Journal of Statistics Applications & Probability
An International Journal

On Sarhan-Balakrishnan Bivariate Distribution

D. Kundu[1], *A. Sarhan*[2] *and Rameshwar D. Gupta*[3]

[1] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur, Pin 208016, INDIA
[2] Department of Mathematics and Statistics, Faculty of Science, Dalhousie University, Halifax NS B3H 3J5, Canada
[3] Department of Computer Science and Applied Statistics. The University of New Brunswick, Saint John, Canada

Received: May 18, 2012; Accepted Jul. 16, 2012
Published online: 1 Nov. 2012

Abstract: Recently Sarhan and Balakrishnan (Journal of Multivariate Analysis, 98, 1508 - 1527, 2007) introduced a new singular bivariate distribution using generalized exponential and exponential distributions. They discussed several interesting properties of this new distribution. Sarhan-Balakrishnan did not discuss any estimation procedure of the unknown parameters. In Sarhan-Balakrishnan model, there is no scale parameter. Unfortunately without the presence of any scale parameter, it is difficult to use it for any data analysis purposes. We introduce a scale parameter in the model and it becomes a four-parameter bivariate model. The usual maximum likelihood calculation involves a four dimensional optimization problem. We discuss the maximum likelihood estimation of the unknown parameters using EM algorithm, and it involves only a one-dimensional optimization calculation at each M-step of the EM algorithm. One data analysis has been performed for illustrative purposes. The performance of the EM algorithm is very satisfactory.

Keywords: Generalized exponential distribution; Absolute continuous distribution; EM algorithm; Hazard function; Monte Carlo simulation.

1. Introduction

Recently Sarhan and Balakrishnan (2007) proposed a new class of bivariate distributions based on generalized exponential and exponential distributions. From now on for brevity we denote this distribution as SBBV distribution. The SBBV is a singular distribution, and its cumulative distribution function (CDF) can be expressed as a mixture of an absolute continuous bivariate distribution and a singular distribution, similar to the Marshall and Olkin (1967) bivariate exponential distribution (MOBE).

Sarhan and Balakrishnan (2007) obtained the joint probability density function (PDF), marginal and conditional probability density functions. It is observed that the moments and moment generating function can not be obtained in explicit form, but they can be obtained using special functions. Although, Sarhan and Balakrishnan (2007) discussed several interesting properties, but they did not discuss any estimation procedure of the unknown parameters. We re-visited the model, discuss further properties of this new distribution and provide maximum likelihood estimates (MLEs) of the unknown parameters.

Sarhan and Balakrishnan (2007) did not introduce any scale parameter in their model. Unfortunately, without the presence of any scale parameter, it may be difficult to use it in practice for any data analysis purposes. We introduce the scale parameter in the model and it makes SBBV as a four-parameter model. It is not easy to compute the maximum likelihood estimators of the four unknown parameters. To compute the MLEs directly, one needs to solve a four dimensional optimization problem. To avoid this problem, we treat this as a missing value problem, and we use the EM algorithm to compute the MLEs of the unknown parameters. For implementing the EM algorithm, it is observed that at each M-step, one needs to solve a one-dimensional optimization problem. It is much easier to solve than the direct four dimensional optimization problem. One real data set has been analyzed for illustrative purposes, and the performance is quite satisfactory.

* Corresponding author: e-mail: kundu@iitk.ac.in

The rest of the paper is organized as follows. In section 2, we provide the SBBV model and briefly discuss its different properties. The implementation of the EM algorithm is provided in section 3. Data analysis are presented in section 4. Finally we conclude the paper in section 5.

2. Sarhan-Balakrishnan Bivariate Model

The random variable X has a generalized exponential (GE) distribution with the shape parameter $\alpha > 0$ and the scale parameter $\lambda > 0$, if it has the following cumulative distribution function (CDF);

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

The corresponding probability density function (PDF) is;

$$f(x;\alpha,\lambda) = \alpha\lambda e^{-\lambda x}\left(1 - e^{-\lambda x}\right)^{\alpha-1}. \tag{2}$$

From now on a generalized exponential random with the PDF (2) and the CDF (1) will be denoted by GE(α,λ). The GE was originally introduced by Gupta and Kundu (1999) for analyzing lifetime data, as an alternative to Weibull and gamma distributions. Extensive work has been done on the GE distribution since its inception. The readers are refereed to the recent review article by Gupta and Kundu (2007) and the references cited there. Note that when α = 1, it coincides with the exponential distribution. In this respect it is a generalization of the exponential distribution similarly as the Weibull and gamma distributions but in different ways.

