International Journal of
Mathematical Modelling & Computations
Vol. 11, No. 04, Fall 2021, 1- 16

The Bivariate Modified Exponential Geometric Distribution: Model, Properties, and Applications

A. R. Zanboori[a], K. Zare[a,*] and Z. Khodadadi[a]

[a]*Department of Statistics, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran.*

Abstract. In this paper, we have introduced a five-parameter bivariate model by taking a geometric minimum of the modified exponential distributions. It is observed that the maximum likelihood estimators of the unknown parameters cannot be obtained in closed form. We propose to use the EM algorithm to compute the maximum likelihood estimators of the unknown parameters. Several simulation experiments have been performed to determine the effectiveness of the proposed EM algorithm. We analyzed two datasets for illustrative purposes, and it is observed that the proposed models and the expectation-maximization algorithm perform at a satisfactory level.

Received: 21 February 2021, Revised: 03 May 2021, Accepted: 22 May 2021.

Keywords: Expectation-Maximization algorithm, Geometric minimum, Maximum likelihood estimation, Bivariate model.

Index to information contained in this paper

1. Introduction

In 1997, Marshall and Olkin developed a general method to introduce a new model by adding an extra parameter to a family of distributions, and they discussed the exponential and Weibull families in detail. Many researchers have investigated the same approach for different other distributions. For instance, Ghitany et al. [7] investigated the properties of a new parametric distribution generated by Marshall and Olkin's [14] extended family of distributions based on the Baldwin–Lomax model. Moreover, Barreto-Souza et al. [2] provided general expansions for the density function, explicit expressions for the moments, and moments of order statistics for the Marshall-Olkin family of distributions. In addition, Ristić and Kundu [18] adopted the Marshall-Olkin approach to introduce the two-parameter generalized exponential distribution. Alizadeh et al. [1] introduced the Kumaraswamy Marshal-Olkin generalized family of distributions using an extension of the Marshall- Olkin family and the Kumaraswamy distribution as a baseline distribution. Using the Marshall-Olkin extended method, Tomy and Gillariose [20] introduced a new class of continuous models, i.e., the Marshall-Olkin extended-inverted Kumaraswamy distribution.

Moreover, Marshall and Olkin [14] introduced a class of distributions, which can be obtained by the minimum or maximum of a sequence of independent and identically distributed (i.i.d.) continuous random variables, where the sample size follows a geometric distribution. Similarly, the modified exponential-Poisson distribution is considered by Preda et al. [17] through a compounding operation using the Marshall-Olkin approach.

*Corresponding author. Email: karim.zare@yahoo.com

http://ijm2c.iauctb.ac.ir

Louzada et al. [12] proposed a new family of distributions, such as the exponentiated exponential-geometric distribution. Bordbar and Nematollahi [4] introduced a modified exponential-geometric distribution and showed that the failure rate of the new distribution could decrease or increase.

Along with introducing the bivariate extension, Marshall and Olkin [14] discussed several exciting properties of the general model. Unfortunately, they did not discuss any estimation procedures for the unknown parameters of the bivariate model, mainly due to the analytical intractability of the general model. Kundu and Gupta [10] and Kundu [9] applied that method for the bivariate Weibull and the bivariate generalized exponential distributions, respectively. Nekoukhou et al. [16] evaluated a three-parameter bivariate distribution obtained by taking the geometric minimum of Rayleigh distributions. The main objective of the current paper is to consider the bivariate modified exponential geometric (BMEG) distribution, which can be obtained by taking a geometric minimum of the modified exponential distributions.; Its marginals include the univariate modified exponential distributions introduced by Bordbar and Nematollahi [4]. The BMEG distribution has five parameters. Due to the presence of these five parameters, the BMEG distribution is a very flexible bivariate distribution. Since the maximum likelihood estimators (MLEs) of the unknown parameters of the BMEG distribution cannot be obtained in closed form, we have proposed to use the EM algorithm to compute the MLEs of the unknown parameters. The rest of this paper is organized as follows:

In Section 2, we will provide the BMEG distribution and discuss its marginals. The estimation of the unknown parameters using the EM algorithm and the statistical inference is provided in Section 3. The results of the simulation experiments and the analyses of the two datasets are presented in Sections 4. Finally, we will conclude the paper in Section 5.

2. Bivariate modified exponential-geometric distribution (minimum)

Consider $G\,(t;\Omega)$ as the representation of the cumulative distribution function (CDF) of a continuous random variable T, which depends on a parameter vector $\Omega = (\omega 1, \dots, \omega r)$. Then according to Marshall and Olkin [14], the corresponding Marshal-Olkin extended distribution would have a cumulative distribution function defined by the following formula:

$$F(t;\alpha,\mathbf{\Omega}) = \frac{G(t;\mathbf{\Omega})}{G(t;\mathbf{\Omega}) + \alpha(1 - G(t;\mathbf{\Omega}))}, \qquad -\infty < t < +\infty,\;\; \alpha > 0. \tag{1}$$

According to this parameterization scheme, the random variable X is concluded to follow a modified exponential distribution with parameters α and β if the CDF of X is as described below:

$$F_{ME}(x;\alpha,\beta) = \frac{1 - e^{-\beta x}}{1 - (1-\alpha)e^{-\beta x}}, \qquad x > 0\,, \alpha,\beta > 0, \tag{2}$$

and 0 in other cases. It will be denoted by $\mathrm{ME}(\alpha,\beta)$ with the survival distribution function as illustrated below:

$$\bar{F}_{ME}(x;\alpha,\beta) = \frac{\alpha\beta \mathrm{e}^{-\beta \mathrm{x}}}{1 - (1-\alpha)\mathrm{e}^{-\beta \mathrm{x}}}, \quad x > 0, \qquad \alpha,\beta > 0, \tag{3}$$

and the probability density function (PDF) is as follows:

$$f_{ME}(x;\alpha,\beta) = \frac{\alpha\beta \mathrm{e}^{-\beta \mathrm{x}}}{(1 - (1-\alpha)\mathrm{e}^{-\beta \mathrm{x}})^2}. \tag{4}$$

Suppose $\{X_i : i = 1,2,\dots\}$ and $\{Y_j : j = 1,2,\dots\}$ are two sequences of independent and identically distributed (i.i.d.) random variables with common distributions $\mathrm{ME}(\alpha_1, \beta_1)$ and $\mathrm{ME}(\alpha_2, \beta_2)$, respectively. It is also assumed that $X_i's$ and $Y_j's$ are independent. Consider N is a geometric random variable with probability mass function of $P(N = n) = p(1 - p)^{n-1}$; for $n \in \{1,2, \dots\}$ and $0 < p < 1$. From now on, it will be denoted by $\mathrm{GM(p)}$. Moreover, N is independent of $X_i's$ and $Y_j's$. Consider two new random variables X and Y, in such a way that

$$X = \min\{X_1, \dots, X_N\}, \qquad Y = \min\{Y_1, \dots, Y_N\}. \tag{5}$$

Then we could say that the bivariate vector (X, Y) has the Bivariate Modified Exponential Geometric (BMEG) distribution, with parameters $\mathbf{\Theta} = (\alpha_1, \alpha_2, \beta_1, \beta_2, p)$ and will be denoted by $BMEG(\alpha_1, \alpha_2, \beta_1, \beta_2, p)$.
The joint survival function (X, Y) is as follows:

$$\begin{aligned}\bar{F}_{BMEG}(x, y; \mathbf{\Theta}) &= P(X > x, Y > y) \\ &= \sum_{n=1}^{\infty} P(X > x, Y > y | N = n) P(N = n) \\ &= \sum_{n=1}^{\infty} \bar{F}_{ME}^{n}(x; \alpha_1, \beta_1) \bar{F}_{ME}^{n}(y; \alpha_2, \beta_2) p(1-p)^{n-1} \\ &= p\bar{F}_{ME}(x; \alpha_1, \beta_1) \bar{F}_{ME}(y; \alpha_2, \beta_2) \\ &\quad \times \sum_{n=1}^{\infty} \left(\bar{F}_{ME}(x; \alpha_1, \beta_1) \bar{F}_{ME}(y; \alpha_2, \beta_2)(1-p)\right)^{n-1} \\ &= \frac{p\bar{F}_{ME}(x; \alpha_1, \beta_1) \bar{F}_{ME}(y; \alpha_2, \beta_2)}{1 - (1-p)\bar{F}_{ME}(x; \alpha_1, \beta_1) \bar{F}_{ME}(y; \alpha_2, \beta_2)}.\end{aligned} \tag{6}$$

