

ON BIVARIATE WEIBULL-GEOMETRIC DISTRIBUTION

Debasis Kundu¹ & Arjun K. Gupta²

Abstract

Marshall and Olkin (1997) provided a general method to introduce a parameter into a family of distributions, and discussed in details about the exponential and Weibull families. They have also briefly introduced the bivariate extension, although not any properties or inferential issues have been explored, mainly due to analytical intractability of the general model. In this paper we consider the bivariate model with a special emphasis on the Weibull distribution. We call this new distribution as the bivariate Weibull-Geometric distribution. We derive different properties of the proposed distribution. This distribution has five parameters, and the maximum likelihood estimators cannot be obtained in closed form. We propose to use the EM algorithm, and it is observed that the implementation of the EM algorithm is quite straight forward. Two data sets have been analyzed for illustrative purposes, and it is observed that the new model and the proposed EM algorithm work quite well in these cases.

KEY WORDS AND PHRASES: Geometric maximum; Weibull distribution; EM algorithm; Fisher information matrix; Monte Carlo simulation.

¹ Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. E-mail: kundu@iitk.ac.in, Phone no. 91-512-2597141, Fax no. 91-512-2597500.

²Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403 - 0221, USA.

1 INTRODUCTION

Marshall and Olkin (1997) proposed a method to introduce a parameter in a family of distributions. It induces an extra parameter to a model, hence brings more flexibility. The model has a nice physical interpretation also. They have explored different properties in case of exponential and Weibull distributions. From now on we call them as univariate exponential geometric (UEG) and univariate Weibull geometric (UWG) distributions, respectively. It is observed that the proposed method brings more flexibility to the well known exponential and Weibull models, and it has several desirable properties. 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 a bivariate extension of the proposed model. They did not provide any properties and also did not discuss any inferential issues. Interestingly, although extensive work has been done on the univariate version of the model, not much work has been done on the bivariate generalization mainly due to its analytical intractability for the general model. We consider the bivariate model with a special emphasis on the Marshall-Olkin bivariate Weibull (MOBW) distribution. We call this new model as the bivariate Weibull-Geometric (BWG) model.

Among several bivariate distributions which can be used to model singular data, Marshall-Olkin's (1967) bivariate exponential (MOBE) model plays the most important role. MOBE distribution has exponential marginals, therefore it has its own limitations. Due to this reason Marshall and Olkin (1967) also introduced MOBW distribution which has Weibull marginals. MOBW distribution has four parameters, see for example Kundu and Dey (2009) or Kundu and Gupta (2013), and it has a singular component. The proposed BWG model

also has a singular component. Moreover, MOBW can be obtained as a special case of the BWG model. Other recent extensions of the MOBE model with a singular component have been discussed in the literature, e.g. see Sarhan and Balakrishnan (2007), Franco and Vivo (2010), Kundu and Gupta (2010) and Franco et al. (2011).

Due to presence of five parameters the joint probability density function of BWG distribution is very flexible, and it can take different shapes depending on the shape parameters. The marginals and conditionals are UWG, and they are also very flexible. It is geometric minimum stable. The generation of random samples from BWG is quite simple, hence simulation experiments can be performed easily.

The maximum likelihood estimators (MLEs) of the unknown parameters of BWG distribution cannot be obtained in closed forms. It involves solving five non-linear equations simultaneously. Finding initial values and the convergence of the algorithm are important issues. To avoid this problem, we propose to use expectation maximization (EM) algorithm in computing the MLEs of the unknown parameters. At each ‘E’-step, of the EM algorithm, the corresponding ‘M’-step involves maximization of one non linear function, similar to the the EM algorithm proposed by Kundu and Dey (2009) in finding the MLEs of the unknown parameters of a MOBW model. For illustrative purposes, we analyze two data sets, and it is observed that the proposed EM algorithm works quite well.

Rest of the paper is organized as follows. In Section 2, we introduce the model. Different properties are discussed in Section 3. EM algorithm is presented in Section 4. In Section 5 we present the analysis of two data sets, one simulated and one real. Finally in Section 6 we conclude the paper .

2 PRELIMINARIES AND MODEL FORMULATION

2.1 PRELIMINARIES & NOTATIONS

We will be using the following notations through out the paper. A two-parameter Weibull distribution with the shape and scale parameter as $\alpha > 0$ and $\lambda > 0$ respectively, will be denoted by $WE(\alpha, \lambda)$, and it has the PDF

$$f(x; \alpha, \lambda) = \alpha \lambda x^{\alpha-1} e^{-\lambda x^\alpha}; \quad x > 0. \quad (1)$$

In (1) when $\alpha = 1$, it becomes the exponential distribution with parameter λ . A geometric random variable N with probability mass function

$$P(N = n) = p(1 - p)^{n-1}; \quad n = 1, 2, \dots, \quad 0 < p < 1 \quad (2)$$

will be denoted by $GM(p)$. If F is any distribution function, we will use \bar{F} to denote its survival function.

Suppose $\{X_n; n = 1, 2, \dots\}$ is a sequence of independent and identically distributed (i.i.d.) non-negative random variables with common distribution function $F(\cdot)$, $N \sim GM(\theta)$ and it is independent of X_n 's. Consider a new random variable U , such that

$$Y = \min\{X_1, \dots, X_N\}. \quad (3)$$

Then the survival function of Y , say $\bar{G}(y)$, is

$$\bar{G}(y) = P(Y \geq y) = \sum_{n=1}^{\infty} P(Y \geq y | N = n) P(N = n) = \frac{\theta \bar{F}(y)}{1 - (1 - \theta) \bar{F}(y)}. \quad (4)$$

Marshall and Olkin (1997) proposed this method to introduce a parameter in a family of distributions. It may be easily verified that the above $G(\cdot; \theta)$ is indeed a survival function for all $0 < \theta \leq 1$. It is interesting to note that \bar{G} as defined in (4) is indeed a proper survival function for $\theta \in \mathbb{R}^+$, although in this paper we consider mainly $0 < \theta \leq 1$.

If the distribution function $F(\cdot)$ has a density function $f(\cdot)$, then the density function corresponding to the survival function $\bar{G}(y; \theta)$ becomes;

$$g(y; \theta) = \frac{\theta f(y)}{(1 - (1 - \theta)\bar{F}(y))^2}, \quad (5)$$

and for $\theta = 1$, $\bar{G} = \bar{F}$.

Marshall and Olkin (1997) considered two special cases, namely the one-parameter exponential distribution and two-parameter Weibull distribution, and discussed their properties in details. When $f(y) = \lambda e^{-\lambda y}$, the PDF (5) takes the following form

$$g(y; \theta, \lambda) = \frac{\theta \lambda e^{-\lambda y}}{(1 - (1 - \theta)e^{-\lambda y})^2}, \quad (6)$$

and it will be denoted by UEG(θ, λ). In case of Weibull distribution, *i.e.* when $f(y)$ has the form (1), the PDF (5) takes the form;

$$g(y; \theta, \lambda, \alpha) = \frac{\theta \alpha \lambda y^{\alpha-1} e^{-\lambda y^\alpha}}{(1 - (1 - \theta)e^{-\lambda y^\alpha})^2}, \quad (7)$$

and it will be denoted by UWG(θ, α, λ). The PDF of UWG can be a decreasing, unimodal or some other shapes. The hazard function can be increasing, decreasing or some other shapes.

