Univariate and Bivariate Geometric Discrete Generalized Exponential Distributions

Debasis Kundu[1] & Vahid Nekoukhou[2]

Abstract

Marshall and Olkin (1997, Biometrika, 84, 641 - 652) introduced a very powerful method to introduce an additional parameter to a class of continuous distribution functions and hence it brings more flexibility to the model. They have demonstrated their method for the exponential and Weibull classes. In the same paper they have briefly indicated regarding its bivariate extension. The main aim of this paper is to introduce the same method, for the first time, to the class of discrete generalized exponential distributions both for the univariate and bivariate cases. We investigate several properties of the proposed univariate and bivariate classes. The univariate class has three parameters, whereas the bivariate class has five parameters. It is observed that depending on the parameter values the univariate class can be both zero inflated as well as heavy tailed. We propose to use EM algorithm to estimate the unknown parameters. Small simulation experiments have been performed to see the effectiveness of the proposed EM algorithm, and a bivariate data set has been analyzed and it is observed that the proposed models and the EM algorithm work quite well in practice.

Key Words and Phrases: Discrete bivariate model; Discrete generalized exponential distribution; EM algorithm; Geometric maximum; Maximum likelihood estimators.

AMS 2000 Subject Classification: Primary 62F10; Secondary: 62H10

1Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur, Pin 208016, India. e-mail: kundu@iitk.ac.in.

[2] Department of Statistics, Khansar Faculty of Mathematics and Computer Science, Khansar, Iran.

1 Introduction

Generalized exponential (GE) distribution originally introduced by Gupta and Kundu (1999) has received considerable attention in recent years. It is an absolute continuous univariate distribution with several interesting properties. It has been used quite successfully as an alternative to a gamma or a Weibull distribution. Although, often in practice we use a continuous random variable mainly due to analytical tractability, discrete data occur in practice quite naturally in various fields. For example, the number of deaths due to a particular cause in a place during a month, the number of attempts needed to crack a computer password or the number of goals scored by a particular team are purely discrete in nature. In these cases it is better to analyze these data using a discrete probability model rather than a continuous probability model. Several attempts have been made to generate various discrete probability distributions and to develop their properties, see for example the book by Johnson et al. (2005), and the references cited therein.

Recently, Nekoukhou et al. (2013) introduced a discrete generalized exponential (DGE) distribution, which can be considered as the discrete analogue of the absolute continuous GE distribution of Gupta and Kundu (1999). The DGE distribution proposed by Nekoukhou et al. (2013) is a very flexible two-parameter distribution. The probability mass function of the DGE distribution can be a decreasing or a unimodal function. Similarly, hazard function of the DGE distribution can be increasing, decreasing or constant depending on the shape parameter. Hence, the geometric distribution can be obtained as a special case of the DGE distribution. A DGE model is appropriate for analyzing both over and under-dispersed discrete data, in view of the fact that the variance can be larger or smaller than the mean. It has been used to analyze various discrete data sets, and the performances were quite satisfactory.

Bivariate distributions are mainly used to analyze the marginals and also to model the dependence structure between the two marginals. Extensive work has been done to propose different bivariate continuous distributions and to develop their properties. Similar to the continuous univariate data, the continuous bivariate data are also analyzed quite often in practice mainly due to analytical tractability. See for example the book by Balakrishnan and Lai (2009) and the references cited therein for different continuous bivariate probability distributions and for their various properties and applications.

Discrete bivariate data also occur quite naturally in practice. For example, the number of goals scored by two competing teams or the number of insurance claims for two different causes are purely discrete in nature. Several bivariate discrete distributions have been proposed in the statistical literature mainly to analyze bivariate discrete data. Recently, Lee and Cha (2015) introduced two general classes of discrete bivariate distributions, and Nekoukhou and Kundu (2017) proposed a four-parameter bivariate discrete generalized exponential distribution. See also the books by Kocherlakota and Kocherlakota (1992), Johnson et al. (1997) and the numerous references cited there in this respect.

Marshall and Olkin (1997) introduced a very efficient mechanism to introduce an extra parameter to a class of continuous univariate distribution functions and hence it brings more flexibility to the existing model. They have illustrated their methods by using exponential and Weibull distribution functions. Since then an extensive amount of work has been done for generalizing several univariate classes of distribution functions, see for example Adamidis and Loukas (1998), Louzada et al. (2014), Ristic and Kundu (2015, 2016) and the references cited therein. Marshall and Olkin (1997) also mentioned about the extension of their method to the bivariate case. Recently, Kundu and Gupta (2014) and Kundu (2015) applied that method for the bivariate Weibull and bivariate generalized exponential distributions, respectively.

Although an extensive amount of work has been done for continuous distributions partic-

ularly for univariate case, no attempt has been made for the discrete distribution except the work by Gómez-Déniz (2010). Gómez-Déniz (2010) adopted the same method as of Marshall and Olkin (1997) and proposed a new generalized version of the geometric distribution. No work has been done, at least not known to the authors about the bivariate case.

The aim of this paper is two fold. First we introduce the univariate geometric discrete generalized exponential (GDGE) distribution. We develop several properties of the proposed univariate GDGE distribution. The proposed univariate GDGE distribution has three parameters. The probability mass function (PMF) of a univariate GDGE distribution can take variety of shapes. It can be zero inflated as well as heavy tailed. It may be mentioned that not too many univariate discrete distributions have these properties. Then we introduce the bivariate GDGE distribution. It has five parameters. Due to the presence of five parameters the bivariate GDGE distribution is a very flexible bivariate discrete distribution. Its marginals are univariate GDGE distributions. We develop several properties of the bivariate GDGE distribution. We provide various dependency measures also. The maximum likelihood estimators (MLEs) of the unknown parameters cannot be obtained in closed forms. One needs to solve five non-linear equations to compute the MLEs of the unknown parameters. We have proposed to use EM algorithm to compute the MLEs of the unknown parameters. Small simulation experiments have been performed to see the effectiveness of the proposed EM algorithm. We have analyzed two univariate and one bivariate data sets to illustrate how the method can be used in practice. It is observed that the performances of the models and the proposed EM algorithm work quite satisfactory.

The rest of the paper is organized as follows. In Section 2 we provide a brief background of the GE and DGE distributions. We introduce and discuss several properties of the univariate GDGE distribution in Section 3. In Section 4 we discuss about the bivariate GDGE distribution. In Section 5, we consider different inferential issues for both the univariate and

bivariate cases. Simulation results and the data analyses have been presented in Section 6. Finally we conclude the paper in Section 7.

2 Preliminaries

2.1 The GE Distribution

The absolute continuous GE distribution was proposed by Gupta and Kundu (1999) as an alternative to the well known gamma and Weibull distributions. The two-parameter GE distribution has the following probability density function (PDF), cumulative distribution function (CDF), and the hazard rate function, respectively;

$$f_{GE}(x;\alpha,\lambda) = \alpha\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\alpha-1}; \quad x>0, \tag{1}$$

$$F_{GE}(x;\alpha,\lambda) = (1-e^{-\lambda x})^{\alpha}; \quad x>0, \tag{2}$$

$$h_{GE}(x;\alpha,\lambda) = \frac{\alpha\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\alpha-1}}{1-(1-e^{-\lambda x})^{\alpha}}; \quad x>0. \tag{3}$$

Here $\alpha>0$ and $\lambda>0$ are the shape and the scale parameters, respectively. From now on a GE distribution with the shape parameter α and the scale parameter λ will be denoted by GE(α,λ). Similar to the gamma and Weibull distributions, it is also a generalization of the exponential distribution and hence, exponential distribution can be obtained as a special case. The PDF (1) and the hazard rate function (3) of a GE distribution can take various shapes. The PDF can be a decreasing or a unimodal function and the hazard rate function can be an increasing, a decreasing or a constant function depending on the shape parameter. It has been observed by several authors during the last fifteen years that this model can be used quite effectively as an alternative to the Weibull and gamma distributions for many practical problems. Interested readers are referred to the review articles by Gupta and Kundu (2007) and Nadarajah (2011) and the recently published monograph by Al-Hussaini and Ahsanullah (2015).