Therefore, the joint survival function of (X, Y) is as follows:

$$\bar{F}_{BMEG}(x, y; \mathbf{\Theta}) = \frac{p\alpha_1\alpha_2 e^{-\beta_1 x} e^{-\beta_2 y}}{(1-(1-\alpha_1)e^{-\beta_1 x})(1-(1-\alpha_2)e^{-\beta_2 y}) - \alpha_1\alpha_2(1-p)e^{-\beta_1 x}e^{-\beta_2 y}}. \tag{7}$$

Using (6), the joint PDF of (X, Y) can be obtained as $f_{X,Y}(x, y) = \partial^2 \bar{F}_{X,Y}(x, y)/\partial x \partial y$ and it is

$$\begin{aligned}&f_{BMEG}(x, y; \mathbf{\Theta}) \\ &= \frac{pf_{ME}(x; \alpha_1, \beta_1) f_{ME}(y; \alpha_2, \beta_2)(1 + (1-p)\bar{F}_{ME}(x; \alpha_1, \beta_1)\bar{F}_{ME}(y; \alpha_2, \beta_2))}{(1 - (1-p)\bar{F}_{ME}(x; \alpha_1, \beta_1)\bar{F}_{ME}(y; \alpha_2, \beta_2))^3};\end{aligned} \tag{8}$$

then the joint PDF of (X, Y) is as follows:

$$\begin{aligned}&f_{BMEG}(x, y; \mathbf{\Theta}) = p\alpha_1\alpha_2\beta_1\beta_2 e^{-\beta_1 x} e^{-\beta_2 y} \\ &\times \frac{\left(1 - (1-\alpha_1)e^{-\beta_1 x}\right)(1 - \alpha_2)e^{-\beta_2 y}) + (1-p)(1 - e^{-\beta_1 x})(1 - e^{-\beta_2 y})}{((1 - (1-\alpha_1)e^{-\beta_1 x})(1 - (1-\alpha_2)e^{-\beta_2 y}) - (1-p)(1 - e^{-\beta_1 x})(1 - e^{-\beta_2 y})^3}.\end{aligned} \tag{9}$$

Regarding (8), when $p = 1$, we have the following result:

$$f_{BMEG}(x, y) = f_{ME}(x; \alpha_1, \beta_1) f_{ME}(y; \alpha_2, \beta_2), \tag{10}$$

meaning that X and Y are independent. The parameter p can thus be taken into account as the correlation coefficient. We have provided the PDF with contour plots of the BMEG distribution for different parameter values in Figure 1. The joint density function of (X, Y, N) is given by

$$f_{X,Y,N}(x,y,n;\mathbf{\Theta}) = n^2 f_{ME}(x;\alpha_1,\beta_1)\bar{F}_{ME}^{n-1}(x;\alpha_1,\beta_1) f_{ME}(y;\alpha_2,\beta_2) \times \bar{F}_{ME}^{n-1}(y;\alpha_2,\beta_2) p(1-p)^{n-1} \tag{11}$$

Now using the joint density function of (X,Y,N) and that of (X,Y) the conditional distribution of N (X,Y) is as follows:

$$\begin{aligned} &P(N=n|X=x,Y=y;\mathbf{\Theta}) \\ &= \frac{f_{X,Y,N}(x,y,n;\mathbf{\Theta})}{f_{BMEG}(x,y;\mathbf{\Theta})} \\ &= n^2(1-p)^{n-1}\bar{F}_{ME}^{n-1}(x;\alpha_1,\beta_1)\bar{F}_{ME}^{n-1}(y;\alpha_2,\beta_2) \\ &\qquad \times \frac{(1-(1-p)\bar{F}_{ME}(x;\alpha_1,\beta_1)\bar{F}_{ME}(y;\alpha_2,\beta_2))^3}{(1+(1-p)^{n-1}\bar{F}_{ME}(x;\alpha_1,\beta_1)\bar{F}_{ME}(y;\alpha_2,\beta_2))}. \end{aligned} \tag{12}$$

Set $g(x,y;\mathbf{\Theta}) = (1-p)\bar{F}_{ME}(x;\alpha_1,\beta_1)\bar{F}_{ME}(y;\alpha_2,\beta_2)$, then the conditional distribution of N, given (X,Y), becomes:

$$P(N=n|X=x,Y=y;\mathbf{\Theta}) = n^2 g^{n-1}(x,y;\mathbf{\Theta})\ \frac{(1-g(x,y;\mathbf{\Theta}))^3}{1+g(x,y;\mathbf{\Theta})} \tag{13}$$

In view of the fact that if $N \sim GM(1-r)$ then $E(N^3) = (r^2+4r+1)/(1-r)^3$, we consequently obtain the following formula:

$$\begin{aligned} E(N|X=x,Y=y;\mathbf{\Theta}) &= \sum_{n=1}^{\infty} n\,P(N=n|X=x,Y=y) \\ &= \textstyle\sum_{n=1}^{\infty} n^3 g^{n-1}(x,y;\mathbf{\Theta})\,\frac{\left(1-g(x,y;\mathbf{\Theta})\right)^3}{1+g(x,y;\mathbf{\Theta})} \\ &= \frac{\left(1-g(x,y;\mathbf{\Theta})\right)^3}{\left(1+g(x,y;\mathbf{\Theta})\right)\left(1-g(x,y;\mathbf{\Theta})\right)} \times \\ &\qquad \textstyle\sum_{n=1}^{\infty} n^3\left(1-g(x,y;\mathbf{\Theta})\right) g^{n-1}(x,y;\mathbf{\Theta}) \\ &= \frac{\left(1-g(x,y;\mathbf{\Theta})\right)^3}{\left(1+g(x,y;\mathbf{\Theta})\right)\left(1-g(x,y;\mathbf{\Theta})\right)} \\ &\qquad \times \frac{g^2(x,y;\mathbf{\Theta})+4g(x,y;\mathbf{\Theta})+1}{\left(1-g(x,y;\mathbf{\Theta})\right)^3} \\ &= \frac{g^2(x,y;\mathbf{\Theta})+4g(x,y;\mathbf{\Theta})+1}{(1-g^2 g(x,y;\Theta))}, \end{aligned} \tag{14}$$

where $E(N|X=x,Y=y)$ will be used for developing the EM algorithm.

