$, $\frac{d^2A(\theta)}{d\theta^2} = A''(\theta)$ and $\frac{d^3A(\theta)}{d\theta^3} = A'''(\theta)$. Now we are in a position to define the BGEP distribution.

Definition 4: Suppose $\{(X_{1n}, X_{2n}); n = 1, 2, \ldots\}$ is a sequence of independent and identically distributed (i.i.d.) BGE$(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha)$ random variables, N is a discrete random variable with the PMF (3) and it is independent of $\{(X_{1n}, X_{2n}); n = 1, 2, \ldots\}$. Let us define

$$Y_1 = \max\{X_{11}, \ldots, X_{1N}\}, \quad \text{and} \quad Y_2 = \max\{X_{21}, \ldots, X_{2N}\}. \tag{4}$$

The bivariate random vector (Y_1, Y_2) is said to have a BGEP distributions with the parameters, α_1, α_2, λ_1, λ_2, α and θ, and it will be denoted by BGEP$(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha, \theta)$.

Before going to derive the joint cumulative distribution function (CDF) and the joint probability density function (PDF) we will briefly mention how it can happen in practice.

Random Stress Model: Suppose a system has two components. The components are subjected to random number of indvidual stresses, say X_i and Y_i, respectively, for $i = 1, 2, \ldots$. Here X_i and Y_i may be dependent, but (X_i, Y_i), for $i = 1, 2, \ldots$ are independent and identically distributed. If N denotes the number of stresses and it may be a random quantity, then the observed stresses at the two components are $X = \max\{X_1, \ldots, X_N\}$ and $Y = \max\{Y_1, \ldots, Y_N\}$, respectively.

Parallel System: Consider two systems, say 1 and 2, each having N number of independent and identical components attached in parallel. Here N is a random variable. If X_i and Y_i denote the lifetime of the i-th component of the system 1 and system 2, respectively, then the lifetime of the two systems become (X, Y), where $X = \max\{X_1, \ldots, X_N\}$ and $Y = \max\{Y_1, \ldots, Y_N\}$.

Theorem 2.1: If $(Y_1, Y_2) \sim$ BGEP$(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha, \theta)$, then the joint CDF of (Y_1, Y_2) is

$$F_{Y_1,Y_2}(y_1, y_2) = \frac{A\left(\theta\left[(1 - e^{-\lambda_1 y_1})^{-\alpha_1} + (1 - e^{-\lambda_2 y_2})^{-\alpha_2} - 1\right]^{-\alpha}\right)}{A(\theta)}. \tag{5}$$

Proof: It can be easily obtained, hence the details are avoided. ∎

Theorem 2.2: If $(Y_1, Y_2) \sim$ BGEP$(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha, \theta)$, then the joint PDF of (Y_1, Y_2) is

$$f_{Y_1,Y_2}(y_1, y_2) = f_1(y_1, y_2) \times f_2(y_1, y_2) \times f_3(y_1, y_2) + f_4(y_1, y_2) \times f_{X_1,X_2}(y_1, y_2). \tag{6}$$

Here for $F_{X_1,X_2}(y_1, y_2)$, the joint CDF of (X_1, X_2) as defined in (2),

$$f_1(y_1, y_2) = \frac{\theta^2 A''(\theta F_{X_1,X_2}(y_1, y_2))}{A(\theta)}$$

$$
\begin{aligned}
f_2(y_1,y_2) &= \frac{\partial F_{X_1,X_2}(y_1,y_2)}{\partial y_1} = \frac{\alpha\alpha_1\lambda_1\left(1-e^{-\lambda_1 y_1}\right)^{-\alpha_1-1}e^{-\lambda_1 y_1}}{\left((1-e^{-\lambda_1 y_1})^{-\alpha_1}+(1-e^{-\lambda_2 y_2})^{-\alpha_2}-1\right)^{\alpha+1}} \\
f_3(y_1,y_2) &= \frac{\partial F_{X_1,X_2}(y_1,y_2)}{\partial y_2} = \frac{\alpha\alpha_2\lambda_2\left(1-e^{-\lambda_2 y_2}\right)^{-\alpha_2-1}e^{-\lambda_2 y_2}}{\left((1-e^{-\lambda_1 y_1})^{-\alpha_1}+(1-e^{-\lambda_2 y_2})^{-\alpha_2}-1\right)^{\alpha+1}}, \\
f_4(y_1,y_2) &= \frac{\theta A'(\theta F_{X_1,X_2}(y_1,y_2))}{A(\theta)}
\end{aligned}
$$

and $f_{X_1,X_2}(y_1,y_2)$ is the joint PDF of (X_1,X_2), and it can be written as

$$
f_{X_1,X_2}(y_1,y_2) = \frac{ce^{-\lambda_1 y_1}e^{-\lambda_2 y_2}(1-e^{-\lambda_1 y_1})^{-\alpha_1-1}(1-e^{-\lambda_2 y_2})^{-\alpha_2-1}}{[(1-e^{-\lambda_1 y_1})^{-\alpha_1}+(1-e^{-\lambda_2 y_2})^{-\alpha_2}-1]^{\alpha+2}}, \tag{7}
$$

when $c=\alpha(\alpha+1)\alpha_1\alpha_2\lambda_1\lambda_2$.

PROOF: The proof can be easily obtained by calculating $\dfrac{\partial^2 F_{Y_1,Y_2}(y_1,y_2)}{\partial y_1 \partial y_2}$, the details are avoided. ∎

THEOREM 2.3: Suppose $(Y_1,Y_2) \sim \text{BGEP}(\alpha_1,\alpha_2,\lambda_1,\lambda_2,\alpha,\theta)$.

(a) The marginal distributions of Y_1 and Y_2 can be obtained as

$$
\begin{aligned}
F_{Y_1}(y_1) &= P(Y_1 \le y_1) = \frac{A\left(\theta\left(1-e^{-\lambda_1 y_1}\right)^{\alpha\alpha_1}\right)}{A(\theta)}, \quad \text{and} \\
F_{Y_2}(y_2) &= P(Y_2 \le y_2) = \frac{A\left(\theta\left(1-e^{-\lambda_2 y_2}\right)^{\alpha\alpha_2}\right)}{A(\theta)},
\end{aligned}
$$

respectively. The associated PDFs of Y_1 and Y_2 are

$$
\begin{aligned}
f_{Y_1}(y_1) &= \frac{\theta A'\left(\theta\left(1-e^{-\lambda_1 y_1}\right)^{\alpha\alpha_1}\right)}{A(\theta)} \times \alpha\alpha_1\lambda_1 e^{-\lambda_1 y_1}\left(1-e^{-\lambda_1 y_1}\right)^{\alpha\alpha_1-1}, \quad \text{and} \\
f_{Y_2}(y_2) &= \frac{\theta A'\left(\theta\left(1-e^{-\lambda_2 y_2}\right)^{\alpha\alpha_2}\right)}{A(\theta)} \times \alpha\alpha_2\lambda_2 e^{-\lambda_2 y_2}\left(1-e^{-\lambda_2 y_2}\right)^{\alpha\alpha_2-1},
\end{aligned}
$$

respectively.

(b) The conditional distributions have the following forms.

$$
\begin{aligned}
P(Y_1 \le y_1 | Y_2 \le y_2) &= \frac{A\left(\theta\left[(1-e^{-\lambda_1 y_1})^{-\alpha_1}+(1-e^{-\lambda_2 y-2})^{-\alpha_2}-1\right]^{-\alpha}\right)}{A\left(\theta\left(1-e^{-\lambda_2 y_2}\right)^{\alpha\alpha_2}\right)} \\
P(Y_2 \le y_2 | Y_1 \le y_1) &= \frac{A\left(\theta\left[(1-e^{-\lambda_1 y_1})^{-\alpha_1}+(1-e^{-\lambda_2 y-2})^{-\alpha_2}-1\right]^{-\alpha}\right)}{A\left(\theta\left(1-e^{-\lambda_1 y_1}\right)^{\alpha\alpha_1}\right)}.
\end{aligned}
$$

PROOF: The proofs can be easily obtained, hence they are avoided. ∎

Now we would like to obtain the joint PDF of Y_1, Y_2, N. It may be useful for developing the EM algorithm.

$$\begin{aligned} P(Y_1 \le y_1, Y_2 \le y_2, N = n) &= P(Y_1 \le y_1, Y_2 \le y_2 | N = n) P(N = n) \\ &= \frac{a_n \theta^n}{A(\theta)} P(\max\{X_{11}, \ldots, X_{1n}\} \le y_1, \max\{X_{21}, \ldots, X_{2n}\} \le y_2) \\ &= \frac{a_n \theta^n}{A(\theta)} \times \left[(1 - e^{-\lambda_1 y_1})^{-\alpha_1} + (1 - e^{-\lambda_2 y_2})^{-\alpha_2} - 1\right]^{-n\alpha}. \end{aligned}$$

Hence, the joint PDF of Y_1, Y_2 and N becomes

$$f_{Y_1,Y_2,N}(y_1, y_2, n) = \frac{a_n \theta^n}{A(\theta)} \times \frac{c_n e^{-\lambda_1 y_1} e^{-\lambda_2 y_2} (1 - e^{-\lambda_1 y_1})^{-\alpha_1 - 1} (1 - e^{-\lambda_2 y_2})^{-\alpha_2 - 1}}{[(1 - e^{-\lambda_1 y_1})^{-\alpha_1} + (1 - e^{-\lambda_2 y_2})^{-\alpha_2} - 1]^{n\alpha + 2}}. \tag{8}$$

Here $c_n = n\alpha(n\alpha + 1)\alpha_1 \alpha_2 \lambda_1 \lambda_2$. Therefore,

$$\begin{aligned} f_{N|Y_1=y_1,Y_2=y_2}(n) &= \frac{f_{Y_1,Y_2,N}(y_1, y_2, n)}{f_{Y_1,Y_2}(y_1, y_2)} \\ &= \frac{a_n \theta^n n(n\alpha + 1) D^n(y_1, y_2)}{[\alpha\theta^2 D^2(y_1, y_2) A''(\theta D(y_1, y_2)) + (\alpha + 1)\theta D(y_1, y_2) A'(\theta D(y_1, y_2))]}. \end{aligned}$$

and for

$$D(y_1, y_2) = \left[(1 - e^{-\lambda_1 y_1})^{-\alpha_1} + (1 - e^{-\lambda_2 y_2})^{-\alpha_2} - 1\right]^{-\alpha}.$$

$$\begin{aligned} &E(N|Y_1 = y_1, Y_2 = y_2) = \\ &\sum_{n=1}^{\infty} \frac{a_n \theta^n n^2 (n\alpha + 1) D^n(y_1, y_2)}{[\alpha\theta^2 D^2(y_1, y_2) A''(\theta D(y_1, y_2)) + (\alpha + 1)\theta D(y_1, y_2) A'(\theta D(y_1, y_2))]} = \\ &\frac{\alpha\theta^2 D^2(y_1, y_2) A^{(3)}(\theta D(y_1, y_2)) + (3\alpha + 1)\theta D(y_1, y_2) A''(\theta D(y_1, y_2)) + (\alpha + 1) A'(\theta D(y_1, y_2))}{\alpha\theta D(y_1, y_2) A''(\theta D(y_1, y_2)) + (\alpha + 1) A'(\theta D(y_1, y_2))}. \end{aligned}$$

Here although $D(y_1, y_2)$ depends on $\alpha_1, \alpha_2, \lambda_1, \lambda_2$, we do not make it explicit for brevity.