Now let us define the SBBV distribution. Suppose, U_0 follows ($\sim$) GE$(1,\lambda\theta)$, $U_1 \sim$ GE(α_1,λ) and $U_2 \sim$ GE(α_2,λ) and they are independently distributed. Define

$$Y_1 = \min\{U_1, U_0\}, \quad Y_2 = \min\{U_2, U_0\}. \tag{3}$$

Then bivariate random vector (Y_1, Y_2) is said to have SBBV distribution, see Sarhan and Balakrishnan (2007). From now on it will be denoted by SBBV($\alpha_1,\alpha_2,\theta,\lambda$). The joint survival function and the joint PDF of (Y_1, Y_2) can be written as

$$S_{Y_1,Y_2}(y_1,y_2) = e^{-\lambda\theta z}\left\{1 - (1 - e^{-\lambda y_1})^{\alpha_1}\right\}\left\{1 - (1 - e^{-\lambda y_2})^{\alpha_2}\right\}, \tag{4}$$

where $z = \max\{y_1, y_2\}$, $y_1 > 0$, $y_2 > 0$. The corresponding joint PDF can be written as

$$f_{Y_1,Y_2}(y_1,y_2) = \begin{cases} f_1(y_1,y_2) & \text{if} \quad y_1 > y_2 > 0 \\ f_2(y_1,y_2) & \text{if} \quad y_2 > y_1 > 0 \\ f_0(y_1,y_2) & \text{if} \quad y_1 = y_2 > 0 \end{cases}, \tag{5}$$

where

$$f_1(y_1,y_2) = \lambda^2\alpha_2 e^{-\lambda(\theta y_1+y_2)}\left(1 - e^{-\lambda y_2}\right)^{\alpha_2-1} \times \left(\theta - \theta(1 - e^{-\lambda y_1})^{\alpha_1} + \alpha_1 e^{-\lambda y_1}(1 - e^{-\lambda y_1})^{\alpha_1-1}\right), \tag{6}$$

$$f_2(y_1,y_2) = \lambda^2\alpha_1 e^{-\lambda(\theta y_2+y_1)}\left(1 - e^{-\lambda y_1}\right)^{\alpha_1-1} \times \left(\theta - \theta(1 - e^{-\lambda y_2})^{\alpha_2} + \alpha_2 e^{-\lambda y_2}(1 - e^{-\lambda y_2})^{\alpha_2-1}\right), \tag{7}$$

and

$$f_0(y,y) = \lambda\theta e^{-\lambda\theta y}\prod_{i=1}^{2}\left[1 - \left(1 - e^{-\lambda y}\right)^{\alpha_i}\right], \tag{8}$$

see Sarhan and Balakrishnan (2007). Note that λ plays the role of a scale parameter, and Sarhan and Balakrishnan (2007) used λ = 1 in their paper.

The function $f_{Y_1,Y_2}(\cdot,\cdot)$ may be considered to be a density function of the SBBV distribution, if it is understood that the first two terms are densities with respect to the two-dimensional Lebesgue measure and the third term is a density function with respect to one dimensional Lebesgue measure, see for example Bemis and Bain (1972).