As a result of the Marshall and Olkin distribution [14], the family of distributions of (3) is the geometric minimum stability, and the BMEG distribution is also closed under the geometric minimum. We thus have the following results:

Proposition 1. If $(X,Y) \sim BMEG(\alpha_1,\alpha_2,\beta_1,\beta_2,p)$, then

i) $X \sim ME(\alpha_1 p,\beta_1)$
ii) $Y \sim ME(\alpha_2 p,\beta_2)$

Proof. We will only prove part (*i*); part (*ii*) can be similarly obtained. To obtain the survival function of random variable X, it is enough to find $\bar{F}_{X,Y}(x,0)$. Using (6), we have the following formula:

$$\bar{F}_X(x) = \bar{F}_{X,Y}(x,0) = \frac{p\bar{F}_{ME}(x;\alpha_1,\beta_1)}{1-(1-p)\bar{F}_{ME}(x;\alpha_1,\beta_1)} = \frac{p\alpha_1 e^{-\beta_1 x}}{1-(1-\alpha_1 p)e^{-\beta_1 x}}. \tag{15}$$

Comparison of survival functions of (15) and (3), yields $X \sim ME(\alpha_1 p, \beta_1)$. ■

Proposition 2. Consider $(U_1,V_1),(U_2,V_2),....$ as a sequence of independent and identically-distributed random vectors with distribution $BMEG(\alpha_1,\alpha_2,\beta_1,\beta_2,p)$. Let $N \sim GM(q)$, and suppose that N and $(U_i,V_i), i=1,2,....$ are independent. Also, define the two new random variables:

$$U = \min\{U_1,U_2,....\} \ \ and \ \ V = \min\{V_1,V_2,...\}; \tag{16}$$

then $(U,V) \sim BMEG(\alpha_1,\alpha_2,\beta_1,\beta_2,pq)$.

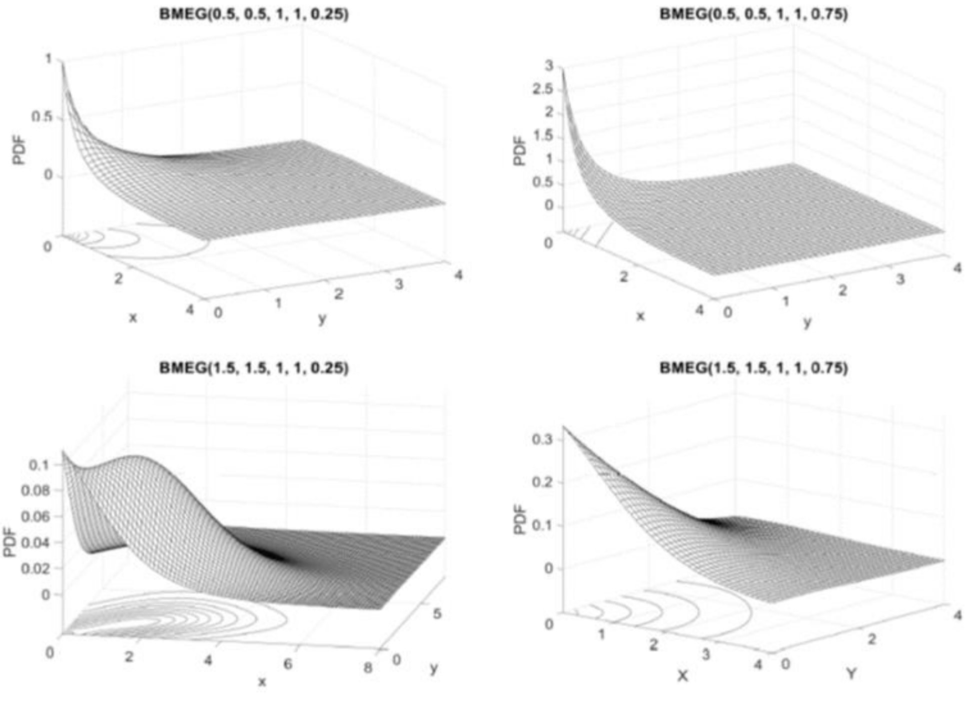

Figure 1. PDF with contour plots of BMEG($\alpha_1,\alpha_2,\beta_1,\beta_2$,$p$) distribution.

Proof The joint survival function (U,V) is written as follows:

$$\begin{aligned}
\bar{F}_{U,V}(u,v) &= p(\min\{U_1,U_2,...U_N\} > u, \min\{V_1,V_2,...\ V_n\} > v) \\
&= \sum_{n=1}^{\infty}\prod_{i=1}^{n} P(\,U_i > u\,, V_i > n\ |N=n)q(1-q)^{n-1} \\
&= \sum_{n=1}^{\infty} \bar{F}_{BMEG}^{n}\ (u,v;\boldsymbol{\Theta})q(1-q)^{n-1}
\end{aligned}$$

$$= q\bar{F}_{BMEG}(u,v;\mathbf{\Theta})\sum_{n=1}^{\infty}(\bar{F}_{BMEG}(u,v;\mathbf{\Theta})\,(1-q))^{n-1}$$
$$= \frac{q\bar{F}_{BMEG}(u,v;\mathbf{\Theta})}{1-(1-q)q\bar{F}_{BMEG}(u,v;\mathbf{\Theta})} \tag{17}$$
$$= \frac{pq\bar{F}_{ME}(u;\alpha_1,\beta_1)\,\bar{F}_{ME}(v;\alpha_2,\beta_2)}{1-(1-pq)\bar{F}_{ME}(u;\alpha_1,\beta_1)\,\bar{F}_{ME}(v;\alpha_2,\beta_2)}$$

Comparison of the joint survival functions of formulas (17) and (6) yield $(U,V)\sim BMEG(\alpha_1,\alpha_2,\beta_1,\beta_2,pq)$. ∎

According to Basu [3], the scalar hazard rate function for a bivariate random variable (X,Y) with the joint PDF $f(x,y)$, and the joint survival function $\bar{F}(x,y)$ is defined as:

$$h_B(x,y) = \frac{f(x,y)}{\bar{F}(x,y)}. \tag{18}$$

The scalar hazard rate function defined in (18) does not uniquely define the joint PDF. Johnson and Kotz [8] introduced a joint bivariate hazard rate function as follows:

$$h(x,y) = \big(h_1(x,y),h_2(x,y)\big) = \left(\frac{-\partial\log\bar{F}(x,y)}{\partial x},\frac{-\partial\log\bar{F}(x,y)}{\partial y}\right). \tag{19}$$

According to Marshall [13], the bivariate hazard function $h(x,y)$ uniquely determines the joint PDF. Now, the scalar hazard rate function of the $BMEG(\alpha_1,\alpha_2,\beta_1,\beta_2,p)$ distribution using (18), is:

$$h_{BMEG}(x,y) = \frac{f_{BMEG}(x,y;\mathbf{\Theta})}{\bar{F}_{BMEG}(x,y;\mathbf{\Theta})}. \tag{20}$$

Figure 2 shows the scalar hazard rate functions of BMEG distributions. The joint hazard rate function of a $BMEG(\alpha_1,\alpha_2,\beta_1,\beta_2,p)$ distribution is as follows:

$$\big(h_1(x,y),h_2(x,y)\big) = \left(-\frac{\partial}{\partial x},-\frac{\partial}{\partial y}\right)\log\bar{F}_{BMEG}(x,y;\mathbf{\Theta}). \tag{21}$$

The plots of the joint hazard rate function of a $BMEG(\alpha_1,\alpha_2,\beta_1,\beta_2,p)$ distribution are illustrated in Figure 3.