Now we provide the copula structure of a BGEP distribution. Observe that $A(\theta)$ as defined in (3) is a strictly increasing function in θ, hence A^{-1} exists. We have the following result.

Theorem 2.3: The BGEP distribution has the following copula structure;

$$C(u_1, u_2) = \frac{1}{A(\theta)} \times A\left[\theta\left\{\left(\frac{1}{\theta}A^{-1}(u_1A(\theta))\right)^{-1/\alpha} + \left(\frac{1}{\theta}A^{-1}(u_2A(\theta))\right)^{-1/\alpha} - 1\right\}^{-\alpha}\right], \quad (9)$$

for $(u_1, u_2) \in [0, 1] \times [0, 1]$.

Proof: Suppose $(Y_1, Y_2) \sim \text{BGEP}(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha, \theta)$, and F_{Y_1}, F_{Y_2}, F_{Y_1,Y_2} are same as defined before. Note that for all $(u_1, u_2) \in [0, 1] \times [0, 1]$,

$$C(u_1, u_2) = F_{Y_1,Y_2}(F_{Y_1}^{-1}(u_1), F_{Y_2}^{-1}(u_2)).$$

The result follows by substituting

$$\begin{aligned} F_{Y_1}^{-1}(u_1) &= -\frac{1}{\lambda_1}\ln\left\{1 - \left[\frac{1}{\theta}A^{-1}(u_1A(\theta))\right]^{1/(\alpha\alpha_1)}\right\} \\ F_{Y_2}^{-1}(u_2) &= -\frac{1}{\lambda_2}\ln\left\{1 - \left[\frac{1}{\theta}A^{-1}(u_2A(\theta))\right]^{1/(\alpha\alpha_2)}\right\} \end{aligned}$$

in (5) and after some simplifications. ∎

The generation from a BGEP distribution is quite straightforward. The following algorithm can be used to generate random samples from a BGEP distribution.

Algorithm

Step 1: First generate N from a power series distributions.

Step 2: Generate $(X_{11}, X_{21}), \ldots, (X_{1N}, X_{2N})$ from a BGE distribution using the method suggested by Kundu et al. (2015).

Step 3: Consider

$$Y_1 = \max\{X_{11}, \ldots, X_{1N}\} \quad \text{and} \quad Y_2 = \max\{X_{21}, \ldots, X_{2N}\}.$$

3 Two Specific Cases

3.1 BGE Geometric Distribution

In this section it is assumed that the power series distributions is a geometric distribution, and we discuss some special properties of the BGE geometric (BGEG) distribution. In this case the BGEP$(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha, \theta)$ will be denoted by BGEG$(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha, \theta)$. The following result provides the joint CDF and joint PDF of a BGEG distribution.

Theorem 3.1: Suppose $(Y_1, Y_2) \sim$ BGEG$(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha, \theta)$.

(a) The joint CDF of (Y_1, Y_2) for $y_1 > 0, y_2 > 0$, is

$$F_{Y_1,Y_2}(y_1, y_2) = \frac{1-\theta}{[(1-e^{-\lambda_1 y_1})^{-\alpha_1} + (1-e^{-\lambda_2 y_2})^{-\alpha_2} - 1]^{\alpha} - \theta} \tag{10}$$

and zero, otherwise.

(b) The joint PDF of (Y_1, Y_2) for $y_1 > 0, y_2 > 0$, is

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) \;&=\; k \times u_1(y_1; \alpha_1, \lambda_1) \times u_2(y_2; \alpha_2, \lambda_2) \times u_3^{\alpha-2}(y_1, y_2; \alpha_1, \lambda_1\alpha_2, \lambda_2) \times \\ &\quad \frac{[(1-\alpha)\,(u_3^{\alpha}(y_1, y_2; \alpha_1, \lambda_1\alpha_2, \lambda_2) - \theta) + 2u_3^{\alpha}(y_1, y_2; \alpha_1, \lambda_1\alpha_2, \lambda_2)]}{(u_3(y_1, y_2; \alpha_1, \lambda_1\alpha_2, \lambda_2) - \theta)^{\,3}}, \end{aligned} \tag{11}$$

and zero otherwise. Here

$$\begin{aligned} k &= \alpha(1-\theta)\alpha_1\alpha_2\lambda_1\lambda_2, \\ u_1(y_1;\alpha_1,\lambda_1) &= e^{-\lambda_1 y_1}(1-e^{-\lambda_1 y_1})^{-\alpha_1-1}, \\ u_2(y_2;\alpha_2,\lambda_2) &= e^{-\lambda_2 y_2}(1-e^{-\lambda_2 y_2})^{-\alpha_2-1} \quad \text{and} \\ u_3(y_1,y_2;\alpha_1,\lambda_1\alpha_2,\lambda_2) &= \left[(1-e^{-\lambda_1 y_1})^{-\alpha_1}+(1-e^{-\lambda_2 y_2})^{-\alpha_2}-1\right]. \end{aligned}$$

PROOF: The proofs can be easily obtained by substituting $A(\theta) = \theta(1-\theta)^{-1}$ in Theorem 2.1 and Theorem 2.2. ∎

In Figure 1 we provide the plots of the joint PDF of (Y_1, Y_2), for different parameter values. It is observed that the joint PDF is an unimodal function, and it can take variety of shapes. It is also observed that it can be very heavy tail also.

The following result provides the CDF, PDF and the hazard functions of the marginal distributions of a BGEG distribution.

THEOREM 3.2: Suppose $(Y_1, Y_2) \sim \text{BGEG}(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha, \theta)$.

(a) The CDFs of Y_1 and Y_2 are

$$F_{Y_1}(y_1) = \frac{1-\theta}{(1-e^{-\lambda_1 y_1})^{-\alpha\alpha_1}-\theta}, \qquad F_{Y_2}(y_2) = \frac{1-\theta}{(1-e^{-\lambda_2 y_2})^{-\alpha\alpha_2}-\theta}.$$

(b) The PDFs of Y_1 and Y_2 are

$$\begin{aligned} f_{Y_1}(y_1) &= \frac{(1-\theta)\alpha\alpha_1\lambda_1 e^{-\lambda_1 y_1}(1-e^{-\lambda_1 y_1})^{-\alpha\alpha_1-1}}{((1-e^{-\lambda_1 y_1})^{-\alpha\alpha_1}-\theta)^2} \\ f_{Y_2}(y_2) &= \frac{(1-\theta)\alpha\alpha_2\lambda_2 e^{-\lambda_2 y_2}(1-e^{-\lambda_2 y_2})^{-\alpha\alpha_2-1}}{((1-e^{-\lambda_2 y_2})^{-\alpha\alpha_2}-\theta)^2}, \end{aligned}$$

respectively.

(c) The hazard functions of Y_1 and Y_2 are

$$h_{Y_1}(y_1) = \frac{(1-\theta)\alpha\alpha_1\lambda_1 e^{-\lambda_1 y_1}(1-e^{-\lambda_1 y_1})^{-\alpha\alpha_1-1}}{((1-e^{-\lambda_1 y_1})^{-\alpha\alpha_1}-1)((1-e^{-\lambda_1 y_1})^{-\alpha\alpha_1}-\theta)},$$

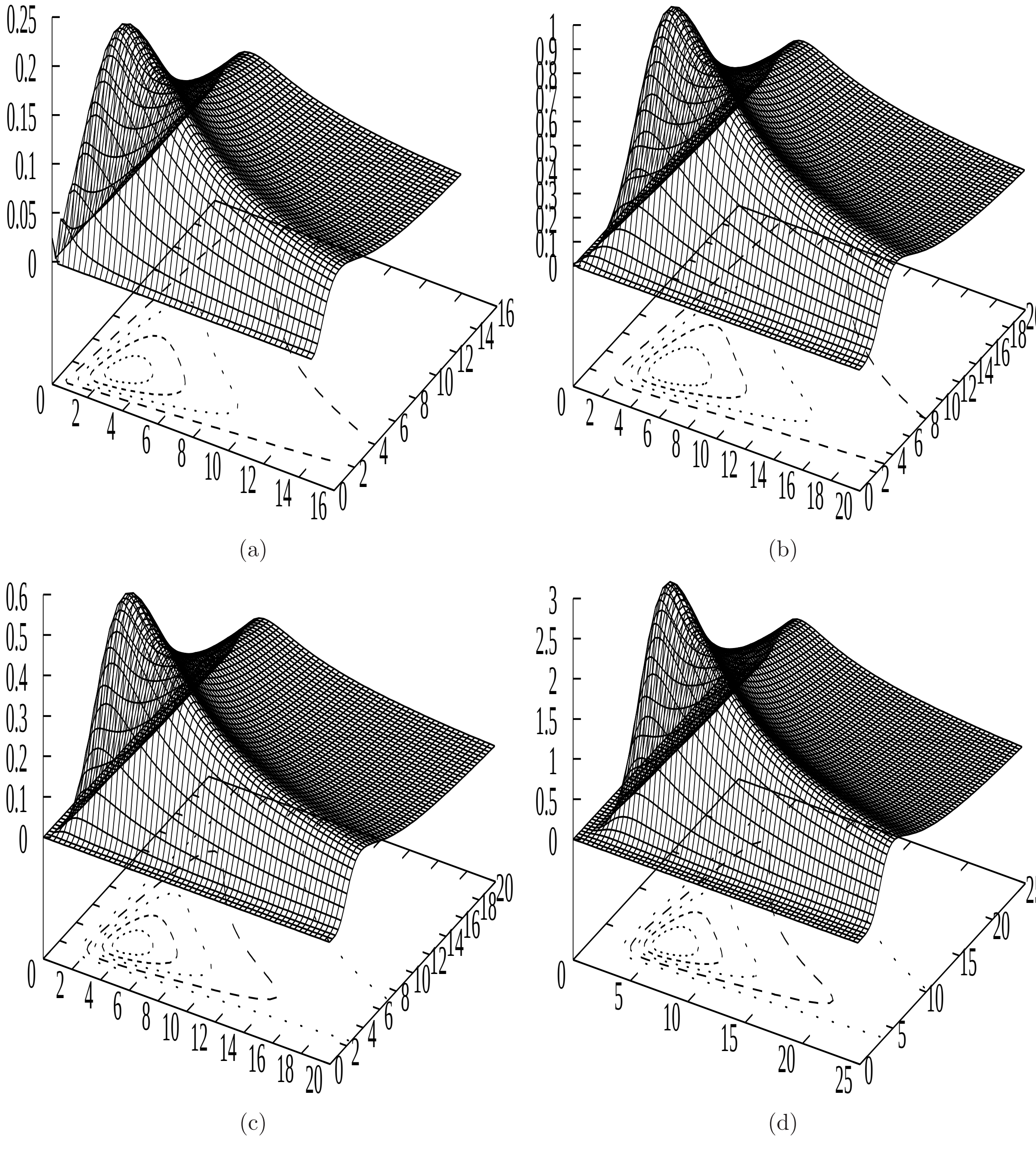


Figure 1: Joint PDF of BGEG distribution for different parameter values (a) $\alpha_1 = \alpha_2 = 1$, $\alpha = 2$, $\theta = 0.25$, (b) $\alpha_1 = \alpha_2 = 1$, $\alpha = 2$, $\theta = 0.75$, (c) $\alpha_1 = \alpha_2 = \alpha = 2$, $\theta = 0.25$, (d) $\alpha_1 = \alpha_2 = \alpha = 2$, $\theta = 0.75$.

$$h_{Y_2}(y_2) = \frac{(1-\theta)\alpha\alpha_2\lambda_2 e^{-\lambda_2 y_2}(1-e^{-\lambda_2 y_2})^{-\alpha\alpha_2-1}}{((1-e^{-\lambda_2 y_2})^{-\alpha\alpha_2}-1)((1-e^{-\lambda_2 y_2})^{-\alpha\alpha_2}-\theta)},$$

respectively.