A bivariate random vector (Y_1, Y_2) has the Marshall-Olkin bivariate Weibull distribution (MOBW) if it has the following joint cumulative survival function

$$\bar{F}_{Y_1, Y_2}(y_1, y_2) = P(Y_1 > y_1, Y_2 > y_2) = \begin{cases} e^{-(\lambda_0 + \lambda_1)y_1^\alpha - \lambda_2 y_2^\alpha} & \text{if } y_1 \geq y_2 \\ e^{-(\lambda_0 + \lambda_2)y_2^\alpha - \lambda_1 y_1^\alpha} & \text{if } y_1 < y_2 \end{cases} \quad (8)$$

the joint PDF

$$f_{Y_1, Y_2}(y_1, y_2) = \begin{cases} f_1(y_1, y_2) & \text{if } 0 < y_2 < y_1 < \infty \\ f_2(y_1, y_2) & \text{if } 0 < y_1 < y_2 < \infty \\ f_0(y) & \text{if } 0 < y_1 = y_2 = y < \infty, \end{cases} \quad (9)$$

where

$$f_1(y_1, y_2) = \alpha^2 y_1^{\alpha-1} y_2^{\alpha-1} (\lambda_0 + \lambda_1) e^{-(\lambda_0 + \lambda_1)y_1^\alpha} \lambda_2 e^{-\lambda_2 y_2^\alpha}, \quad (10)$$

$$f_2(y_1, y_2) = \alpha^2 y_1^{\alpha-1} y_2^{\alpha-1} \lambda_1 e^{-\lambda_1 y_1^\alpha} (\lambda_0 + \lambda_2) e^{-(\lambda_0 + \lambda_2)y_2^\alpha}, \quad (11)$$

$$f_0(y) = \alpha y^{\alpha-1} \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} e^{-(\lambda_0 + \lambda_1 + \lambda_2)y^\alpha}, \quad (12)$$

and it will be denoted by MOBW($\alpha, \lambda_0, \lambda_1, \lambda_2$). Note that if V_0, V_1 and V_2 are independent Weibull random variables with parameters (α, λ_0) , (α, λ_1) and (α, λ_2) respectively, then

$$(Y_1, Y_2) \stackrel{d}{=} \{\min\{V_0, V_1\}, \min\{V_0, V_2\}\}. \quad (13)$$

A special case of MOBW distribution is the Marshall-Olkin bivariate exponential (MOBE) distribution, where $\alpha = 1$.

2.2 MARSHALL-OLKIN EXTENDED BIVARIATE DISTRIBUTION

Marshall-Olkin extended bivariate distribution can be described as follows. In defining this distribution we restrict the parameter θ as $0 < \theta \leq 1$. Suppose $\{(X_{1n}, X_{2n}); n = 1, 2, \dots\}$ is a sequence of i.i.d. non-negative bivariate random variables with common joint distribution function $F(\cdot, \cdot)$. N is a geometric random variable independent of $\{(X_{1n}, X_{2n}), n = 1, 2, \dots\}$. Consider the following bivariate random variable (Y_1, Y_2)

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

The joint survival function of (Y_1, Y_2) becomes

$$P(Y_1 > y_1, Y_2 > y_2) = \sum_{n=1}^{\infty} \bar{F}^n(y_1, y_2) \theta (1 - \theta)^{n-1} = \frac{\theta \bar{F}(y_1, y_2)}{1 - (1 - \theta) \bar{F}(y_1, y_2)}. \quad (14)$$

Suppose F_1 and F_2 are the marginal distribution functions of F , *i.e.* $F_1(x) = F(x, \infty)$, $F_2(x) = F(\infty, x)$. Similarly, suppose \bar{F}_1 and \bar{F}_2 are the marginal survival functions of \bar{F} , *i.e.* $\bar{F}_1(x) = \bar{F}(x, 0)$ and $\bar{F}_2 = \bar{F}(0, x)$. Then survival function of Y_1 and Y_2 are

$$P(Y_1 > y_1) = \frac{\theta \bar{F}_1(y_1)}{1 - (1 - \theta) \bar{F}_1(y_1)} \quad \text{and} \quad P(Y_2 > y_2) = \frac{\theta \bar{F}_2(y_2)}{1 - (1 - \theta) \bar{F}_2(y_2)} \quad (15)$$

respectively.

In the rest of the paper we mainly consider the special case when $F(u_1, u_2)$ follows MOBW distribution. We discuss different properties and consider several inferential issues.

3 BWG DISTRIBUTION

3.1 BASIC PROPERTIES

The random variable (Y_1, Y_2) is said to have a bivariate Weibull geometric distribution with parameters $\theta, \alpha, \lambda_0, \lambda_1, \lambda_2$, if the distribution F in (14) is $\text{MOBW}(\alpha, \lambda_0, \lambda_1, \lambda_2)$. Therefore, the joint survival function of (Y_1, Y_2) becomes;

$$\bar{G}(y_1, y_2) = P(Y_1 > y_1, Y_2 > y_2) = \begin{cases} \frac{\theta e^{-(\lambda_0 + \lambda_1)y_1^\alpha - \lambda_2 y_2^\alpha}}{1 - (1 - \theta)e^{-(\lambda_0 + \lambda_1)y_1^\alpha - \lambda_2 y_2^\alpha}} & \text{if } y_1 \geq y_2 \\ \frac{\theta e^{-(\lambda_0 + \lambda_2)y_2^\alpha - \lambda_1 y_1^\alpha}}{1 - (1 - \theta)e^{-(\lambda_0 + \lambda_2)y_2^\alpha - \lambda_1 y_1^\alpha}} & \text{if } y_1 < y_2 \end{cases} \quad (16)$$

and it will be denoted by $(Y_1, Y_2) \sim \text{BWG}(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$.

THEOREM 3.1: Let $(Y_1, Y_2) \sim \text{BWG}(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$, then the joint PDF of (Y_1, Y_2) is

$$g(y_1, y_2) = \begin{cases} g_1(y_1, y_2) & \text{if } y_1 > y_2 \\ g_2(y_1, y_2) & \text{if } y_1 < y_2 \\ g_0(y_1, y_2) & \text{if } y_1 = y_2 = y, \end{cases} \quad (17)$$

where

$$g_1(y_1, y_2) = \frac{\alpha^2 y_1^{\alpha-1} y_2^{\alpha-1} \theta (\lambda_0 + \lambda_1) \lambda_2 e^{-(\lambda_0 + \lambda_1)y_1^\alpha - \lambda_2 y_2^\alpha} (1 + (1 - \theta)e^{-(\lambda_0 + \lambda_1)y_1^\alpha - \lambda_2 y_2^\alpha})}{(1 - (1 - \theta)e^{-(\lambda_0 + \lambda_1)y_1^\alpha - \lambda_2 y_2^\alpha})^3} \quad (18)$$

$$g_2(y_1, y_2) = \frac{\alpha^2 y_1^{\alpha-1} y_2^{\alpha-1} \theta (\lambda_0 + \lambda_2) \lambda_1 e^{-(\lambda_0 + \lambda_2)y_2^\alpha - \lambda_1 y_1^\alpha} (1 + (1 - \theta)e^{-(\lambda_0 + \lambda_2)y_2^\alpha - \lambda_1 y_1^\alpha})}{(1 - (1 - \theta)e^{-(\lambda_0 + \lambda_2)y_2^\alpha - \lambda_1 y_1^\alpha})^3} \quad (19)$$

$$g_0(y) = \frac{\alpha y^{\alpha-1} \theta \lambda_0 e^{-(\lambda_0 + \lambda_1 + \lambda_2)y^\alpha}}{(\lambda_0 + \lambda_1 + \lambda_2)(1 - (1 - \theta)e^{-(\lambda_0 + \lambda_1 + \lambda_2)y^\alpha})^2}. \quad (20)$$