3. Maximum likelihood estimation

This section demonstrates how EM-type algorithms are employed for estimating unknown parameters of BMEG distributions. Consider $(x_1,y_1),\dots,(x_m,y_m)$ as a bivariate sample of size m from BMEG with parameters $\mathbf{\Theta}=(\alpha_1,\alpha_2,\beta_1,\beta_2,p)$. The log-likelihood function based on the observation is as follows:

$$l(\mathbf{\Theta}) = \log\prod_{i=1}^{n} f_{BMEG}(x_i,y_i;\mathbf{\Theta})$$
$$= \log\prod_{i=1}^{n}(p\alpha_1\alpha_2\beta_1\beta_2\,e^{-\beta_1 x_i}e^{-\beta_2 y_i} \tag{22}$$
$$\times\frac{(1-(1-\alpha_1)e^{-\beta_1 x_i}\big(1-(1-\alpha_2)e^{-\beta_2 y_i}\big)+(1-p)(1-e^{-\beta_1 x_i})(1-e^{-\beta_2 y_i})}{((1-(1-\alpha_1)e^{-\beta_1 x_i}\big(1-(1-\alpha_2)e^{-\beta_2 y_i}\big)+(1-p)\big(1-e^{-\beta_1 x_i}\big)\big(1-e^{-\beta_2 y_i}\big))^3}).$$

Hence, the MLEs of the unknown parameters can be obtained by maximizing (22) with respect to the unknown parameters. MLEs cannot be obtained in closed forms. Since the proposed BMEG model has five parameters, the MLEs of the unknown parameters can be obtained through solving a five-dimensional optimization problem. Therefore, numerical methods can be applied both to solve non-linear equations and to obtain the MLEs.

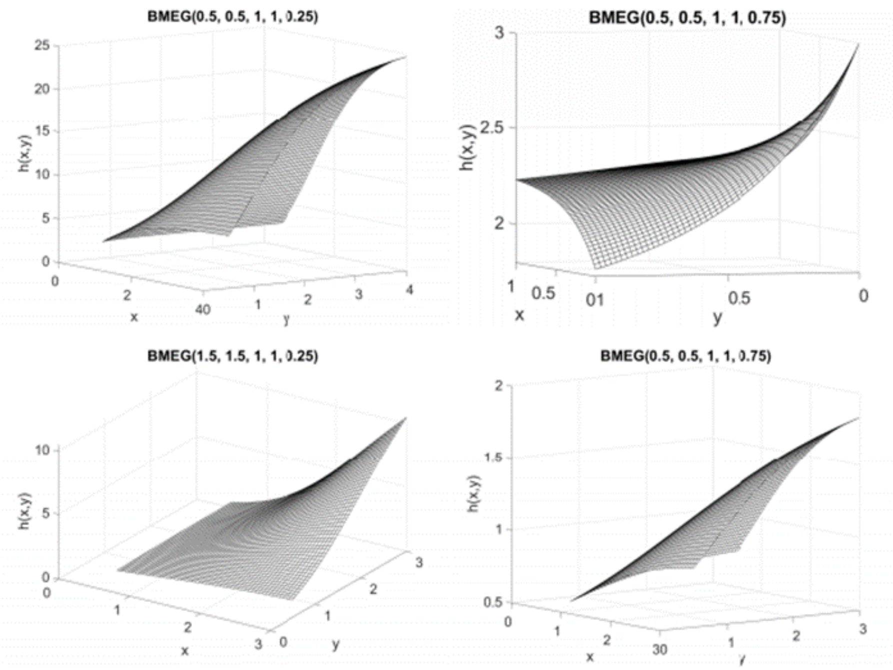

Figure 2. The scalar hazard rate function plots of $BMEG(\alpha_1, \alpha_2, \beta_1, \beta_2, p)$ distributions.

To avoid solving the five-dimensional optimization problem, we propose to use the EM-type algorithm to compute the MLEs. The EM algorithm is a well-known technique for ML estimation when unobserved (or missing) data or latent variables are present in the process of modeling. For more detail, please refer to McLachlan and Krishnan [15].

To pose this model to an incomplete data problem, it is conceivable to introduce the hypothetical random variable $N_1, \dots, N_m$ corresponding to the random vectors $(x_1, y_1), \dots, (x_m, y_m)$. According to (11), based on observed data $(x_1, y_1), \dots, (x_m, y_m)$ and latent data $n_1, \dots, n_m$the complete data log-likelihood function of the unknown parameters $\Theta = (\alpha_1, \alpha_2, \beta_1, \beta_2, p)$, given $(x_i, y_i, n_i), i = 1, \dots, m$ ignoring constant, is as follows:

$$\begin{aligned} l_c(\boldsymbol{\Theta}) &= m \log \alpha_1 + m \log \beta_1 - \beta_1 \textstyle\sum_{i=1}^{m} \log(1 - (1 - \alpha_1) e^{-\beta_1 x_i}) \\ &\quad + \textstyle\sum_{i=1}^{m} (n_i - 1)(\log \alpha_1 e^{-\beta_1 x_i}) - \log(1 - \alpha_1) e^{-\beta_1 x_i})) \\ &\quad + m \log \alpha_2 + m \log \beta_2 - \beta_2 \textstyle\sum_{i=1}^{m} \log\big(1 - (1 - \alpha_2) e^{-\beta_2 y_i}\big) \\ &\quad + \textstyle\sum_{i=1}^{m} (n_i - 1)(\log \alpha_1 e^{-\beta_2 y_i}) - \log\big(1 - (1 - \alpha_2) e^{-\beta_2 y_i}\big)) \\ &\quad + m \log p + \textstyle\sum_{i=1}^{m} (n_i - 1) \log(1 - p) . \end{aligned} \tag{23}$$