The SBBV model may be used as a competing risk model or a shock model similarly as the MOBE model. The marginals of the SBBV model can take different shapes. The hazard functions of the marginals can be increasing ($\alpha > 1$), decreasing ($\alpha < 1$) or constant ($\alpha = 1$). It may be easily observed that for all $y_1 > 0$ and $y_2 > 0$

$$S_{Y_1,Y_2}(y_1,y_2) \geq S_{Y_1}(y_1)S_{Y_2}(y_2) \quad \Leftrightarrow \quad F_{Y_1,Y_2}(y_1,y_2) \geq F_{Y_1}(y_1)F_{Y_2}(y_2). \tag{9}$$

Since (9) is true, Y_1 and Y_2 are positive quadrant dependent, $i.e.$ for every pair of increasing functions $h_1(\cdot)$ and $h_2(\cdot)$

$$\text{Cov}(h_1(Y_1), h_2(Y_2)) \geq 0. \tag{10}$$

3. Maximum Likelihood Estimators

In this section we discuss the maximum likelihood estimators (MLEs) of the unknown parameters of the SBBV($\alpha_1, \alpha_2, \theta, \lambda$) model, based on the following sample $\{(y_{11}, y_{12}), \ldots, (y_{n1}, y_{n2})\}$. We use the following notation;

$$I_1 = \{i; y_{i1} > y_{i2}\}, \quad I_2 = \{i; y_{i1} < y_{i2}\}, \quad I_0 = \{i; y_{i1} = y_{i2} = y_i\}, \quad I = I_1 \cup I_2 \cup I_3,$$

$$n_0 = |I_0|, \quad n_1 = |I_1|, \quad n_2 = |I_2|.$$

Based on the above sample the log-likelihood function becomes;

$$l(\gamma) = \sum_{i \in I_1} \ln f_1(y_{i1}, y_{i2}) + \sum_{i \in I_2} \ln f_2(y_{i1}, y_{i2}) + \sum_{i \in I_0} \ln f_0(y_i, y_i), \tag{11}$$

here $\gamma = (\alpha_1, \alpha_2, \theta, \lambda)$. The MLEs of γ can be obtained by maximizing (11) with respect to γ. It does not have any explicit solutions and the solutions can be obtained only by solving a four dimensional optimization problem, which is clearly not a trivial issue. To avoid that we propose to use the EM algorithm, by treating this problem as a missing value problem.

First let us identify the complete observations as well as the missing observations. It will help us to formulate the EM algorithm. Suppose, instead of (Y_1, Y_2), (U_0, U_1, U_2) are known. For, example we observe $\{u_{i0}, u_{i1}, u_{i2}\}$, for $i = 1, \ldots, n$. The log-likelihood function based on u_i's becomes

$$\begin{aligned} l(\gamma) &= n(\ln \alpha_1 + \ln \alpha_2) + (\alpha_1 - 1) \sum_{i=1}^{n} \ln(1 - e^{-\lambda u_{i1}}) + (\alpha_2 - 1) \sum_{i=1}^{n} \ln(1 - e^{-\lambda u_{i2}}) \\ &\quad + 3n \ln \lambda - \lambda\theta \sum_{i=1}^{n} u_{i0} - \lambda(\sum_{i=1}^{n} u_{i1} + \sum_{i=1}^{n} u_{i2}). \end{aligned} \tag{12}$$

For a given λ, the MLEs of α_1, α_2 and θ can be obtained as

$$\widehat{\alpha}_1(\lambda) = -\frac{n}{\sum_{i=1}^{n} \ln(1 - e^{-\lambda u_{i1}})}, \quad \widehat{\alpha}_2(\lambda) = -\frac{n}{\sum_{i=1}^{n} \ln(1 - e^{-\lambda u_{i2}})}, \quad \widehat{\theta}(\lambda) = \frac{n}{\lambda \sum_{i=1}^{n} u_{i0}}, \tag{13}$$

and finally by maximizing the profile log-likelihood function $l(\widehat{\alpha}_1(\lambda), \widehat{\alpha}_2(\lambda), \widehat{\theta}(\lambda), \lambda)$ with respect to λ, the MLE of λ can be obtained. Therefore, it is clear that if all the u_i's are known, the MLEs of the unknown parameters can be obtained by solving a one dimensional optimization problem. We exploit this property to formulate the EM algorithm.