PROOF: The expressions of $g_1(\cdot, \cdot)$ and $g_2(\cdot, \cdot)$ can be obtained by taking $\{\partial^2 / \partial y_1 \partial y_2 \bar{G}(y_1, y_2)\}$ for $y_1 > y_2$ and $y_1 < y_2$ respectively. Although g_0 cannot be obtained similarly, it can be obtained from the following relation.

$$\int_0^\infty \int_{y_2}^\infty g_1(y_1, y_2) dy_1 dy_2 + \int_0^\infty \int_{y_1}^\infty g_2(y_1, y_2) dy_2 dy_1 + \int_0^\infty g_0(y) dy = 1. \quad (21)$$

Using the same method as in Sarhan and Balakrishnan (2007) or Kundu and Gupta (2009), note that

$$I_1 = \int_{y_2}^\infty g_1(y_1, y_2) dy_1 \quad \text{and} \quad I_2 = \int_{y_1}^\infty g_2(y_1, y_2) dy_2$$

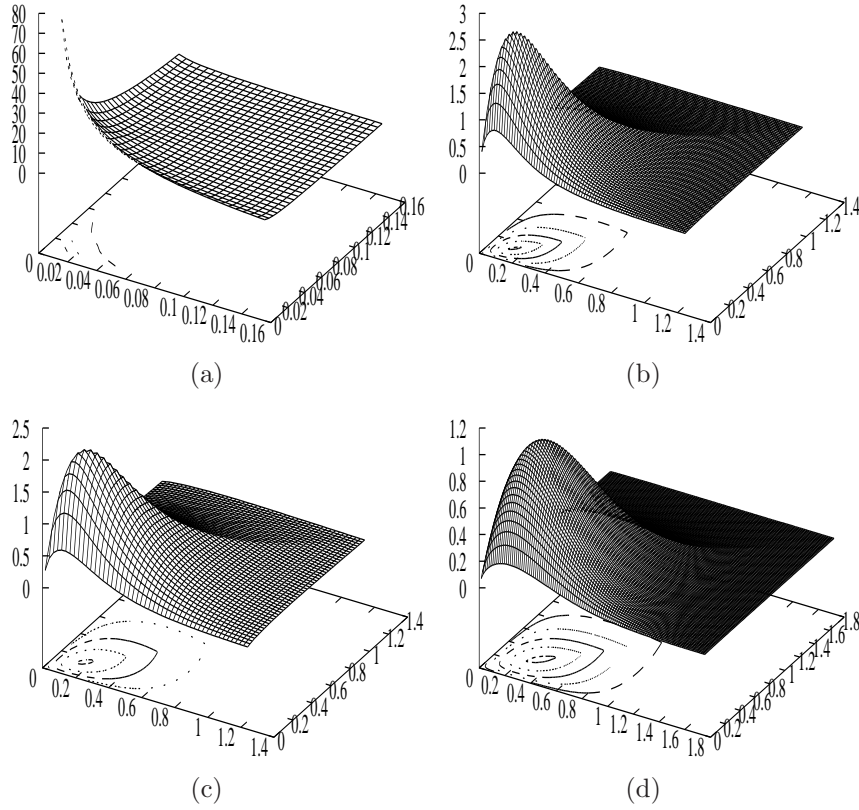


Figure 1: Contour plots of the absolute continuous part of the joint PDF of the BWG distribution for different parameter values: $(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$: (a) $(0.5, 0.75, 2.0, 1, 1)$, (b) $(1.5, 0.5, 2.0, 1, 1)$, (c) $(1.5, 0.5, 1.0, 1, 1)$, (d) $(1.5, 1.0, 2.0, 1, 1)$.

can be obtained in explicit forms. Using I_1, I_2 and using the identity (21), the results follow after some lengthy calculations. ■

The joint PDF of the BWG distribution can take different shapes. We have provided the contour plots of the absolutely continuous part of the PDF of BWG distribution in Figure 1 for different parameter values.

The joint PDF of (Y_1, Y_2) as provided in Theorem 3.1, can be written as follows:

$$g(y_1, y_2) = \frac{\lambda_1 + \lambda_2}{\lambda_0 + \lambda_1 + \lambda_2} g_a(y_1, y_2) + \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} g_s(y), \quad (22)$$

here

$$g_a(y_1, y_2) = \frac{\lambda_0 + \lambda_1 + \lambda_2}{\lambda_1 + \lambda_2} \times \begin{cases} g_1(y_1, y_2) & \text{if } y_1 > y_2 \\ g_2(y_1, y_2) & \text{if } y_1 < y_2 \end{cases} \quad (23)$$

and

$$g_s(y) = \frac{\theta\alpha(\lambda_0 + \lambda_1 + \lambda_2)y^{\alpha-1}e^{-(\lambda_0+\lambda_1+\lambda_2)y^\alpha}}{(1 - (1 - \theta)e^{-(\lambda_0+\lambda_1+\lambda_2)y^\alpha})^2} \quad \text{if } y_1 = y_2 = y \quad (24)$$

and 0 otherwise. Clearly, $g_a(\cdot, \cdot)$ is the absolute continuous part, and $g_s(\cdot)$ is the singular part. If $\lambda_0 = 0$, it does not have any singular part, and it becomes an absolute continuous distribution. For $\theta = 1$, it is immediate that MOBE and MOBW can be obtained as special cases of BWG.