Proof: The proofs can be easily obtained from Theorem 3.1, the details are avoided. ∎

It can be seen that the hazard function of Y_1 can be (a) increasing if $\alpha\alpha_1 > 1$ for all values of θ, (b) decreasing if $\alpha\alpha_1 \leq 1$ and $\theta = 0$ and (c) bathtub shaped if $\alpha\alpha_1 \leq 1$ and $\theta > 0$. Similarly, the hazard function of Y_2 can be obtained.

Now using the moment generating function of a GE random variable, Gupta and Kundu (1999), the moment generating functions of Y_1 and Y_2 can be obtained as follows:

$$M_{Y_1}(t) = \frac{(1-\theta)\Gamma(1-t/\lambda_1)}{\theta}\sum_{n=1}^{\infty}\frac{\theta^n\Gamma(n\alpha\alpha_1+1)}{\Gamma(n\alpha\alpha_1-t/\lambda_1+1)}$$

and

$$M_{Y_2}(t) = \frac{(1-\theta)\Gamma(1-t/\lambda_2)}{\theta}\sum_{n=1}^{\infty}\frac{\theta^n\Gamma(n\alpha\alpha_2+1)}{\Gamma(n\alpha\alpha_2-t/\lambda_2+1)},$$

for $t < \lambda_1$ and $t < \lambda_2$, respectively. Hence all the moments of the marginals can be easily obtained. The following result provides the copula of a BGEG distribution.

Theorem 3.3: The BGEG distribution has the following copula structure;

$$C(u,v) = \frac{1-\theta}{\left[\left(\frac{1-(1-u)\theta}{u}\right)^{1/\alpha}+\left(\frac{1-(1-v)\theta}{v}\right)^{1/\alpha}-1\right]^{\alpha}-\theta}, \qquad (12)$$

for $(u,v) \in [0,1]\times[0,1]$.

Proof: By using $A(\theta) = \theta(1-\theta)^{-1}$ in Theorem 2.3, the result can be easily obtained. ∎

Now we will present the medial correlation and the bivariate tail dependence using the copula property. The population version of the medial correlation coefficient of a pair (Y_1, Y_2) of continuous random variables was defined by Blomqvist (1950). If M_{Y_1} and M_{Y_2} denote the medians of Y_1 and Y_2, respectively, then M_{Y_1,Y_2}, the medial correlation coefficient of Y_1 and Y_2 is

$$M_{Y_1,Y_2} = P((Y_1 - M_{Y_1})(Y_2 - M_{Y_2}) > 0) - P((Y_1 - M_{Y_1})(Y_2 - M_{Y_2}) < 0).$$

It has been shown by Nelsen (2006) that the median correlation coefficient is a copula property, and $M_{Y_1,Y_2} = 4C\left(\frac{1}{2}, \frac{1}{2}\right)$. Therefore, if (Y_1, Y_2) follows BGEG distribution, the median correlation coefficient between Y_1 and Y_2 is $\frac{4(1-\theta)}{\{2(2-\theta)^{1/\alpha} - 1\}^{\alpha} - \theta}$.

The concept of bivariate tail dependence relates to the amount of dependence in the upper quadrant (or lower quadrant) tail of a bivariate distribution, see Joe (1997, page 33). In terms of the original random variables Y_1 and Y_2, the upper tail dependence is defined as

$$\chi = \lim_{z \to \infty} P(Y_2 \geq F_{Y_2}^{-1}(z) | Y_1 \geq F_{Y_1}^{-1}(z)).$$

Intuitively, the upper tail dependence exists, when there is a positive probability that some positive outliers may occur jointly. If $\chi \in (0, 1]$, then Y_1 and Y_2 are said to be asymptotically dependent, and if $\chi = 0$, they are asymptotically independent. Coles, Hefferman and Tawn (1999) showed using copula function that

$$\chi = \lim_{u \to 1} \frac{1 - 2u + C(u, u)}{1 - u}.$$

In case of BGEG it can be shown that $\chi = 0$, i.e. Y_1 and Y_2 are asymptotically independent.

It may be mentioned that the Kendall's tau and Spearman's rho provide important measures of dependence between two correlated random variables. They provide the measures of association between two random variables. It is known that they are copula functions. In case of BGEG they cannot be computed analytically. They need to be computed numerically. In Tables 2 and 3 we provide Spearman's rho and Kendall's tau values, respectively, for different α and θ. From Tables 2 and 3, it is observed that both Spearman's rho and Kendall's tau are always non-negative. When $\alpha = 1$, they are more or less constant for all

Table 2: Spearman's rho for the BGEG copula function

$\alpha \downarrow \theta \rightarrow$	0.1	0.2	0.4	0.6	0.8	0.9
0.5	0.67	0.66	0.64	0.61	0.56	0.52
1.0	0.48	0.48	0.48	0.48	0.48	0.47
1.5	0.37	0.38	0.40	0.41	0.44	0.45
2.0	0.30	0.32	0.35	0.38	0.42	0.44
2.5	0.26	0.28	0.31	0.35	0.40	0.43
3.0	0.23	0.25	0.29	0.34	0.39	0.43
5.0	0.16	0.18	0.24	0.30	0.38	0.42
10.0	0.10	0.13	0.20	0.27	0.37	0.41
20.0	0.07	0.10	0.17	0.25	0.35	0.41
50.0	0.05	0.08	0.16	0.24	0.35	0.40

θ in both the cases. For any fixed $\alpha < 1$, as θ increases, Spearman's rho and Kendal's tau decrease, and for $\alpha > 1$, it is the other way.

3.2 BGE Poisson distribution

In this section it is assumed that the power series distributions is a Poisson distribution, and we discuss some properties of the BGE Poisson (BGEPo) distribution. In this case the BGEP$(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha, \theta)$ will be denoted by BGEPo$(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha, \theta)$. The following result provides the joint CDF and joint PDF of a BGEPo distribution.

Theorem 3.4: Suppose $(Y_1, Y_2) \sim$ BGEPo$(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha, \theta)$.

(a) The joint CDF of (Y_1, Y_2) for $y_1 > 0$, $y_2 > 0$, is

$$F_{Y_1,Y_2}(y_1, y_2) = \frac{\exp\left\{\theta\left[(1-e^{-\lambda_1 y_1})^{-\alpha_1} + (1-e^{-\lambda_2 y_2})^{-\alpha_2} - 1\right]^{-\alpha}\right\} - 1}{e^{\theta} - 1}$$

and zero, otherwise.

Table 3: Kendall's tau for the BGEG copula function

$\alpha \downarrow \theta \rightarrow$	0.1	0.2	0.4	0.6	0.8	0.9
0.5	0.44	0.44	0.41	0.40	0.35	0.33
1.0	0.34	0.33	0.33	0.32	0.32	0.32
1.5	0.24	0.26	0.26	0.28	0.28	0.29
2.0	0.21	0.22	0.24	0.25	0.28	0.29
2.5	0.18	0.19	0.19	0.22	0.26	0.28
3.0	0.16	0.17	0.19	0.21	0.26	0.28
5.0	0.11	0.12	0.16	0.19	0.25	0.27
10.0	0.07	0.08	0.13	0.18	0.24	0.26
20.0	0.05	0.07	0.12	0.16	0.23	0.26
50.0	0.04	0.05	0.11	0.16	0.22	0.25

(b) The joint PDF of (Y_1, Y_2) for $y_1 > 0$, $y_2 > 0$, is

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) \quad = \quad & k \times u_1(y_1; \alpha_1, \lambda_1) \times u_2(y_2; \alpha_2, \lambda_2) \times u_3^{-\alpha-2}(y_1, y_2; \alpha_1, \lambda_1, \alpha_2, \lambda_2) \times \\ & \exp\left\{\theta u_3^{-\alpha}(y_1, y_2; \alpha_1, \lambda_1, \alpha_2, \lambda_2)\right\} \left[(1+\alpha) + \alpha\theta u_3^{-\alpha}(y_1, y_2; \alpha_1, \lambda_1, \alpha_2, \lambda_2)\right], \end{aligned}$$

and zero otherwise. Here $u_1(y_1; \alpha_1, \lambda_1)$, $u_2(y_2; \alpha_2, \lambda_2)$ and $u_3(y_1, y_2; \alpha_1, \lambda_1, \alpha_2, \lambda_2)$ are same as defined in Theorem 3.1 (b), and $k = \dfrac{\alpha\theta\alpha_1\alpha_2\lambda_1\lambda_2}{e^\theta - 1}$.

PROOF: The proofs can be easily obtained by substituting $A(\theta) = e^\theta - 1$ in Theorem 2.1 and Theorem 2.2. ∎

In Figure 2 we provide the plots of the joint PDF of BGEPo for different parameter values. It is observed that joint PDF of a BGEPo is unimodal but it can take variety of shapes. It can be heavy tail in this case also.

The following result provides the CDF, PDF and the hazard functions of the marginal distributions of a BGEPo distribution.

THEOREM 3.5: Suppose $(Y_1, Y_2) \sim \text{BGEPo}(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \alpha, \theta)$.

