

ON BIVARIATE GENERALIZED EXPONENTIAL-POWER SERIES CLASS OF DISTRIBUTIONS

Ali Akbar Jafari[†], Rasool Roozegar^{†*} and Debasis Kundu^{††}

[†]Department of Statistics, Yazd University, P.O. Box 89195-741, Yazd, Iran

^{††}Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India

Abstract

In this paper, we introduce a new class of bivariate distributions by compounding the bivariate generalized exponential and power-series distributions. This new class contains the bivariate generalized exponential-Poisson, bivariate generalized exponential-logarithmic, bivariate generalized exponential-binomial and bivariate generalized exponential-negative binomial distributions as special cases. We derive different properties of the proposed class of distributions. It is observed that the proposed class of bivariate distributions is a very flexible class of distribution functions. The joint probability density functions can have variety of shapes. It can be bimodal as well as heavy tail also. This distribution has five parameters. The maximum likelihood estimators of the parameters cannot be obtained in closed form. We propose to use EM algorithm to compute the maximum likelihood estimators of the unknown parameters. It is observed that the proposed EM algorithm can be implemented very easily in practice. One data set has been analyzed for illustrative purposes. It is observed that the proposed model and the EM algorithm work quite well in practice.

Keywords: Generalized exponential; Power series distributions; Bivariate distribution; Maximum likelihood estimator; EM algorithm.

2010 AMS Subject Classification: 62E15, 62H10.

*Corresponding author: roozegar@yazd.ac.ir

1 INTRODUCTION

Marshall and Olkin (1997) proposed a method to introduce an extra parameter to a family of distributions to bring more flexibility. The model has an interesting physical interpretation also. The basic idea is very simple. They have considered geometric maximum or minimum of a sequence of independent and identically distributed (i.i.d.) random variables. They have explored several properties in case of exponential and Weibull distributions. Since then extensive work has been done on this method exploring different properties and extending to some other lifetime distributions, see for example Ghitany et al. (2005, 2007), Pham and Li (2007), Barreto-Souza (2012) and the references cited therein. In the same paper Marshall and Olkin (1997) briefly introduced the bivariate extension of their model. They did not discuss any detailed properties, but indicated that it will be a powerful tool to incorporate in a bivariate class of distributions. Although, an extensive amount of work has been done in the univariate set up, not much attention has been paid for the bivariate class of distributions. This is an attempt towards that direction.

In the last few years an attempt has been going on to generalize the original idea of Marshall and Olkin (1997) to incorporate more flexibility to a class of univariate distributions. Instead of using the geometric distribution, as it has been proposed by Marshall and Olkin (1997), a more general power series distribution has been used. Hence, the model of Marshall and Olkin (1997) can be obtained as a special case. Kus (2007) first proposed the exponential power series distribution, Tahmasbi and Rezaei (2008) introduced the exponential logarithmic distribution. Chahkandi and Ganjali (2009) first combined all these procedures and introduced the exponential power series distribution. Later, Morais and Barreto-Souza (2011) and Mahmoudi and Jafari (2012) proposed Weibull-power series and generalized exponential-power series (GEPS) distributions, respectively. It is observed that the GEPS class of distributions is a three-parameter distribution with a very flexible probability density functions (PDFs) and hazard functions. The hazard function of the GEPS distribution can be an increasing, decreasing or bathtub shaped. The PDFs can be unimodal and bimodal also. It may be mentioned that not too many three parameter distributions have these properties.

Extensive work has been in developing several continuous bivariate distributions and analyzing their properties, see for example Balakrishnan and Lai (2009) for a comprehensive review on this topic. Among several bivariate distributions with continuous marginals and which can be used to model singular data, Marshall-Olkin's (1967) bivariate exponential (MOBE) distribution plays the most important role. The MOBE has exponential marginals, hence it has decreasing PDF and constant hazard function. Due to these limitations, it is difficult to use it in many real life data sets. Marshall and Olkin (1967) also introduced Marshall-Olkin bivariate Weibull (MOBW) distribution with Weibull marginals. The MOBW has four parameters, and has some interesting properties, see for example Kundu and Dey (2009) or Kundu and Gupta (2013). Some of the other bivariate distributions with continuous marginals and which can be used to analyze bivariate data with ties have been proposed by Sarhan and Balakrishnan (2007), Kundu and Gupta (2009, 2010, 2014), Kundu (2015) and see the references cited therein. Unfortunately none of these models can have bimodal joint PDFs or heavy tails of marginals.

The main aim of this paper is to introduce a bivariate distribution with continuous marginals and having a singular component which can be used to analyze data with ties. The main advantage of the proposed bivariate distribution is that it can have marginals with heavy tails, moreover the joint PDF of the absolute continuous part can be bimodal also. The proposed model has some interesting physical interpretations also. Hence, it may be more flexible than the existing models and it will give the practitioner one more option to choose a model among the possible class of bivariate models for analyzing data with ties.

The basic idea of the proposed model is quite simple. Let us recall that the two-parameter generalized exponential (GE) distribution has been introduced by Gupta and Kundu (1999) and the PDF of a GE distribution can be a decreasing or an unimodal function. The hazard function of a GE distribution can be increasing, decreasing or constant. Mahmoudi and Jafari (2012) introduced the GEPS distributions by compounding GE distribution with the power series class of distributions. The proposed class of GEPS distributions is very flexible. It can have bimodal PDF and it can be heavy tails also. This motivates us to compound the bivariate generalized exponential (BGE) distribu-

tion, proposed by Kundu and Gupta (2009), with the power series class of distributions. We call the new class of distributions as the BGE power series (BGEPS) class of distributions.

The proposed class of bivariate distributions has several interesting properties. It is a five-parameter bivariate distribution with a singular component similarly as the MOBE or MOBW distribution. The joint CDF and the joint PDF can be written in explicit form, hence it can be used quite conveniently in practice. The joint PDF of the absolute continuous part of the BGEPS distribution can take various shapes. It can be unimodal, bimodal and heavy tails. The marginals have the GEPS distribution, hence it enjoys all the nice properties of the GEPS distribution as has been observed by Mahmoudi and Jafari (2012). The hazard functions of the marginals can be increasing, decreasing, constant and bathtub shaped. It is further observed that the proposed BGEPS has some interesting physical interpretations, and the BGE of Kundu and Gupta (2009) can be obtained as a special case.

First we derive the BGEPS by compounding the BGE distribution with the power series class of distributions. It may be mentioned that the BGE distribution has several interesting properties. We obtain several properties and also provide the marginal and conditional distributions of the BGEPS distribution. It is observed that it has a very convenient copula structure. Hence this can be used in deriving various dependence structure and evaluating different dependence measures. We discuss some of the special cases in details. The proposed class has five unknown parameters. The maximum likelihood estimators (MLEs) cannot be obtained in closed form. The MLEs can be obtained by solving five non-linear equations simultaneously. The standard Newton-Raphson algorithm may be used for this purpose, but it requires a very good choice of the initial guesses of the five parameters. Otherwise, it has the standard problem of converging to local maximum rather the global maximum. To avoid that, we have used an EM algorithm which involves solving at most two one-dimensional non-linear equations depending on the candidate of the power family of distributions. It is observed that the implementation of the proposed EM algorithm is very simple in practice. We have analyzed one data set for illustrative purposes. The performance is quite satisfactory.

The rest of the paper is organized as follows. In Section 2, we briefly describe the

BGE distribution and mention some of the recent developments on this distribution. The BGEPS is defined in Section 3, and we provide some physical interpretations and describe several properties in the same section. In Section 4, we have discussed some of the special cases. The inference procedure is provided in Section 5. The analysis of a data set is provided in Section 6, and finally we conclude the paper in Section 7.

2 BGE DISTRIBUTIONS

In this section we briefly describe about the BGE distributions. The two-parameter GE distribution with the shape parameter $\alpha > 0$ and the scale parameter $\lambda > 0$ has the following cumulative distribution function (CDF) and the PDF, respectively

$$\begin{aligned} F_{GE}(x; \alpha, \lambda) &= (1 - e^{-\lambda x})^\alpha, \quad x > 0, \quad \text{and} \\ f_{GE}(x; \alpha, \lambda) &= \alpha \lambda e^{-\lambda x} (1 - e^{-\lambda x})^{\alpha-1}, \quad x > 0. \end{aligned}$$

Kundu and Gupta (2009) introduced four-parameter BGE distribution based on three independent GE distributions. The CDF of the bivariate generalized exponential (BGE) with parameters $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_3 > 0$ and $\lambda > 0$ is given by

$$.F_{BGE}(x_1, x_2; \alpha_1, \alpha_2, \alpha_3, \lambda) = \begin{cases} (1 - e^{-\lambda x_1})^{\alpha_1 + \alpha_3} (1 - e^{-\lambda x_2})^{\alpha_2} & \text{if } x_1 \leq x_2 \\ (1 - e^{-\lambda x_1})^{\alpha_1} (1 - e^{-\lambda x_2})^{\alpha_2 + \alpha_3} & \text{if } x_1 > x_2. \end{cases} \quad (2.1)$$

The joint PDF of a BGE associated with the above CDF can be written as follows:

$$f_{BGE}(x_1, x_2; \alpha_1, \alpha_2, \alpha_3, \lambda) = \begin{cases} f_{GE}(x_1; \alpha_1 + \alpha_3, \lambda) f_{GE}(x_2; \alpha_2, \lambda) & \text{if } x_1 < x_2 \\ f_{GE}(x_1; \alpha_1, \lambda) f_{GE}(x_2; \alpha_2 + \alpha_3, \lambda) & \text{if } x_1 > x_2 \\ \frac{\alpha_3}{\alpha_1 + \alpha_2 + \alpha_3} f_{GE}(x_1; \alpha_1 + \alpha_2 + \alpha_3, \lambda) & \text{if } x_1 = x_2. \end{cases}$$

Here the first two terms are the densities with respect to two dimensional Lebesgue measure and the third term is the density with respect to one dimensional Lebesgue measure, see for example Kundu and Gupta (2009). Note that the BGE distribution has both an absolutely continuous part and a singular part similar to the MOBE, MOBW or the model proposed by Sarhan and Balakrishnan (2007). The BGE has several interesting properties. The marginals of the BGE distributions are GE distributions, and the conditional distributions are also GE distributions in a certain range. Moreover, it also has the positive quadrant dependence property.