Now we would like to obtain the joint PDF of Y_1, Y_2 and N where (Y_1, Y_2) has BWG distribution and N is same as defined in Section 2.2. Note that

$$\begin{aligned} P(Y_1 > y_1, Y_2 > y_2, N = n) &= P(Y_1 > y_1, Y_2 > y_2 | N = n)P(N = n) \\ &= \begin{cases} \theta(1 - \theta)^{n-1}e^{-n(\lambda_0+\lambda_1)y_1^\alpha - n\lambda_2y_2^\alpha} & \text{if } y_1 \geq y_2 \\ \theta(1 - \theta)^{n-1}e^{-n(\lambda_0+\lambda_2)y_2^\alpha - n\lambda_1y_1^\alpha} & \text{if } y_1 < y_2. \end{cases} \end{aligned} \quad (25)$$

The joint PDF of Y_1, Y_2 and N becomes;

$$f_{Y_1, Y_2, N}(y_1, y_2, n) = \begin{cases} \theta(1 - \theta)^{n-1}f_{1n}(y_1, y_2) & \text{if } y_2 < y_1 \\ \theta(1 - \theta)^{n-1}f_{2n}(y_1, y_2) & \text{if } y_1 < y_2 \\ \theta(1 - \theta)^{n-1}f_{0n}(y) & \text{if } y_1 = y_2 = y \end{cases} \quad (26)$$

where

$$\begin{aligned} f_{1n}(u_1, u_2) &= \alpha^2 n^2 u_1^{\alpha-1} u_2^{\alpha-1} (\lambda_0 + \lambda_1) e^{-n(\lambda_0+\lambda_1)u_1^\alpha} \lambda_2 e^{-n\lambda_2 u_2^\alpha}, \\ f_{2n}(u_1, u_2) &= \alpha^2 n^2 u_1^{\alpha-1} u_2^{\alpha-1} \lambda_1 e^{-n\lambda_1 u_1^\alpha} (\lambda_0 + \lambda_2) e^{-n(\lambda_0+\lambda_2)u_2^\alpha}, \\ f_{0n}(u) &= \alpha u^{\alpha-1} \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} e^{-n(\lambda_0+\lambda_1+\lambda_2)u^\alpha}. \end{aligned} \quad (27)$$

The conditional probability mass function of N given $Y_1 = y_1$ and $Y_2 = y_2$ becomes;

$$f_N(n|y_1, y_2) = \begin{cases} c_1(y_1, y_2)(1 - \theta)^{n-1}n^2 e^{-(n-1)(\lambda_0+\lambda_1)y_1^\alpha} e^{-(n-1)\lambda_2 y_2^\alpha} & \text{if } y_2 < y_1 \\ c_2(y_1, y_2)(1 - \theta)^{n-1}n^2 e^{-(n-1)(\lambda_0+\lambda_2)y_2^\alpha} e^{-(n-1)\lambda_1 y_1^\alpha} & \text{if } y_1 < y_2 \\ c_0(y)(1 - \theta)^{n-1}e^{-(n-1)(\lambda_0+\lambda_1+\lambda_2)u^\alpha} & \text{if } y_1 = y_2 = y, \end{cases} \quad (28)$$

where

$$\begin{aligned} c_1(y_1, y_2) &= \frac{(1 - (1 - \theta)e^{-(\lambda_0 + \lambda_1)y_1^\alpha - \lambda_2 y_2^\alpha})^3}{(1 + (1 - \theta)e^{-(\lambda_0 + \lambda_1)y_1^\alpha - \lambda_2 y_2^\alpha})} \\ c_2(y_1, y_2) &= \frac{(1 - (1 - \theta)e^{-(\lambda_0 + \lambda_2)y_2^\alpha - \lambda_1 y_1^\alpha})^3}{(1 + (1 - \theta)e^{-(\lambda_0 + \lambda_2)y_2^\alpha - \lambda_1 y_1^\alpha})} \\ c_0(y) &= (1 - (1 - \theta)e^{-(\lambda_0 + \lambda_1 + \lambda_2)y^\alpha}). \end{aligned}$$

We will be using the following notations for (28).

$$f_N(n|y_1, y_2) = \begin{cases} \frac{(1 - \xi_1(y_1, y_2, \theta, \gamma))^3}{1 + \xi_1(y_1, y_2, \theta, \gamma)} \xi_1^{n-1}(y_1, y_2, \theta, \gamma) n^2 & \text{if } y_2 < y_1 \\ \frac{(1 - \xi_2(y_1, y_2, \theta, \gamma))^3}{1 + \xi_2(y_1, y_2, \theta, \gamma)} \xi_2^{n-1}(y_1, y_2, \theta, \gamma) n^2 & \text{if } y_1 < y_2 \\ (1 - \xi_0(y_1, y_2, \theta, \gamma)) \xi_0^{n-1}(y_1, y_2, \theta, \gamma) & \text{if } y_1 = y_2 = y, \end{cases}$$

where $\gamma = (\alpha, \lambda_0, \lambda_1, \lambda_2)$ and

$$\begin{aligned} \xi_1(y_1, y_2, \theta, \eta) &= (1 - \theta)e^{-(\lambda_0 + \lambda_1)y_1^\alpha - \lambda_2 y_2^\alpha} \\ \xi_2(y_1, y_2, \theta, \eta) &= (1 - \theta)e^{-(\lambda_0 + \lambda_2)y_2^\alpha - \lambda_1 y_1^\alpha} \\ \xi_0(y_1, y_2, \theta, \eta) &= 1 - c_0(y). \end{aligned}$$

We immediately obtain

$$E(N|y_1, y_2) = \begin{cases} \frac{(1 - \xi_1(y_1, y_1, \theta, \eta))^2 - 6(1 - \xi_1(y_1, y_2, \theta, \eta)) + 6}{(1 - \xi_1(y_1, y_2, \theta, \eta))^3} & \text{if } y_2 < y_1 \\ \frac{(1 - \xi_2(y_1, y_1, \theta, \eta))^2 - 6(1 - \xi_2(y_1, y_2, \theta, \eta)) + 6}{(1 - \xi_2(y_1, y_2, \theta, \eta))^3} & \text{if } y_1 < y_2 \\ \frac{1}{c_0(y)} & \text{if } y_1 = y_2 = y. \end{cases} \quad (29)$$

From the marginal and joint survival functions the following result follows.

THEOREM 3.2 Let $(Y_1, Y_2) \sim \text{BWG}(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$, then

- (a) $Y_1 \sim \text{UWG}(\theta, \alpha, \lambda_0 + \lambda_1)$
- (b) $Y_2 \sim \text{UWG}(\theta, \alpha, \lambda_0 + \lambda_2)$
- (c) $\min\{Y_1, Y_2\} \sim \text{UWG}(\theta, \alpha, \lambda_0 + \lambda_1 + \lambda_2)$
- (d) $P(Y_1 < Y_2) = \frac{\lambda_1}{\lambda_0 + \lambda_1 + \lambda_2}$.

PROOF: The proofs of (a), (b) and (c) are immediate from (15) with \bar{F}_i 's having Weibull distributions. The proof of (d) can be obtained as follows:

$$\begin{aligned}
P(Y_1 < Y_2) &= \sum_{n=1}^{\infty} P(Y_1 < Y_2, N = n) \\
&= \theta \sum_{n=1}^{\infty} (1 - \theta)^{n-1} \int_0^{\infty} \int_{y_1}^{\infty} f_{2n}(y_1, y_2) dy_2 dy_1 \\
&= \theta \sum_{n=1}^{\infty} (1 - \theta)^{n-1} \times \frac{\lambda_1}{\lambda_0 + \lambda_1 + \lambda_2} = \frac{\lambda_1}{\lambda_0 + \lambda_1 + \lambda_2}
\end{aligned}$$