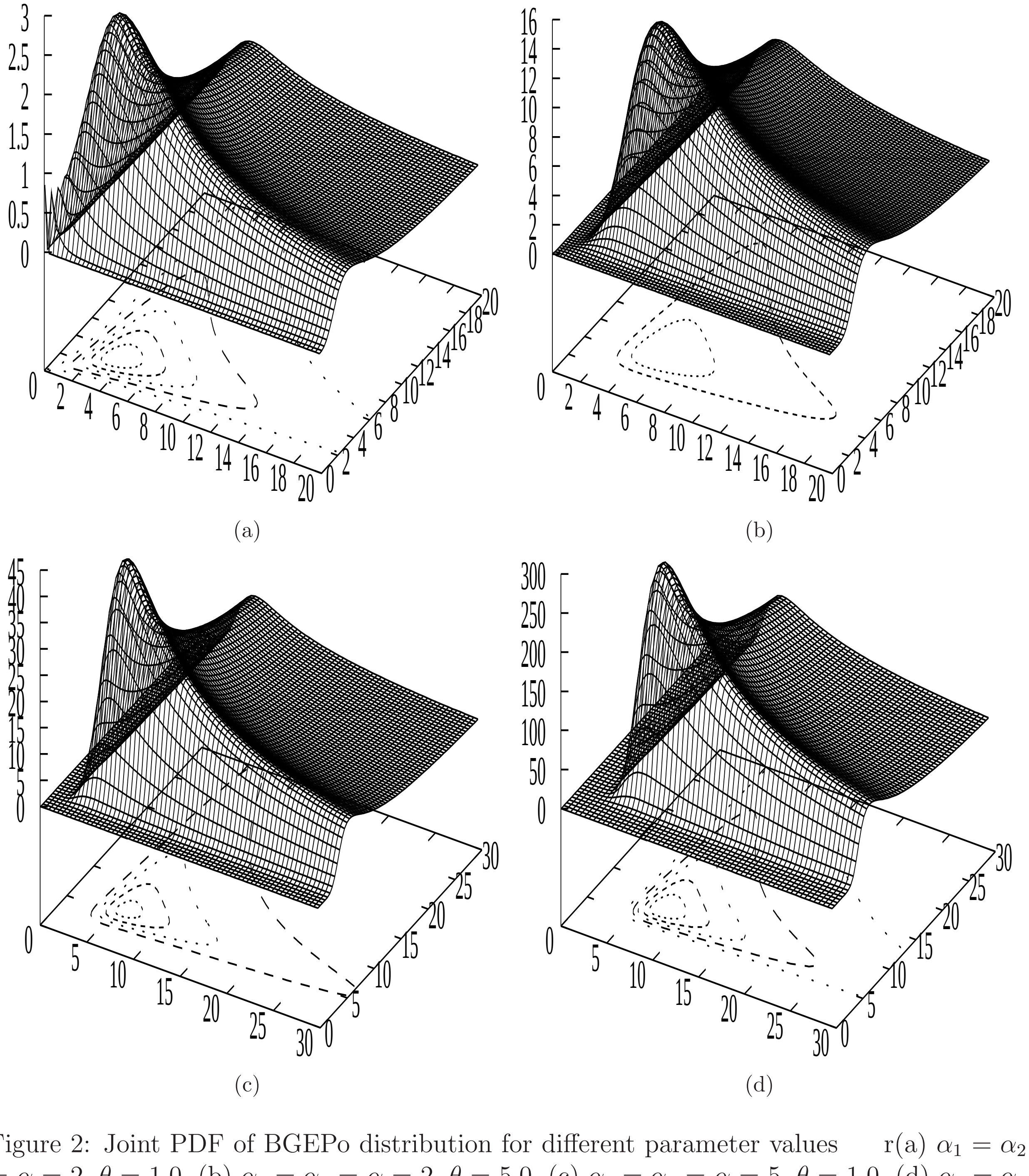


Figure 2: Joint PDF of BGEPo distribution for different parameter values r(a) $\alpha_1 = \alpha_2 = \alpha = 2$, $\theta = 1.0$, (b) $\alpha_1 = \alpha_2 = \alpha = 2$, $\theta = 5.0$, (c) $\alpha_1 = \alpha_2 = \alpha = 5$, $\theta = 1.0$, (d) $\alpha_1 = \alpha_2 = \alpha = 5$, $\theta = 5.0$.

(a) The CDFs of Y_1 and Y_2 are

$$\begin{aligned}F_{Y_1}(y_1) &= \frac{\exp\left\{\theta(1-e^{-\lambda_1 y_1})^{\alpha\alpha_1}\right\}-1}{e^{\theta}-1}\\ F_{Y_2}(y_2) &= \frac{\exp\left\{\theta(1-e^{-\lambda_2 y_2})^{\alpha\alpha_2}\right\}-1}{e^{\theta}-1}.\end{aligned}$$

(b) The PDFs of Y_1 and Y_2 are

$$\begin{aligned}f_{Y_1}(y_1) &= \frac{\alpha\alpha_1\theta\lambda_1 e^{-\lambda_1 y_1}(1-e^{-\lambda_1 y_1})^{\alpha\alpha_1-1}}{e^{\theta}-1}\times\exp\left\{\theta(1-e^{-\lambda_1 y_1})^{\alpha\alpha_1}\right\}\\ f_{Y_2}(y_2) &= \frac{\alpha\alpha_2\theta\lambda_2 e^{-\lambda_2 y_2}(1-e^{-\lambda_2 y_2})^{\alpha\alpha_2-1}}{e^{\theta}-1}\times\exp\left\{\theta(1-e^{-\lambda_2 y_2})^{\alpha\alpha_2}\right\},\end{aligned}$$

respectively.

(c) The hazard functions of Y_1 and Y_2 are

$$\begin{aligned}h_{Y_1}(y_1) &= \alpha\alpha_1\theta\lambda_1 e^{-\lambda_1 y_1}(1-e^{-\lambda_1 y_1})^{\alpha\alpha_1-1}\times\frac{\exp\left\{\theta(1-e^{-\lambda_1 y_1})^{\alpha\alpha_1}\right\}}{e^{\theta}-\exp\left\{\theta(1-e^{-\lambda_1 y_1})^{\alpha\alpha_1}\right\}}\\ h_{Y_2}(y_2) &= \alpha\alpha_2\theta\lambda_2 e^{-\lambda_2 y_2}(1-e^{-\lambda_2 y_2})^{\alpha\alpha_2-1}\times\frac{\exp\left\{\theta(1-e^{-\lambda_2 y_2})^{\alpha\alpha_2}\right\}}{e^{\theta}-\exp\left\{\theta(1-e^{-\lambda_2 y_2})^{\alpha\alpha_2}\right\}},\end{aligned}$$

respectively.

PROOF: The proofs can be easily obtained using Theorem 3.4, the details are avoided. ∎

It can be seen that the hazard function of Y_1 can be (a) increasing if $\alpha\alpha_1 > 1$ for all values of θ, (b) decreasing if $\alpha\alpha_1 \leq 1$ and $\theta = 0$ and (c) bathtub shaped if $\alpha\alpha_1 \leq 1$ and $\theta > 0$. Similarly, the hazard function of Y_2 can be obtained.

The moment generating functions of Y_1 and Y_2 can be obtained as follows:

$$M_{Y_1}(t) = \frac{\Gamma(1-t/\lambda_1)}{e^{\theta}-1}\sum_{n=1}^{\infty}\frac{1}{n!}\frac{\theta^n\Gamma(n\alpha\alpha_1+1)}{\Gamma(n\alpha\alpha_1-t/\lambda_1+1)}$$

for $t < \lambda_1$ and

$$M_{Y_2}(t) = \frac{\Gamma(1-t/\lambda_2)}{e^{\theta}-1}\sum_{n=1}^{\infty}\frac{1}{n!}\frac{\theta^n\Gamma(n\alpha\alpha_2+1)}{\Gamma(n\alpha\alpha_2-t/\lambda_2+1)},$$

for $t < \lambda_2$. Also, the above moment generating functions can be expressed as

$$M_{Y_1}(t) = \sum_{n=1}^{\infty} \sum_{j=0}^{\infty} (-1)^j \binom{n\alpha\alpha_1 - 1}{j} \frac{\theta^n}{n!(e^\theta - 1)} \frac{n\alpha\alpha_1}{j + 1 - t/\lambda_1}$$

for $t < \lambda_1$ and

$$M_{Y_2}(t) = \sum_{n=1}^{\infty} \sum_{j=0}^{\infty} (-1)^j \binom{n\alpha\alpha_2 - 1}{j} \frac{\theta^n}{n!(e^\theta - 1)} \frac{n\alpha\alpha_2}{j + 1 - t/\lambda_2},$$

for $t < \lambda_2$. Here all the moments of the marginals can be easily obtained. The following result provides the copula of a BGEG distribution.

THEOREM 3.6: The BGEPo distribution has the following copula structure;

$$C(u, v) = \frac{\exp\left[\theta\left\{\left(\frac{1}{\theta}\ln(u(e^\theta - 1) + 1)\right)^{-1/\alpha} + \left(\frac{1}{\theta}\ln(v(e^\theta - 1) + 1)\right)^{-1/\alpha} - 1\right\}^{-\alpha}\right] - 1}{e^\theta - 1}, \tag{13}$$

for $(u, v) \in [0, 1] \times [0, 1]$.

PROOF: By using $A(\theta) = e^\theta - 1$ in Theorem 2.3, the result can be easily obtained. ∎

Based on the BGEPo copula, the median correlation coefficient between Y_1 and Y_2 can be written as

$$M_{Y_1,Y_2} = \frac{4\exp\left[\theta\left\{2\left(\frac{1}{\theta}\ln(\frac{1}{2}(e^\theta + 1))\right)^{-1/\alpha} - 1\right\}^{-\alpha}\right] - 1}{e^\theta - 1}.$$

In this case also after some calculation it can be seen that the upper tail dependence $\chi = 0$. Therefore, if (Y_1, Y_2) follows BGEPo, then Y_1 and Y_2 are asymptotically independent.

In case of BGEPo also Kendall's tau and Spearman's rho cannot be computed analytically, and they need to be obtained numerically. In Tables 4 and 5 we provide Spearman's rho and Kendall's tau values, respectively, for different α and θ. From Tables 4 and 5, it is observed that Spearman's rho and Kendall's tau are always non-negative in case of BGEPo. In both the cases for any fixed α, as θ increases, Spearman's rho decreases.

Table 4: Spearman's rho for the BGEPo copula function

$\alpha \downarrow \theta \rightarrow$	1.0	2.0	3.0	4.0	10.0	20.0
0.5	0.63	0.57	0.50	0.44	0.21	0.11
1.0	0.46	0.43	0.38	0.32	0.15	0.08
1.5	0.38	0.36	0.32	0.28	0.13	0.06
2.0	0.32	0.32	0.29	0.25	0.11	0.06
2.5	0.29	0.29	0.27	0.24	0.11	0.05
3.0	0.27	0.26	0.26	0.23	0.10	0.05
5.0	0.25	0.24	0.23	0.21	0.09	0.05
10.0	0.22	0.21	0.21	0.19	0.09	0.04
20.0	0.20	0.19	0.19	0.18	0.08	0.04
50.0	0.19	0.19	0.19	0.17	0.08	0.04

Table 5: Kendall's tau for the BGEPo copula function

$\alpha \downarrow \theta \rightarrow$	1.0	2.0	3.0	4.0	10.0	20.0
0.5	0.44	0.38	0.33	0.29	0.15	0.08
1.0	0.31	0.30	0.26	0.22	0.11	0.05
1.5	0.26	0.25	0.22	0.19	0.08	0.04
2.0	0.22	0.21	0.18	0.17	0.07	0.04
2.5	0.20	0.19	0.18	0.15	0.07	0.04
3.0	0.18	0.18	0.17	0.15	0.06	0.04
5.0	0.16	0.16	0.15	0.14	0.06	0.04
10.0	0.14	0.14	0.13	0.13	0.05	0.03
20.0	0.14	0.14	0.12	0.12	0.05	0.03
50.0	0.12	0.12	0.11	0.11	0.04	0.03

4 Estimation

In this section we provide the estimation procedures of the unknown parameters for both BGEG and BGEP$_o$ distributions based on a complete sample. It is assumed that we have a random sample of size m, say $\{(y_{11}, y_{12}), \ldots, (y_{m1}, y_{m2})\}$ either from a BGEG or from a BGEP$_o$ distribution, and we would like to estimate the unknown parameters based on the above sample. Note that to compute the maximum likelihood estimators of the unknown parameters, it involves solving a six dimensional optimization problem. Let us denote the unknown parameter vector as $\mathbf{\Gamma} = (\alpha, \theta, \lambda_1, \alpha_1, \lambda_2, \alpha_2)$.