2.2 The DGE Distribution

Recently, the DGE distribution was proposed by Nekoukhou et al. (2013). A discrete random variable X is said to have a DGE distribution with parameters α and p $(= e^{-\lambda})$, if the PMF of X can be written as follows:

$$f_{DGE}(x;\alpha,p) = P(X = x) = (1-p^{x+1})^{\alpha} - (1-p^{x})^{\alpha}, \qquad x \in \mathbb{N}_0 = \{0,1,2,\ldots\}. \tag{4}$$

The corresponding CDF becomes

$$F_{DGE}(x;\alpha,p) = P(X \leq x) = \begin{cases} 0 & \text{if} \quad x < 0 \\ (1-p^{[x]+1})^{\alpha} & \text{if} \quad x \geq 0. \end{cases} \tag{5}$$

Here $[x]$ denotes the largest integer less than or equal to x. From now on a DGE distribution with parameters α and p will be denoted by $\mathrm{DGE}(\alpha, p)$. The PMF and the hazard rate function of a DGE distribution can take various shapes. The PMF can be a decreasing or a unimodal, and the hazard rate function can be an increasing or a decreasing function. Different moments and the distribution of the order statistics were obtained by the authors in the same paper. A DGE model is appropriate for modeling both over and under-dispersed data since, in this model, the variance can be larger or smaller than the mean which is not the case with most of the standard classical discrete distributions.

The following representation of a DGE random variable becomes very useful. Suppose $X \sim \mathrm{DGE}(\alpha, p)$, then for $\lambda = -\ln p$,

$$Y \sim \mathrm{GE}(\alpha,\lambda) \Longrightarrow X = [Y] \sim \mathrm{DGE}(\alpha, p). \tag{6}$$

Using (6), the generation of a random sample from a $\mathrm{DGE}(\alpha, p)$ becomes very simple. For example, first we can generate a random sample Y from a $\mathrm{GE}(\alpha, \lambda)$ distribution, and then by considering $X = [Y]$, we can obtain a generated sample from $\mathrm{DGE}(\alpha, p)$.

3 Univariate GDGE Distribution

Suppose $X_1, X_2, \ldots$ are independent identically distributed (i.i.d.) DGE(α, p) random variables, where $0 < \alpha < \infty$ and $0 < p < 1$ and N is a geometric random variable with the following PMF for $0 < \theta < 1$,

$$P(N = n) = \theta(1-\theta)^{n-1}; \quad n \in \mathbb{N} = \{1, 2, ...\}. \tag{7}$$

The above geometric distribution will be denoted by GM(θ), in the rest of the paper. It is further assumed that N is independent of X_i's. Let us define a new random variable

$$X = \max\{X_1, \ldots, X_N\}. \tag{8}$$

Then the distribution of X is said to have the univariate GDGE distribution and it will be denoted by UGDGE(α, p, θ).

Note that the univariate GDGE distribution can be used quite effectively to analyze a parallel system with random number of components. In this case the number of components is a random quantity and it follows a geometric distribution where as the lifetime of each component follows a DGE distribution. It may be mentioned that the analysis of a parallel system with random number of components has received considerable attention for quite sometime in the statistical reliability literature, see for example Bartoszewicz (2001), Shaked and Wong (1997) and the references cited therein.

The CDF of X for $x \geq 0$, can be obtained as follows.

$$\begin{aligned} F(x) = P(X \leq x) &= \sum_{n=1}^{\infty} P(X \leq x, N = n) = \sum_{n=1}^{\infty} P(X \leq x | N = n) P(N = n) \\ &= \theta(1 - p^{[x]+1})^{\alpha} \sum_{n=0}^{\infty} (1 - p^{[x]+1})^{n\alpha} (1-\theta)^{\alpha} \\ &= \frac{\theta(1 - p^{[x]+1})^{\alpha}}{1 - (1-\theta)(1 - p^{[x]+1})^{\alpha}}. \end{aligned} \tag{9}$$

Hence, the PMF of X becomes

$$P(X=x)=f_X(x)=\frac{\theta\left[(1-p^{x+1})^\alpha-(1-p^x)^\alpha\right]}{\left[1-(1-\theta)(1-p^{x+1})^\alpha\right]\left[1-(1-\theta)(1-p^x)^\alpha\right]};\qquad x\in\mathbb{N}_0. \tag{10}$$

It is clear that if $\theta=1$, then univariate GDGE becomes univariate DGE distribution, and if $\theta=\alpha=1$, then it becomes a geometric distribution. Therefore, clearly the proposed univariate GDGE distribution is a generalization of the geometric and univariate DGE distributions. In fact from (9) it is clear that for fixed x as $\theta\to 0$, then $F(x)\to 0$. Hence, for any fixed x, as $\theta\to 0$, $P(X>x)\to 1$. It implies that as $\theta\to 0$, the univariate GDGE becomes a heavy tailed distribution. In Figures 1 to 3 we have provided the plots of the PMFs of the univariate GDGE distribution for different parameter values. It is clear from the plots of the PMFs that it can take different shapes depending on the parameter values. It can be a decreasing or a unimodal function, and it can be heavy tailed also.

Note that the PMF of X for $x\in\mathbb{N}_0$ can be written as

$$f_X(x)=F_X(x)-F_X(x-1)=w(x)f_{DGE}(x;\alpha,p), \tag{11}$$

where

$$w(x)=\frac{\theta}{\left[1-(1-\theta)F_{DGE}(x;\alpha,p)\right]\left[1-(1-\theta)F_{DGE}(x-1;\alpha,p)\right]}.$$

Therefore, the PMF of X can be written as the weighted DGE probability mass functions with the weight function $w(x)$. The hazard rate function of X is also given by

$$h_X(x)=w^*(x)h_{DGE}(x;\alpha,p), \tag{12}$$

where

$$w^*(x)=\frac{\theta}{1-(1-\theta)F_{DGE}(x-1;p,\alpha)},$$

where h_{DGE} denotes the hazard rate function of a DGE distribution.

The γ-th percentile point of a GDGE(α,θ,p) distribution is given by

$$\xi_\gamma=\frac{\ln\left\{1-\left[\frac{\gamma}{\theta+\gamma(1-\theta)}\right]^{1/\alpha}\right\}}{\ln p}-1.$$

Now we will show that the univariate GDGE can be written as an infinite mixture of DGE distributions.

THEOREM 1: The PMF of a UGDGE(α, θ, p) distribution can be written as an infinite mixture of DGE distributions for $0 < \theta < 1$.