An extensive amount of work has been done on the BGE distributions and some related topics in the last few years. Dey and Kundu (2012) discussed about discriminating between a bivariate generalized exponential distribution and a bivariate Weibull distribution based on the likelihood ratio statistics. Recently, Al-Turk et al. (2017) discussed the inference procedures of the unknown parameters of a BGE distribution based on copula functions. Some of the other work related to BGE distribution can be found in Achcar et al. (2015), Ibrahim et al. (2017), Kundu and Gupta (2011), Kundu et al. (2015), Muhammad (2016), Shoaee and Khorram (2012) and the see the references cited therein.

3 BGEPS CLASS OF DISTRIBUTIONS

A random variable N follows the power series distribution if it has the following probability mass function

$$P(N = n) = \frac{a_n \theta^n}{C(\theta)}, \quad n = 1, 2, \dots,$$

where $a_n \geq 0$ depends only on n , $C(\theta) = \sum_{n=1}^{\infty} a_n \theta^n$ and $\theta \in (0, s)$ (s can be ∞) is such that $C(\theta)$ is finite. The power series distributions contains some truncated (at zero) distributions such as geometric, Poisson, logarithmic, binomial and negative binomial. The quantity a_n for these distributions are 1 , $n!^{-1}$, n^{-1} , $\binom{k}{n}$ and $\binom{n-1}{k-1}$, respectively. Also, $C(\theta)$ for these distributions are $\theta(1 - \theta)^{-1}$, $e^\theta - 1$, $-\log(1 - \theta)$, $(1 + \theta)^k - 1$ and $\binom{n-1}{k-1}$, respectively. Detailed properties of the power series distribution can be found in Noack (1950).

Now, suppose $\{(X_{1n}, X_{2n}); n = 1, 2, \dots\}$ is a sequence of independent and identically distributed (i.i.d.) non-negative bivariate random variables with common joint distribution function $F_{\mathbf{X}}(\cdot, \cdot)$, where $\mathbf{X} = (X_1, X_2)$. Take N to be a power series random variable independent of (X_{1i}, X_{2i}) . Let

$$Y_i = \max\{X_{i1}, \dots, X_{iN}\}, \quad i = 1, 2.$$

Note that for the random vector $\mathbf{Y} = (Y_1, Y_2)$ we have

$$P(Y_1 \leq y_1, Y_2 \leq y_2, N = n) = F_{Y_1, Y_2 | N}(y_1, y_2 | n) P(N = n) = (F_{\mathbf{X}}(y_1, y_2))^n \frac{a_n \theta^n}{C(\theta)}. \quad (3.1)$$

Therefore, the joint CDF of $\mathbf{Y} = (Y_1, Y_2)$ becomes

$$F_{\mathbf{Y}}(y_1, y_2) = \sum_{n=1}^{\infty} (F_{\mathbf{X}}(y_1, y_2))^n \frac{a_n \theta^n}{C(\theta)} = \frac{C(\theta F_{\mathbf{X}}(y_1, y_2))}{C(\theta)}.$$

In this case, we call \mathbf{Y} has a bivariate F-power series (BFPS) distribution. The corresponding marginal distribution function of Y_i is

$$F_{Y_i}(y_i) = \frac{C(\theta F_{\mathbf{X}_i}(y_i))}{C(\theta)}, \quad i = 1, 2.$$

As we had already mentioned that in recent years many authors have considered this univariate class, see for example Mahmoudi and Jafari (2012), Flores et al. (2013) and the references cited therein.

Remark 2.1. If we consider $Z_i = \min \{X_{i1}, \dots, X_{iN}\}$, $i = 1, 2$, another class of bivariate distributions is obtained with the following joint cumulative survival function:

$$\bar{F}_{Z_1, Z_2}(y_1, y_2) = P(Z_1 > y_1, Z_2 > y_2) = \frac{C(\theta \bar{F}_{\mathbf{X}}(y_1, y_2))}{C(\theta)},$$

where $\bar{F}_{\mathbf{X}}(y_1, y_2) = P(X_1 > y_1, X_2 > y_2)$. Several papers have studied the univariate case of this class: see for example Morais and Barreto-Souza (2011) and Roozgar and Nadarajah (2016) among others.

In this paper we take F to be the bivariate generalized exponential given in (2.1). Therefore, we consider the bivariate generalized exponential-power series (BGEPS) class of distributions which is defined by the following CDF:

$$\begin{aligned} F_{\mathbf{Y}}(y_1, y_2) &= \begin{cases} \frac{C(\theta(1-e^{-\lambda y_1})^{\alpha_1+\alpha_3}(1-e^{-\lambda y_2})^{\alpha_2})}{C(\theta)} & \text{if } y_1 \leq y_2 \\ \frac{C(\theta(1-e^{-\lambda y_1})^{\alpha_1}(1-e^{-\lambda y_2})^{\alpha_2+\alpha_3})}{C(\theta)} & \text{if } y_1 > y_2, \end{cases} \\ &= \begin{cases} \frac{C(\theta F_{GE}(y_1; \alpha_1 + \alpha_3, \lambda) F_{GE}(y_2; \alpha_2, \lambda))}{C(\theta)} & \text{if } y_1 \leq y_2 \\ \frac{C(\theta F_{GE}(y_1; \alpha_1, \lambda) F_{GE}(y_2; \alpha_2 + \alpha_3, \lambda))}{C(\theta)} & \text{if } y_1 > y_2, \end{cases} \end{aligned} \quad (3.2)$$

We denote it by BGEPS($\alpha_1, \alpha_2, \alpha_3, \lambda, \theta$). The following interpretations can be provided of the above random variables.

RANDOM STRESS MODEL: Suppose, a system has two components. Each component is subject to random number of individual stresses. Suppose the i -th stress at the two components are X_{1i} and X_{2i} , respectively. If N is the number of stresses, which is not observable, then the observed stresses at the two components are $Y_1 = \max\{X_{11}, X_{12}, \dots, X_{1N}\}$

and $Y_2 = \max\{X_{21}, X_{22}, \dots, X_{2N}\}$, respectively. If N has a Power series distribution and (X_{1i}, X_{2i}) for $i = 1, 2, \dots$ are i.i.d. BGE random variables and they are independent of N , then (Y_1, Y_2) has a BGEPS distribution.

PARALLEL SYSTEMS: 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_{1i} and X_{2i} denote the lifetimes of the i -th components of system 1 and system 2, respectively, then the lifetime of the two systems become (Y_1, Y_2) , where $Y_1 = \max\{X_{11}, X_{12}, \dots, X_{1N}\}$ and $Y_2 = \max\{X_{21}, X_{22}, \dots, X_{2N}\}$. In this case also, if N has a Power series distribution and (X_{1i}, X_{2i}) for $i = 1, 2, \dots$ are i.i.d. BGE random variables and they are independent of N , then (Y_1, Y_2) has a BGEPS distribution.

Proposition 1. *Let $F_{\mathbf{Y}}(y_1, y_2)$ be the CDF of BGEPS distributions given in (3.2). Then*

$$F_{\mathbf{Y}}(y_1, y_2) = \sum_{n=1}^{\infty} p_n F_{BGE}(y_1, y_2; n\alpha_1, n\alpha_2, n\alpha_3, \lambda),$$

where $p_n = P(N = n) = \frac{a_n \theta^n}{C(\theta)}$.

Proof. It is straightforward by the total probability law and conditioning the $F_{\mathbf{Y}}(y_1, y_2)$ on the random variable N . □

We note that for the bivariate generalized exponential distribution in (1.2), the parameter α_3 is a very important parameter because when $\alpha_3 = 0$, the random variables are independent. By Proposition 1, we see this is not true for the bivariate generalized exponential power series distribution.

Proposition 2. *Let (Y_1, Y_2) follows BGEPS $(\alpha_1, \alpha_2, \alpha_3, \lambda, \theta)$ distribution. Then*

1. *Each Y_i has a GEPS distribution with parameters $\alpha_i + \alpha_3$, λ and θ .*
2. *The random variable $U = \max(Y_1, Y_2)$ has a GEPS distribution with parameters $\alpha_1 + \alpha_2 + \alpha_3$, λ and θ .*
3. *If $C(\theta) = \theta$, then \mathbf{Y} has a BGE distribution with parameters $\alpha_1, \alpha_2, \alpha_3$ and λ .*
4. *$P(Y_1 < Y_2) = \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_3}$.*

Proof. The proofs of (1) and (3) are immediate from (3.2). To prove (2), we consider the following:

$$F_U(u) = P(U \leq u) = P(\max(Y_1, Y_2) \leq u) = P(Y_1 \leq u, Y_2 \leq u)$$

$$\begin{aligned}
&= \sum_{n=1}^{\infty} P(Y_1 \leq u, Y_2 \leq u | N = n) P(N = n) \\
&= \sum_{n=1}^{\infty} (F_{X_1, X_2}(u, u))^n P(N = n) \\
&= \frac{C(\theta(1 - e^{-\lambda u})^{\alpha_1 + \alpha_2 + \alpha_3})}{C(\theta)},
\end{aligned}$$

which is a GEPS distributions with parameters $\alpha_1 + \alpha_2 + \alpha_3$, λ and θ . The proof of (4) can be obtained as follows:

$$\begin{aligned}
P(Y_1 < Y_2) &= \sum_{n=1}^{\infty} P(Y_1 < Y_2, N = n) \\
&= \sum_{n=1}^{\infty} \frac{a_n \theta^n}{C(\theta)} \int_0^{\infty} \int_{y_1}^{\infty} f_{1n}(y_1, y_2) dy_2 dy_1 \\
&= \sum_{n=1}^{\infty} \frac{a_n \theta^n}{C(\theta)} \times \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_3} \\
&= \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_3}.
\end{aligned}$$