The standard Newton-Raphson type algorithm can be used to solve the normal equations simultaneously. It is well known that the standard Newton-Raphson algorithm requires a very good set of initial values. Moreover, it has a problem of converging to a local maximum rather than a global maximum, particularly in a multidimensional problem like this. In case of BGEG, the log-likelihood function of the unknown parameters can be written as

$$\begin{aligned} l_G(\mathbf{\Gamma}) &= m \ln \alpha + m \ln(1-\theta) + m \ln \alpha_1 + m \ln \lambda_1 + m \ln \alpha_2 + m \ln \lambda_2 \\ & \sum_{i=1}^{m} \ln u_1(y_{i1}; \alpha_1, \lambda_1) + \sum_{i=1}^{m} \ln u_2(y_{i2}; \alpha_2, \lambda_2) + (\alpha - 2) \sum_{i=1}^{m} \ln u_3(y_{i1}, y_{i2}; \alpha_1, \lambda_1, \alpha_2, \lambda_2) \\ & + \sum_{i=1}^{m} \ln \left([u_3^{\alpha}(y_{i1}, y_{i2}; \alpha_1, \lambda_1, \alpha_2, \lambda_2) - \theta] + 2u_3^{\alpha}(y_{i1}, y_{i2}; \alpha_1, \lambda_1, \alpha_2, \lambda_2) \right) \\ & -3 \ln \left[u_3^{\alpha}(y_{i1}, y_{i2}; \alpha_1, \lambda_1, \alpha_2, \lambda_2) - \theta \right]. \end{aligned}$$

Similarly, in case of BGEP$_o$ it becomes

$$\begin{aligned} l_P(\mathbf{\Gamma}) &= m \ln \alpha + m \ln \theta + m \ln \alpha_1 + m \ln \lambda_1 + m \ln \alpha_2 + m \ln \lambda_2 - m \ln(e^{\theta} - 1) \\ & \sum_{i=1}^{m} \ln u_1(y_{i1}; \alpha_1, \lambda_1) + \sum_{i=1}^{m} \ln u_2(y_{i2}; \alpha_2, \lambda_2) - (\alpha + 2) \sum_{i=1}^{m} \ln u_3(y_{i1}, y_{i2}; \alpha_1, \lambda_1, \alpha_2, \lambda_2) \\ & + \theta \sum_{i=1}^{m} u_3^{-\alpha}(y_{i1}, y_{i2}; \alpha_1, \lambda_1, \alpha_2, \lambda_2) + \sum_{i=1}^{m} \ln \left((1+\alpha) + \alpha\theta u_3^{-\alpha}(y_{i1}, y_{i2}; \alpha_1, \lambda_1, \alpha_2, \lambda_2) \right). \end{aligned}$$

We propose to treat this problem as a missing value problem and will use EM algorithm to compute the MLEs of the unknown parameters. Using the idea of Song et al. (2005), it is observed that at each 'M'-step of the EM algorithm, the 'pseudo' log-likelihood function can be maximized by solving three one dimensional non-linear equations. It helps to reduce the computational burden significantly and it also avoids to converge to a local maximum rather than a global maximum. The EM algorithm is based on the following observations.

First we will show that if N is also observed along with Y_1 and Y_2, then the MLEs of the unknown parameters can be obtained by solving three one dimensional non-linear equations, see Theorem 1. This is the main motivation behind the proposed EM algorithm. Suppose the complete observations are as follows: $\{(y_{i1}, y_{i2}, n_i); i = 1, \ldots, m\}$. Therefore, using (8) the log-likelihood function based on the complete observations can be written as

$$l_G^c(\mathbf{\Gamma}) = \widetilde{n} \ln \theta - m \ln \theta + m \ln(1-\theta) + \sum_{i=1}^{m} \ln f_{BGE}(y_{i1}, y_{i2}; \alpha_1, \lambda_1, \alpha_2, \lambda_2, n_i\alpha), \qquad (14)$$

in case of BGEG distribution and

$$l_P^c(\mathbf{\Gamma}) = \widetilde{n} \ln \theta - m \ln(e^{\theta} - 1) + \sum_{i=1}^{m} \ln f_{BGE}(y_{i1}, y_{i2}; \alpha_1, \lambda_1, \alpha_2, \lambda_2, n_i\alpha) \qquad (15)$$

in case of BGEP$_o$ distribution. Here $\widetilde{n} = \sum_{i=1}^{m} n_i$. Therefore, the MLEs of the unknown parameters can be obtained in case of BGEG and BGEP$_o$ by maximizing (14) and (15), respectively. Therefore, in case of BGEG, the MLE of θ can be obtained as

$$\widehat{\theta} = \frac{\widetilde{n} - m}{\widetilde{n}}$$

and in case of BGEP$_o$, it can be obtained by solving the following non-linear equation

$$\widetilde{n}(1 - e^{-\theta}) = m\theta.$$

In both the cases by maximizing $\sum_{i=1}^{m} \ln f_{BGE}(y_{i1}, y_{i2}; \alpha_1, \lambda_1, \alpha_2, \lambda_2, n_i\alpha)$ with respect to the unknown parameters the MLEs of $\alpha, \alpha_1, \lambda_1, \alpha_2, \lambda_2$ can be obtained. We use the following Theorem for further development.

Theorem 5.1: The maximization of $\sum_{i=1}^{m} \ln f_{BGE}(y_{i1}, y_{i2}; \alpha_1, \lambda_1, \alpha_2, \lambda_2, n_i\alpha)$ with respect to the unknown parameters can be obtained by solving three one dimensional non-linear equations.

Proof: See in the Appendix A.

Now based on the above observations we propose the following EM algorithm. Suppose at the j-th step the estimate of $\mathbf{\Gamma}$ is $\mathbf{\Gamma}^{(j)}$.

Step 1: Compute $n_i^{(j)} = E(N|y_{i1}, y_{i2}, \mathbf{\Gamma}^{(j)})$.

Step 2: Obtain the 'pseudo' log-likelihood function from (14) and (15) by replacing n_i with $n_i^{(j)}$.

Step 3: Obtain $\mathbf{\Gamma}^{(j+1)}$ by maximizing the 'pseudo' log-likelihood function.

Step 4: Check the convergence, otherwise go back to Step 1.

5 Simulation Experiments

In this section we have performed some small simulation experiments mainly to see how the proposed EM algorithm works in practice. We have considered both the models namely (a) Model 1: BGEG and (b) Model 2: BGEP$_o$. For Model 1, we have considered the following set of parameters: $\alpha_1 = \alpha_2 = \lambda_1 = \lambda_2 = 1.0$, $\alpha = 2.0$, $\theta = 0.5$ and for Model 2: $\alpha_1 = \alpha_2 = \lambda_1 = \lambda_2 = 1.0$, $\alpha = 2.0$, $\theta = 1.5$. We have taken different sample sizes, and each case we have computed the MLEs based on the proposed EM algorithm. We report the average MLEs and the associated square root of the mean squared errors (MSEs) over 1000 replications. The results are reported in Tables 6 and 7.

Some of the points are quite clear from the simulation results. It is observed that in both

n	α_1	λ_1	α_2	λ_2	α	θ
20	1.1756 (0.0621)	1.1143 (0.0415)	1.1698 (0.0656)	1.1097 (0.0427)	2.2315 (0.0956)	0.5216 (0.0398)
40	1.1331 (0.0319)	1.0853 (0.0242)	1.1398 (0.0365)	1.0887 (0.0217)	2.1767 (0.0511)	0.5117 (0.0213)
60	1.0767 (0.0218)	1.0515 (0.0141)	1.0761 (0.0241)	1.0551 (0.0131)	2.0711 (0.0319)	0.5081 (0.0128)
80	1.0668 (0.0181)	1.0439 (0.0156)	1.0616 (0.0171)	1.0431 (0.0179)	2.0412 (0.0281)	0.5023 (0.0103)

Table 6: The average MLEs and the associated square root of the MSEs (within brackets) are reported for Model 1.

n	α_1	λ_1	α_2	λ_2	α	θ
20	1.1854 (0.0710)	1.1179 (0.0521)	1.1821 (0.0728)	1.1188 (0.0552)	2.1654 (0.0887)	1.5319 (0.0987)
40	1.1445 (0.0365)	1.0976 (0.0261)	1.1467 (0.0378)	1.0988 (0.0258)	2.1328 (0.0413)	1.5175 (0.0472)
60	1.0887 (0.0245)	1.0691 (0.0178)	1.0856 (0.0251)	1.0624 (0.0169)	2.0612 (0.0305)	1.5043 (0.0298)
80	1.0710 (0.0161)	1.0512 (0.0143)	1.0694 (0.0159)	1.0489 (0.0137)	2.0117 (0.0217)	1.5017 (0.0225)

Table 7: The average MLEs and the associated square root of the MSEs (within brackets) are reported for Model 2.

Subject	1	2	3	4	5	6	7	8	9	10	11
Stage 1	143	142	109	123	276	235	208	267	183	245	324
Stage 2	368	155	167	135	216	386	175	358	193	268	507
Subject	12	13	14	15	16	17	18	19	20	21	22
Stage 1	148	130	119	102	279	244	196	279	167	345	524
Stage 2	378	142	171	94	204	365	168	358	183	238	507

Table 8: Left brow EMG amplitudes from 22 subjects.

the cases as the sample size increases the average biases and square root of the MSEs decrease. It indicates the consistency property of the MLEs. Moreover, in all our replications it is observed that the EM algorithm converges, hence it shows that the proposed EM algorithm is working well for this model.

6 Data Analysis

In this section we present the analysis of two data sets mainly for illustrative purposes. The main idea is to see how the present models and the proposed EM algorithm work in practice.

6.1 EMG Data:

This data set arises from a study of affective facial expressions conducted on 22 subjects. In this study several pieces of music were played to each individual to elicit selected affective states. Here two stages are reported, Stage 1: relaxing music condition, and Stage 2: it was designed to create positive effects. In this case each trial lasted for 90 seconds, and the response variable at each stage was the mean electromyographic (EMG) amplitudes (μV) from the left eye brow region. The original data set is available in Davis (2002) and we are reporting the data set in Table 8 for easy reference.