The following Table 1 will be needed for further development. The exact expressions for p_{ijk}'s are presented in the

Case No.	Different Cases	Prob	Y_1 & Y_2	Set
1	$U_0 < U_1 < U_2$	p_{012}	$U_0 = Y_1 = Y_2$	I_0
2	$U_0 < U_2 < U_1$	p_{021}	$U_0 = Y_1 = Y_2$	I_0
3	$U_1 < U_0 < U_2$	p_{102}	$U_1 = Y_1 < Y_2 = U_0$	I_2
4	$U_1 < U_2 < U_0$	p_{120}	$U_1 = Y_1 < Y_2 = U_2$	I_2
5	$U_2 < U_0 < U_1$	p_{201}	$U_2 = Y_2 < Y_1 = U_0$	I_1
6	$U_2 < U_1 < U_0$	p_{210}	$U_2 = Y_2 < Y_1 = U_1$	I_1

Table 1 Different cases, the associated probabilities and the corresponding Y_1 and Y_2 are presented.

Appendix A. The following observations are useful. In set I_0, only U_0 is observable, both U_1 and U_2 are not observable. In set I_1, U_2 is always observable, but U_0 (U_1) is observable for Case No. 5 (6). Similarly, in set I_2, U_1 is always observable, and U_0 (U_2) is observable for case no. 3 (4). Further, we need the following notation also.

$$b_1 = P(U_2 < U_0 < U_1 | Y_2 < Y_1) = \frac{p_{201}}{p_{201} + p_{210}}, \quad b_2 = P(U_2 < U_1 < U_0 | Y_2 < Y_1) = \frac{p_{210}}{p_{201} + p_{210}},$$

$$c_1 = P(U_1 < U_0 < U_2 | Y_1 < Y_2) = \frac{p_{102}}{p_{102} + p_{120}}, \quad c_2 = P(U_1 < U_2 < U_0 | Y_1 < Y_2) = \frac{p_{120}}{p_{102} + p_{120}}.$$

We also need the following;

$$a_0(u) = E(U_0|U_0 > u) = u + \frac{1}{\lambda\theta}, \tag{14}$$

$$a_1(u) = E(U_1|U_1 > u) = \frac{\int_0^{e^{-\lambda u}} (1-y)^{\alpha_1 - 1} \ln y dy}{\lambda(1-(1-e^{-\lambda u})^{\alpha_1})}, \tag{15}$$

$$a_2(u) = E(U_2|U_2 > u) = \frac{\int_0^{e^{-\lambda u}} (1-y)^{\alpha_2 - 1} \ln y dy}{\lambda(1-(1-e^{-\lambda u})^{\alpha_2})}. \tag{16}$$

With these notation, we are now able to write the 'E'-step of the EM algorithm. Note that the 'E'-step of the 'EM' algorithm can be obtained by replacing the missing values with their expected values. The corresponding log-likelihood function is known as the 'pseudo-log-likelihood' function.

If $i \in I_0$, the 'pseudo-log-likelihood' contribution of (y_i, y_i) is

$$\ln f(a_1(y_i); \alpha_1, \lambda) + \ln f(a_2(y_i); \alpha_2, \lambda) + \ln f(y_i; 1, \lambda\theta).$$

If $i \in I_1$, the 'pseudo-log-likelihood' contribution of (y_{i1}, y_{i2}) is

$$b_1 \left[\ln f(a_1(y_{i1}); \alpha_1, \lambda) + \ln f(y_{i2}; \alpha_2, \lambda) + \ln f(y_{i1}; 1, \lambda\theta)\right] +$$
$$b_2 \left[\ln f(y_{i1}; \alpha_1, \lambda) + \ln f(y_{i2}; \alpha_2, \lambda) + \ln f(a_0(y_{1i}); 1, \lambda\theta)\right],$$

and if $i \in I_2$, the 'pseudo-log-likelihood' contribution of (y_{i1}, y_{i2}) is

$$c_1 \left[\ln f(y_{i1}; \alpha_1, \lambda) + \ln f(a_2(y_{i2}); \alpha_2, \lambda) + f(y_{i2}; 1, \lambda\theta)\right] +$$
$$c_2 \left[\ln f(y_{i1}; \alpha_1, \lambda) + \ln f(y_{i2}; \alpha_2, \lambda) + \ln f(a_0(y_{i2}); 1, \lambda\theta)\right].$$