PROOF: By using the series representation

$$(1-z)^{-1} = \sum_{k=0}^{\infty} z^k; \quad \text{for} \quad |z| < 1,$$

the CDF of a UGDGE(α, θ, p) can be written as

$$F(x) = \theta \sum_{k=0}^{\infty} (1-\theta)^k (1-p^{[x]+1})^{\alpha(k+1)}. \tag{13}$$

Therefore,

$$P(X=x) = \theta \sum_{k=0}^{\infty} (1-\theta)^k \left\{ (1-p^{x+1})^{\alpha(k+1)} - (1-p^x)^{\alpha(k+1)} \right\}; \quad x \in \mathbb{N}_0.$$

Hence, the result follows. ∎

Now using the probability generating function (PGF) and the moment generating function (MGF) of a DGE(α, p), see Nekoukhou and Kundu (2017), the PGF and MGF of UGDGE(α, θ, p) can be easily obtained as

$$G_X(z) = \theta \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} (-1)^j \binom{\alpha(k+1)}{j} (1-\theta)^k \frac{1-p^j}{1-zp^j}; \quad |z| < 1,$$

and

$$M_X(t) = \theta \sum_{j=1}^{\infty} \sum_{n=1}^{\infty} (-1)^{j+1} \binom{n\alpha}{j} \frac{(1-p^j)(1-\theta)^{n-1}}{1-p^j e^t}, \quad t < -\ln p,$$

respectively. Both the proofs can be easily obtained, and hence the details are avoided. It must be mentioned that for an integer α, $\sum_{j=1}^{\infty}$ should be replaced by $\sum_{j=1}^{\alpha}$; see Nekoukhou et al. (2013).

Different moments of the univariate GDGE distribution can be obtained as follows. If $X \sim \text{UGDGE}(\alpha, \theta, p)$, then the rth moment of X can be obtained as

$$E(X^r) = \sum_{x=1}^{\infty} \{x^r - (x-1)^r\} P(X \geq x) = \sum_{x=1}^{\infty} \{x^r - (x-1)^r\} \frac{1-(1-p^x)^\alpha}{1-(1-\theta)(1-p^x)^\alpha}. \quad (14)$$

Therefore, in particular the mean and the second moment of X become

$$E(X) = \sum_{x=1}^{\infty} \frac{1-(1-p^x)^\alpha}{1-(1-\theta)(1-p^x)^\alpha}$$

and

$$E(X^2) = \sum_{x=1}^{\infty} (2x-1) \frac{1-(1-p^x)^\alpha}{1-(1-\theta)(1-p^x)^\alpha}.$$

We have the following results regarding the stochastic ordering of the family of univariate GDGE distributions.

Result 1: If $X_1 \sim \text{UGDGE}(\alpha_1, \theta, p)$ and $X_2 \sim \text{UGDGE}(\alpha_2, \theta, p)$, then for $\alpha_1 > \alpha_2$, X_1 is stochastically larger than X_2.

Result 2: If $X_1 \sim \text{UGDGE}(\alpha, \theta, p_1)$ and $X_2 \sim \text{UGDGE}(\alpha, \theta, p_2)$, then for $p_1 > p_2$, X_1 is stochastically larger than X_2.

Result 3: If $X_1 \sim \text{UGDGE}(\alpha, \theta_1, p)$ and $X_2 \sim \text{UGDGE}(\alpha, \theta_2, p)$, then for $\theta_1 < \theta_2$, X_1 is stochastically larger than X_2.

Proof: The proofs of Results 1 & 2 can be obtained using (13), and the proof of Result 3 can be obtained using (9). ∎

Let us recall the Marshall-Olkin generalized exponential (MOGE) distribution as introduced by Ristic and Kundu (2015). Suppose $Y_1, Y_2, \ldots$ are i.i.d. $\text{GE}(\alpha, \lambda)$ random variables, and N is a geometric random variable with the PMF as defined in (7). It is also assumed that N and Y_i's are independently distributed. Then

$$Y = \max\{Y_1, \ldots, Y_N\}$$

is said to have MOGE distribution with the parameters α and $p = e^{-\lambda}$, and it will be denoted by MOGE(α, θ, p). Now we have the following result.

THEOREM 2: If $Y \sim$ MOGE(α, θ, p), then $X = [Y] \sim$ UGDGE(α, θ, p).

PROOF: Since,

$$P(Y < y) = P(Y \leq y) = \theta \sum_{j=1}^{\infty} (1 - p^y)^{\alpha j} (1 - \theta)^j = \frac{\theta(1 - p^y)^\alpha}{1 - (1 - \theta)(1 - p^y)^\alpha},$$

therefore, for $y \in \mathbb{N}_0$

$$P(X \leq y) = P(Y < y + 1) = \frac{\theta(1 - p^{y+1})^\alpha}{1 - (1 - \theta)(1 - p^{y+1})^\alpha}.$$

Hence, for $y \geq 0$,

$$P(X \leq y) = \frac{\theta(1 - p^{[y]+1})^\alpha}{1 - (1 - \theta)(1 - p^{[y]+1})^\alpha}.$$

∎

Theorem 2 can be used quite effectively to generate samples from a UGDGE(α, θ, p). The following algorithm can be used for that purpose.

ALGORITHM 1:

1. First generate N from a GM(θ).
2. Generate Y from GE$(N\alpha, -\ln p)$ using the inverse transformation method.
3. Obtain the required random variable X as $X = [Y]$.

Now, we will show that the univariate GDGE distributions are closed under geometric maximum. More precisely, we have the following result.

THEOREM 3: Let $\{U_i : i \geq 1\}$ be a sequence of i.i.d. UGDGE(α, θ, p) random variables, and $M \sim$ GM(q), $0 < q < 1$. In addition, suppose that U_i's and M are independent, then

we have the following result

$$U = \max\{U_1, U_2, ..., U_M\} \sim \text{UGDGE}(\alpha, q\theta, p). \tag{15}$$

PROOF: Note that

$$\begin{aligned} P(U \leq u) &= \sum_{m=1}^{\infty} P(U_1 \leq u, ..., U_M \leq u | M = m) P(M = m) \\ &= q \sum_{m=1}^{\infty} \left[\frac{\theta(1-p^{u+1})^{\alpha}}{1-(1-\theta)(1-p^{u+1})^{\alpha_1}} \right]^m (1-q)^{m-1} \\ &= \frac{q\theta(1-p^{u+1})^{\alpha}}{1-(1-q\theta)(1-p^{u+1})^{\alpha}}. \end{aligned}$$

∎

The following results will be useful for future development, mainly for developing the EM algorithm to compute the MLEs of the unknown parameters. First, note that the joint PMF of (X, N), say $f_{X,N}$, is given by

$$f_{X,N}(x,n) = \theta(1-\theta)^{n-1} \left[(1-p^{x+1})^{n\alpha} - (1-p^x)^{n\alpha} \right], \quad x \in \mathbb{N}_0, n \in \mathbb{N}. \tag{16}$$

Therefore, the joint CDF of (X, N) is also given by

$$\begin{aligned} F_{X,N}(x,n) &= \sum_{j=1}^{n} P(X \leq x | N = j) P(N = j) \\ &= \theta \sum_{j=1}^{n} (1-p^{[x]+1})^{j\alpha} (1-\theta)^{j-1} \\ &= \frac{\theta(1-p^{[x]+1})^{\alpha} \left[1 - (1-p^{[x]+1})^{n\alpha} (1-\theta)^n \right]}{1-(1-\theta)(1-p^{[x]+1})^{\alpha}}. \end{aligned} \tag{17}$$

Clearly, the CDF of X also can be obtained as follows

$$F_X(x) = \lim_{n\to\infty} F_{X,N}(x,n) = \frac{\theta(1-p^{[x]+1})^{\alpha}}{1-(1-\theta)(1-p^{[x]+1})^{\alpha}}. \tag{18}$$

The conditional PMF of N given $X = x$ is given by

$$\begin{aligned} f_{N|X}(n|x) &= P(N = n | X = x) \\ &= (1-\theta)^{n-1} \frac{f_{DGE}(x; n\alpha, p)}{f_{DGE}(x; \alpha, p)} \\ &\times \left[1 - (1-\theta) F_{DGE}(x; \alpha, p) \right] \left[1 - (1-\theta) F_{DGE}(x-1; \alpha, p) \right]. \end{aligned} \tag{19}$$