Preliminary data analysis suggests that the marginals are not symmetric and the empirical hazard functions are not constant. We have used BGEG and BGEP$_o$ for the analysis of this data set. Before progressing further we have divided all the data points by 100, it is not going to affect in the inference issues.

BGEG Model:

From the marginals of Y_1 we obtain the initial estimates of λ_1, $\beta_1 = \alpha\alpha_1$ and θ as 1.23, 6.50 and 0.22, respectively. Similarly, from the marginal of Y_2, we obtain the initial estimates of λ_2, $\beta_2 = \alpha\alpha_2$ and θ as 0.7765, 3.90 and 0.10, respectively. We have used the above initial estimates of λ_1, λ_2, β_1 and β_2, and for θ we have used $(0.22+0.10)/2 = 0.16$. We have used these initial estimates to obtain an initial estimate of α from the $\max\{Y_1, Y_2\}$, and it is obtained as 0.11. We have used these initial estimates for our EM algorithm and obtain the MLEs of the unknown parameters and the associated 95% confidence intervals as

$$\begin{aligned}
\widehat{\lambda}_1 &= 1.2083(\mp 0.1878), \quad \widehat{\alpha}_1 = 31.5221(\mp 2.6783) \\
\widehat{\lambda}_2 &= 0.7998(\mp 0.0898), \quad \widehat{\alpha}_2 = 19.2061(\mp 1.1156) \\
\widehat{\alpha} &= 0.2031(\mp 0.0012), \quad \widehat{\theta} = 0.1582(\mp 0.0005).
\end{aligned}$$

The associated log-likelihood value is -138.324. Now the natural question is whether BGEG provides a good fit to the bivariate data set or not. We have computed the Kolmogorov-Smirnov distances and the associated p values between the fitted cumulative distribution function (CDF) based on the model and the empirical distribution functions of the two marginals and the results are reported in Table 9. We have also used the multivariate Kolmogorov-Smirnov test of goodness of fit as proposed by Justel et al. (1997). We obtain the value of the test statistic as 0.2981. Based on 10000 replications, we obtain 10% critical values as 0.3227. Hence $p > 0.1$. Therefore, based on the p value, we cannot reject the null hypothesis that the data are coming from BGEG distribution. Based on all these, it is clear that the proposed BGEG provides a good fit to the data set.

Variable	K-S Distance	p-value
Y_1	0.0960	0.9873
Y_2	0.1293	0.8556

Table 9: Kolmogorov-Smirnov Distances and the associated p values for EMG Data Set for BGEG model.

Variable	K-S Distance	p-value
Y_1	0.0945	0.9893
Y_2	0.1258	0.8770

Table 10: Kolmogorov-Smirnov Distances and the associated p values for EMG Data Set for BGEP model.

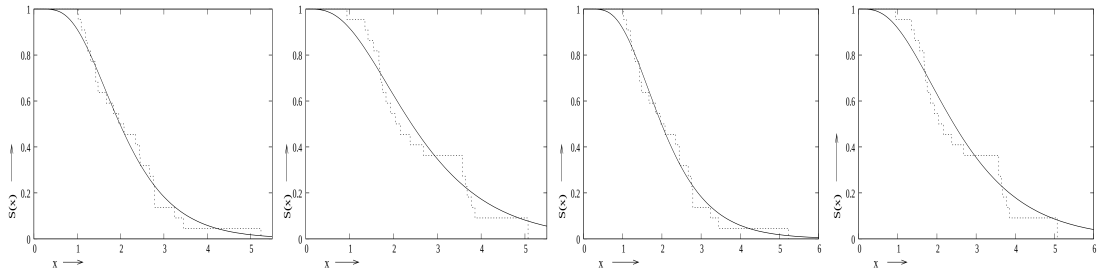
BGEP$_o$ Model

Using the same method as in the previous example we have obtained the initial estimates of λ_1, α_1, λ_2, α_2, β and θ as 1.17, 31.12, 0.67, 23.56, 0.12 and 0.21, respectively. We have used these initial estimates for our EM algorithm and obtain the MLEs of the unknown parameters and the associated 95% confidence intervals as

$$\begin{aligned}
\widehat{\lambda}_1 &= 1.2047(\mp 0.1922), \quad \widehat{\alpha}_1 = 35.3521(\mp 3.1153) \\
\widehat{\lambda}_2 &= 0.7765(\mp 0.0915), \quad \widehat{\alpha}_2 = 20.5528(\mp 1.6861) \\
\widehat{\alpha} &= 0.1864(\mp 0.0009), \quad \widehat{\theta} = 0.2682(\mp 0.0006).
\end{aligned}$$

The associated log-likelihood value is -143.452. We have computed the Kolmogorov-Smirnov distances and the associated p values between the fitted CDF and the empirical distribution functions of the marginals and the results are reported in Table 10. In Figure 3 we have provided the plots for the empirical survival functions for the two marginals and the associated fitted survival functions for GE-Geometric (GEG) and GE-Poisson (GEPo) distributions. In this case the multivariate Kolmogorov-Smirnov test statistic becomes 0.2342, and based on 10000 replications we obtain 10% critical value as 0.2756. Hence, in this case also $p > 0.1$.

It is clear that both the models provide good fit to the EMG data set. Based on the



(a) (b) (c) (d)

Figure 3: Empirical survival and fitted survival functions for the two marginals for EMG data set (a) Fitted GEG model for Stage 1, (b) Fitted GEG model for Stage 2, (c) Fitted GEPo model for Stage 1, (d) Fitted GEPo model for Stage 2.

Subject	1	2	3	4	5	6	7	8	9	10
Y_1	2.139	1.873	1.887	1.739	1.734	1.509	1.695	1.740	1.811	1.954
Y_2	2.138	1.741	1.809	1.547	1.715	1.474	1.656	1.777	1.759	2.009
Subject	11	12	13	14	15	16	17	18	19	20
Y_1	1.624	2.204	1.508	1.786	1.902	1.743	1.863	2.028	1.390	2.187
Y_2	1.657	1.846	1.458	1.811	1.606	1.794	1.869	2.032	1.342	2.087
Subject	21	22	23	24	25					
Y_1	1.650	2.334	1.037	1.509	1.971					
Y_2	1.378	2.225	1.268	1.422	1.869					

Table 11: Bone mineral density of Dominant humerus and humerus of 25 subjects.

log-likelihood value we can say that the BGEG provides a better fit than the BGEPo for the EMG data set.

6.2 Bone Mineral Data:

This bivariate data set has been obtained from Johnson and Wichern (1999, page 34), representing the bone mineral density (BMD) measured in g/cm^2 for 25 individuals. Here Y_1 and Y_2 denote the BMD of the dominant humerus and humerus, respectively. The data set is presented in Table 11

In this case also the preliminary data analysis indicates that the marginals are not sym-

Variable	K-S Distance	p-value
Y_1	0.1335	0.8288
Y_2	0.1193	0.9129

Table 12: Kolmogorov-Smirnov Distances and the associated p values for bone-mineral Data Set for BGEG model.

metric and the empirical hazard functions of the marginals are not constant. We have used BGEG and BGEP$_o$ to analyze this data set. In this case we have subtracted one from each data point, and again it is not going affect in the inference issues.

BGEG Model

Similarly as the previous case from the two marginals we have obtained the initial estimates of λ_1, β_1 and θ as 4.61, 9.90 and 0.71 , respectively. The initial estimates of λ_2, β_2 and θ are 4.23, 10.21 and 0.54, respectively. The initial estimate of α becomes 0.01. We have used these initial estimates for our EM algorithm and obtain the MLEs of the unknown parameters and the associated 95% confidence intervals as

$$\begin{aligned}
\widehat{\lambda}_1 &= 5.8431(\mp 0.9878), \quad \widehat{\alpha}_1 = 657.8641(\mp 28.6751) \\
\widehat{\lambda}_2 &= 6.2659(\mp 0.1.0412), \quad \widehat{\alpha}_2 = 589.1842(\mp 26.7781) \\
\widehat{\alpha} &= 0.0076(\mp 0.0008), \quad \widehat{\theta} = 0.9511(\mp 0.0487).
\end{aligned}$$

The associated log-likelihood value is -237.322. Similar to the previous example we have reported different K-S distances and the associated p-values in Table 12. The multivariate Kolmogorov-Smirnov test statistic becomes 0.1943, and based on 10000 replications we obtain 10% critical value as 0.2313. Hence, in this case also $p > 0.1$.

BGEP$_o$ Model

Using the same method as in the previous example we have obtained the initial estimates of λ_1, α_1, λ_2, α_2, α and θ as 4.01, 331.67, 4.11, 299.76, 0.0076 and 5.13, respectively. We

Variable	K-S Distance	p-value
Y_1	0.0889	0.9889
Y_2	0.0970	0.9726

Table 13: Kolmogorov-Smirnov Distances and the associated p values for bone-mineral Data Set for BGEP model.

have used these initial estimates for our EM algorithm and obtain the MLEs of the unknown parameters and the associated 95% confidence intervals as

$$\begin{aligned}
\widehat{\lambda}_1 &= 4.3428(\mp 0.5987), \quad \widehat{\alpha}_1 = 374.3793(\mp 31.7727) \\
\widehat{\lambda}_2 &= 4.6052(\mp 0.6127), \quad \widehat{\alpha}_2 = 311.6293(\mp 28.9818) \\
\widehat{\alpha} &= 0.0116(\mp 0.0006), \quad \widehat{\theta} = 5.4310(\mp 0.9878).
\end{aligned}$$

The associated log-likelihood value is -223.451. We have computed the Kolmogorov-Smirnov distances and the associated p values between the fitted model and the empirical distribution functions of the marginals and the maximum of the two marginals and the results are reported in Table 13. In Figure 4 we have provided the plots for the empirical survival functions for the two marginals and the associated fitted survival functions for GEG and GEPo distributions. The multivariate Kolmogorov-Smirnov test statistic becomes 0.2012, and based on 10000 replications we obtain 10% critical value as 0.2547. Hence, in this case also $p > 0.1$.

It is clear that both the models provide good fit to the BMD data set, and based on the log-likelihood value we can say that BGEPo provides a better fit than the BGEG for the BMD data set.

7 Multivariate GE power series distributions

In this section we would like to define multivariate GE power series distribution. Recently, Kundu et al. (2015) defined the MGE distribution as follows. A p-variate random vector

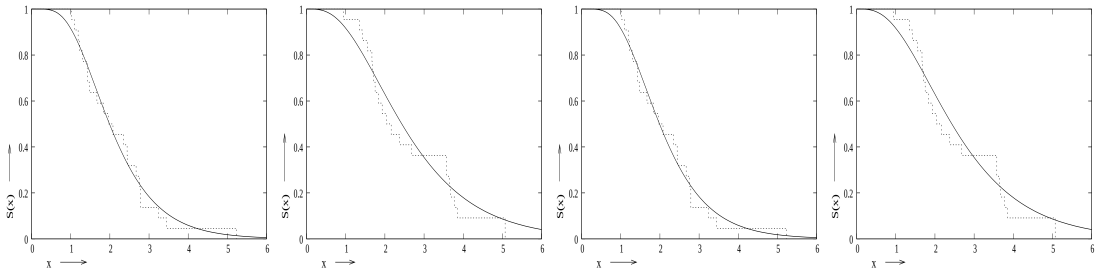


Figure 4: Empirical survival and fitted survival functions for the two marginals for the BMD data set (a) Fitted GEG model for BMD of dominant humerus, (b) Fitted GEG model for BMD of humerus, (c) Fitted GEPo model for BMD of dominant humerus, (d) Fitted GEPo model for BMD of humerus.