□

Theorem 3.1. *Let \mathbf{Y} has a BGEPS $(\alpha_1, \alpha_2, \alpha_3, \lambda, \theta)$ distribution. Then the joint PDF of \mathbf{Y} is*

$$f_{\mathbf{Y}}(y_1, y_2) = \begin{cases} f_1(y_1, y_2) & \text{if } 0 < y_1 < y_2 \\ f_2(y_1, y_2) & \text{if } 0 < y_2 < y_1 \\ f_0(y) & \text{if } 0 < y_1 = y_2 = y, \end{cases}$$

where

$$\begin{aligned}
f_1(y_1, y_2) &= \frac{\theta}{C(\theta)} f_{GE}(y_1; \alpha_1 + \alpha_3, \lambda) f_{GE}(y_2; \alpha_2, \lambda) [\theta F_{GE}(y_1; \alpha_1 + \alpha_3, \lambda) \\
&\quad \times F_{GE}(y_2; \alpha_2, \lambda) C'''(\theta F_{GE}(y_1; \alpha_1 + \alpha_3, \lambda) F_{GE}(y_2; \alpha_2, \lambda)) \\
&\quad + C'(\theta F_{GE}(y_1; \alpha_1 + \alpha_3, \lambda) F_{GE}(y_2; \alpha_2, \lambda))], \tag{3.3}
\end{aligned}$$

$$\begin{aligned}
f_2(y_1, y_2) &= \frac{\theta}{C(\theta)} f_{GE}(y_1; \alpha_1, \lambda) f_{GE}(y_2; \alpha_2 + \alpha_3, \lambda) [\theta F_{GE}(y_1; \alpha_1, \lambda) \\
&\quad \times F_{GE}(y_2; \alpha_2 + \alpha_3, \lambda) C'''(\theta F_{GE}(y_1; \alpha_1, \lambda) F_{GE}(y_2; \alpha_2 + \alpha_3, \lambda)) \\
&\quad + C'(\theta F_{GE}(y_1; \alpha_1, \lambda) F_{GE}(y_2; \alpha_2 + \alpha_3, \lambda))], \tag{3.4}
\end{aligned}$$

$$f_0(y) = \frac{\theta \alpha_3}{C(\theta)(\alpha_1 + \alpha_2 + \alpha_3)} f_{GE}(y; \alpha_1 + \alpha_2 + \alpha_3, \lambda) C'(\theta F_{GE}(y; \alpha_1 + \alpha_2 + \alpha_3, \lambda)). \tag{3.5}$$

Proof. The proof can be obtained along the same line as the proof of Theorem 3.1 of Kundu and Gupta (2014), the details are avoided. \square

Just to show how different shapes the absolute continuous part of the joint PDF of BGEP can take, we have considered $C(\theta) = \theta + \theta^{20}$, $\lambda = 1$, the joint PDF are depicted in Figure 1. It is clear that the joint PDF can take various shapes, it can have heavy tails and it can be bimodal also.

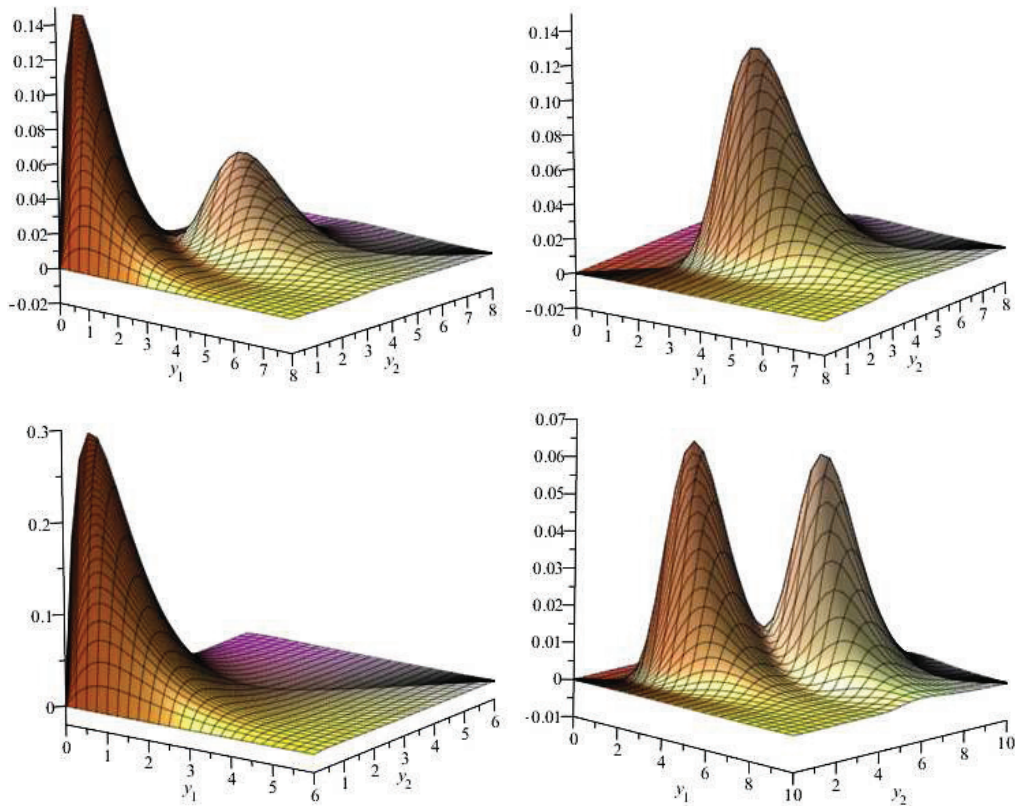


Figure 1: The PDF of the BGEPs class of distributions for some values of parameters: $\alpha_1 = \alpha_2 = \alpha_3 = 1$, $\theta = 1$ (left top), $\alpha_1 = \alpha_2 = \alpha_3 = 1$, $\theta = 2$ (right top), $\alpha_1 = \alpha_2 = \alpha_3 = 1$, $\theta = 0.5$ (left bottom), $\alpha_1 = \alpha_2 = \alpha_3 = 10$, $\theta = 1$ (right bottom).

Remark 2.2. Since $\theta C'''(\theta) + C'(\theta) = \sum_{n=1}^{\infty} n^2 a_n \theta^{n-1}$ and $C'(\theta) = \sum_{n=1}^{\infty} n a_n \theta^{n-1}$, we have

$$f_1(y_1, y_2) = \sum_{n=1}^{\infty} p_n f_{GE}(y_1; n\alpha_1 + n\alpha_3, \lambda) f_{GE}(y_2; n\alpha_2, \lambda),$$

$$f_2(y_1, y_2) = \sum_{n=1}^{\infty} p_n f_{GE}(y_1; n\alpha_1, \lambda) f_{GE}(y_2; n\alpha_2 + n\alpha_3, \lambda),$$

$$f_0(y) = \frac{\alpha_3}{\alpha_1 + \alpha_2 + \alpha_3} \sum_{n=1}^{\infty} p_n f_{GE}(y; n\alpha_1 + n\alpha_2 + n\alpha_3, \lambda),$$

where $p_n = P(N = n) = \frac{a_n \theta^n}{C(\theta)}$ and $f_{GE}(\cdot; n\alpha, \lambda)$ is the PDF of GE distribution with parameters $n\alpha$ and λ . Note that $f_{GE}(\cdot; n\alpha, \lambda)$ is the PDF of random variable $\max(U_1, \dots, U_n)$ where U_i 's are independent random variables from a GE distribution with parameters α and λ .

Corollary 3.1. *The joint PDF of the BGEPS distributions provided in Theorem 3.1 can be written as*

$$f_{\mathbf{Y}}(y_1, y_2) = \frac{\alpha_1 + \alpha_2}{\alpha_1 + \alpha_2 + \alpha_3} g_a(y_1, y_2) + \frac{\alpha_3}{\alpha_1 + \alpha_2 + \alpha_3} g_s(y),$$

where

$$g_a(y_1, y_2) = \frac{\alpha_1 + \alpha_2 + \alpha_3}{\alpha_1 + \alpha_2} \begin{cases} f_1(y_1, y_2) & \text{if } y_1 < y_2 \\ f_2(y_1, y_2) & \text{if } y_2 < y_1, \end{cases}$$

$$g_s(y) = \frac{\theta}{C(\theta)} f_{GE}(y; \alpha_1 + \alpha_2 + \alpha_3, \lambda) C'(\theta F_{GE}(y; \alpha_1 + \alpha_2 + \alpha_3, \lambda)) \quad \text{if } y_1 = y_2 = y,$$

and 0 otherwise. Clearly, $g_a(\cdot, \cdot)$ is the absolute continuous part and $g_s(\cdot)$ is the singular part. If $\alpha_3 = 0$, it does not have any singular part and it becomes an absolute continuous density function. Note that $g_s(\cdot)$ is the PDF of GEPS distribution with parameters $\alpha_1 + \alpha_2 + \alpha_3$ and λ .