The conditional expectation of N can be obtained as

$$\begin{aligned} E(N|X=x) &= \frac{F_{DGE}(x;\alpha,p)\left[1-(1-\theta)F_{DGE}(x-1;\alpha,p)\right]}{f_{DGE}(x;\alpha,p)\left[1-(1-\theta)F_{DGE}(x;\alpha,p)\right]} \\ &- \frac{F_{DGE}(x-1;\alpha,p)\left[1-(1-\theta)F_{DGE}(x;\alpha,p)\right]}{f_{DGE}(x;\alpha,p)\left[1-(1-\theta)F_{DGE}(x-1;\alpha,p)\right]}. \end{aligned} \tag{20}$$

4 Bivariate GDGE Distribution

Suppose $X_1, X_2, \ldots$ are i.i.d. DGE(α_1, p_1) random variables, $Y_1, Y_2, \ldots$ are i.i.d DGE(α_2, p_2) random variables and N is a GM(θ) random variable. All the random variables are independently distributed. Now consider the bivariate discrete random variable (X, Y), where

$$X = \max\{X_1, X_2, ..., X_N\} \quad \text{and} \quad Y = \max\{Y_1, Y_2, ..., Y_N\}. \tag{21}$$

The joint distribution of (X, Y) is said to have the bivariate GDGE distribution. The following interpretations can be provided for a bivariate GDGE model.

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 $X_1, X_2, \ldots$ denote the lifetime of the components of system 1 which are reported in a discrete scale, and in a same manner, $Y_1, Y_2, \ldots$ denote the lifetime of the components of system 2, then the lifetime of the two systems becomes (X, Y).

Random Stress Model: Suppose a system has two components. Each component is subject to random number of individual independent discrete stresses, say $\{X_1, X_2, ...\}$ and $\{Y_1, Y_2, \ldots\}$, respectively. If N is the number of stresses, then the observed stresses at the two components are $X = \max\{X_1, ..., X_N\}$ and $Y = \max\{Y_1, ..., Y_N\}$, respectively.

The joint CDF of a bivariate GDGE distribution for $x \geq 0$ and $y \geq 0$ can be obtained as

$$F_{X,Y}(x,y) = \sum_{n=1}^{\infty} P(X \leq x, Y \leq y|N=n)P(N=n)$$

$$= \theta \sum_{n=1}^{\infty} (1-p_1^{[x]+1})^{n\alpha_1} (1-p_2^{[y]+1})^{n\alpha_2} (1-\theta)^{n-1}$$
$$= \frac{\theta (1-p_1^{[x]+1})^{\alpha_1} (1-p_2^{[y]+1})^{\alpha_2}}{1-(1-\theta)(1-p_1^{[x]+1})^{\alpha_1}(1-p_2^{[y]+1})^{\alpha_2}}.$$

The above bivariate distribution with the parameter vector $\boldsymbol{\Omega} = (\alpha_1, \alpha_2, p_1, p_2, \theta)^T$ will be denoted by BGDGE$(\alpha_1, \alpha_2, p_1, p_2, \theta)$.

It is interesting to note that if $\theta = 1$, then we have

$$F_{X,Y}(x,y) = F_{DGE}(x;\alpha_1,p_1) F_{DGE}(y;\alpha_2,p_2),$$

i.e., X and Y become independent. Therefore, the parameter θ plays the role of the correlation parameter. The joint PMF of (X,Y) for $x \in \mathbb{N}_0$ and $y \in \mathbb{N}_0$ can be obtained as

$$f_{X,Y}(x,y) = F_{X,Y}(x,y) - F_{X,Y}(x-1,y) - F_{X,Y}(x,y-1) + F_{X,Y}(x-1,y-1).$$

More precisely, the joint PMF is given by

$$f_{X,Y}(x,y) = g(x,y;\boldsymbol{\Omega}) - g(x,y-1;\boldsymbol{\Omega}), \tag{22}$$

where

$$g(x,y;\boldsymbol{\Omega}) = \frac{\theta F_{DGE}(y;\alpha_2,p_2) f_{DGE}(x;\alpha_1,p_1)}{[1-(1-\theta)F_{DGE}(x;\alpha_1,p_1)F_{DGE}(y;\alpha_2,p_2)]\,[1-(1-\theta)F_{DGE}(x-1;\alpha_1,p_1)F_{DGE}(y;\alpha_2,p_2)]}.$$

Let us recall the bivariate geometric generalized exponential (BGGE) distribution as introduced by Kundu (2015). Suppose $U_1, U_2, \ldots$ are i.i.d. GE(α_1, λ_1) random variables, $V_1, V_2, \ldots$ are i.i.d. GE(α_2, λ_2) random variables, N is GM(θ) random variable and all are independently distributed. The bivariate absolute continuous random variable (U,V), where

$$U = \max\{U_1, \ldots, U_N\} \quad \text{and} \quad V = \max\{V_1, \ldots, V_N\}$$

is said to have BGGE distribution with parameters $\alpha_1, \alpha_2, \lambda_1, \lambda_2$ and θ. From now on it will be denoted by BGGE$(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \theta)$. It can be shown similarly as Theorem 2 that

if $(U,V) \sim \text{BGGE}(\alpha_1,\alpha_2,\lambda_1,\lambda_2,\theta)$, then $(X,Y) \sim \text{BGDGE}(\alpha_1,\alpha_2,p_1,p_2,\theta)$, where $X = [U], Y = [V]$, $p_1 = e^{-\lambda_1}$ and $p_2 = e^{-\lambda_2}$. Therefore, it is clear that the bivariate GDGE distribution is the natural discrete version of the BGGE distribution.

Generation from a bivariate GDGE distribution is also quite straight forward. First N is generated from a $\text{GM}(\theta)$, and once $N = n$ is observed, U and V, can be generated from $\text{GE}(n\alpha_1,\lambda_1)$ and $\text{GE}(n\alpha_2,\lambda_2)$, respectively. Then $([U],[V])$ becomes the required bivariate random variable. Now, we have the following results for the bivariate GDGE distribution.

The PGF for $|u| < 1, |v| < 1$, and MGF for $t_1 < -\ln p_1, t_2 < -\ln p_2$ of (X,Y) can be obtained as

$$G_{X,Y}(u,v) = \theta \sum_{k=0}^{\infty}\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}(-1)^{i+j}\binom{\alpha_1(k+1)}{j}\binom{\alpha_2(k+1)}{i}(1-\theta)^k \frac{1-p_1^j}{1-up_1^j}\frac{1-p_2^i}{1-vp_2^i}.$$

and

$$\begin{aligned} M_{X,Y}(t_1,t_2) &= E(e^{t_1X+t_2Y}) = E_N\left\{E_{X,Y|N}(e^{t_1X+t_2Y})|N=n\right\} \\ &= \theta\sum_{n=1}^{\infty}\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}(-1)^{j+k}\binom{n\alpha_1}{j}\binom{n\alpha_2}{k}\frac{(1-p_1^j)(1-p_2^k)}{(1-p_1^je^{t_1})(1-p_2^ke^{t_2})}, \end{aligned}$$

respectively.

THEOREM 4: If $(X,Y) \sim \text{BGDGE}(\alpha_1,\alpha_2,p_1,p_2,\theta)$, then we have the following results:

(a) $X \sim \text{UGDGE}(\alpha_1,p_1,\theta)$ and $Y \sim \text{UGDGE}(\alpha_2,p_2,\theta)$.

(b) $X \le x|Y \le y \sim \text{UGDGE}(\alpha_1,p_1,p^*)$, where $p^* = 1-(1-\theta)(1-p_2^{y+1})^{\alpha_2}$.