Combining the three, the 'pseudo-log-likelihood' function of the observed data can be written as;

$$l_{pseudo}(\gamma) = g_1(\alpha_1, \lambda) + g_2(\alpha_2, \lambda) + g_3(\theta, \lambda), \tag{17}$$

where

$$g_1(\alpha_1, \lambda) = \sum_{i \in I_0} \ln f(a_1(y_i); \alpha_1, \lambda) + b_1 \sum_{i \in I_1} \ln f(a_1(y_{i1}); \alpha_1, \lambda) + b_2 \sum_{i \in I_1} \ln f(y_{i1}; \alpha_1, \lambda) +$$
$$\sum_{i \in I_2} \ln f(y_{i1}; \alpha_1, \lambda) \tag{18}$$

$$g_2(\alpha_2, \lambda) = \sum_{i \in I_0} \ln f(a_2(y_i); \alpha_2, \lambda) + \sum_{i \in I_1} \ln f(y_{i2}; \alpha_2, \lambda) + c_1 \sum_{i \in I_2} \ln f(a_2(y_{i2}); \alpha_2, \lambda) +$$
$$c_2 \sum_{i \in I_2} \ln f(y_{i2}; \alpha_2, \lambda) \tag{19}$$

$$g_0(\theta, \lambda) = \sum_{i \in I_0} \ln f(y_i; 1, \theta\lambda) + b_1 \sum_{i \in I_1} \ln f(y_{i1}; 1, \theta\lambda) + b_2 \sum_{i \in I_1} \ln f(a_0(y_{i1}); 1, \theta\lambda) +$$
$$c_1 \sum_{i \in I_2} \ln f(y_{i2}; 1, \theta\lambda) + c_2 \sum_{i \in I_2} \ln f(a_0(y_{i2}); 1, \theta\lambda). \tag{20}$$

Now the 'M'-step of the 'EM' algorithm involves maximizing (17) with respect to the unknown parameters. It can be performed in two stages. Note that for fixed λ, the maximization of $g_1(\alpha_1, \lambda)$, $g_2(\alpha_2, \lambda)$ and $g_3(\theta, \lambda)$ with respect to α_1, α_2 and θ respectively can be obtained as

$$\widehat{\alpha}_1(\lambda) = -\frac{n}{h_1(\lambda)}, \quad \widehat{\alpha}_2(\lambda) = -\frac{n}{h_2(\lambda)}, \quad \widehat{\theta}(\lambda) = \frac{n}{h_3(\lambda)},$$

where

$$h_1(\lambda) = \sum_{i \in I_0} \ln(1 - e^{-\lambda a_1(y_i)}) + b_1 \sum_{i \in I_1} \ln(1 - e^{-\lambda a_1(y_{i1})}) + b_2 \sum_{i \in I_1} \ln(1 - e^{-\lambda y_{i1}}) +$$
$$\sum_{i \in I_2} \ln(1 - e^{-\lambda y_{i1}})$$

$$h_2(\lambda) = \sum_{i\in I_0} \ln(1-e^{-\lambda a_2(y_i)}) + \sum_{i\in I_1} \ln(1-e^{-\lambda y_{i2}}) + c_1 \sum_{i\in I_2} \ln(1-e^{-\lambda a_2(y_{i2})}) + c_2 \sum_{i\in I_2} \ln(1-e^{-\lambda y_{i2}})$$

$$h_0(\lambda) = \lambda \left[\sum_{i\in I_0} y_i + b_1 \sum_{i\in I_1} y_{i1} + b_2 \sum_{i\in I_1} a_0(y_{i1}) + c_1 \sum_{i\in I_2} y_{i2} + c_2 \sum_{i\in I_2} a_0(y_{i2}) \right].$$

The maximization of the $l_{pseudo}(\gamma)$ can be obtained by maximizing

$$g_1(\widehat{\alpha}_1(\lambda),\lambda) + g_2(\widehat{\alpha}_2(\lambda),\lambda) + g_3(\widehat{\theta}(\lambda),\lambda) \tag{21}$$

with respect to λ. Although, we could not prove it theoretically that (21) is an unimodal function, but in our experiments it is always an unimodal function.