■

4 MAXIMUM LIKELIHOOD ESTIMATION

In this section we discuss the estimation of the unknown parameters of the BWG distribution using EM algorithm. Suppose $\{(y_{11}, y_{21}), \dots, (y_{1m}, y_{2m})\}$ is a bivariate random sample of size m from $BWG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$. We use the following notation

$$I_0 = \{i : y_{1i} = y_{2i} = y_i\}, \quad I_1 = \{i : y_{1i} > y_{2i}\}, \quad I_2 = \{i : y_{1i} < y_{2i}\}, \quad (30)$$

and suppose m_0 , m_1 and m_2 denote the number of observations in the set I_0 , I_1 and I_2 , respectively. The log-likelihood function can be written as

$$l(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2) = \sum_{i \in I_0} \ln g_0(y_i) + \sum_{i \in I_1} \ln g_1(y_{1i}, y_{2i}) + \sum_{i \in I_2} \ln g_2(y_{1i}, y_{2i}), \quad (31)$$

where g_1 , g_2 and g_0 are defined in (18), (19) and (20) respectively. Clearly, the MLEs cannot be obtained in closed form. They can be obtained by solving five non-linear equations simultaneously, which is a difficult problem. It has to be obtained using a five dimensional optimization problem. It involves finding the initial guesses which may not be very trivial in general. To avoid that we treat this as a missing value problem and use EM algorithm. First it is assumed that θ is known in advance. We obtain the MLEs of α , λ_0 , λ_1 and λ_2 ,

hence obtain the profile likelihood function of θ . Finally maximizing the profile likelihood function of θ , the MLE of θ also can be obtained.

Now, let us see how we can treat this as a missing value problem. Suppose along with (Y_1, Y_2) , we observe N also. The joint PDF of (Y_1, Y_2, N) is provided in (26). Note that

$$(Y_1, Y_2 | N = n) \sim \text{MOBW}(\alpha, n\lambda_0, n\lambda_1, n\lambda_2).$$

The complete data is as follows: $\{(y_{1i}, y_{2i}, n_i); i = 1, \dots, m\}$. Since θ is assumed to be known, the MLEs of α , λ_0 , λ_1 and λ_2 can be obtained by maximizing the conditional log-likelihood function, and it is as follows;

$$l_1(\alpha, \lambda_0, \lambda_1, \lambda_2) = \sum_{i \in I_0} \ln f_{0n_i}(y_i) + \sum_{i \in I_1} \ln f_{1n_i}(y_{1i}, y_{2i}) + \sum_{i \in I_2} \ln f_{2n_i}(y_{1i}, y_{2i}). \quad (32)$$

The functions, f_{0n_i} , f_{1n_i} and f_{2n_i} are defined in (27). It can be easily checked that the MLEs of α , λ_0 , λ_1 and λ_2 can be obtained by solving a four dimensional optimization problem.

We treat this as a missing value problem as in Kundu and Dey (2009). The details are as follows. Consider a new set of random variables

$$\{V_0 | N = n\} \sim \text{WE}(\alpha, n\lambda_0), \quad \{V_1 | N = n\} \sim \text{WE}(\alpha, n\lambda_1), \quad \{V_2 | N = n\} \sim \text{WE}(\alpha, n\lambda_2), \quad (33)$$

and they are assumed to be conditionally independent. It is well known that

$$\{Y_1 | N = n\} = \min\{V_0, V_1\} | N = n \quad \text{and} \quad \{Y_2 | N = n\} = \min\{V_0, V_2\} | N = n.$$

It is further assumed that for the bivariate random vector (Y_1, Y_2) , there is an associated random vector (Δ_1, Δ_2) , where (Δ_1, Δ_2) is defined as follows

$$\Delta_1 = \begin{cases} 0 & \text{if } Y_1 = V_0 \\ 1 & \text{if } Y_1 = V_1 \end{cases} \quad \text{and} \quad \Delta_2 = \begin{cases} 0 & \text{if } Y_2 = V_0 \\ 1 & \text{if } Y_2 = V_2. \end{cases} \quad (34)$$

Here Y_i 's are same as defined above. It can be easily seen that if instead of (Y_1, Y_2) we had observed $(Y_1, Y_2, \Delta_1, \Delta_2, N)$, then the MLEs of the unknown parameters can be obtained by solving a one dimensional optimization problem.

Note that even if we know (Y_1, Y_2) , the associated (Δ_1, Δ_2) may not be known always. For example, if $Y_1 = Y_2$, then $\Delta_1 = \Delta_2 = 0$, is known. But if $Y_1 \neq Y_2$, then (Δ_1, Δ_2) , is not known. If $(Y_1, Y_2) \in I_1$, then the possible values of (Δ_1, Δ_2) are $(0, 1)$ or $(1, 1)$, and if $(Y_1, Y_2) \in I_2$, then the possible values of (Δ_1, Δ_2) are $(1, 0)$ or $(1, 1)$, with positive probabilities, see for example Kundu and Dey (2009).

We adopt the same procedure as in Dinse (1982) or Kundu (2004) to compute the 'pseudo' log-likelihood function for implementation of the EM algorithm,. 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, we kept the log-likelihood contribution of all the observations belonging to I_0 intact, as in this case the corresponding (Δ_1, Δ_2) are known completely. If the observations belong to I_1 or I_2 , we treat them as missing observations. If $(y_1, y_2) \in I_1$, we form 'pseudo observation' as in Dinse (1982) or Kundu (2004), by fractioning (y_1, y_2) to two partially complete 'pseudo observations' of the form $(y_1, y_2, u_1(\gamma))$ and $(y_1, y_2, u_2(\gamma))$. Here $\gamma = (\alpha, \lambda_0, \lambda_1, \lambda_2)$, and the fractional mass $u_1(\gamma)$ and $u_2(\gamma)$ assigned to the 'pseudo observation' are the conditional probabilities that (Δ_1, Δ_2) takes values $(0, 1)$ or $(1, 1)$, respectively, given that $(Y_1, Y_2) \in I_1$. Similarly, if $(Y_1, Y_2) \in I_2$, 'pseudo observation's are formed as $(y_1, y_2, v_1(\gamma))$ and $(y_1, y_2, v_2(\gamma))$, where $v_1(\gamma)$ and $v_2(\gamma)$ are the conditional probabilities that (Δ_1, Δ_2) takes values $(1, 0)$ and $(1, 1)$ respectively. Since, in this case

$$P(V_2 < V_1 < V_0 | N = n) = \frac{\lambda_1 \lambda_2}{(\lambda_0 + \lambda_1)(\lambda_0 + \lambda_1 + \lambda_2)},$$

$$P(V_2 < V_0 < V_1 | N = n) = \frac{\lambda_0 \lambda_2}{(\lambda_0 + \lambda_1)(\lambda_0 + \lambda_1 + \lambda_2)},$$

therefore,

$$u_1(\gamma) = \frac{\lambda_0}{\lambda_1 + \lambda_0} \quad \text{and} \quad u_2(\gamma) = \frac{\lambda_1}{\lambda_1 + \lambda_0}.$$

Similarly,

$$v_1(\gamma) = \frac{\lambda_0}{\lambda_2 + \lambda_0} \quad \text{and} \quad v_2(\gamma) = \frac{\lambda_2}{\lambda_2 + \lambda_0}.$$