Proposition 3. *The conditional distribution of Y_1 given $Y_2 \leq y_2$ is an absolute continuous distribution function with the following CDF:*

$$P(Y_1 \leq y_1 | Y_2 \leq y_2) = \begin{cases} \frac{C(\theta(1-e^{-\lambda y_1})^{\alpha_1 + \alpha_3} (1-e^{-\lambda y_2})^{\alpha_2})}{C(\theta(1-e^{-\lambda y_2})^{\alpha_2 + \alpha_3})} & \text{if } y_1 < y_2 \\ \frac{C(\theta(1-e^{-\lambda y_2})^{\alpha_2 + \alpha_3} (1-e^{-\lambda y_1})^{\alpha_1})}{C(\theta(1-e^{-\lambda y_2})^{\alpha_2 + \alpha_3})} & \text{if } y_2 < y_1. \end{cases}$$

Proposition 4. *The limiting distribution of BGEPS when $\theta \rightarrow 0^+$ is*

$$\begin{aligned} \lim_{\theta \rightarrow 0^+} F_{\mathbf{Y}}(y_1, y_2) &= \lim_{\theta \rightarrow 0^+} \frac{C(\theta F_{\mathbf{X}}(y_1, y_2))}{C(\theta)} = \lim_{\theta \rightarrow 0^+} \frac{\sum_{n=1}^{\infty} a_n \theta^n (F_{\mathbf{X}}(y_1, y_2))^n}{\sum_{n=1}^{\infty} a_n \theta^n} \\ &= \lim_{\theta \rightarrow 0^+} \frac{a_c (F_{\mathbf{X}}(y_1, y_2))^c + \sum_{n=c+1}^{\infty} a_n \theta^{n-c} (F_{\mathbf{X}}(y_1, y_2))^n}{a_c + \sum_{n=c+1}^{\infty} a_n \theta^{n-c}} \\ &= (F_{\mathbf{X}}(y_1, y_2))^c \\ &= \begin{cases} (1 - e^{-\lambda y_1})^{c(\alpha_1 + \alpha_3)} (1 - e^{-\lambda y_2})^{c\alpha_2} & \text{if } y_1 \leq y_2 \\ (1 - e^{-\lambda y_1})^{c\alpha_1} (1 - e^{-\lambda y_2})^{c(\alpha_2 + \alpha_3)} & \text{if } y_1 > y_2, \end{cases} \end{aligned}$$

which is the PDF of a BGE distribution with parameters $c\alpha_1$, $c\alpha_2$, $c\alpha_3$ and λ , where $c = \min\{n \in \mathbb{N} : a_n > 0\}$.

The following results will be useful for further development. For the joint random variables (Y_1, Y_2, N) , consider equation (3.1) when \mathbf{X} has a BGE distribution. Since $(Y_1, Y_2|N = n)$ has a BGE with parameters $n\alpha_1, n\alpha_2, n\alpha_3$, and λ , the joint PDF of (Y_1, Y_2, N) is

$$f_{Y_1, Y_2, N}(y_1, y_2, n) = \begin{cases} \frac{a_n \theta^n}{C(\theta)} f_{1n}(y_1, y_2) & \text{if } y_1 < y_2 \\ \frac{a_n \theta^n}{C(\theta)} f_{2n}(y_1, y_2) & \text{if } y_2 < y_1 \\ \frac{a_n \theta^n}{C(\theta)} f_{0n}(y) & \text{if } y_1 = y_2 = y, \end{cases}$$

where

$$\begin{aligned} f_{1n}(y_1, y_2) &= n^2 \lambda^2 (\alpha_1 + \alpha_3) \alpha_2 e^{-\lambda y_1 - \lambda y_2} (1 - e^{-\lambda y_1})^{n(\alpha_1 + \alpha_3) - 1} (1 - e^{-\lambda y_2})^{n\alpha_2 - 1}, \\ f_{2n}(y_1, y_2) &= n^2 \lambda^2 (\alpha_2 + \alpha_3) \alpha_1 e^{-\lambda y_1 - \lambda y_2} (1 - e^{-\lambda y_1})^{n\alpha_1 - 1} (1 - e^{-\lambda y_2})^{n(\alpha_2 + \alpha_3) - 1}, \\ f_{0n}(y) &= n \lambda \alpha_3 e^{-\lambda y} (1 - e^{-\lambda y})^{n(\alpha_1 + \alpha_2 + \alpha_3) - 1}. \end{aligned}$$

Theorem 3.2. *The conditional expectation of N given $Y_1 = y_1$ and $Y_2 = y_2$ is*

$$E(N|y_1, y_2) = \begin{cases} \frac{R_1(y_1, y_2)}{h_1(y_1, y_2)} & \text{if } y_1 < y_2 \\ \frac{R_2(y_1, y_2)}{h_2(y_1, y_2)} & \text{if } y_2 > y_1 \\ \frac{h_2(y_1, y_2)}{\theta K_0(y) C''(\theta K_0(y)) + C'(\theta K_0(y))} & \text{if } y_1 = y_2 = y, \end{cases}$$

where

$$\begin{aligned} R_i(y_1, y_2) &= (\theta K_i(y_1, y_2))^2 C''''(\theta K_i(y_1, y_2)) + 3\theta K_i(y_1, y_2) C'''(\theta K_i(y_1, y_2)) \\ &\quad + C''(\theta K_i(y_1, y_2)), \quad i = 1, 2. \end{aligned}$$

Proof. The conditional probability mass function of N given $Y_1 = y_1$ and $Y_2 = y_2$ is

$$f_{N|Y_1, Y_2}(n|y_1, y_2) = \begin{cases} \frac{n^2 a_n (\theta K_1(y_1, y_2))^{n-1}}{h_1(y_1, y_2)} & \text{if } y_1 < y_2 \\ \frac{n^2 a_n (\theta K_2(y_1, y_2))^{n-1}}{h_2(y_1, y_2)} & \text{if } y_2 < y_1 \\ \frac{n a_n (\theta K_0(y))^{n-1}}{h_0(y)} & \text{if } y_1 = y_2 = y, \end{cases}$$

where

$$\begin{aligned} h_1(y_1, y_2) &= \theta F_{GE}(y_1; \alpha_1 + \alpha_3, \lambda) F_{GE}(y_2; \alpha_2, \lambda) C'''(\theta F_{GE}(y_1; \alpha_1 + \alpha_3, \lambda) F_{GE}(y_2; \alpha_2, \lambda)) \\ &\quad + C''(\theta F_{GE}(y_1; \alpha_1 + \alpha_3, \lambda) F_{GE}(y_2; \alpha_2, \lambda)), \end{aligned}$$

$$\begin{aligned} h_2(y_1, y_2) &= \theta F_{GE}(y_1; \alpha_1, \lambda) F_{GE}(y_2; \alpha_2 + \alpha_3, \lambda) C'''(\theta F_{GE}(y_1; \alpha_1, \lambda) F_{GE}(y_2; \alpha_2 + \alpha_3, \lambda)) \\ &\quad + C''(\theta F_{GE}(y_1; \alpha_1, \lambda) F_{GE}(y_2; \alpha_2 + \alpha_3, \lambda)), \end{aligned}$$

$$h_0(y) = C'(\theta F_{GE}(y; \alpha_1 + \alpha_2 + \alpha_3, \lambda)),$$

and

$$K_1(y_1, y_2) = (1 - e^{-\lambda y_1})^{\alpha_1 + \alpha_3} (1 - e^{-\lambda y_2})^{\alpha_2} = F_{GE}(y_1; \alpha_1 + \alpha_3, \lambda) F_{GE}(y_2; \alpha_2, \lambda),$$

$$K_2(y_1, y_2) = (1 - e^{-\lambda y_1})^{\alpha_1} (1 - e^{-\lambda y_2})^{\alpha_2 + \alpha_3} = F_{GE}(y_1; \alpha_1, \lambda) F_{GE}(y_2; \alpha_2 + \alpha_3, \lambda),$$

$$K_0(y) = (1 - e^{-\lambda y})^{\alpha_1 + \alpha_2 + \alpha_3} = F_{GE}(y; \alpha_1 + \alpha_2 + \alpha_3, \lambda).$$

Since $\theta^2 C'''(\theta) + 3\theta C''(\theta) + C'(\theta) = \sum_{n=1}^{\infty} n^3 a_n \theta^{n-1}$, $\theta C'''(\theta) + C'(\theta) = \sum_{n=1}^{\infty} n^2 a_n \theta^{n-1}$ and $C'(\theta) = \sum_{n=1}^{\infty} n a_n \theta^{n-1}$, therefore, we can obtain the conditional expectation of N given $Y_1 = y_1$ and $Y_2 = y_2$ as

$$E(N|y_1, y_2) = \begin{cases} \frac{R_1(y_1, y_2)}{h_1(y_1, y_2)} & \text{if } y_1 < y_2 \\ \frac{R_2(y_1, y_2)}{h_2(y_1, y_2)} & \text{if } y_2 > y_1 \\ \frac{\theta K_0(y) C''(\theta K_0(y)) + C'(\theta K_0(y))}{h_0(y)} & \text{if } y_1 = y_2 = y, \end{cases}$$

□

3.1 COPULA AND MEDIAN CORRELATION COEFFICIENT

Every bivariate distribution function, F_{X_1, X_2} , with continuous marginals F_{X_1} and F_{X_2} , corresponds to a unique function $Co: [0, 1]^2 \rightarrow [0, 1]$, called a copula such that

$$F_{X_1, X_2}(x_1, x_2) = Co\{F_{X_1}(x_1), F_{X_2}(x_2)\} \text{ for } (x_1, x_2) \in \mathbb{R}^2,$$

and

$$Co(u, v) = F_{X_1, X_2}(F_{X_1}^{-1}(u), F_{X_2}^{-1}(v)) \text{ for } 0 < u, v < 1.$$