(c) If $p_1 = p_2 = p$, then $\max\{X,Y\} \sim \text{UGDGE}(\alpha_1+\alpha_2,p,\theta)$.

(d) $P(X \le x|Y=y) = \frac{[1-(1-\theta)F_{DGE}(y;\alpha_2,p_2)][1-(1-\theta)F_{DGE}(y-1;\alpha_2,p_2)][F_{X,Y}(x,y)-F_{X,Y}(x,y-1)]}{\theta f_{DGE}(y;\alpha_2,p_2)}$

PROOF. The proof of (a)-(c) is straight forward and we provide the proof of (d) as follows:

$$P(X \le x|Y=y) = \sum_{n=1}^{\infty} P(X \le x|Y=y,N=n)P(N=n|Y=y)$$

$$= \sum_{n=1}^{\infty}(1-p_1^{x+1})^{n\alpha_1}P(N=n|Y=y).$$

By substitution the $P(N=n|Y=y)$, given by (19), the result is obtained. ∎

The joint PMF of (X,Y,N) is also given by

$$f_{X,Y,N}(x,y,n)=\theta(1-\theta)^{n-1}\left[(1-p_1^{x+1})^{n\alpha_1}-(1-p_1^{x})^{n\alpha_1}\right]\left[(1-p_2^{y+1})^{n\alpha_2}-(1-p_2^{y})^{n\alpha_2}\right].$$

Therefore, we see that

$$\begin{aligned} P(N=n|X=x,Y=y) &= \left[(1-p_1^{x+1})^{n\alpha_1}-(1-p_1^{x})^{n\alpha_1}\right]\left[(1-p_2^{y+1})^{n\alpha_2}-(1-p_2^{y})^{n\alpha_2}\right] \\ &\times \frac{\theta(1-\theta)^{n-1}}{g(x,y;\boldsymbol{\Omega})-g(x,y-1;\boldsymbol{\Omega})}, \end{aligned} \tag{23}$$

and hence we obtain the following conditional expectation,

$$\begin{aligned} E(N|X=x,Y=y) &= \frac{F_{X,Y}(x,y)a(x,y)-F_{X,Y}(x-1,y)a(x-1,y)}{g(x,y;\boldsymbol{\Omega})-g(x,y-1;\boldsymbol{\Omega})} \\ &- \frac{F_{X,Y}(x,y-1)a(x,y-1)-F_{X,Y}(x-1,y-1)a(x-1,y-1)}{g(x,y;\boldsymbol{\Omega})-g(x,y-1;\boldsymbol{\Omega})}, \end{aligned}$$

where $a(x,y)=\left[1-(1-\theta)(1-p_1^{x+1})^{\alpha_1}(1-p_2^{y+1})^{\alpha_2}\right]^{-1}$.

The BGDGE distribution, similar to its marginals, is closed under geometric maximum. More precisely, we have the following result whose proof is avoided.

THEOREM 5: Let $\{(U_i,V_i): i\geq 1)\}$ be a sequence of i.i.d. BGDGE$(\alpha_1,\alpha_2,p_1,p_2,\theta)$ random variables, and $M\sim$ GM(q), $0<q<1$. In addition, M is independent of (U_i,V_i)'s. If we consider the random variables

$$U=\max\{U_1,U_2,...,U_M\} \quad \text{and} \quad \mathrm{V}=\max\{\mathrm{V}_1,\mathrm{V}_2,...,\mathrm{V}_\mathrm{M}\}, \tag{24}$$

then $(U,V)\sim$ BGDGE$(\alpha_1,\alpha_2,p_1,p_2,q\theta)$.

5 Statistical Inference

5.1 Maximum Likelihood Estimation

In this section we consider the maximum likelihood estimation of the unknown parameters of a BGDGE$(\alpha_1, \alpha_2, p_1, p_2, \theta)$ model based on a random sample of size m, namely $\mathcal{D} = \{(x_1, y_1), \ldots, (x_m, y_m)\}$. The proposed bivariate GDGE model has five parameters. It is observed that the MLEs of the unknown parameters can be obtained by solving a five-dimensional optimization problem. To avoid that we have used an EM algorithm and in this case it is observed that at each 'E'-step the corresponding 'M'-step can be obtained by maximizing two two-dimensional optimization problems. Hence, the implementation of the proposed EM algorithm is quite simple in practice. The same EM algorithm with obvious modifications can also be used to compute the MLEs of the unknown parameters of a UGDGE(α, θ, p) model. We will indicate that towards the end of this section.

From now on we will indicate the parameter vector by $\boldsymbol{\Omega} = (\alpha_1, \alpha_2, p_1, p_2, \theta)$. Based on the random sample $\mathcal{D}$ as mentioned above and using (22), the log-likelihood function can be written as

$$l(\boldsymbol{\Omega}|\mathcal{D}) = \sum_{i=1}^{m} \ln \left\{ g(x_i, y_i; \boldsymbol{\Omega}) - g(x_i, y_i - 1; \boldsymbol{\Omega}) \right\}. \tag{25}$$

Here $g(x, y)$ is same as defined in (22). The MLEs of the unknown parameters can be obtained by maximizing (25) with respect to $\boldsymbol{\Omega}$. It needs solving five non-linear equations

$$\dot{l}_{\alpha_1}(\boldsymbol{\Omega}|\mathcal{D}) = 0, \quad \dot{l}_{\alpha_2}(\boldsymbol{\Omega}|\mathcal{D}) = 0, \quad \dot{l}_{p_1}(\boldsymbol{\Omega}|\mathcal{D}) = 0, \quad \dot{l}_{p_2}(\boldsymbol{\Omega}|\mathcal{D}) = 0, \quad \dot{l}_{\theta}(\boldsymbol{\Omega}|\mathcal{D}) = 0, \tag{26}$$

simultaneously. Clearly, they cannot be obtained in explicit forms. Newton-Raphson method may be used to solve these non-linear equations but it has the usual problem of convergence and the choice of the initial guesses.

To avoid that problems we propose to use an EM algorithm to compute the MLEs of

the unknown parameters. The main idea about the EM algorithm is based on the following observations. It is observed that if along with (x, y), the associated n is also known then the MLEs of $\boldsymbol{\Omega}$ can be obtained in a more convenient manner. Suppose we have the following observations $\mathcal{D}^* = \{(x_1, y_1, n_1), \ldots, (x_m, y_m, n_m)\}$. Then based on $\mathcal{D}^*$, the log-likelihood function without the additive constant can be written as

$$l_{complete}(\boldsymbol{\Omega}|\mathcal{D}^*) = mln\theta + (k-m)\ln(1-\theta) + g_1(\alpha_1, p_1) + g_2(\alpha_2, p_2). \tag{27}$$

Here $k = \sum_{i=1}^{m} n_i$ and

$$g_1(\alpha_1, p_1) = \sum_{i=1}^{m} \ln\left[(1-p_1^{x_i+1})^{n_i\alpha_1} - (1-p_1^{x_i})^{n_i\alpha_1}\right],$$

$$g_2(\alpha_2, p_2) = \sum_{i=1}^{m} \ln\left[(1-p_2^{y_i+1})^{n_i\alpha_2} - (1-p_2^{y_i})^{n_i\alpha_2}\right].$$

The following result will be useful for further development.

LEMMA 1: (a) For any $0 < p_1 < 1$, $g_1(\alpha_1, p_1)$ is a unimodal function of α_1. (b) For any $0 < p_2 < 1$, $g_2(\alpha_2, p_2)$ is a unimodal function of α_2.

PROOF: See in the Appendix.