Now let us look at the 'pseudo' log-likelihood contribution of an observation $(y, y) \in I_0$. Conditioning on $N = n$, the log-likelihood contribution becomes

$$\ln \alpha + \ln n + \ln \lambda_0 + (\alpha - 1) \ln y - (\lambda_0 + \lambda_1 + \lambda_2)ny^\alpha. \quad (35)$$

It simply follows from the fact (33) and in this case $V_0 = y$, $V_1 > y$, $V_2 > y$. Similarly, when $(y_1, y_2) \in I_1$ and $(y_1, y_2) \in I_2$, the 'pseudo' log-likelihood contributions, conditioning on $N = n$ become

$$2 \ln \alpha + 2 \ln n + \ln \lambda_2 + (\alpha - 1)(\ln y_1 + \ln y_2) - (\lambda_0 + \lambda_1)ny_1^\alpha - \lambda_2ny_2^\alpha + u_1 \ln \lambda_0 + u_2 \ln \lambda_1 \quad (36)$$

and

$$2 \ln \alpha + 2 \ln n + \ln \lambda_1 + (\alpha - 1)(\ln y_1 + \ln y_2) - \lambda_1ny_1^\alpha - (\lambda_0 + \lambda_2)ny_2^\alpha + v_1 \ln \lambda_0 + v_2 \ln \lambda_2, \quad (37)$$

respectively. Since $\ln n$ and n are missing, they will be replaced by $E(\ln N|y_1, y_2)$ and $E(N|y_1, y_2)$, respectively to compute the 'pseudo' log-likelihood function at the E-step of the EM algorithm. To describe the EM algorithm, we will be using the following notations. At the k -th stage of the EM algorithm the estimates of the parameters $\gamma = (\alpha, \lambda_0, \lambda_1, \lambda_2)$ will be denoted by $\gamma^{(k)} = (\alpha^{(k)}, \lambda_0^{(k)}, \lambda_1^{(k)}, \lambda_2^{(k)})$. Further

$$\begin{aligned} E(N|y_{1i}, y_{2i}, \theta, \gamma) &= a_i(\theta, \gamma), \quad E(N|y_{1i}, y_{2i}, \theta, \gamma^{(k)}) = a_i^{(k)}, \\ u_1(\gamma^{(k)}) &= u_1^{(k)}, u_2(\gamma^{(k)}) = u_2^{(k)}, v_1(\gamma^{(k)}) = v_1^{(k)}, v_2(\gamma^{(k)}) = v_2^{(k)}. \end{aligned}$$

Now we are ready to provide the EM algorithm.

E-Step: At the k -stage of the EM algorithm, the 'pseudo' log-likelihood function without the additive constant can be written as follows:

$$\begin{aligned} l_p(\gamma|\gamma^{(k)}, \theta) &= (m_0 + 2m_1 + 2m_2) \ln \alpha + (m_0 + m_1u_2^{(k)} + m_2v_1^{(k)}) \ln \lambda_0 + (m_1u_1^{(k)} + m_2) \ln \lambda_1 \\ &\quad + (m_1 + m_2v_2^{(k)}) \ln \lambda_2 + (\alpha - 1) \left(\sum_{i \in I_0} \ln y_i + \sum_{i \in I_1 \cup I_2} (\ln y_{1i} + \ln y_{2i}) \right) \end{aligned}$$

$$\begin{aligned}
& -\lambda_0 \left(\sum_{i \in I_0} a_i^{(k)} y_i^\alpha + \sum_{i \in I_1} a_i^{(k)} y_{1i}^\alpha + \sum_{i \in I_2} a_i^{(k)} y_{2i}^\alpha \right) \\
& -\lambda_1 \left(\sum_{i \in I_0} a_i^{(k)} y_i^\alpha + \sum_{i \in I_1 \cup I_2} a_i^{(k)} y_{1i}^\alpha \right) - \lambda_2 \left(\sum_{i \in I_0} a_i^{(k)} y_i^\alpha + \sum_{i \in I_1 \cup I_2} a_i^{(k)} y_{2i}^\alpha \right). \quad (38)
\end{aligned}$$

M-Step: It involves maximizing (38) with respect to the unknown parameters, namely α , λ_0 , λ_1 and λ_2 . Note that for fixed α , the maximization with respect to the unknown parameters can be obtained as follows:

$$\widehat{\lambda}_0(\alpha) = \frac{m_0 + m_1 u_2^{(k)} + m_2 v_1^{(k)}}{\sum_{i \in I_0} a_i^{(k)} y_i^\alpha + \sum_{i \in I_1} a_i^{(k)} y_{1i}^\alpha + \sum_{i \in I_2} a_i^{(k)} y_{2i}^\alpha} \quad (39)$$

$$\widehat{\lambda}_1(\alpha) = \frac{m_1 u_1^{(k)} + m_2}{\sum_{i \in I_0} a_i^{(k)} y_i^\alpha + \sum_{i \in I_1 \cup I_2} a_i^{(k)} y_{1i}^\alpha} \quad (40)$$

$$\widehat{\lambda}_2(\alpha) = \frac{m_1 + m_2 v_2^{(k)}}{\sum_{i \in I_0} a_i^{(k)} y_i^\alpha + \sum_{i \in I_1 \cup I_2} a_i^{(k)} y_{2i}^\alpha}. \quad (41)$$

Substituting $\widehat{\lambda}_i(\alpha)$, for $i = 1, 2, 3$, in (38) we can obtain the profile 'pseudo' log-likelihood function $l_p(\alpha, \widehat{\lambda}_0(\alpha), \widehat{\lambda}_1(\alpha), \widehat{\lambda}_2(\alpha))$, and by taking the second derivative with respect to α , it can be easily shown that the profile 'pseudo' log-likelihood function is an unimodal function of α . Hence it has a unique maximum. The unique maximum $\widehat{\alpha}$ at the k -th stage can be obtained by solving the following fixed point type equation

$$g(\alpha) = \alpha, \quad (42)$$

where

$$\begin{aligned}
h(\alpha) = & \left[\widehat{\lambda}_0(\alpha) \left(\sum_{i \in I_0} a_i^{(k)} y_i^\alpha \ln y_i + \sum_{i \in I_1} a_i^{(k)} y_{1i}^\alpha \ln y_{1i} + \sum_{i \in I_2} a_i^{(k)} y_{2i}^\alpha \ln y_{2i} \right) \right. \\
& + \widehat{\lambda}_1(\alpha) \left(\sum_{i \in I_0} a_i^{(k)} y_i^\alpha \ln y_i + \sum_{i \in I_1 \cup I_2} a_i^{(k)} y_{1i}^\alpha \ln y_{1i} \right) \\
& + \widehat{\lambda}_2(\alpha) \left(\sum_{i \in I_0} a_i^{(k)} y_i^\alpha \ln y_i + \sum_{i \in I_1 \cup I_2} a_i^{(k)} y_{2i}^\alpha \ln y_{2i} \right) \\
& \left. - \left(\sum_{i \in I_0} \ln y_i + \sum_{i \in I_1 \cup I_2} (\ln y_{1i} + \ln y_{2i}) \right) \right], \quad (43)
\end{aligned}$$

then

$$g(\alpha) = \frac{m_0 + 2m_1 + 2m_2}{h(\alpha)}. \quad (44)$$

Very simple iterative procedure can be used to solve the non-linear equation (42), for example $\alpha^{(j+1)} = g(\alpha^{(j)})$, where $\alpha^{(j)}$ is the estimate of α at the j -th iterate, works very well in this case. Hence, we suggest use the following algorithm:

- Step 1: At the k -th stage using $\gamma^{(k)}$, obtain $u_1^{(k)}, u_2^{(k)}, v_1^{(k)}, v_2^{(k)}$.
- Step 2: Solve (42), to obtain $\alpha^{(k+1)}$.
- Step 3: Once $\alpha^{(k+1)}$ is obtained, obtain $\lambda_0^{(k+1)}, \lambda_1^{(k+1)}$ and $\lambda_2^{(k+1)}$. Hence we obtain $\gamma^{(k+1)}$, and go back to Step 1.