The copulas provide a natural way to measure the dependence between two random variables. Now we obtain the copula for BGEPS class of distributions, and then for each submodels of this class. Since $C(\theta)$ is an increasing function, C^{-1} exists. Hence, BGEPS class of distributions has the following copula structure;

$$Co(u, v) = \frac{C(\theta \times F(u; \theta, \theta_1) \times G(v; \theta, \theta_2) \times H(u, v; \theta, \theta_1, \theta_2))}{C(\theta)},$$

where $\theta_i = \frac{\alpha_3}{\alpha_i + \alpha_3}$, $i = 1, 2$ and

$$F(u; \theta, \theta_1) = [C^{-1}(uC(\theta)) / \theta]^{1-\theta_1},$$

$$G(v; \theta, \theta_2) = [C^{-1}(vC(\theta)) / \theta]^{1-\theta_2},$$

$$H(u, v; \theta, \theta_1, \theta_2) = \min \left\{ [C^{-1}(uC(\theta)) / \theta]^{\theta_1} [C^{-1}(vC(\theta)) / \theta]^{\theta_2} \right\},$$

The respective copulas for BGEG, BGEP, BGEL, BGEB distributions are

$$Co(u, v) = \frac{(1-\theta) [u/(1+u\theta-\theta)]^{1-\theta_1} [v/(1+v\theta-\theta)]^{1-\theta_2} \times G}{1-\theta [u/(1+u\theta-\theta)]^{1-\theta_1} [v/(1+v\theta-\theta)]^{1-\theta_2} \times G},$$

$$Co(u, v) = \frac{\exp \left[\theta (\log(ue^\theta - u + 1)/\theta)^{1-\theta_1} (\log(ve^\theta - v + 1)/\theta)^{1-\theta_2} \times P \right] - 1}{e^\theta - 1},$$

$$Co(u, v) = \frac{\log 1 - \theta ([1 - (1-\theta)^u] / \theta)^{1-\theta_1} ([1 - (1-\theta)^v] / \theta)^{1-\theta_2} \times L}{\log 1 - \theta},$$

$$Co(u, v) = \frac{\left(\theta \left[\left((u(\theta+1)^m - u - 1)^{1/m} - 1 \right) / \theta \right]^{1-\theta_1} \times B_1 \times B_2 + 1 \right)^m - 1}{(\theta+1)^m - 1},$$

where

$$G = \min \left\{ [u/(1+u\theta-\theta)]^{\theta_1}, [v/(1+v\theta-\theta)]^{\theta_2} \right\},$$

$$P = \min \left\{ (\log(ue^\theta - u + 1)/\theta)^{\theta_1}, (\log(ve^\theta - v + 1)/\theta)^{\theta_2} \right\},$$

$$L = \min \left\{ ([1 - (1-\theta)^u] / \theta)^{\theta_1}, ([1 - (1-\theta)^v] / \theta)^{\theta_2} \right\},$$

$$B_1 = \left[\left((v(\theta+1)^m - v - 1)^{1/m} - 1 \right) / \theta \right]^{1-\theta_2}, B_2 =$$

$$\min \left\{ \left[\left((u(\theta+1)^m - u - 1)^{1/m} - 1 \right) / \theta \right]^{\theta_1}, \left[\left((v(\theta+1)^m - v - 1)^{1/m} - 1 \right) / \theta \right]^{\theta_2} \right\}.$$

Remark 2.3. When $C(\theta) = \theta$, we have

$$Co(u, v) = u^{1-\theta_1} v^{1-\theta_2} \min\{u^{\theta_1}, v^{\theta_2}\},$$

where $\theta_i = \frac{\alpha_3}{\alpha_i + \alpha_3}$, $i = 1, 2$ which is the copula for BGE distribution. In the following, we compute a measure of dependency, namely the median correlation. The population version of the median correlation coefficient for a pair (Y_1, Y_2) of continuous random variables was defined by Blomqvist (1950). Suppose M_{Y_1} and M_{Y_2} denote the medians of Y_1 and Y_2 , respectively. The median correlation of Y_1 and Y_2 , say M_{Y_1, 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 (1999) and Domma (2010) that the median correlation coefficient is also a copula properly, and $M_{Y_1, Y_2} = 4Co\left(\frac{1}{2}, \frac{1}{2}\right) - 1$. Therefore, for each

submodels of BGEPS distributions the median correlation coefficient between Y_1 and Y_2 is written in Table 1.

Table 1: Median correlation coefficient for submodels of BGEPS distributions.

Distribution	Median correlation coefficient
BGEG	$M_{Y_1, Y_2} = \frac{4(1 - \theta)}{(2 - \theta)^{2 - \max\{\theta_1, \theta_2\}} - \theta} - 1$
BGEP	$M_{Y_1, Y_2} = \frac{4 \left[\exp \left\{ \theta^{\max\{\theta_1, \theta_2\} - 1} (\log(e^\theta + 1/2))^{2 - \max\{\theta_1, \theta_2\}} \right\} - 1 \right]}{e^\theta - 1} - 1$
BGEL	$M_{Y_1, Y_2} = \frac{4 \log \left(1 - \theta^{\max\{\theta_1, \theta_2\} - 1} [1 - \sqrt{1 - \theta}]^{2 - \max\{\theta_1, \theta_2\}} \right)}{\log(1 - \theta)} - 1$
BGEB	$M_{Y_1, Y_2} = \frac{4 [(\theta T + 1)^m - 1]}{(\theta + 1)^m - 1} - 1$

where

$$T = \left[\frac{1}{\theta} \left\{ ((\theta + 1)^m / 2 - 3/2)^{1/m} - 1 \right\} \right]^{2 - \max\{\theta_1, \theta_2\}}.$$

4 SPECIAL CASES

In this section, we consider some special cases of BGEPS distributions.

4.1 BIVARIATE GENERALIZED EXPONENTIAL-GEOMETRIC DISTRIBUTION

When $a_n = 1$ and $C(\theta) = \frac{\theta}{1 - \theta}$ ($0 < \theta < 1$), the power series distribution becomes the geometric distribution (truncated at zero). Therefore, the CDF of bivariate generalized exponential-geometric (BGEG) distribution is given by

$$F_{\mathbf{Y}}(y_1, y_2) = \begin{cases} \frac{(1 - \theta)(1 - e^{-\lambda y_1})^{\alpha_1 + \alpha_3} (1 - e^{-\lambda y_2})^{\alpha_2}}{1 - \theta(1 - e^{-\lambda y_1})^{\alpha_1 + \alpha_3} (1 - e^{-\lambda y_2})^{\alpha_2}} & \text{if } y_1 \leq y_2 \\ \frac{(1 - \theta)(1 - e^{-\lambda y_1})^{\alpha_1} (1 - e^{-\lambda y_2})^{\alpha_2 + \alpha_3}}{1 - \theta(1 - e^{-\lambda y_1})^{\alpha_1} (1 - e^{-\lambda y_2})^{\alpha_2 + \alpha_3}} & \text{if } y_1 > y_2. \end{cases}$$

Remark 3.1. When $\theta^* = 1 - \theta$, we have

$$F_{\mathbf{Y}}(y_1, y_2) = \begin{cases} \frac{\theta^*(1-e^{-\lambda y_1})^{\alpha_1+\alpha_3}(1-e^{-\lambda y_2})^{\alpha_2}}{1-(1-\theta^*)(1-e^{-\lambda y_1})^{\alpha_1+\alpha_3}(1-e^{-\lambda y_2})^{\alpha_2}} & \text{if } y_1 \leq y_2 \\ \frac{\theta^*(1-e^{-\lambda y_1})^{\alpha_1}(1-e^{-\lambda y_2})^{\alpha_2+\alpha_3}}{1-(1-\theta^*)(1-e^{-\lambda y_1})^{\alpha_1}(1-e^{-\lambda y_2})^{\alpha_2+\alpha_3}} & \text{if } y_1 > y_2. \end{cases}$$

It is also a CDF for all $\theta^* > 0$, see Marshall and Olkin (1997). In fact, this is in Marshall-Olkin bivariate class of distributions.

4.2 BIVARIATE GENERALIZED EXPONENTIAL-POISSON DISTRIBUTION

When $a_n = \frac{1}{n!}$ and $C(\theta) = e^\theta - 1$ ($\theta > 0$), the power series distribution becomes the Poisson distribution (truncated at zero). Therefore, the CDF of bivariate generalized exponential- Poisson (BGEP) distribution is given by

$$F_{\mathbf{Y}}(y_1, y_2) = \begin{cases} \frac{e^{\theta(1-e^{-\lambda y_1})^{\alpha_1+\alpha_3}}(1-e^{-\lambda y_2})^{\alpha_2} - 1}{e^\theta - 1} & \text{if } y_1 \leq y_2 \\ \frac{e^{\theta(1-e^{-\lambda y_1})^{\alpha_1}}(1-e^{-\lambda y_2})^{\alpha_2+\alpha_3} - 1}{e^\theta - 1} & \text{if } y_1 > y_2. \end{cases}$$

4.3 BIVARIATE GENERALIZED EXPONENTIAL-BINOMIAL DISTRIBUTION

When $a_n = \binom{k}{n}$ and $C(\theta) = (\theta + 1)^k - 1$ ($\theta > 0$), where $k(n \leq k)$ is the number of replicas, the power series distribution becomes the binomial distribution (truncated at zero). Therefore, the CDF of bivariate generalized exponential- binomial (BGEB) distribution is given by

$$F_{\mathbf{Y}}(y_1, y_2) = \begin{cases} \frac{\theta^k(1-e^{-\lambda y_1})^{k(\alpha_1+\alpha_3)}(1-e^{-\lambda y_2})^{k\alpha_2} - 1}{(\theta+1)^k - 1} & \text{if } y_1 \leq y_2 \\ \frac{\theta^k(1-e^{-\lambda y_1})^{k\alpha_1}(1-e^{-\lambda y_2})^{k(\alpha_2+\alpha_3)} - 1}{(\theta+1)^k - 1} & \text{if } y_1 > y_2. \end{cases}$$