Therefore, based on the $\mathcal{D}^*$, the MLE of θ can be obtained as

$$\widehat{\theta} = \frac{m}{k}, \tag{28}$$

and the MLEs of α_1 and p_1 can be obtained by maximizing $g_1(\alpha_1, p_1)$ with respect to the unknown parameters. Similarly, the MLEs of α_2 and p_2 also can be obtained from $g_2(\alpha_2, p_2)$. The maximization of $g_1(\alpha_1, p_1)$ and $g_2(\alpha_2, p_2)$ can be obtained by using the profile likelihood method. Therefore, based on the complete data $\mathcal{D}^*$ the MLEs can be obtained quite conveniently and we will denote them as $\widehat{\boldsymbol{\Omega}} = (\widehat{\alpha}_1, \widehat{\alpha}_2, \widehat{p}_1, \widehat{p}_2, \widehat{\theta})$.

Now, to implement the EM algorithm we treat this problem as a missing value problem. It is assumed that the complete data set is $\mathcal{D}^*$ and the observed data set is $\mathcal{D}$. Therefore,

for each (x_i, y_i) the associated n_i is missing in this case, and we need to estimate n_i from the observed data. Let us use the following notations. At the jth iterate of the EM algorithm the estimates of the parameters are denoted by $\mathbf{\Omega}^{(j)} = (\alpha_1^{(j)}, \alpha_2^{(j)}, p_1^{(j)}, p_2^{(j)}, \theta^{(j)})$. At the jth iteration step the missing n_i is estimated by maximizing the conditional probability $P(N = n|X = x_i, Y = y_i)$ as in (23). To compute $P(N = n|X = x_i, Y = y_i)$, the parameter vector on the right hand side is replaced by $\mathbf{\Omega}^{(j)}$. We denote the estimated n_i as $\widetilde{n}_i^{(j)}$. Therefore, we see that

$$\widetilde{n}_i^{(j)} = \arg\max_n P(N = n|X = x_i, Y = y_i, \mathbf{\Omega}^{(j)}). \tag{29}$$

Let us denote $k^{(j)} = \sum_{i=1}^{m} \widetilde{n}_i^{(j)}$. Now at the $(j+1)$th step the EM algorithm takes the following form.

'E'-Step: The 'E'-step of the EM algorithm can be obtained by replacing n_i with $\widetilde{n}_i^{(j)}$ in (27). At this stage the *pseudo* log-likelihood function becomes

$$l_{pseudo}(\mathbf{\Omega}|\mathcal{D}^*, \mathbf{\Omega}^{(j)}) = m \ln\theta + (k^{(j)} - m)\ln(1-\theta) + g_1(\alpha_1, p_1|\mathbf{\Omega}^{(j)}) + g_2(\alpha_2, p_2|\mathbf{\Omega}^{(j)}), \tag{30}$$

where

$$g_1(\alpha_1, p_1|\mathbf{\Omega}^{(j)}) = \sum_{i=1}^{m} \ln\left[(1-p_1^{x_i+1})^{\widetilde{n}_i^{(j)}\alpha_1} - (1-p_1^{x_i})^{\widetilde{n}_i^{(j)}\alpha_1}\right], \tag{31}$$

$$g_2(\alpha_2, p_2|\mathbf{\Omega}^{(j)}) = \sum_{i=1}^{m} \ln\left[(1-p_2^{y_i+1})^{\widetilde{n}_i^{(j)}\alpha_2} - (1-p_2^{y_i})^{\widetilde{n}_i^{(j)}\alpha_2}\right], \tag{32}$$

and $\widetilde{n}_i^{(j)}$ is obtained as in (29).

'M'-Step: The 'M'-step involves maximizing (30) with respect to the unknown parameters and they can be obtained as follows: $\theta^{(j+1)} = m/k^j$ and

$$(\alpha_1^{(j+1)}, p_1^{(j+1)}) = \arg\max_{(\alpha_1,p_1)} g_1(\alpha_1, p_1|\mathbf{\Omega}^{(j)}), \tag{33}$$

$$(\alpha_2^{(j+1)}, p_2^{(j+1)}) = \arg\max_{(\alpha_2,p_2)} g_2(\alpha_2, p_2|\mathbf{\Omega}^{(j)}). \tag{34}$$

Now it is clear that the same EM algorithm can be used even for univariate GDGE distribution with obvious modification. In this case if we denote $\boldsymbol{\Omega} = (\alpha, p, \theta)$, $\mathcal{D}^* = \{(x_1, n_1), \ldots, (x_m, n_m)\}$,

$$\widetilde{n}_i^{(j)} = \arg\max_n P(N = n | X = x_i, \boldsymbol{\Omega}^{(j)}),$$

and all the other notations are same as before, then the EM algorithm takes the following form.

'E'-Step: The *pseudo* log-likelihood function becomes

$$l_{pseudo}(\boldsymbol{\Omega}|\mathcal{D}^*, \boldsymbol{\Omega}^{(j)}) = m \ln \theta + (k^{(j)} - m) \ln(1 - \theta) + g_1(\alpha_1, p_1|\boldsymbol{\Omega}^{(j)}), \tag{35}$$

where $g_1(\alpha_1, p_1|\boldsymbol{\Omega}^{(j)})$ is same as defined in (31).

'M'-Step: The 'M'-step involves maximizing (35) with respect to the unknown parameters and they can be obtained as follows: $\theta^{(j+1)} = m/k^j$ and

$$(\alpha_1^{(j+1)}, p_1^{(j+1)}) = \arg\max_{(\alpha_1, p_1)} g_1(\alpha_1, p_1|\boldsymbol{\Omega}^{(j)}).$$

5.2 Testing of Hypotheses

In this section we discuss two different testing of hypotheses problems which have some practical relevance. In both the cases we have mainly used the likelihood ratio test (LRT). In each case the MLE of any arbitrary parameter δ will be denoted by $\widehat{\delta}$ and under the null hypothesis it will be denoted by $\widetilde{\delta}$.

Test 1: We want to test the following null hypothesis

$$H_0 : \alpha_1 = \alpha_2 = \alpha \text{ (unknown)} \quad \text{and} \quad p_1 = p_2 = p \text{ (unknown)}. \tag{36}$$

This is an important problem as it tests that the marginals are equal. Under H_0, the MLEs of α, θ and p can be obtained by using EM algorithm similarly as defined for bivariate GDGE

distribution. In this case at the 'M'-step we need to maximize the function

$$g(\alpha, p|\boldsymbol{\Omega}^{(j)}) = \sum_{i=1}^{m} \ln \left[(1-p^{x_i+1})^{\widetilde{n}_i^{(j)}\alpha} - (1-p^{x_i})^{\widetilde{n}_i^{(j)}\alpha} \right],$$

with respect to α and p. Under H_0,

$$2(l(\widehat{\alpha}_1, \widehat{\alpha}_2, \widehat{p}_1, \widehat{p}_2, \widehat{\theta}|\mathcal{D}) - l(\widetilde{\alpha}, \widetilde{\alpha}, \widetilde{p}, \widetilde{p}, \widehat{\theta}|\mathcal{D})) \longrightarrow \chi^2_2, \tag{37}$$

see for example Casella and Berger (2001).