$\boldsymbol{X} = (X_1, \ldots, X_p)^T$ is said to have a MGE distribution with parameters $\alpha_1 > 0, \ldots, \alpha_p > 0$, $\lambda_1 > 0, \ldots, \lambda_p > 0$, $\alpha > 0$, if the joint PDF of $\boldsymbol{X}$ for $x_1 > 0, \ldots, x_p > 0$, is

$$f_{\boldsymbol{X}}(x_1, \ldots, x_p) = \frac{c e^{-\sum_{i=1}^{p} \lambda_i x_i} \prod_{i=1}^{p} \left(1 - e^{-\lambda_i x_i}\right)^{-\alpha_i - 1}}{\left[\sum_{i=1}^{p} \left(1 - e^{-\lambda_i x_i}\right)^{-\alpha_i} - (p-1)\right]^{\alpha + p}}, \tag{16}$$

and zero otherwise. Here c is the normalizing constant, and $c = \prod_{i=1}^{p} \alpha_i \lambda_i (\alpha + i - 1)$. The joint CDF of $(X_1, \ldots, X_p)^T$ becomes

$$F_{\boldsymbol{X}}(x_1, \ldots, x_p) = \frac{1}{\left[\sum_{i=1}^{p} (1 - e^{-\lambda_i x_i})^{-\alpha_i} - (p-1)\right]^{\alpha}}, \tag{17}$$

for $x_1 > 0, \ldots, x_p > 0$, and zero otherwise. From now on a multivariate random variable with the joint CDF (17) and joint PDF (16) will be denoted by $\text{MGE}(\alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_p, \alpha)$.

Suppose $\{(X_{1n}, X_{2n}, \ldots, X_{pn}); n = 1, 2, \ldots\}$ is a sequence of independent and identically distributed (i.i.d.) $\text{MGE}(\alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_p, \alpha)$ random variables, N is a discrete random variable with the PMF (6) and it is independent of $\{(X_{1n}, X_{2n}, \ldots, X_{pn}); n = 1, 2, \cdots\}$. Let us define

$$Y_j = \max\{X_{j1}, \cdots, X_{jN}\}, \quad j = 1, 2, \ldots, p.$$

Then $(Y_1, \ldots, Y_p)$ is said to have a MVGEP distribution with parameters $\alpha_1, \ldots, \alpha_p$, $\lambda_1, \ldots, \lambda_p$, α, θ, and it will be denoted by $\text{MVGEP}(\alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_p, \alpha)$.

Theorem 6.1: If $(Y_1, \ldots, Y_p) \sim \text{MVGEP}(\alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_p, \alpha)$, then the joint CDF of $(Y_1, \cdots, Y_p)$ is

$$F_{Y_1,\ldots,Y_p}(y_1, \ldots, y_p) = \frac{A\left(\theta\left[\sum_{i=1}^{p}(1-e^{-\lambda_i y_i})^{-\alpha_i} - (p-1)\right]^{-\alpha}\right)}{A(\theta)}.$$

Proof:

$$\begin{aligned}
F_{Y_1,\ldots,Y_p}(y_1, \ldots, y_p) &= P(Y_1 \le y_1,\, \ldots, Y_p \le y_p) \\
&= \sum_{n=1}^{\infty} P(Y_1 \le y_1,\, \ldots, Y_p \le y_p | N = n) P(N = n) \\
&= \sum_{n=1}^{\infty} F_{\mathbf{X}}^{n}(y_1, \ldots, y_p) \frac{a_n \theta^n}{A(\theta)} \\
&= \frac{A(\theta F_{\mathbf{X}}(y_1, \ldots, y_p))}{A(\theta)} \\
&= \frac{A\left(\theta\left[\sum_{i=1}^{p}(1-e^{-\lambda_i y_i})^{-\alpha_i} - (p-1)\right]^{-\alpha}\right)}{A(\theta)}. \qquad \blacksquare
\end{aligned}$$

Theorem 6.2: The MVGEP distribution has the following copula structure

$$C(u_1, \ldots, u_p) = \frac{A\left[\theta\left\{\sum_{i=1}^{p}\left(\frac{1}{\theta}A^{-1}(u_i A(\theta))\right)^{-1/\alpha} - (p-1)\right\}^{-\alpha}\right]}{A(\theta)},$$

for all $(u_1, \ldots, u_p) \in [0, 1]^p$.

Proof: The proof can be obtained similarly as the proof of the Theorem 2.3. ∎

Theorem 6.3: If $\mathbf{Y} = (Y_1, \ldots, Y_p)^T \sim \text{MVGEP}(\alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_p, \alpha, \theta)$, then

- (a) $Y_1 \sim \text{GEP}(\alpha_1\alpha, \lambda_1, \theta), \cdots, Y_p \sim \text{GEP}(\alpha_p\alpha, \lambda_p, \theta)$.
- (b) For any non-empty subset $I_q = (i_1, \ldots, i_q) \subset (1, \ldots, p)$, the q-dimensional marginal $\mathbf{Y}_{I_q} = (Y_{i_1}, \ldots, Y_{i_q})^T \sim \text{MVGEP}_q(\alpha_{i_1}, \ldots, \alpha_{i_q}, \lambda_{i_1}, \ldots, \lambda_{i_q}, \alpha, \theta)$.

- (c) The conditional distribution function of $(\mathbf{Y}_{I_q}|\mathbf{Y}_{I-I_q} \leq \mathbf{y}_{I-I_q})$, where the set $I - I_q = \{i \in I, i \neq i_1, \ldots, i_q\}$, is

$$P(\mathbf{Y}_{I_q} \leq \mathbf{y}_{I_q}|\mathbf{Y}_{I-I_q} \leq \mathbf{y}_{I-I_q}) = \frac{A\left(\theta\left[\sum_{i\in I}(1-e^{-\lambda_i y_i})^{-\alpha_i} - (p-1)\right]^{-\alpha}\right)}{A\left(\theta\left[\sum_{i\in I-I_q}(1-e^{-\lambda_i y_i})^{-\alpha_i} - (p-q-1)\right]^{-\alpha}\right)}. \quad (18)$$

- (d) The survival function of $\mathbf{Y} = (Y_1, \ldots, Y_p)^T$ is

$$\begin{aligned} S_{\mathbf{Y}}(\mathbf{y}) &= 1 - \sum_{i=1}^{p} \frac{A\left(\theta(1-e^{-\lambda_i y_i})^{\alpha\alpha_i}\right)}{A(\theta)} \\ &+ \sum_{1\leq i<j\leq p} \frac{A\left(\theta\left[(1-e^{-\lambda_i y_i})^{-\alpha_i} + (1-e^{-\lambda_j y_j})^{-\alpha_j} - 1\right]^{-\alpha}\right)}{A(\theta)} \\ &+ \cdots + (-1)^{p+1} \frac{A\left(\theta\left[\sum_{i=1}^{p}(1-e^{-\lambda_i y_i})^{-\alpha_i} - (p-1)\right]^{-\alpha}\right)}{A(\theta)}. \end{aligned}$$

Proof The proofs of (a), (b) and (c) can be obtained using the CDF of $(Y_1, \ldots, Y_p)$. The proof of (d) can be obtained using the following relation:

$$\begin{aligned} P(\mathbf{Y} > \mathbf{y}) &= 1 - P((\mathbf{Y} > \mathbf{y})^c) \\ &= 1 - P(\{Y_1 \leq y_1\} \cup \{Y_2 \leq y_2\} \cup \cdots \cup \{Y_p \leq y_p\}) \\ &= 1 - \sum_{i=1}^{p} P(Y_i \leq y_i) + \sum_{1\leq i<j\leq p} P(Y_i \leq y_i, Y_j \leq y_j) \\ &+ \cdots + (-1)^{p+1} P(\prod_{i=1}^{p}\{Y_i \leq y_i\}). \end{aligned}$$

8 Conclusion

In this paper we have introduced a very flexible absolute continuous bivariate model by compounding the absolute continuous bivariate generalized exponential with the power series distributions. We have discussed several properties of the proposed model. It is observed that the model has a very convenient copula structure and it can be used to establish different dependency properties and also to compute different dependency measures. Two special cases have been discussed in details to show flexibility of the proposed model. The MLEs of the unknown parameters cannot be obtained in closed form. We have proposed to use EM algorithm to compute the MLEs. Two data sets have been analyzed and it is observed the proposed models and the EM algorithm work quite well in practice. Finally we have generalized the bivariate distribution to the multivariate case, and several of its properties have been established. It will be interesting to develop proper inference procedures for the multivariate case. More work is needed in that direction.

Appendix A

Proof of Theorem 5.1:

In this appendix we will use the following notations for convenient purposes.

$$\theta_1 = \alpha_1, \theta_2 = \lambda_1, \theta_3 = \alpha_2, \theta_4 = \lambda_2, \theta_5 = \alpha$$

and

$$f_{BGE}(y_{i1}, y_{i2}; \theta_1, \theta_2, \theta_3, \theta_4, n_i\theta_5) = f_{BGE}(\theta_1, \theta_2, \theta_3, \theta_4, n_i\theta_5).$$

After some simplification it can be seen that

$$\begin{aligned}\sum_{i=1}^{m} \ln f_{BGE}(\theta_1, \theta_2, \theta_3, \theta_4, n_i\theta_5) &= l_1(\theta_1, \theta_2, \theta_3, \theta_4) + l_2(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5) \\ &= l_{11}(\theta_1, \theta_2) + l_{12}(\theta_3, \theta_4) + l_2(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5), \qquad (19)\end{aligned}$$

where

$$\begin{aligned} l_{11}(\theta_1, \theta_2) &= -\theta_2 \sum_{i=1}^{m} y_{i1} + m\theta_1 \sum_{i=1}^{m} \ln(1 - e^{-\theta_2 y_{i1}}) + m \ln \theta_1 + m \ln \theta_2 \\ l_{12}(\theta_3, \theta_4) &= -\theta_4 \sum_{i=1}^{m} y_{i2} + m\theta_3 \sum_{i=1}^{m} \ln(1 - e^{-\theta_4 y_{i4}}) + m \ln \theta_3 + m \ln \theta_4 \end{aligned}$$

$$
\begin{aligned}
l_1(\theta_1,\theta_2,\theta_3,\theta_4) &= l_{11}(\theta_1,\theta_2) + l_{12}(\theta_3,\theta_4) \\
l_2(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5) &= \theta_1\theta_5 \sum_{i=1}^{m} n_i \ln(1-e^{-\theta_2 y_{i1}}) + \theta_3\theta_5 \sum_{i=1}^{m} n_i \ln(1-e^{-\theta_4 y_{i2}}) + m\ln\theta_5 + \\
&\quad \sum_{i=1}^{m} \ln n_i + \sum_{i=1}^{n} \ln(1+n_i\alpha) - \sum_{i=1}^{m} (n_i\theta_5+2)\ln\big[(1-e^{-\theta_2 y_{i1}})^{\theta_1} + \\
&\qquad (1-e^{-\theta_3 y_{i2}})^{\theta_3} - (1-e^{-\theta_2 y_{i1}})^{\theta_1}(1-e^{-\theta_3 y_{i2}})^{\theta_3}\big].
\end{aligned}
$$

We use the method of Song et al. (2005) to maximize $l_1(\theta_1,\theta_2,\theta_3,\theta_4) + l_2(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5)$ as follows. The maximization can be obtained by solving the following five non-linear equations

$$
\frac{\partial l_1}{\partial \theta_i} + \frac{\partial l_2}{\partial \theta_i} = 0; \qquad i = 1,\ldots,5. \tag{20}
$$

The following algorithm can be used to provide the solutions of (20), see Song et al. (2005) for details.