Continue the process unless convergence is obtained.

Note that in case of $BEG(\theta, \lambda_0, \lambda_1, \lambda_2)$, the above EM algorithm can be easily modified by replacing $\alpha = 1$, throughout. In this case at the M-step, the solutions can be obtained explicitly.

5 DATA ANALYSIS

5.1 SIMULATED DATA

In this section we have analyzed one simulated data set to show how the proposed EM algorithm works in practice. The generation of random samples from a $BWG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$ is simple. The following algorithm can be used for that purpose.

ALGORITHM:

1. Generate n from $GM(\theta)$.

S.N.		S.N.		S.N.	
1.	0.5405 0.3133	2.	0.1657 0.2830	3.	1.2877 0.9139
4.	0.4777 0.3620	5.	0.5065 0.9563	6.	0.4009 0.6170
7.	0.2147 0.2361	8.	0.5282 0.5282	9.	0.2760 0.2760
10.	0.1006 0.1006	11.	0.5793 0.4274	12.	0.3073 0.3073
13.	0.3395 0.6090	14.	0.2888 0.2888	15.	0.3262 0.2634
16.	0.3919 0.9232	17.	0.4929 0.2192	18.	0.7873 1.3645
19.	0.4425 0.4425	20.	0.8439 0.5858	21.	0.1734 0.5460
22.	0.5660 0.6563	23.	0.2925 0.8118	24.	0.1710 0.2258
25.	0.9225 0.8032				

Table 1: Simulated data from BWG(0.5,2.0,1.0,1.0,1).

2. Generate $\{v_{01}, \dots, v_{0n}\}$ from $\text{WE}(\alpha, \lambda_0)$, similarly, $\{v_{11}, \dots, v_{1n}\}$ from $\text{WE}(\alpha, \lambda_1)$ and $\{v_{21}, \dots, v_{2n}\}$ from $\text{WE}(\alpha, \lambda_2)$.
3. Obtain $u_{1k} = \min\{v_{0k}, v_{1k}\}$ and $u_{2k} = \min\{v_{0k}, v_{2k}\}$, for $k = 1, \dots, n$.
4. Compute the desired (y_1, y_2) as $y_1 = \min\{u_{11}, \dots, u_{1n}\}$ and $y_2 = \min\{u_{21}, \dots, u_{2n}\}$.

We have generated a sample of size 25 from a BWG distribution, with $\theta = 0.5$, $\alpha = 2.0$, $\lambda_0 = 1.0$, $\lambda_1 = 1.0$, $\lambda_2 = 1.0$, and they are presented in Table 1. In this case $m_0 = 6$, $m_1 = 8$ and $m_2 = 11$.

We have started the iteration with $\alpha^{(0)} = \lambda_0^{(0)} = \lambda_1^{(0)} = \lambda_2^{(0)} = 1.0$. For fixed θ , we have used the EM algorithm as described in the previous section. For each fixed θ , the profile 'pseudo' log-likelihood function is an unimodal function of α . The maximum of the profile 'pseudo' log-likelihood function is obtained by solving the fixed point type algorithm (42). The iteration stops when the relative difference between two iterates is less than 10^{-5} . A typical profile 'pseudo' log-likelihood function for $\theta = 0.6$, is provided in Figure 2. Finally we obtain the profile log-likelihood function of θ , and the profile log-likelihood function of θ is provided in Figure 3.

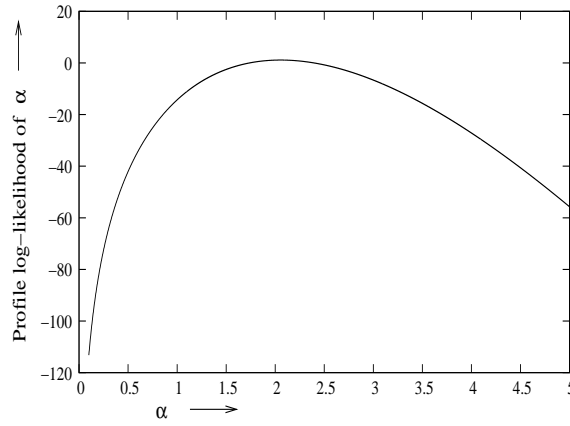


Figure 2: Profile 'pseudo' log-likelihood function of α , for $\theta = 0.6$, at the second iteration for the simulated data.

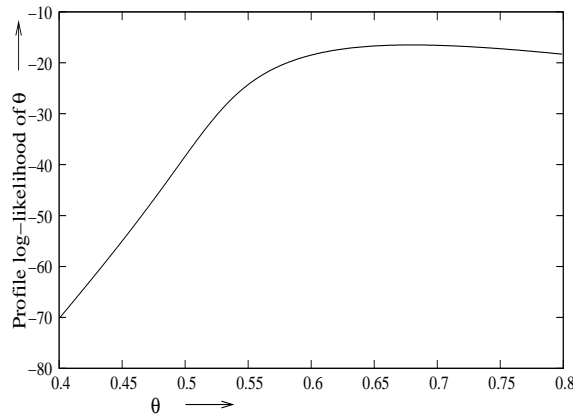


Figure 3: Profile log-likelihood function of θ for the simulated data

The final estimates are as follows:

$$\hat{\theta} = 0.6805, \quad \hat{\alpha} = 2.2302, \quad \hat{\lambda}_0 = 0.9124, \quad \hat{\lambda}_1 = 1.3461 \quad \hat{\lambda}_2 = 0.9883.$$

The associated 95% confidence intervals are

$$0.6805 \mp 0.2115, \quad 2.2302 \mp 0.7312, \quad 0.9124 \mp 0.3312, \quad 1.3461 \mp 0.4011, \quad 0.9883 \mp 0.3127,$$

respectively. For all θ , the EM algorithm converges within 15 iterations, and it seems that the proposed EM algorithm works quite well for this model.

5.2 REAL DATA SET

In this section we have analyzed one real data set to see how the proposed model works for analyzing real bivariate data set. This data set was first analyzed by Meintanis (2007) using Marshall-Olkin bivariate exponential model and later it was analyzed by Kundu and Dey (2009) using Marshall-Olkin bivariate Weibull model. The data set represents the English soccer data played during 2004-2006, and it is available in Meintanis (2007). It represents soccer data where at least one goal has been scored by a kick goal (like penalty kick, foul kick or any other direct kick) by any team and one goal has been scored by the home team. Here the bivariate data (Y_1, Y_2) , where Y_1 represents the time in minutes of the first kick goal scored by any team, and Y_2 represents the first goal scored by the home team. Note that all possibilities are there, namely (i) $Y_1 > Y_2$, (ii) $Y_1 < Y_2$ and (iii) $Y_1 = Y_2$.