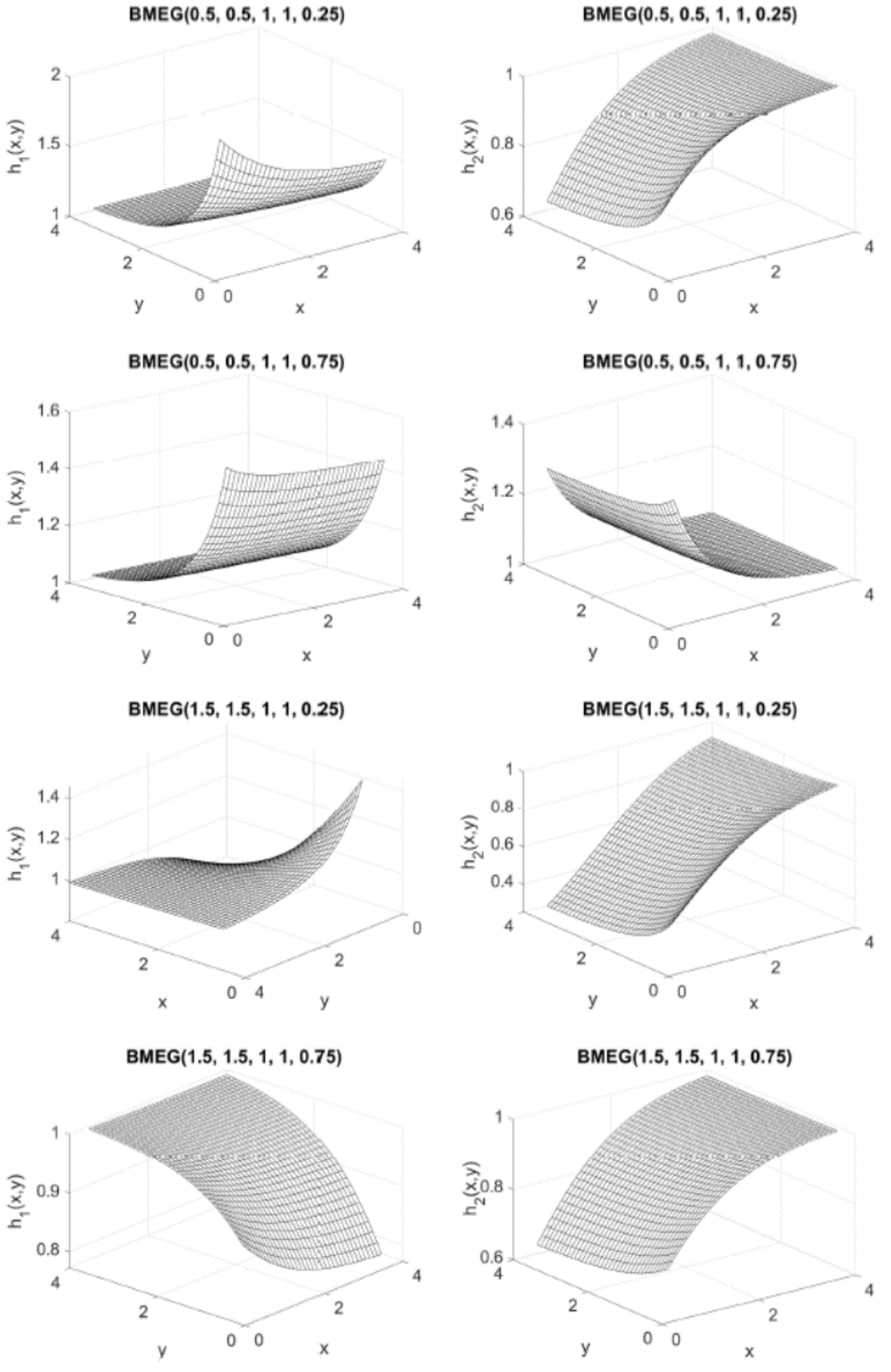

Figure 3. The joint hazard rate function of a BMEG $(\alpha_1, \alpha_2, \beta_1, \beta_2, p)$ distribution.

Using (23), given the current estimate of $\widehat{\mathbf{\Theta}}^{(k)} = (\hat{\alpha}_1{}^{(k)}, \hat{\alpha}_2{}^{(k)}, \hat{\beta}_1{}^{(k)}, \hat{\beta}_2{}^{(k)}, \hat{p}^{(k)})$ at the k^{th} iteration, the expected complete data log-likelihood function or the Q-function as asserted in Dempster et al. [6] is as follows:

$$
\begin{aligned}
Q\left(\mathbf{\Theta}\middle|\widehat{\mathbf{\Theta}}^{(k)}\right) = {} & m\log\alpha_1 + m\log\beta_1 - \beta_1\sum_{i=1}^{m} x_i - 2\sum_{i=1}^{m}\log\left(1-(1-\alpha_1)e^{-\beta_1 x_i}\right) \\
& + \sum_{i=1}^{m}\left(\eta_i^{(k)} - 1\right)(\log\left(\alpha_1 e^{-\beta_1 x_i}\right) - \log(1-(1-\alpha_1)e^{-\beta_1 x_i})) \\
& + m\log\alpha_2 + m\log\beta_2 - \beta_2\sum_{i=1}^{m} y_i - 2\sum_{i=1}^{m}\log(1-(1-\alpha_2)e^{-\beta_2 y_i})) \\
& + \sum_{i=1}^{m}\left(\eta_i^{(k)} - 1\right)(\log\left(\alpha_2 e^{-\beta_2 x_i}\right) - \log(1-(1-\alpha_2)e^{-\beta_2 x_i})) \\
& + m\log p + \sum_{i=1}^{m}\left(\eta_i^{(k)} - 1\right)\log(1-p).
\end{aligned}
$$

where

$$
\begin{aligned}
\eta_i^{(k)} &= E\left(N_i\middle|X_i = x_i, Y_i = y_i;\ \widehat{\Theta}^{(k)}\right) \\
&= \frac{g^2\left(x,y;\widehat{\Theta}^{(k)}\right) + 4g\left(x,y;\widehat{\Theta}^{(k)}\right) + 1}{(1-g^2\left(x,y;\widehat{\Theta}^{(k)}\right))}.
\end{aligned} \tag{25}
$$

In summary, the implementation of the EM-type (ECME) algorithm is as follows:

E-step: Given $\Theta = \widehat{\Theta}^{(k)}$ compute $\eta_i^{(k)}$ using (25) for $i = 1, \dots, m$.

CM-step 1: Update $\hat{p}^{(k)}$ by maximizing (24) over p, which leads to

$$
\hat{p}^{(k+1)} = \frac{m}{\sum_{i=1}^{m}\eta_i^{(k)}}. \tag{26}
$$

CM-step 2: Obtain $\hat{\alpha}_1{}^{(k+1)}$ as the solution of

$$
\begin{aligned}
\frac{m}{\alpha_1} - 2\sum_{i=1}^{m}\frac{e^{-\hat{\beta}_1{}^{(k)}x_i}}{1-(1-\alpha_1)e^{-\hat{\beta}_1{}^{(k)}x_i}} + \sum_{i=1}^{m}(\eta_i^{(k)} \\
- 1)\left(\frac{1}{\alpha_1} - \frac{e^{-\hat{\beta}_1{}^{(k)}x_i}}{1-(1-\alpha_1)e^{-\hat{\beta}_1{}^{(k)}x_i}}\right) = 0
\end{aligned} \tag{27}
$$

with respect to α_1,

CM-step 3: Obtain $\hat{\beta}_1{}^{(k+1)}$ as the solution of

$$
\begin{aligned}
\frac{m}{\beta_1} - \sum_{i=1}^{m} x_i - 2\sum_{i=1}^{m}\frac{\left(1-\hat{\alpha}_1{}^{(k+1)}\right)x_i e^{-\beta_1 x_i}}{1-\left(1-\hat{\alpha}_1{}^{(k+1)}\right)x_i e^{-\beta_1 x_i}} \\
+ \sum_{i=1}^{m}\left(\eta_i^{(k)} - 1\right)\left(-x_i - \frac{\left(1-\hat{\alpha}_1{}^{(k+1)}\right)x_i e^{-\beta_1 x_i}}{1-\left(1-\hat{\alpha}_1{}^{(k+1)}\right)x_i e^{-\beta_1 x_i}}\right) = 0
\end{aligned} \tag{28}
$$

with respect to β_1.

CM-step 4: Obtain $\hat{\alpha}_2{}^{(k+1)}$ as the solution of

$$\frac{m}{\alpha_2} - 2\sum_{i=1}^{m} \frac{e^{-\hat{\beta}_2{}^{(k)} y_i}}{1-(1-\alpha_2)e^{-\hat{\beta}_2{}^{(k)} y_i}} + \sum_{i=1}^{m} (\eta_i^{(k)} - 1) \times \left(\frac{1}{\alpha_2} - \frac{e^{-\hat{\beta}_2{}^{(k)} y_i}}{1-(1-\alpha_2)e^{-\hat{\beta}_2{}^{(k)} y_i}} \right) = 0 \tag{29}$$

with respect to α_2.

CM-step 5: Obtain $\hat{\beta}_2{}^{(k+1)}$ as the solution of

$$\frac{m}{\beta_2} - \sum_{i=1}^{m} y_i - 2\sum_{i=1}^{m} \frac{\left(1-\hat{\alpha}_2{}^{(k+1)}\right) y_i e^{-\beta_2 y_i}}{1-\left(1-\hat{\alpha}_2{}^{(k+1)}\right) e^{-\beta_2 y_i}} + \sum_{i=1}^{m} \left(\eta_i^{(k)} - 1\right) \left(-y_i - \frac{\left(1-\hat{\alpha}_2{}^{(k+1)}\right) y_i e^{-\beta_2 y_i}}{1-\left(1-\hat{\alpha}_2{}^{(k+1)}\right) e^{-\beta_2 y_i}} \right) = 0 \tag{30}$$

with respect to β_2.

CM-Steps 2, 3, 4 and 5 require a one-dimensional search for the root of $\alpha_1, \beta_1, \alpha_2$, and β_2, respectively, which can be easily achieved using the MATLAB $f\,solve$ built-in function.
The iterations of the above algorithm are repeated until a suitable convergence rule is satisfied, e.g., the iteration stops when the absolute difference between the two consecutive log-likelihood values is less than 10^{-4} .