4.4 BIVARIATE GENERALIZED EXPONENTIAL-LOGARITHMIC DISTRIBUTION

When $a_n = \frac{1}{n}$ and $C(\theta) = -\log(1 - \theta)$ ($0 < \theta < 1$), the power series distribution becomes the logarithmic distribution (truncated at zero). Therefore, the CDF of bivariate generalized exponential- logarithmic (BGEL) distribution is given by

$$F_{\mathbf{Y}}(y_1, y_2) = \begin{cases} \frac{\log(1-\theta(1-e^{-\lambda y_1})^{\alpha_1+\alpha_3}(1-e^{-\lambda y_2})^{\alpha_2})}{\log(1-\theta)} & \text{if } y_1 \leq y_2 \\ \frac{\log(1-\theta(1-e^{-\lambda y_1})^{\alpha_1}(1-e^{-\lambda y_2})^{\alpha_2+\alpha_3})}{\log(1-\theta)} & \text{if } y_1 > y_2. \end{cases}$$

4.5 BIVARIATE GENERALIZED EXPONENTIAL- NEGATIVE BINOMIAL DISTRIBUTION

When $a_n = \binom{n-1}{k-1}$ and $C(\theta) = (\frac{\theta}{1-\theta})^k$ ($0 < \theta < 1$), the power series distribution becomes the negative binomial distribution (truncated at zero). Therefore, the CDF of bivariate generalized exponential- negative binomial (BGENB) distribution is given by

$$F_{\mathbf{Y}}(y_1, y_2) = \begin{cases} \frac{(1-\theta)^k (1-e^{-\lambda y_1})^{k\alpha_1+k\alpha_3} (1-e^{-\lambda y_2})^{k\alpha_2}}{(1-\theta(1-e^{-\lambda y_1})^{\alpha_1+\alpha_3} (1-e^{-\lambda y_2})^{\alpha_2})^k} & \text{if } y_1 \leq y_2 \\ \frac{(1-\theta)^k (1-e^{-\lambda y_1})^{k\alpha_1} (1-e^{-\lambda y_2})^{k\alpha_2+k\alpha_3}}{(1-\theta(1-e^{-\lambda y_1})^{\alpha_1} (1-e^{-\lambda y_2})^{\alpha_2+\alpha_3})^k} & \text{if } y_1 > y_2. \end{cases}$$

5 ESTIMATION

In this section, we consider the estimation of the unknown parameters of the BGEPS distributions. Let $(y_{11}, y_{12}), \dots, (y_{m1}, y_{m2})$ be an observed sample with size m from a BGEPS distribution with parameters $\Theta = (\alpha_1, \alpha_2, \alpha_3, \lambda, \theta)$. We will use the following notations.

$$I_0 = \{i : y_{1i} = y_{2i} = y_i\}, \quad I_1 = \{i : y_{1i} < y_{2i}\}, \quad I_2 = \{i : y_{1i} > y_{2i}\},$$

and

$$m_0 = |I_0|, \quad m_1 = |I_1|, \quad m_2 = |I_2|, \quad m = m_0 + m_1 + m_2.$$

Therefore, the log-likelihood function can be written as

$$\ell(\Theta) = \sum_{i \in I_0} \log(f_0(y_i)) + \sum_{i \in I_1} \log(f_1(y_{1i}, y_{2i})) + \sum_{i \in I_2} \log(f_2(y_{1i}, y_{2i})), \quad (5.1)$$

where f_1 , f_2 and f_0 are given in (3.3), (3.4) and (3.5), respectively. We can obtain the MLEs of the parameters by maximizing $\ell(\Theta)$ in (5.1) with respect to the unknown parameters when we know the explicit form of $C(\theta)$. This is clearly a five-dimensional optimization problem, and they can be obtained by solving five non-linear equations simultaneously. Clearly, no explicit expressions are available for the MLEs. Alternatively, the maximization can be performed using a command like the *nlminb* routine in the R software (R Development Core Team, 2014). But both the methods need very good choice of the initial estimates. It is well known that without the proper choice of the initial guesses, often one finds the local maximum rather than the global maximum.

To avoid that problem we propose to use an expectation-maximization (EM) algorithm similar to Kundu and Dey (1990) to find the MLEs of parameters.

For given n , consider that independent random variables $\{Z_i|N = n\}$, $i = 1, 2, 3$ have the GE distribution with parameters $n\alpha_i$ and λ . It is well-known that

$$\{Y_1|N = n\} = \{\max(Z_1, Z_3)|N = n\}, \quad \{Y_2|N = n\} = \{\max(Z_2, Z_3)|N = n\}.$$

Assumed that for the bivariate random vector (Y_1, Y_2) , there is an associated random vectors

$$\Lambda_1 = \begin{cases} 0 & Y_1 = Z_1 \\ 1 & Y_1 = Z_2 \end{cases} \quad \text{and} \quad \Lambda_2 = \begin{cases} 0 & Y_2 = Z_1 \\ 1 & Y_2 = Z_3. \end{cases}$$

Note that if $Y_1 = Y_2$, then $\Lambda_1 = \Lambda_2 = 0$. But if $Y_1 < Y_2$ or $Y_1 > Y_2$, then (Λ_1, Λ_2) is missing. If $(Y_1, Y_2) \in I_1$ then the possible values of (Λ_1, Λ_2) are $(1, 0)$ or $(1, 1)$, and If $(Y_1, Y_2) \in I_2$ then the possible values of (Λ_1, Λ_2) are $(0, 1)$ or $(1, 1)$ with non-zero probabilities.

We form the conditional ‘pseudo’ log-likelihood function, conditioning on N , and then replace N by $E(N|Y_1, Y_2)$. In the E-step of the EM-algorithm, we treat it as complete observation when they belong to I_0 . If the observation belongs to I_1 , we form the ‘pseudo’ log-likelihood function by fractioning (y_1, y_2) to two partially complete ‘pseudo’ observations of the form $(y_1, y_2, u_1(\Theta))$ and $(y_1, y_2, u_2(\Theta))$, where $u_1(\Theta)$ and $u_2(\Theta)$ are the conditional probabilities that (Λ_1, Λ_2) takes values $(1, 0)$ and $(1, 1)$, respectively. Since

$$P(Z_3 < Z_1 < Z_2|N = n) = \frac{\alpha_1\alpha_2}{(\alpha_1 + \alpha_3)(\alpha_1 + \alpha_2 + \alpha_3)},$$

$$P(Z_1 < Z_3 < Z_2|N = n) = \frac{\alpha_2\alpha_3}{(\alpha_1 + \alpha_3)(\alpha_1 + \alpha_2 + \alpha_3)},$$

therefore,

$$u_1(\Theta) = \frac{\alpha_1}{\alpha_1 + \alpha_3}, \quad u_2(\Theta) = \frac{\alpha_3}{\alpha_1 + \alpha_3}.$$

Similarly, If the observation belong to I_2 , we form the ‘pseudo’ log-likelihood function of the from $(y_1, y_2, v_1(\Theta))$ and $(y_1, y_2, v_2(\Theta))$, where $v_1(\Theta)$ and $v_2(\Theta)$ are the conditional probabilities that (Λ_1, Λ_2) takes values $(0, 1)$ and $(1, 1)$, respectively. Therefore,

$$v_1(\Theta) = \frac{\alpha_2}{\alpha_2 + \alpha_3}, \quad v_2(\Theta) = \frac{\alpha_3}{\alpha_2 + \alpha_3}.$$

For brevity, we write $u_1(\Theta)$, $u_2(\Theta)$, $v_1(\Theta)$, $v_2(\Theta)$ as u_1 , u_2 , v_1 , v_2 , respectively.

E-step: Consider $b_i = E(N|y_{1i}, y_{2i}, \Theta)$. The log-likelihood function without the additive constant can be written as follows:

$$\begin{aligned}
\ell_{\text{pseudo}}(\Theta) &= \log(\theta) \sum_{i=1}^m b_i - m \log(C(\theta)) + (m_0 + 2m_1 + 2m_2) \log(\lambda) \\
&\quad + (m_1 u_1 + m_2) \log(\alpha_1) + (m_1 + m_2 v_1) \log(\alpha_2) \\
&\quad + (m_0 + m_1 u_2 + m_2 v_2) \log(\alpha_3) - \lambda \left(\sum_{i \in I_0} y_i + \sum_{i \in I_1 \cup I_2} (y_{1i} + y_{2i}) \right) \\
&\quad + \alpha_1 \left(\sum_{i \in I_0} b_i Q(y_i) + \sum_{i \in I_1 \cup I_2} b_i Q(y_{1i}) \right) + \alpha_2 \left(\sum_{i \in I_0} b_i Q(y_i) + \sum_{i \in I_1 \cup I_2} b_i Q(y_{2i}) \right) \\
&\quad + \alpha_3 \left(\sum_{i \in I_0} b_i Q(y_i) + \sum_{i \in I_1} b_i Q(y_{1i}) + \sum_{i \in I_2} b_i Q(y_{2i}) \right) \\
&\quad - \sum_{i \in I_0} Q(y_i) - \sum_{i \in I_1 \cup I_2} Q(y_{1i}) - \sum_{i \in I_1 \cup I_2} b_i Q(y_{2i}),
\end{aligned}$$

where $Q(y) = \log(1 - e^{-\lambda y})$.