We suggest the following algorithm to obtain the MLEs of α_1, α_2, θ and λ;

Algorithm

–Step 1: Take some initial guesses of γ, say $\gamma^{(0)} = (\alpha_1^{(0)}, \alpha_2^{(0)}, \theta^{(0)}, \lambda^{(0)})$.
–Step 2: Compute $a_0, a_1, a_2, b_1, b_2, c_1, c_2$.
–Step 3: Obtain $\lambda^{(1)}$ by maximizing (21) with respect to λ.
–Step 4: Obtain $\gamma^{(1)} = (\alpha_1^{(1)}, \alpha_2^{(1)}, \theta^{(1)}, \lambda^{(1)})$, where

$$\alpha_1^{(1)} = \widehat{\alpha}_1(\lambda^{(1)}), \quad \alpha_2^{(1)} = \widehat{\alpha}_2(\lambda^{(1)}), \quad \theta^{(1)} = \widehat{\theta}(\lambda^{(1)}).$$

–Step 5: Compare $\gamma^{(0)}$ and $\gamma^{(1)}$, if they are close to each other stop the iterative process, otherwise replace $\gamma^{(0)}$ by $\gamma^{(1)}$ and continue the process.

4. Data Analysis

The following data represent the American Football (National Football League) League data and they are obtained from the matches played on three consecutive weekends in 1986. The data were first published in 'Washington Post' and they are also available in Csorgo and Welsh (1989).

It is a bivariate data set, and the variables Y_1 and Y_2 are as follows: Y_1 represents the 'game time' to the first points scored by kicking the ball between goal posts, and Y_2 represents the 'game time' to the first points scored by moving the ball into the end zone. These times are of interest to a casual spectator who wants to know how long one has to wait to watch a touchdown or to a spectator who is interested only at the beginning stages of a game.

The data (scoring times in minutes and seconds) are represented in Table 2. The data set was first analyzed by Csorgo and Welsh (1989), by converting the seconds to the decimal minutes, $i.e.$ 2:03 has been converted to 2.05, 3:59 to 3.98 and so on. We have also adopted the same procedure.

The variables Y_1 and Y_2 have the following structure: (i) $Y_1 < Y_2$ means that the first score is a field goal, (ii) $Y_1 = Y_2$ means the first score is a converted touchdown, (iii) $Y_1 > Y_2$ means the first score is an unconverted touchdown or safety. In this case the ties are exact because no 'game time' elapses between a touchdown and a point-after conversion attempt. Therefore, here ties occur quite naturally and they can not be ignored. It should be noted that the possible scoring times are restricted by the duration of the game but it has been ignored similarly as in Csorgo and Welsh (1989).

If we define the following random variables:

$U_1 =$ time to first field goal

$U_2 =$ time to first safety or unconverted touchdown

$U_3 =$ time to first converted touchdown,

then, Y_1 = min$\{U_1, U_3\}$ and Y_2 = min$\{U_2, U_3\}$. Therefore, (Y_1, Y_2) has a similar structure as the MOBE model or the proposed SBBV model. Csorgo and Welsh (1989) analyzed the data using the MOBE model but concluded that it does not work well. They claimed that Y_2 may be exponential but Y_1 is not.

We would like to examine the behavior of the hazard function of Y_1. We first look at the scaled TTT plot as suggested by Aarset (1987), which provides an idea of the shape of the hazard function of a distribution. For a family with the survival function $S(y) = 1 - F(y)$, the scaled TTT transform, with $H^{-1}(u) = \int_0^{F^{-1}(u)} S(y)dy$ defined for $0 < u < 1$