4. Simulation and data analysis

4.1 *Simulation study*

This section provides the results of the simulation study. Some simulation experiments have been performed to investigate how the proposed EM algorithm works in computing the MLEs.
It may be observed that it is simple to use the definition of the model to generate from a BMEG$(\alpha_1, \alpha_2, \beta_1, \beta_2, p)$ distribution. To this end, the following algorithm can be used:

1. Generate n from $GM(p)$
2. Generate $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ from $ME(\alpha_1, \beta_1)$ and $ME(\alpha_2, \beta_2)$, respectively
3. Compute the desired (x, y) as $x = min\{x_1, \ldots, x_n\}$ and $y = min\{y_1, \ldots, y_n\}$

The samples of size m = 100, 200, 500, and 1000 are simulated from the BMEG distributions with parameter $\alpha_1 = 0.2, \alpha_2 = 0.2, \beta_1 = 1, \beta_2 = 1$; $p = 0.25, 0.5, 0.75$, and 0.99 chosen arbitrarily. The process is repeated 1000 times to compute the average estimate (AE) and the mean squared errors (MSE) of the MLEs. The results are reported in Table 1.

Table 1. The AEs (MSEs) of the MLEs based on 1000 simulations of the BMEG distribution.

BMEG	$m = 100$	$m = 200$	$m = 500$	$m = 1000$
$\alpha_1 = 0.2$	0.37666 (0.27771)	0.26158 (0.03563)	0.22678 (0.00864)	0.21336 (0.00374)
$\alpha_2 = 0.2$	0.38748 (0.28435)	0.26624 (0.03645)	0.22480 (0.00837)	0.21292 (0.00395)
$\beta_1 = 1$	1.41938 (0.96642)	1.18956 (0.35947)	1.09412 (0.11719)	1.04009 (0.05024)
$\beta_2 = 1$	1.47413 (1.12661)	1.21158 (0.3839)	1.08233 (0.11328)	1.03752 (0.05562)
$p = 0.025$	0.27588 (0.02089)	0.26402 (0.00941)	0.25616 (0.00374)	0.25134 (0.0017)
$\alpha_1 = 0.2$	0.28661 (0.06429)	0.24486 (0.01824)	0.21763 (0.00482)	0.20790 (0.002)
$\alpha_2 = 0.2$	0.28953 (0.10453)	0.24234 (0.01676)	0.22153 (0.0053)	0.20879 (0.00215)
$\beta_1 = 1$	1.23865 (0.43498)	1.13195 (0.17733)	1.04935 (0.05613)	1.01614 (0.02617)
$\beta_2 = 1$	1.21370 (0.41962)	1.11941 (0.18845)	1.06872 (0.06174)	1.01977 (0.02627)
$p = 0.5$	0.52796 (0.03886)	0.51378 (0.02125)	0.50406 (0.00848)	0.49941 (0.00406)
$\alpha_1 = 0.2$	0.28630 (0.04763)	0.23503 (0.01215)	0.21313 (0.00332)	0.20785 (0.00168)
$\alpha_2 = 0.2$	0.28084 (0.04742)	0.23208 (0.01119)	0.21397 (0.00365)	0.20434 (0.0016)
$\beta_1 = 1$	1.17021 (0.25565)	1.09274 (0.11943)	1.03263 (0.03845)	1.02489 (0.02017)
$\beta_2 = 1$	1.15461 (0.23019)	1.09036 (0.11618)	1.03059 (0.03823)	1.00968 (0.01777)
$p = 0.75$	0.73298 (0.04113)	0.75574 (0.02494)	0.75042 (0.01202)	0.75447 (0.0064)
$\alpha_1 = 0.2$	0.23370 (0.01232)	0.21250 (0.00437)	0.20560 (0.00186)	0.20321 (0.00081)
$\alpha_2 = 0.2$	0.23784 (0.01324)	0.21047 (0.00462)	0.20402 (0.0016)	0.20268 (0.00078)
$\beta_1 = 1$	1.10852 (0.15729)	1.03901 (0.06975)	1.01964 (0.02964)	1.00952 (0.01263)
$\beta_2 = 1$	1.12702 (0.17959)	1.03344 (0.06977)	1.01141 (0.02591)	1.00619 (0.01258)
$p = 0.99$	0.98894 (0.00041)	0.98999 (0.00002)	0.99038 (0.00001)	0.98899 (0.00001)

Some of the points are quite clear in the simulation results: (i) generally, as the sample size increases, the MSEs decrease in each case. These results suggest that the EM estimates have performed consistently. MLEs show better performances as p increases because when $p = 1$, the two marginals are independent. As a result, the model becomes simple.

Based on the simulation results, it can be concluded that the proposed EM algorithm works quite well in small and medium-size samples as well and can be used quite effectively for data analysis purposes.

4.2 *Data analysis*

This section analyzes a data set using the BMEG model for illustrative purposes. We illustrate our proposed methods with a data set obtained from the study conducted by Kundu and Gupta [11]. A sample of size 30 is generated from a singular bivariate modified Sarhan–Balakrishnan (SBVMSB) distribution by Kundu and Gupta [11]. The data are presented in Table 2 for ease of reference.

Table 2. Simulated data from SBVMSB distribution (see [11]).

No.	x	y	No.	x	y	No.	x	y
1	0.4180	0.4240	11	0.3620	0.6450	21	0.9450	0.9450
2	0.1060	0.8510	12	0.2570	1.4640	22	0.8500	0.8500
3	1.1470	1.1470	13	1.6080	1.6080	23	0.3540	0.3540
4	0.5290	1.7950	14	0.6280	0.6280	24	0.3450	2.3080
5	0.4460	0.4460	15	0.3510	0.3510	25	1.1980	0.6200
6	0.3260	0.3260	16	0.8850	0.7910	26	0.5250	0.5040
7	0.2050	0.2050	17	0.0490	0.2000	27	0.5480	0.5480
8	1.1060	1.1060	18	1.0880	1.1280	28	2.8370	1.0570
9	0.4350	0.9730	19	1.4530	1.1550	29	0.2120	1.6970
10	0.9350	0.8500	20	0.8780	0.8780	30	2.3560	1.3480

The BMEG distribution has been fitted to these data. The EM algorithm has been used to compute the MLEs of the unknown parameters. Different starting values have been utilized, which have provided the same estimates in all cases. Figure 4 shows that the initial estimates of the unknown parameters are not problematic regarding the convergence of the EM algorithm.

Table 3 demonstrates the MLEs, the associated 95% bootstrap confidence intervals, the Akaike information criterion (AIC), and estimated log-likelihood value for fitted BMEG and SBVMSB distribution. Based on the results provided in Table 3, it is evident that AIC and log-likelihood are lower for the BMEG distribution as compared with the SBVMSB distributions. Figure 5 displays the scatter plot of simulated data from the SBVMSB (data in Table 1) distribution and probability density contour plot for the fitted BMEG distribution.

Table 3. Summary of results obtained from fitting the BMEG and SBVMSB distributions to data in Table 2.