M-step: At this step, $\ell_{\text{pseudo}}(\Theta)$ is maximized with respect to $\alpha_1, \alpha_2, \alpha_3, \lambda$ and θ . For fixed λ , the maximization occurs at

$$\hat{\alpha}_1(\lambda) = \frac{-(m_1 u_1 + m_2)}{\sum_{i \in I_0} b_i Q(y_i) + \sum_{i \in I_1 \cup I_2} b_i Q(y_{1i})}, \quad (5.2)$$

$$\hat{\alpha}_2(\lambda) = \frac{-(m_1 + m_2 v_1)}{\sum_{i \in I_0} b_i Q(y_i) + \sum_{i \in I_1 \cup I_2} b_i Q(y_{2i})}, \quad (5.3)$$

$$\hat{\alpha}_3(\lambda) = \frac{-(m_0 + m_1 u_2 + m_2 v_2)}{\sum_{i \in I_0} b_i Q(y_i) + \sum_{i \in I_1} b_i Q(y_{1i}) + \sum_{i \in I_2} b_i Q(y_{2i})}, \quad (5.4)$$

and solving the following non-linear equation with respect to θ :

$$\frac{\theta C'(\theta)}{C(\theta)} = \frac{\sum_{i=1}^m b_i}{m}. \quad (5.5)$$

Remark 4.1. When $C(\theta) = \frac{\theta}{1-\theta}$, then the solution of equation in (5.5) is $\hat{\theta} = 1 - \frac{m}{\sum_{i=1}^m b_i}$.

Remark 4.2. We do not need to solve the equation in (5.5), when $C(\theta) = \theta$. In fact, the BGEPS distribution reduces to the BGE distribution.

Finally, $\hat{\lambda}$ can be obtained as a solution of the following equation:

$$\frac{m_0 + 2m_1 + 2m_2}{g(\lambda)} = \lambda, \quad (5.6)$$

where

$$\begin{aligned}
g(\lambda) &= \sum_{i \in I_0} y_i + \sum_{i \in I_1 \cup I_2} (y_{1i} + y_{2i}) - \sum_{i \in I_0} (b_i \hat{\alpha}_1 + b_i \hat{\alpha}_2 + b_i \hat{\alpha}_3 - 1) \frac{y_i e^{-\lambda y_i}}{1 - e^{-\lambda y_i}} \\
&\quad - \sum_{i \in I_1} (b_i \hat{\alpha}_1 + b_i \hat{\alpha}_3 - 1) \frac{y_{1i} e^{-\lambda y_{1i}}}{1 - e^{-\lambda y_{1i}}} - \sum_{i \in I_1} (b_i \hat{\alpha}_2 - 1) \frac{y_{2i} e^{-\lambda y_{2i}}}{1 - e^{-\lambda y_{2i}}} \\
&\quad - \sum_{i \in I_2} (b_i \hat{\alpha}_1 - 1) \frac{y_{1i} e^{-\lambda y_{1i}}}{1 - e^{-\lambda y_{1i}}} - \sum_{i \in I_2} (b_i \hat{\alpha}_2 + b_i \hat{\alpha}_3 - 1) \frac{y_{2i} e^{-\lambda y_{2i}}}{1 - e^{-\lambda y_{2i}}}.
\end{aligned}$$

The following steps can be used to compute the MLEs of the parameters via the EM algorithm:

Step 1: Take some initial value of Θ , say $\Theta^{(0)} = (\alpha_1^{(0)}, \alpha_2^{(0)}, \alpha_3^{(0)}, \lambda^{(0)}, \theta^{(0)})'$.

Step 2: compute $b_i = E(N|y_{1i}, y_{2i}; \Theta^{(0)})$

Step 3: Compute u_1, u_2, v_1 , and v_2 .

Step 4: Find $\hat{\lambda}$ by solving the equation (5.6), say $\hat{\lambda}^{(1)}$.

Step 5: Compute $\hat{\alpha}_i^{(1)} = \hat{\alpha}_i(\hat{\lambda}^{(1)})$, $i = 1, 2, 3$ from (5.2)-(5.4).

Step 6: Find $\hat{\theta}$ by solving the equation (5.5), say $\hat{\theta}^{(1)}$.

Step 7: Replace $\Theta^{(0)}$ by $\Theta^{(1)} = (\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_3^{(1)}, \lambda^{(1)}, \theta^{(1)})'$, go back to step 1 and continue the process until convergence take place.

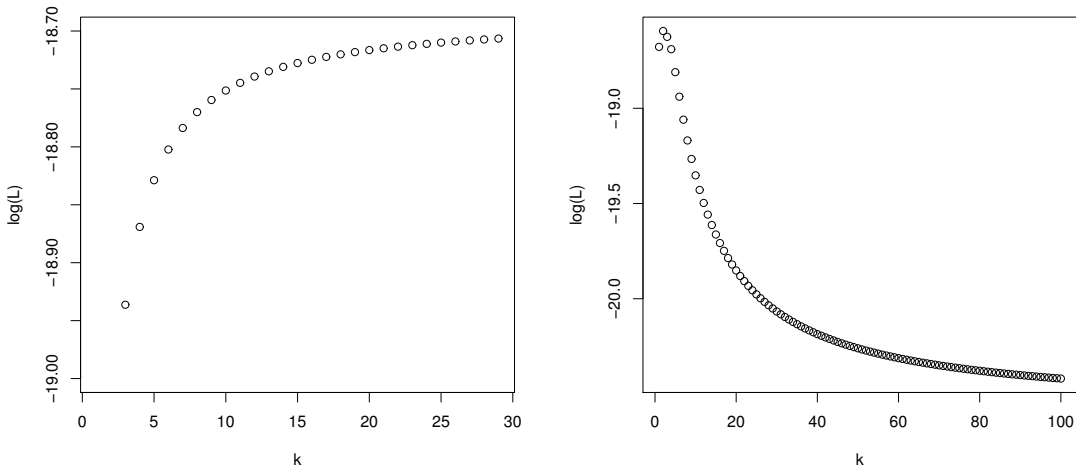


Figure 2: The log-likelihood values of BGEb (left) and BGENb (right) distributions for different values of k .

6 AN EXAMPLE

In this section we present the analysis of a data set to show how the model and the proposed EM algorithm work in practice. The data set is available in Meintanis (2007), and it is obtained from the group stage of the UEFA Champion's League for the years 2004-05 and 2005-2006. Several authors have analyzed this data set using different models; for example Kundu and Gupta (2009) and Kundu and Dey (2009) analyzed this data by using BGE and Marshall-Olkin bivariate Weibull distributions, respectively. The data represent the football (soccer) data where at least one goal scored by the home team and at least one goal scored directly from a kick goal (like penalty kick, foul kick or any other direct kick) by any team have been considered. Here Y_1 represents the time in minutes of the first kick goal scored by any team and Y_2 represents the first goal of any type scored by the home team. There are some ties in the data set, hence the proposed BGEPS may be used to analyze this data set.

We divided all the data point by 100 mainly for computational purpose. It is not going to affect the inference procedure. Then six special cases of BGEPS distributions are considered: BGE, BGE G , BGE P , BGE B , BGE NB , and BGE L . Using the proposed EM algorithm, these models are fitted to the bivariate data set, and the MLEs and their corresponding log-likelihood values are calculated. The standard errors (se) based on the observed information matrix are obtained. The results are given in Table 2.

For choosing the best value for k in BGE B distribution, we considered $k = 2, \dots, 30$, and obtained the corresponding log-likelihood values for all models. The results are given in Figure 2. It can be concluded that the log-likelihood values increase when k increases. But there is negligible variation for large k . Therefore, the results for $k = 30$ are given in Table 2. Similarly, for choosing the best value for k in BGE NB distribution, we consider $k = 1, \dots, 100$. The results are given in Figure 2. It can be concluded that the largest log-likelihood value occurs in the case of $k = 2$. Therefore, we present the results for $k = 2$ in Table 2.

One natural question is how good are these models to fit this data set. For each fitted model, the Akaike Information Criterion (AIC), the corrected Akaike information criterion (AICC) and the Bayesian information criterion (BIC) are calculated. We also

obtain the Kolmogorov-Smirnov (K-S) distances between the fitted distribution and the empirical distribution function and the corresponding p-values (in brackets) for Y_1 , Y_2 and $\max(Y_1, Y_2)$. The marginals and the maximum fit the univariate distributions quite well in all the cases considered. Now we want to test the copula based goodness of fit procedures as given in Genest et al. (2008). When dealing with the 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. Based on extensive simulations it has been observed, see Genest et al. (2008) for example, that the Cramer-von Mises distance between the empirical copula and the fitted copula can be used as a good discrepancy measure. Based on bootstrapping (1000 replications) Cramer-von mises statistics (CVS) and the associated p values for each case have been obtained and they are reported in 2. Therefore, based on the copula goodness fit all the models provide good fit to the data, but BGENB provides the best fit to the data set. Finally, we make use the likelihood ratio test (LRT) for testing the BGE against other models. The statistics and the corresponding p-values are given in Table 2. Since in each case the p value is not very large, it is recommended not to use BGE for this data set.

7 CONCLUSIONS

In this paper we have studied a new class of bivariate distributions namely bivariate generalized exponential power series class of distributions. This bivariate distributions has an absolute continuous and a singular part similar to the MOBE model. It is a five parameter model and the joint PDF of the absolute continuous part can take variety of shapes. The most important feature about the joint PDF is that it can be bimodal also. We have derived several properties of the proposed class of distributions and provided a very efficient EM algorithm to compute the MLEs of unknown parameters. In this paper we have only considered the bivariate model, but it should be possible to generalize to the multivariate case also. It will be interesting to develop different properties and the EM algorithm. More work is needed in that direction.

Acknowledgements

The authors would like to thank the Editor, Associate Editor and the anonymous referees for valuable comments and suggestions which greatly improved the paper. The first and second authors thank to the Yazd University for supporting this research.