Meintanis (2007) observed that the MOBE model does not fit well as one of marginals does not have constant empirical hazard function. Kundu and Dey (2009) observed, using the scaled TTT plots, see Aarset (1987), that both the marginals have increasing failure rates. Therefore, the proposed BWG model may be used for analyzing this data set. Before analyzing the data we have divided all the data points by 100, mainly for computational purposes, and it is not going to affect the inferential procedure.

We have started EM algorithm with the same initial guesses as before, namely $\lambda_0 = \lambda_1 = \lambda_2 = 1$, and we have used the same stopping criterion as in the previous example. The profile log-likelihood function is an unimodal function, and it has been provided in Figure 4. For each θ , the EM algorithm converges within 10 iterations. The final estimates of the unknown parameters are

$$\hat{\theta} = 0.6307, \quad \hat{\alpha} = 1.7221, \quad \hat{\lambda}_0 = 1.3010, \quad \hat{\lambda}_1 = 0.8759 \quad \hat{\lambda}_2 = 2.0448,$$

the corresponding log-likelihood value is -5.294, and the associated 95% confidence intervals

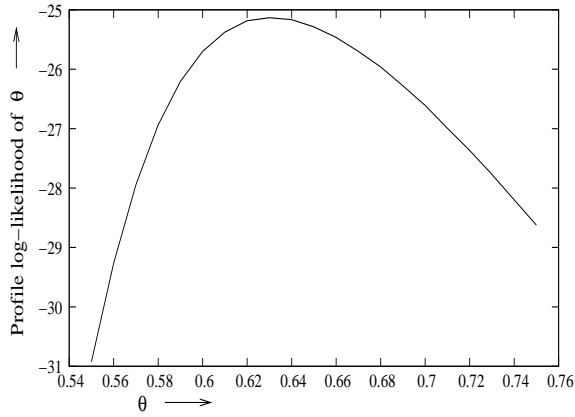


Figure 4: Profile log-likelihood function of θ for the soccer data

are as follows;

$$0.6307 \mp 0.1988 \quad 1.7221 \mp 0.6238 \quad 1.3010 \mp 0.5467 \quad 0.8759 \mp 0.3879 \quad 2.0448 \mp 0.7854.$$

One natural question is how good the proposed BWG model fits the data. Although, there are several goodness of fit tests available for any arbitrary univariate distribution function, not much is available for a general bivariate distribution function. Based on Theorem 3.2, we fit UWG distribution to the two marginals and the minimum. We report the Kolmogorov-Smirnov distances and the associated p values in Table 2. It indicates BWG can be used for analyzing this bivariate data set. Now to compare the results with the fitted MOBW model,

Distribution	KS Distance	p Value
Y_1	0.1524	0.3566
Y_2	0.1628	0.2806
$\min \{Y_1, Y_2\}$	0.1355	0.5056

Table 2: Kolmogorov-Smirnov distances and the associated p values

we compare the maximized log-likelihood values of the two fitted models. From Kundu and Dey (2009), it is observed that the maximized log-likelihood value for the MOBW model is

-13.118. Therefore, based on either AIC or BIC, we prefer BWG than MOBW for this data set.

6 CONCLUSIONS

In this paper we have proposed a new five-parameter bivariate distribution which is a generalization of the Marshall-Olkin bivariate exponential and Weibull distributions. The proposed distribution is a very flexible distribution, and it has a physical interpretation also. EM algorithm has been used to compute the MLEs of the unknown parameters, and it is observed that the proposed algorithm works quite effectively in this case.

It should be mentioned that in this paper we have proposed the bivariate model, but along the same line multivariate generalization is also possible. Development of different properties and proper estimation procedure will be of interest. More work is needed along that direction.

ACKNOWLEDGEMENTS:

The authors would like to thank two unknown referees and the associate editor for their constructive suggestions, which has helped to improve the presentation of the paper significantly.

References

- [1] Aarset, M. V. (1987), "How to identify a bathtub hazard rate", *IEEE Transactions on Reliability*, vol. 36, 106 - 108.

- [2] Barreto-Souza, W. (2012), "Bivariate gamma-geometric law and its induce Levy process", *Journal of Multivariate Analysis*, vol. 109, 130 -145.
- [3] Dinse, G.E. (1982), "Non-parametric estimation of partially incomplete time and types of failure data", *Biometrics*, vol. 38, 417431.
- [4] Franco, M. and Vivo, J-M (2010), "A multivariate extension of Sarhan and Balakrishnan bivariate distribution and its ageing and dependence properties", *Journal of Multivariate Analysis*, vol. 101, 491 - 499.
- [5] Franco, M., Kundu, D. and Vivo, J-M (2011), " Multivariate Extension of Modified Sarhan-Balakrishnan Bivariate Distribution", *Journal of Statistical Planning and Inference*, vol. 141, 3400 - 3412.
- [6] 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*, vol. 32, 1025 - 1034.
- [7] 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*, vol. 36, 1855 - 1866.
- [8] Kundu, D. (2004), "Parameter estimation for partially complete time and type of failure data", *Biometrical Journal*, vol. 46, 165 - 179.
- [9] Kundu, D. and Dey, A. (2009), "Estimating the parameters of the Marshall-Olkin bivariate Weibull distribution by EM algorithm", *Computational Statistics and Data Analysis*, vol. 53, 956 - 965.
- [10] Kundu, D. and Gupta, R. D. (2009), "Bivariate generalized exponential distribution", *Journal of Multivariate Analysis*, vol. 100, 581 - 593.

- [11] Kundu, D. and Gupta, R.D. (2010), “Modified Sarhan-Balakrishnan singular bivariate distribution”, *Journal of Statistical Planning and Inference*, vol. 140, 526 - 538.
- [12] Kundu, D. and Gupta, A. (2013), “Bayes estimation for the Marshall-Olkin bivariate Weibull distribution”, *Computational Statistics and Data Analysis*, vol. 57, 271 - 281.
- [13] Marshall, A.W. and Olkin, I. (1967), “A multivariate exponential distribution”, *Journal of the American Statistical Association*, vol. 62, 3044.
- [14] Marshall, A.W. and Olkin, I. (1997), “ A new method of adding a parameter to a family of distributions with application to the exponential and Weibull families”, *Biometrika*, vol. 84, 641 - 652.
- [15] Meintanis, S.G. (2007), “Test of fit for Marshall-Olkin distributions with applications”, *Journal of Statistical Planning and Inference*, vol. 137, 3954 - 3963.
- [16] Pham, H. and Lai, C-D. (2007), “On recent generalizations of the Weibull distribution”, *IEEE Transactions on Reliability*, vol. 56, 454 - 458.
- [17] Sarhan, A.M. and Balakrishnan, N. (2007), “A new class of bivariate distribution”, *Journal of Multivariate Analysis*, vol. 98, 1508 - 1527.