Test 2: We want to test the following null hypothesis

$$H_0 : \theta = 1. \tag{38}$$

This is an important problem as it tests that the two marginals are independent and both of them follow univariate DGE distributions. In this case the MLEs of α_1, α_2, p_1 and p_2 can also be obtained using the proposed EM algorithm with the obvious modification, i.e. replacing $\widehat{\theta} = 1$ at each stage. Since θ is in the boundary, the standard asymptotic result does not work. In this case, using Theorem 3 of Self and Liang (1987) yields that

$$2(l(\widehat{\alpha}_1, \widehat{\alpha}_2, \widehat{p}_1, \widehat{p}_2, \widehat{\theta}|\mathcal{D}) - l(\widetilde{\alpha}_1, \widetilde{\alpha}_2, \widetilde{p}_1, \widetilde{p}_2, 1|\mathcal{D})) \longrightarrow \frac{1}{2} + \frac{1}{2}\chi^2_1. \tag{39}$$

6 Simulation and Data Analysis

6.1 Simulation

We have performed some simulation experiments to see how the proposed EM algorithm performs in computing the MLEs. We have taken different sample sizes and two different θ values. We have taken $\alpha_1 = \alpha_2 = 2.0$, $p_1 = p_2 = 0.25$, $\theta = 0.25, 0.50$, $n = 25, 50, 75$ and

100. In each case we have generated a random sample from the bivariate GDGE distribution with the given sample size and the parameter values. We have estimated the parameters using the proposed EM algorithm. We have reported the average estimates and the mean squared errors (MSEs) over 1000 replications. In each box the top figure represents the average estimate (AE) and the associated MSEs is reported below within a bracket. The results are reported in Tables 1 to 2.

n	$\alpha_1 = 2.0$	$p_1 = 0.25$	$\alpha_2 = 2.0$	$p_2 = 0.25$	$\theta = 0.25$
25	1.7124 (0.5716)	0.1987 (0.0581)	1.7215 (0.5618)	0.1921 (0.0534)	0.2011 (0.0611)
50	1.7618 (0.3011)	0.2041 (0.0312)	1.7691 (0.2987)	0.2116 (0.0349)	0.2278 (0.0289)
75	1.8312 (0.2018)	0.2287 (0.0211)	1.8579 (0.2111)	0.2318 (0.0228)	0.2410 (0.0198)
100	1.9891 (0.1439)	0.2510 (0.0143)	2.0104 (0.1411)	0.2498 (0.0137)	0.2501 (0.0114)

Table 1: The AEs and the associated MSEs of the MLEs when $\alpha_1 = \alpha_2 = 2.0$, $p_1 = p_2 = 0.25$, $\theta = 0.25$.

n	$\alpha_1 = 2.0$	$p_1 = 0.25$	$\alpha_2 = 2.0$	$p_2 = 0.25$	$\theta = 0.50$
25	1.7422 (0.5218)	0.2012 (0.0487)	1.7519 (0.5198)	0.2065 (0.0446)	0.2117 (0.0576)
50	1.7776 (0.2567)	0.2145 (0.0265)	1.7890 (0.2514)	0.2208 (0.0276)	0.2365 (0.0245)
75	1.9676 (0.1676)	0.2376 (0.0167)	1.9786 (0.1632)	0.2406 (0.0187)	0.2473 (0.0141)
100	2.0015 (0.1256)	0.2504 (0.0110)	2.0011 (0.1198)	0.2501 (0.0116)	0.2500 (0.0101)

Table 2: The AEs and the associated MSEs of the MLEs when $\alpha_1 = \alpha_2 = 2.0$, $p_1 = p_2 = 0.25$, $\theta = 0.50$.

Some of the points are quite clear from the simulation experiments that as the sample size increases the biases and MSEs decrease in each case. Moreover, it is also observed that

as θ increases the biases and MSEs decrease for each estimators.

6.2 Data Analysis

In this section we would present the analysis of a bivariate data set to show how the proposed model and the EM algorithm work in practice. This bivariate data set represents Italian Series A football match score data between Italian giants 'ACF Firontina' (X_1) and 'Juventus' (X_2) during the time period 1996 to 2011. The data set is available in Lee and Cha (2015) and it is presented in Table 3 for easy reference. It is presented in the contingency table form in Table 4.

Obs.	ACF Firontina (X_1)	Juventus (X_2)	Obs.	ACF Firontina (X_1)	Juventus (X_2)
1	1	2	14	1	2
2	0	0	15	1	1
3	1	1	16	1	3
4	2	2	17	3	3
5	1	1	18	0	1
6	0	1	19	1	1
7	1	1	20	1	2
8	3	2	21	1	0
9	1	1	22	3	0
10	2	1	23	1	2
11	1	2	24	1	1
12	3	3	25	0	1
13	0	1	26	0	1

Table 3: Italian Series A data

First we would like to fit univariate GDGE distribution to both the marginals. We have used the EM algorithm to compute the MLEs. We present the MLEs and the associated 95% confidence intervals in Table 5. Now to see whether univariate GDGE fits the marginals or not we have calculated the χ^2 values and also the associated p-values for both the marginals.

$X_1 \downarrow X_2 \rightarrow$	0	1	2	3	Total
0	1	5	0	0	6
1	1	7	5	1	14
2	0	1	1	0	2
3	1	0	1	2	4
Total	3	13	7	3	26

Table 4: Italian Series A data

The results are presented in the same Table 5. Since the p-values are greater than 0.1 for both the marginals we conclude that univariate GDGE fits both the marginals well.

Variable	α	θ	p	LL	χ^2	p-value
X_1	4.6587 $\mp$ 0.8756	0.9987 $\mp$ 0.0014	0.2618 $\mp$ 0.0541	-33.4193	6.1486	> 0.1
X_2	6.8029 $\mp$ 1.1562	0.3288 $\mp$ 0.0087	0.1683 $\mp$ 0.0465	-31.8832	0.6853	> 0.1

Table 5: MLEs of the unknown parameters and the associated 95% confidence intervals

Now we would like to fit the bivariate GDGE distribution to the bivariate data set. We have used the EM algorithm for the bivariate GDGE distribution proposed in the previous section. Based on the fitted univariate GDGE marginals we have used the following initial values for the unknown parameters:

$$\alpha_1^{(0)} = 4.6587, \quad p_1^{(0)} = 0.2618, \quad \alpha_2^{(0)} = 6.8029, \quad p_2^{(0)} = 0.1683, \quad \theta^{(0)} = (0.9987+0.3288)/2.0 = 0.6638.$$

The EM algorithm stops after nineteen iteration and the MLEs and the associated 95% confidence bounds are presented within parenthesis below

$$\widehat{\alpha}_1 = 4.5519(\mp 1.1101) \quad \widehat{p}_1 = 0.2570(\mp 0.0721)$$

$$\widehat{\alpha}_2 = 8.3892(\mp 1.9767) \quad \widehat{p}_2 = 0.2250(\mp 0.0518)$$

$$\widehat{\theta} = 0.9211(\mp 0.0312).$$

Now to check whether the proposed BDGE fits the bivariate data or not we have obtained the observed and expected values and they are presented in Table 6. The chi-square value is 7.79 and the corresponding p-value with 9 degrees of freedom is greater than 0.1. Hence we cannot reject the hypothesis that the data are coming from a bivariate GDGE distribution.

$X_1 \downarrow X_2 \rightarrow$	0	1	2	3
0	1(0.64)	5(3.31)	0(1.75)	0 (0.30)
1	1(1.17)	7(6.32)	5(3.51)	1 (0.98)
2	0(0.88)	1(2.67)	1(1.54)	0(0.44)
3	1(0.84)	0(0.69)	1(0.78)	2(1.16)

Table 6: Observed and expected frequencies for Italian Series A data

7 Conclusions

In this paper we have introduced univariate and bivariate GDGE distributions using the method proposed by Marshall and Olkin (1997). Apparently this is the first time the method of Marshall and Olkin (1997) has been used for the discrete bivariate distributions. We have derived different properties of the proposed distributions. It is observed that both the univariate and bivariate distributions are very flexible. The MLEs of the unknown parameters cannot be obtained in explicit forms. We have proposed to use a new EM algorithm which is applicable for the discrete distributions. We have performed some simulation experiments to see the effectiveness of the proposed EM algorithm, and analyzed one data set for illustrative purposes. It is observed that the proposed method and the new EM algorithm work quite well in practice.