References

- [1] Achcar, J.A., Moala, F.A., Tarumoto, M.H. and Coladello, L.F. (2015), “A bivariate generalized exponential distribution derived from copula functions in the presence of censored data and covariates”, *Pesquisa Operacional*, 35 (1), 165 – 186.
- [2] Al-Turk, L.I., Elaal, M.K.A. and Jarwan, R.S. (2017), “Inference of bivariate generalized exponential distribution based on copula functions”, *Applied Mathematical Sciences*, 11(24) 1155 – 1186.
- [3] Balakrishnan, N. and Lai, C-D. (2009), *Continuous Bivariate Distributions*, 2nd edition, Springer, New York.
- [4] Barreto-Souza, W. (2012), “Bivariate gamma-geometric law and its induce Levy process”, *Journal of Multivariate Analysis*, 109, 130 -145.
- [5] Blomqvist, N. (1950), “On a measure of dependence between two random variables”, *The Annals of Mathematical Statistics*, 21(4):593–600.
- [6] Cancho, V. G., Louzada-Neto, F., and Barriga, G. D. C. (2011), “The Poissonexponential lifetime distribution”, *Computational Statistics and Data Analysis*, 55(1):677–686.
- [7] Chahkandi, M. and Ganjali, M. (2009), “On some lifetime distributions with decreasing failure rate”, *Computational Statistics and Data Analysis*, 53(12):4433–4440.
- [8] Dey, A.K. and Kundu, D. (2012), “Discriminting between bivariate bivariate generalized exponential and bivariate Weibull distributions”, *Chilean Journal of Statistics*, 3(1), 93 – 110.

- [9] Domma, F. (2010), “Some properties of the bivariate Burr type III distribution”, *Statistics*, 44(2):203–215.
- [10] Elaal, M.K. and Jarwan, R.S. (2017), “Inference of bivaiate generalized exponential distribution based on copula functions”, *Applied Mathematical Sciences*, 11(24), 1155 – 1186.
- [11] Flores, J., Borges, P., Cancho, V. G., and Louzada, F. (2013), “The complementary exponential power series distribution”, *Brazilian Journal of Probability and Statistics*, 27(4):565–584.
- [12] Genest, C., Remillard, B. and Beaudoin, D. (2008), “Goodness of fit tests for copulas: A review and power study”, *Insurance: mathematics and Economics*, 42, 199 - 213.
- [13] Ghitany, M.E., Al-Hussaini, E.K. and Al-Jarallah, R.A. (2005), “Marshall-Olkin extended Weibull distribution and its application to censored data”, *Journal of Applied Statistics*, 32, 1025 - 1034.
- [14] Ghitany, M.E., Al-Awadhi, F.A. and Alkhalfan, L.A. (2007), “Marshall-Olkin extended Lomax distribution and its application to censored data”, *Communications in Statistics - Theory and Methods*, 36, 1855 - 1866
- [15] Gupta, R. D. and Kundu, D. (1999), “Generalized exponential distributions”, *Australian & New Zealand Journal of Statistics*, 41(2): 173–188.
- [16] Ibrahim, M., Eliwa, M.S. and El-Morshedy, M. (2017), “Bivariate exponentiated generalized linear exponential distribution with applications in reliability analysis”, arXiv preprint, arXiv:1710.00502.
- [17] Kundu, D. (2015), “Bivariate Geometric (Maximum) Generalized Exponential Distribution”, *Journal of Data Science*, 13(4):693-712.
- [18] Kundu, D. and Dey, A. K. (2009), “Estimating the parameters of the Marshall-Olkin bivariate Weibull distribution by EM algorithm”, *Computational Statistics and Data Analysis*, 53(4):956–965.

- [19] Kundu, D. and Gupta, R. D. (2009), “Bivariate generalized exponential distribution”, *Journal of Multivariate Analysis*, 100(4):581–593.
- [20] Kundu, D. and Gupta, R.D. (2010), “Modified Sarhan-Balakrishnan singular bivariate distribution”, *Journal of Statistical Planning and Inference*, 140, 526 – 538.
- [21] Kundu, D. and Gupta, R.D. (2011), “Absolute continuous bivariate generalized exponential distribution”, *Advances in Statistical Analysis*, 95, 169 – 185.
- [22] Kundu, D. and Gupta, A. K. (2013), “Bayes estimation for the Marshall-Olkin bivariate Weibull distribution”, *Computational Statistics and Data Analysis*, 57, 271 – 281.
- [23] Kundu, D. and Gupta, A.K. (2014), “On bivariate Weibull Geometric distribution”, *Journal of Multivariate Analysis*, 123, 19 – 29.
- [24] Kundu, D. and Gupta, R.D. (2011), “Absolute continuous bivariate generalized exponential distribution”, *Advances in Statistical Analysis*, 95, 169 – 185.
- [25] Kundu, D., Kumar, A. and Gupta, A.K. (2015), “Absolute continuous multivariate generalized exponential distribution”, *Sankhya, Ser. B*, 77(2), 175 – 206.
- [26] Kus, C. (2007), “A new lifetime distribution”, *Computational Statistics and Data Analysis*, 51, 4497 – 4509.
- [27] Louzada, F., Roman, M., and Cancho, V. G. (2011), “The complementary exponential geometric distribution: Model, properties, and a comparison with its counterpart”, *Computational Statistics and Data Analysis*, 55(8):2516–2524.
- [28] Mahmoudi, E. and Jafari, A. A. (2012), “Generalized exponential power series distributions”, *Computational Statistics and Data Analysis*, 56(12):4047–4066.
- [29] Marshall, A. W. and Olkin, I. (1967), “A multivariate exponential distribution”, *Journal of the American Statistical Association*, 62(317):30–44.
- [30] Marshall, A. W. and 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.

- [31] Meintanis, S. G. (2007), “Test of fit for Marshall-Olkin distributions with applications”, *Journal of Statistical Planning and Inference*, 137(12):3954–3963.
- [32] Morais, A. L. and Barreto-Souza, W. (2011), “A compound class of Weibull and power series distributions”, *Computational Statistics and Data Analysis*, 55(3):1410–1425.
- [33] Muhammad, M. (2016), “Poisson-odd generalized exponential family of distributions: theory and applications”, *Hacettepe Journal of Mathematics and Statistics*, 47, 1 – 20.
- [34] Nelsen, R. B. (1999), *An Introduction to Copulas*, Springer, New York.
- [35] Noack, A. (1950), “A class of random variables with discrete distributions”, *The Annals of Mathematical Statistics*, 21(1):127–132.
- [36] Pham, H. and Lai, C-D. (2007), “On recent generalizations of the Weibull distribution”, *IEEE Transactions on Reliability*, 56, 454 - 458.
- [37] R Development Core Team (2014), “A Language and Environment for Statistical Computing”, *R Foundation for Statistical Computing*, Vienna, Austria.
- [38] Roozegar, R. and Nadarajah, S. (2016), “Block The quadratic hazard rate power series distribution”, *Journal of Testing and Evaluation*, 45(3).
- [39] Sarhan, A. M. and Balakrishnan, N. (2007), “A new class of bivariate distributions and its mixture”, *Journal of Multivariate Analysis*, 98(7):1508–1527.
- [40] Shoaee, S. and Khorram, E. (2012), “A new absolute continuous bivariate generalized exponential distributions”, *Journal of Statistical Planning and Inference*, 142 (7), 2203 – 2220.
- [41] Tahmasbi, R. and Rezaei, S. (2008), “A two-parameter lifetime distribution with decreasing failure rate”, *Computational Statistics and Data Analysis*, 52, 3889 – 3901.

Table 2: The MLEs, log-likelihood, AIC, AICC, BIC, K-S, and LRT statistics for six sub-models of BGEPS distribution.

Statistic	Distribution					
	BGE	BGEG	BGEP	BGEB($k = 30$)	BGENB($k = 2$)	BGEL
$\hat{\alpha}_1$ (s.e.)	1.4452 (0.4160)	0.9964 (0.4938)	0.5644 (0.5758)	0.5980 (0.5442)	0.2538 (0.3340)	1.1871 (0.4140)
$\hat{\alpha}_2$ (s.e.)	0.4681 (0.1879)	0.3047 (0.1825)	0.1676 (0.1792)	0.1780 (0.1717)	0.0755 (0.1044)	0.3646 (0.1659)
$\hat{\alpha}_3$ (s.e.)	1.1704 (0.2866)	0.7205 (0.3811)	0.4009 (0.4053)	0.4264 (0.3843)	0.1792 (0.2443)	0.8615 (0.2929)
$\hat{\lambda}$ (s.e.)	3.8994 (0.5603)	5.0600 (0.8472)	4.6565 (0.6698)	4.6252 (0.6620)	4.8877 (0.7579)	5.3466 (1.0009)
$\hat{\theta}$ (s.e.)	— —	0.6932 (0.2318)	4.1980 (3.7768)	0.1484 (0.1351)	0.8212 (0.2739)	0.8844 (0.1250)
$\log(\ell)$	-20.5904	-18.6775	-18.6864	-18.7065	-18.5941	-18.4599
AIC	49.1809	47.3549	47.3728	47.4130	47.1882	46.9198
AICC	51.0856	49.2904	49.3083	49.3485	49.1237	48.8553
BIC	54.2133	55.4095	55.4274	55.4676	55.2428	54.9744
K-S (Y_1) (p-value)	0.1034 (0.8240)	0.0952 (0.8906)	0.0865 (0.9448)	0.0871 (0.9418)	0.0837 (0.9576)	0.1011 (0.8435)
K-S (Y_2) (p-value)	0.1001 (0.8527)	0.0900 (0.9255)	0.0957 (0.8870)	0.0961 (0.8844)	0.0920 (0.9132)	0.0843 (0.9551)
K-S ($\max(Y_1, Y_2)$) (p-value)	0.1431 (0.4344)	0.1405 (0.4587)	0.1329 (0.5301)	0.1331 (0.5289)	0.1344 (0.5161)	0.1518 (0.3613)
CVS (p-value)	0.0317 (0.31)	0.0310 (0.36)	0.0301 (0.38)	0.0305 (0.37)	0.0297 (0.39)	0.0315 (0.33)
LRT (p-value)	— —	3.8258 (0.0504)	3.8079 (0.0510)	3.7677 (0.0522)	3.9925 (0.0457)	4.2609 (0.0389)