Model	MLEs (Confidence Intervals)	$\hat{l}$	AIC
BMEG	α_1=2.8622 (2.6617, 3.0626) α_2=12.4245 (11.5142, 13.3349) β_1= 2.0660 (2.0169, 2.1152) β_2=3.0192 (2.9682, 3.0703) p= 0.9517(0.9216, 0.9818)	−41.2054	92.4109
SBV MSB	α_1= 1.5725 (1.0367, 2.8692) α_2= 3.1405 (1.7187,4.7924) α_3= 1.9585 (1.3833,3.8889) λ= 1.0932 (0.7640,1.4650)	−64.8979	137.7958

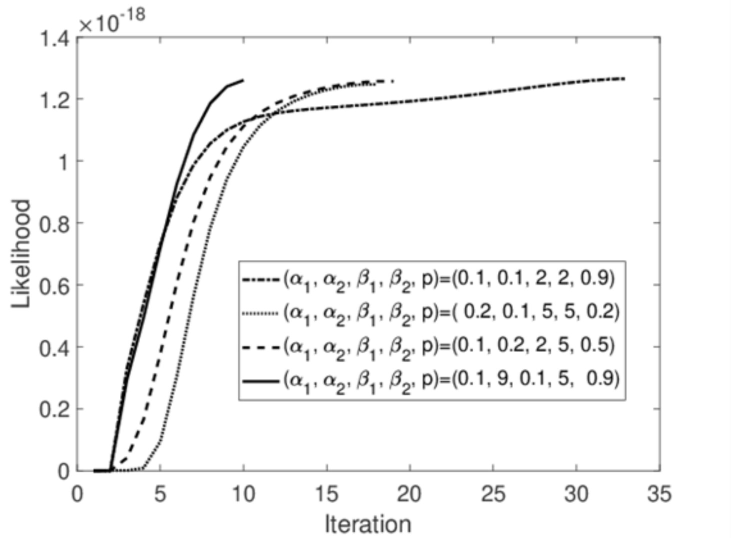

Figure 4. Plots of the iteration number versus likelihood with respect to different initial values for EM algorithm used to obtain MLEs of the BMEG model for data in Table 2.

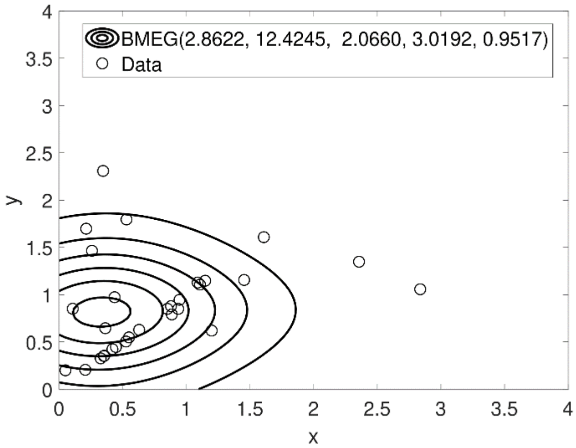

Figure 5. Scatter plot of simulated data from SBVMSB distribution (data of Table 2) and probability density contour plot for fitted the BMEG distribution.

The marginals are tested. The results for the Kolmogorov-Smirnov distances and the associated p-values are reported in Table 4. Regarding Proposition 1, the marginal distribution of X and Y is $ME(\alpha_1 p, \beta_1)$ and $ME(\alpha_2 p, \beta_2)$, respectively. According to the p-values in Table 4, it is clear that ME fits X and Y very well. Due to very high p-values, we cannot reject the null hypothesis that the data come from the BMEG model. However, the data have been generated from the SBVMSB distribution. The plots of the densities and cumulative distribution functions are given in Figure 6.

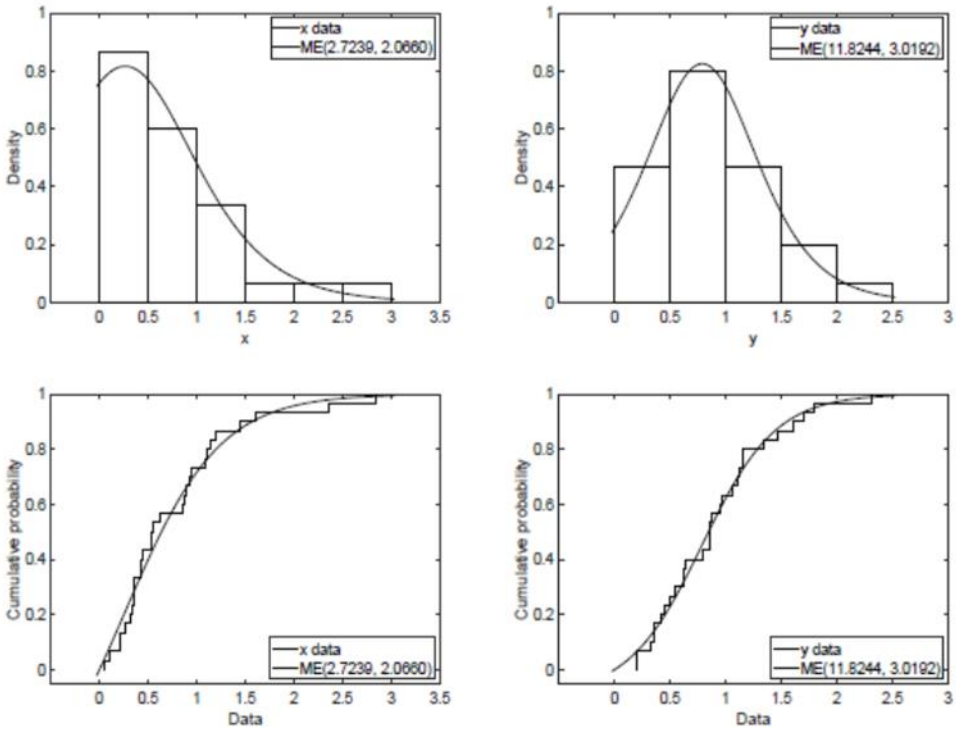

Figure 6. The fitted PDF and CDF of the ME model for data generated from SBVMSB distribution (data of Table 2).

Table 4. The results of the goodness-of-fit test of data of Table 1.

Data	α	β	KS-stat	Critical value	p-value
X	2.7239	2.06604	0.0977	0.2417	0.9103
Y	11.8244	3.01922	0.0717	0.2417	0.9948

4.3 *Real data analysis: Electromyographic (EMG) data*

In this section, the BMEG distribution is fitted to a dataset obtained from the study conducted by Davis [5]. This dataset has been derived from studying affective facial expressions conducted on 22 individuals. In this study, several pieces of music were played for each individual in two stages for 90 seconds. The first stage is relaxation music condition, and the second stage is designed to create positive effects. The response variable at each stage is the average EMG amplitudes of the left eyebrow region. The dataset is presented in Table 5.

Table 5. Average electromyography (EMG) amplitudes from the left eyebrow for 22 subjects.

No.	Stage 1	Stage 2	No.	Stage 1	Stage 2
1	143	368	12	148	378
2	142	155	13	130	142
3	109	167	14	119	171
4	123	135	15	102	94
5	276	216	16	279	204
6	235	386	17	244	365
7	208	175	18	196	168
8	267	358	19	279	358
9	183	193	20	167	183
10	245	268	21	345	238
11	324	507	22	524	507

Roozegar and Kundu [19] fit the bivariate generalized exponential-geometric (BGEG) and bivariate generalized exponential-Poisson (BGEP) distributions to this data; they demonstrated that the BGEG distribution provides a better fit than the BGEP distribution for this dataset, based on the log-likelihood value. Therefore, the results are compared only with the BGEG model. Similar to the study conducted by Roozegar and Kundu [19], we divided all the data points by 100 and then fit the BMEG distribution to this dataset. Figure 7 illustrates the scatter plot of EMG data and the contour plot of the probability density of the fitted BMEG distribution.