In this paper, we have considered the univariate and bivariate cases. It is important to see how it can be generalized to the multivariate case. More work is needed along this direction.

Acknowledgements:

The authors would like to thank the reviewers and the associate editor for their constructive comments which have helped us to improve the manuscript significantly.

Appendix

PROOF OF LEMMA 1:

It is enough to show that the function

$$h(\alpha|p) = \ln\left[(1-p^{j+1})^{\alpha} - (1-p^{j})^{\alpha}\right],$$

for $0 < p < 1$, and for any $j = 0, 1, \ldots$, is a log-concave function of α. By straight forward calculation it can be seen that

$$\frac{d^2}{d\alpha^2} h(\alpha|p) = \frac{h_1(\alpha|p)}{h_2(\alpha|p)} < 0,$$

where

$$\begin{aligned} h_1(\alpha|p) &= -(1-p^{j})^{\alpha}(1-p^{j+1})^{\alpha}\left\{\ln(1-p^{j+1}) - \ln(1-p^{j})\right\}^2, \\ h_2(\alpha|p) &= \left\{(1-p^{j+1})^{\alpha} - (1-p^{j})^{\alpha}\right\}^2. \end{aligned}$$

∎

References

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

[2] Al-Hussaini, E.K. and Ahsanullah, M. (2015), *Exponentiated Distributions*, Atlantis Press, Paris, France.

[3] Balakrishnan, N. and Lai, C-D (2009), *Continuous bivariate distributions*, Springer, New York.

[4] Bartoszewicz, J. (2001), "Stochastic comparisons of random minima and maxima from life distributions", *Statistics and Probability Letters*, vol. 55, 107V112.

[5] Casella, G. and Berger, R. L. (2001). Statistical Inference (2nd. ed.), Duxbury.

[6] Gómez-Déniz, E. (2010), "Another generalization of the geometric distribution", *Test*, vol. 19, 399-415.

[7] Gupta, R.D. and Kundu, D. (1999), "Generalized exponential distributions", *Australian and New Zealand Journal of Statistics*, vol. 41, 173-188.

[8] Gupta, R.D. and Kundu, D. (2007), "Generalized exponential distribution: existing results and some recent developments", *Journal of the Statistical Planning and Inference*, vol. 137, 3537 - 3547.

[9] Johnson, N.L., Kemp, A.W. and Kotz, S. (2005), *Univariate discrete distributions*, 3rd. edition, Wiley & Sons, New York.

[10] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1997), *Discrete multivariate distributions*, Wiley and Sons, New York.

[11] Kocherlakota, S. and Kocherlakota, K. (1992), *Bivariate discrete distributions*, Marcel and Dekker, New York.

[12] Kundu, D. (2015), "Bivariate Geometric (Maximum) Generalized Exponential Distribution", *Journal of Data Sciences*, vol. 13, 693 - 712.

[13] Kundu, D. and Gupta, A.K. (2014), "On bivariate Weibull Geometric distribution", *Journal of Multivariate Analysis*, vol. 123, 19 - 29.

[14] Lee, H. and Cha, J.H. (2015), "On two general classes of discrete bivariate distributions", *American Statistician*, vol. 69, 221 – 230.

[15] Louzada, F., Marchi, V.A.A. and Roman, M. (2014), "The exponentiated exponential-geometric distribution: with decreasing, increasing and unimodal failure rate", *Statistics*, vol. 48, 167 - 181.

[16] Marshall, A.W., Olkin, I., (1997), "A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families", *Biometrika*, vol. 84, 641 - 652.

[17] Nadarajah, S. (2011), "The exponentiated exponential distribution; a survey", *Advances in Statistical Analysis*, vol. 95, 219 - 251.

[18] Nekoukhou, V., Alamatsaz, M.H. and Bidram, H. (2013), "Discrete generalized exponential distribution of a second type", *Statistics*, vol. 47, 876 - 887.

[19] Nekoukhou, V. and Kundu, D. (2017), "Bivariate discrete generalized exponential distribution", *Statistics*, DOI:10.1080/02331888.2017.1289534.

[20] Ristic, M.M. and Kundu, D. (2015), "Marshall-Olkin generalized exponential distribution", *Metron*, vol. 73, 317-333.

[21] Ristic, M.M. and Kundu, D. (2016), "Generalized Exponential Geometric Extreme Distributions", *Journal of Statistical Theory and Practice*, vol. 10, no. 1, 179 - 201.

[22] Self, S.G. and Liang, K-Y. (1987), "Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under non-standard condition", *Journal of the American Statistical Association*, vol. 82, 605–610.

[23] Shaked, M. and Wong, T. (1997), "Stochastic comparisons of random minima and maxima", *Journal of Applied Probabability*, vol. 34, 420–425.

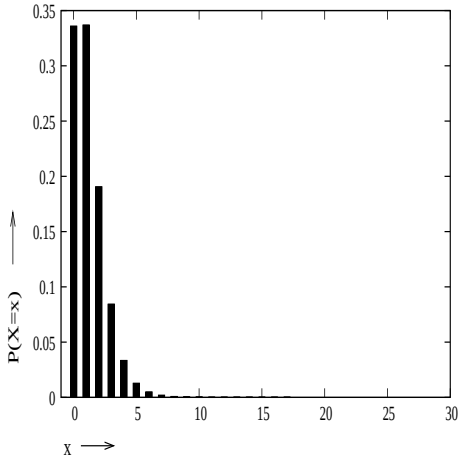

Figure 1: The PMF of a univariate GDGE distribution when $\alpha = 1.5$, $\theta = 0.5$, $p = e^{-1.0}$.

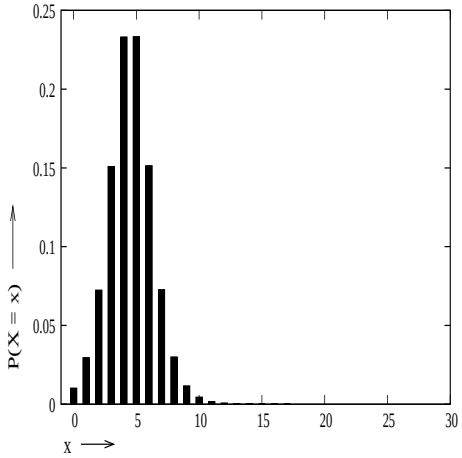

Figure 2: The PMF of a univariate GDGE distribution when $\alpha = 1.5$, $\theta = 0.01$, $p = e^{-1.0}$.

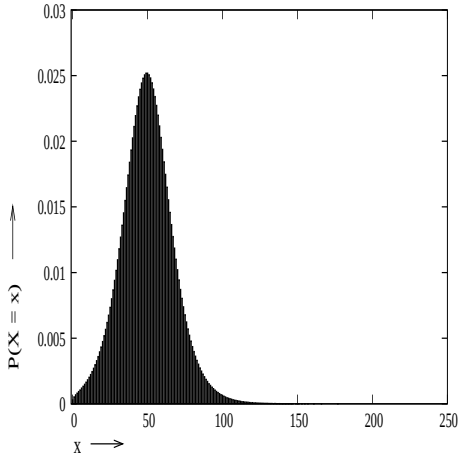

Figure 3: The PMF of a univariate GDGE distribution when $\alpha = 1.5$, $\theta = 0.01$, $p = e^{-0.1}$.