Algorithm:

Step 1: First solve

$$
\frac{\partial l_1}{\partial \theta_i} = 0; \qquad i = 1,\ldots,4, \tag{21}
$$

and obtain $\theta_1^{(1)}$, $\theta_2^{(1)}$, $\theta_3^{(1)}$ and $\theta_4^{(1)}$.

Step 2: Obtain $\theta_5^{(1)}$ by solving

$$
\frac{\partial l_2(\theta_1^{(1)},\theta_2^{(1)},\theta_3^{(1)},\theta_4^{(1)},\theta_5)}{\partial \theta_5} = 0; \tag{22}
$$

Step 3: Let us denote $c_i^{(1)}$ as

$$
c_i^{(1)} = \frac{\partial l_2(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5^{(1)})}{\partial \theta_i}; \quad i = 1,\ldots,4,
$$

and obtain $\theta_i^{(2)}$, for $i = 1,\ldots,4$ by solving

$$
\frac{\partial l_1}{\partial \theta_i} = -c_i^{(1)}; \qquad i = 1,\ldots,4, \tag{23}
$$

Step 4: Obtain $\theta_5^{(2)}$ by solving

$$\frac{\partial l_2(\theta_1^{(2)}, \theta_2^{(2)}, \theta_3^{(2)}, \theta_4^{(2)}, \theta_5)}{\partial \theta_5} = 0; \tag{24}$$

Check the convergence. If it is not satisfied go back to Step 2 to obtain the next set of iterates.

Since $l_1(\theta_1, \theta_2, \theta_3, \theta_4) = l_{11}(\theta_1, \theta_2) + l_{12}(\theta_3, \theta_4)$, it can be easily seen that (21) and (23) can be obtained by solving two one-dimensional non-linear equations.

Acknowledgements:

The authors would like to thank two unknown referees for their constructive suggestions which have helped to improve the manuscript considerably.

References

[1] Adamidis, K., & Loukas, S. (1998), "A lifetime distribution with decreasing failure rate", *Statistics & Probability Letters*, 39(1), 35 – 42.

[2] Barreto-Souza, W., de Morais, A. L., & Cordeiro, G. M. (2011), "The Weibull-geometric distribution", *Journal of Statistical Computation and Simulation*, 81(5), 645 – 657.

[3] Barreto-Souza, W., Lemonte, A. J., & Cordeiro, G. M. (2013), "General results for the Marshall and Olkin's family of distributions", *Anais da Academia Brasileira de Cincias*, 85(1), 3 – 21.

[4] Block, H. W., & Basu, A. P. (1974), "A continuous, bivariate exponential extension", *Journal of the American Statistical Association*, 69(348), 1031 – 1037.

[5] Blomqvist, N. (1950), “On a measure of dependence between two random variables”, *Annals of Mathematical Statistics*, , vol. 21, 593 – 600.

[6] Chahkandi, M., & Ganjali, M. (2009), “On some lifetime distributions with decreasing failure rate”, *Computational Statistics & Data Analysis*, 53(12), 4433 – 4440.

[7] Coles, S., Hefferman, J. & Tawn, J. (1999), “Dependence measures for extreme value analysis”, *Extremes*, vol 2, 339 – 365.

[8] Cordeiro, G. M. & Silva, R. B. (2014), “The complementary extended Weibull power series class of distributions”, *Ciancia e' Natura*, 36(3), 1 – 13.

[9] Davis, C. (2002), *Statistical methods for the analysis of repeated measures*, Springer, New York.

[10] Flores, J., Borges, P., Cancho, V. G., & Louzada, F. (2013), “The complementary exponential power series distribution”, *Brazilian Journal of Probability and Statistics*, 27(4), 565 – 584.

[11] Freund, J. E. (1961), “A bivariate extension of the exponential distribution”, *Journal of the American Statistical Association*, 56(296), 971 – 977.

[12] Ghitany, M. E., Al-Hussaini, E. K., & Al-Jarallah, R. A. (2005), “Marshall Olkin extended Weibull distribution and its application to censored data”, *Journal of Applied Statistics*, 32(10), 1025 – 1034.

[13] Ghitany, M. E., Al-Awadhi, F. A., & Alkhalfan, L. A. (2007), “Marshall Olkin extended Lomax distribution and its application to censored data”, *Communications in Statistics-Theory and Methods*, 36(10), 1855 – 1866.

[14] Gumbel, E. J. (1960), “Bivariate exponential distributions”, *Journal of the American Statistical Association*, 55(292), 698 – 707.

[15] Gupta, R. D., & Kundu, D. (1999), "Generalized exponential distributions", *Australian and New Zealand Journal of Statistics*, 41(2), 173 – 188.

[16] Harandi, S. S. and Alamatsaz, M. H. (2016), "Generalized linear failure rate power series distribution", *Communications in Statistics - Theory and Methods*, 45, 2204 – 2227.

[17] Johnson, N.L., Kemp, A.W. and Kotz, S. (2005), *Univariate discrete distribution*, 3-rd edition, John Wiley and Sons, New Jersey.

[18] Johnson, R. A. and Wichern, D. W. (1999), *Applied multivariate analysis*, 4-th ed., Prentice Hall, New Jersey.

[19] Justel, A., Pena, D. and Zamar, R. (1997), "A multivariate Kolmogorov-Smirnov test of goodness of fit", *Statistics and Probability Letters*, vol. 35, 251 - 291.

[20] Kundu, D. & Dey, A. K. (2009), "Estimating the parameters of the Marshall Olkin bivariate Weibull distribution by EM algorithm", *Computational Statistics & Data Analysis*, 53(4), 956 – 965.

[21] Kundu, D., & Gupta, R. D. (2009), "Bivariate generalized exponential distribution", *Journal of Multivariate Analysis*, 100(4), 581 – 593.

[22] Kundu, D., & Gupta, R. D. (2010), "Modified Sarhan Balakrishnan singular bivariate distribution", *Journal of Statistical Planning and Inference*, 140(2), 526-538.

[23] Kundu, D., & Gupta, R. D. (2011), "Absolute continuous bivariate generalized exponential distribution", *AStA Advances in Statistical Analysis*, 95(2), 169 – 185.

[24] Kundu, D. (2015), "Bivariate Geometric (Maximum) Generalized Exponential Distribution", *Journal of Data Science*, 13, 693 – 712.

[25] Kundu, D., & Gupta, A. K. (2014), "On Bivariate Weibull-Geometric Distribution", *Journal of Multivariate Analysis*, 123, 19 – 29.

[26] Kundu, D., Kumar, A. & Gupta, A.K. (2015), "Absolute continuous multivariate generalized exponential distribution. *Sankhya, Ser. B*, 77(2), 175 - 206.

[27] Kus, C. (2007), "A new lifetime distribution", *Computational Statistics & Data Analysis*, 51(9), 4497 – 4509.

[28] Louzada, F., Marchi, V., & Roman, M. (2014), "The exponentiated exponential geometric distribution: a distribution with decreasing, increasing and unimodal failure rate", *Statistics*, 48(1), 167 – 181.

[29] Mahmoudi, E. & Jafari, A. A. (2012), "Generalized exponential power series distributions", *Computational Statistics and Data Analysis*, 56(12), 4047 – 4066.

[30] Mahmoudi, E. & Jafari, A. A. (2015), "The compound class of linear failure rate-power series distributions: Model, properties and applications", *Communications in Statistics-Simulation and Computation*, in press.

[31] Marshall, A. W., & Olkin, I. (1967), "A multivariate exponential distribution", *Journal of the American Statistical Association*, 62(317), 30 – 44.

[32] Marshall, A. W., & Olkin, I. (1997), "A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families", *Biometrika*, 84(3), 641 – 652.

[33] Meintanis, S. G. (2007), "Test of fit for Marshall Olkin distributions with applications", *Journal of Statistical Planning and Inference*, 137(12), 3954 – 3963.

[34] Mohsin, M., Kazianka, H., Pilz, J., & Gebhardt, A. (2014), "A new bivariate exponential distribution for modeling moderately negative dependence", *Statistical Methods and Applications*, 23(1), 123 – 148.

[35] Morais, A. L. & Barreto-Souza, W. (2011), "A compound class of Weibull and power series distributions", *Computational Statistics and Data Analysis*, 55(3), 1410 – 1425.

[36] Nelsen, R. B. (2006), *An Introduction to Copulas*, second edition. Springer Verlag, New York.

[37] Noack, A. (1950), "A class of random variables with discrete distributions", *Annals of Mathematical Statistics*, 127 – 132.

[38] Pham, H., & Lai, C. D. (2007), "On recent generalizations of the Weibull distribution", *IEEE Transactions on Reliability*, 56(3), 454 – 458.

[39] Ristic, N.M. and Kundu, D. (2015), "Marshall Olkin generalized exponential distribution", *METRON*, 73, 317 – 333.

[40] Ristic, M. M., & Kundu, D. (2016), "Generalized Exponential Geometric Extreme Distribution", *Journal of Statistical Theory and Practice*, 10, 179 – 201.

[41] Sarhan, A. M., & Balakrishnan, N. (2007), "A new class of bivariate distributions and its mixture", *Journal of Multivariate Analysis*, 98(7), 1508 – 1527.

[42] Shoaee, S., & Khorram, E. (2012), "A new absolute continuous bivariate generalized exponential distribution", *Journal of Statistical Planning and Inference*, 142(7), 2203 – 2220.

[43] Song, P. X.-K., Fan, Y. & Kalbfleish, J.D. (2005), "Maximization by parts in likelihood inference" (with discussions), *Journal of the American Statistical Association*, 100, 1145 – 1167.

[44] Tahmasbi, R., & Rezaei, S. (2008), "A two-parameter lifetime distribution with decreasing failure rate", *Computational Statistics & Data Analysis*, 52(8), 3889 – 3901.