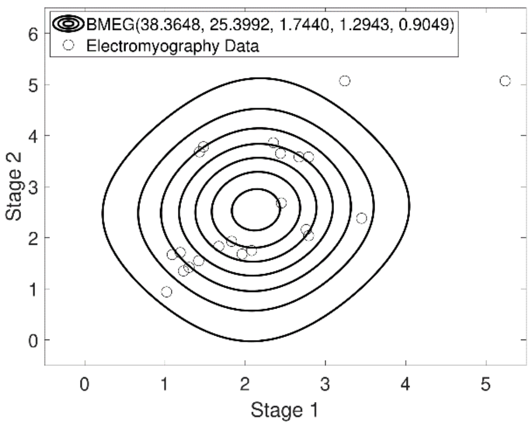

Figure 7. Scatter plot of electromyographic data and probability density contour plot for fitted BMEG distribution.

Table 6 presents the results of fitting the BMEG and BGEG distributions to EMG data. Table 7 depicts the results of the Kolmogorov-Smirnov goodness-of-fit test for fitting the ME model to the marginal of fitted BMEG distribution. Figure 8 also displays the fitted PDFs and CDFs of the ME distribution for EMG data.

Table 6. A summary of the results of fitting the BMEG and BGEG distributions to EMG data.

Model	MLEs (Confidence intervals)	$\hat{l}$	AIC
BMEG	α_1=38.3648 (37.1427, 39.5831)	-64.0818	138.1636
	α_2=25.3992 (24.6048, 26.1912)		
	β_1=1.7440 (1.7322, 1.7558)		
	β_1=1.2943 (1.2849, 1.3037)		
	p=0.9049 (0.8952, 0.9146)		
BGEG	λ_1= 1.2083 (1.0205, 1.3961)	-138.324	288.6480
	λ_2=0.7998 (0.7100, 0.8896)		
	α_1= 31.5221 (28.8438, 34.2004)		
	α_2=19.2061 (18.0905, 20.3217)		
	α=0.2031 (0.2019, 0.2043)		
	θ=0.1582 (0.1577, 0.1587)		

Table 7. The results of the Kolmogorov-Smirnov (KS) goodness-of-fit test for EMG data.

Data	α	β	KS-stat	Critical value	p-value
Stage 1	34.7163	1.7440	0.1242	0.2809	0.8457
Stage 2	22.9837	1.2943	0.1796	0.2809	0.4272

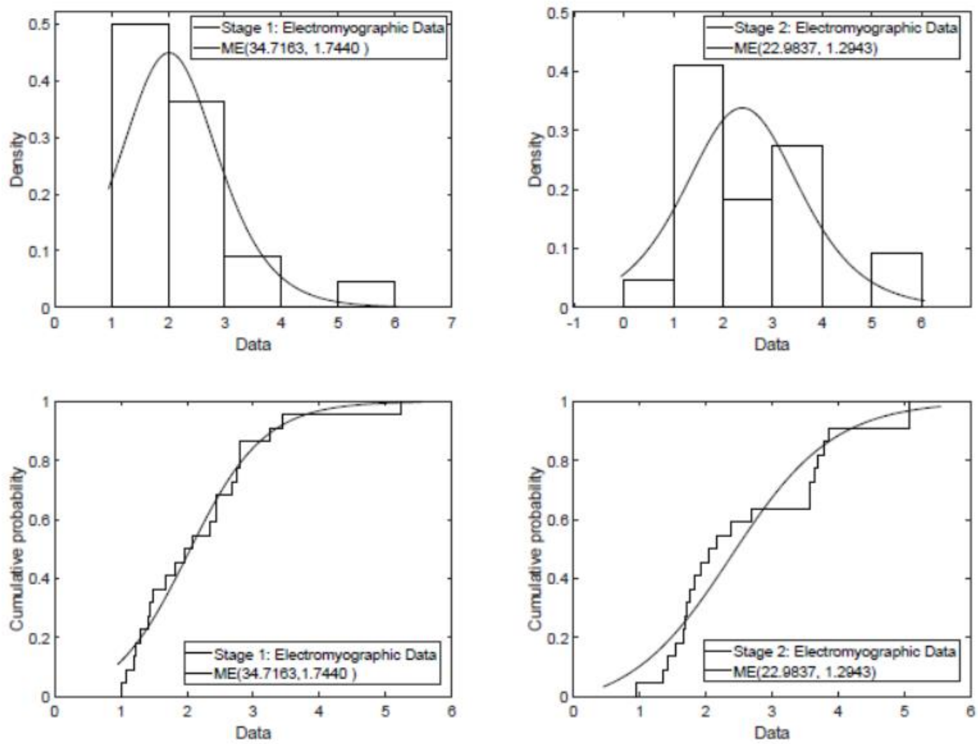

Figure 8. The fitted PDF and CDF of the ME model for electromyographic data.

5. Conclusions

This paper proposed a five-parameter bivariate model called the bivariate modified exponential-geometric (BMEG) distribution, whose marginals are univariate modified exponential distributions. The MLEs of the unknown parameters cannot be obtained in closed forms. The expectation-maximization (EM) algorithm was introduced to compute the MLEs of the unknown parameters quite effectively. Two datasets were analyzed, indicating the excellent performance of the proposed models and the EM algorithm.

References

[1] M. Alizadeh, M. Tahir, G. M. Cordeiro, M. Mansoor, M. Zubair and G. Hamedani, The Kumaraswamy Marshal-Olkin family of distributions, Journal of the Egyptian Mathematical Society, **23** (2015) 546–557.

[2] W. Barreto-souza, A. J. Lemonte and G. M. Cordeiro, General results for the Marshall and Olkin's family of distributions, Anais da Academia Brasileir de Cienc, **85** (2013) 3–21.

[3] A. P. Basu, Bivariate failure rate, Journal of the American Statistical Association, **66** (1971)103–104.

[4] F. Bordbar and A. R. Nematollahi, The modified exponential-geometric distribution, Communications in Statistics - Theory and Methods, **45** (2016) 173–181.

[5] C. S. Davis, Statistical Methods for the Analysis of Repeated Measurements, Springer, (2002).

[6] A. P. Dempster, N. M. Laird and D. B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, Journal of the Royal Statistical Society: Series B (Methodological), **39** (1977) 1–38.

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

[8] N. Johnson and S. Kotz, A vector multivariate hazard rate, Journal of Multivariate Analysis, **5** (1975) 53–66.

[9] D. Kundu, Bivariate geometric (maximum) generalized exponential distribution, Journal of Data Science, **13** (2015) 693–712.

[10] D. Kundu and A. K. Gupta, On bivariate Weibull-geometric distribution, Journal of Multivariate Analysis, **123** (2014) 19–29.

[11] D. Kundu and R. D. Gupta, Modified Sarhan–Balakrishnan singular bivariate distribution, Journal of Statistical Planning and Inference, **140** (2010) 526–538.

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

[13] A. W. Marshall, Some comments on the hazard gradient, Stochastic Processes and Their Applications, **3** (1975) 293–300.

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

[15] G. McLachlan and T. Krishnan, The EM Algorithm and Extensions, Wiley Series in Probability and Statistics, Wiley, (1996).

[16] V. Nekoukhou, A. Khalifeh and E. Mahmoudi, Bivariate Rayleigh geometric distribution, Journal of Statistical Sciences, **13** (2020) 539–555.

[17] V. Preda, E. Panaitescu and R. Cimumara, Marshall-Olkin generalized exponential distribution, Proceedings of the Romanian Academy, Series A, of the Romanian Academy, **12** (2011) 22–29.

[18] M. M. Ristić and D. Kundu, Marshall-Olkin generalized exponential distribution, METRON, **73** (2015) 317–333.

[19] R. Roozegar and D. Kundu, On absolutely continuous bivariate generalized exponential power series distribution, Communications in Statistics – Simulation and Computation, **49** (2018) 1678–1703.

[20] L. Tomy and J. Gillariose, The Marshall-Olkin distribution, Biometrics & Biostatistics International Journal, **7** (2018) 10–16.