Y_1	Y_2	Y_1	Y_2	Y_1	Y_2
2:03	3:59	5:47	25:59	10:24	14:15
9:03	9:03	13:48	49:45	2:59	2:59
0:51	0:51	7:15	7:15	3:53	6:26
3:26	3:26	4:15	4:15	0:45	0:45
7:47	7:47	1:39	1:39	11:38	17:22
10:34	14:17	6:25	15:05	1:23	1:23
7:03	7:03	4:13	9:29	10:21	10:21
2:35	2:35	15:32	15:32	12:08	12:08
7:14	9:41	2:54	2:54	14:35	14:35
6:51	34:35	7:01	7:01	11:49	11:49
32:27	42:21	6:25	6:25	5:31	11:16
8:32	14:34	8:59	8:59	19:39	10:42
31:08	49:53	10:09	10:09	17:50	17:50
14:35	20:34	8:52	8:52	10:51	38:04

Table 2 American Football League (NFL) data

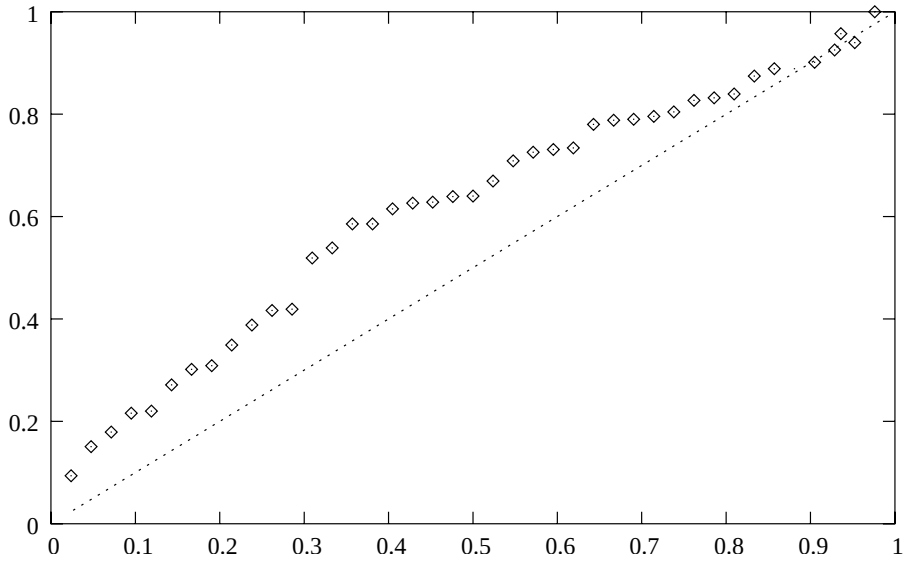

Figure 1 Scaled TTT transform for Y_1.

is $g(u) = H^{-1}(u)/H^{-1}(1)$. The corresponding empirical version of the scaled TTT transform is given by $g_n(r/n) = H_n^{-1}(r/n)/H_n^{-1}(1) = [(\sum_{i=1}^{r} y_{i:n}) + (n-r)y_{r:n}]/(\sum_{i=1}^{n} y_{i:n})$, where $r = 1, \ldots, n$ and $y_{i:n}, i = 1, \ldots, n$ represent the order statistics of the sample. It has been shown by Aarset (1987) that the scaled TTT transform is convex (concave) if the hazard rate is decreasing (increasing), and for bathtub (unimodal) hazard rates, the scaled TTT transform is first convex (concave) and then concave (convex). We plot the scaled TTT transform for Y_1 in Figure 1. From the Figure 1 it is quite apparent that Y_1 has increasing hazard function. Therefore, SBBV may be used to analyze this data set.

We analyze the data using the SBBV model. We have taken the initial guesses of α_1, α_2, θ and λ are all equal to 1. The profile log-likelihood function of λ as given by (21) is provided in Figure 2. It is an unimodal function. The maximization at each step of the EM algorithm is very simple. The EM algorithm converges after 4 steps and the estimates of α_1, α_2, θ and λ become 5.5919, 6.6921, 0.8173 and 1.1112 respectively. The corresponding log-likelihood value is -165.8216. The 95% bootstrap confidence intervals of α_1, α_2, θ and λ are (3.1395, 7.0714), (3.6401, 9.1141), (0.7211, 1.1581) and (0.8208, 1.3412) respectively.

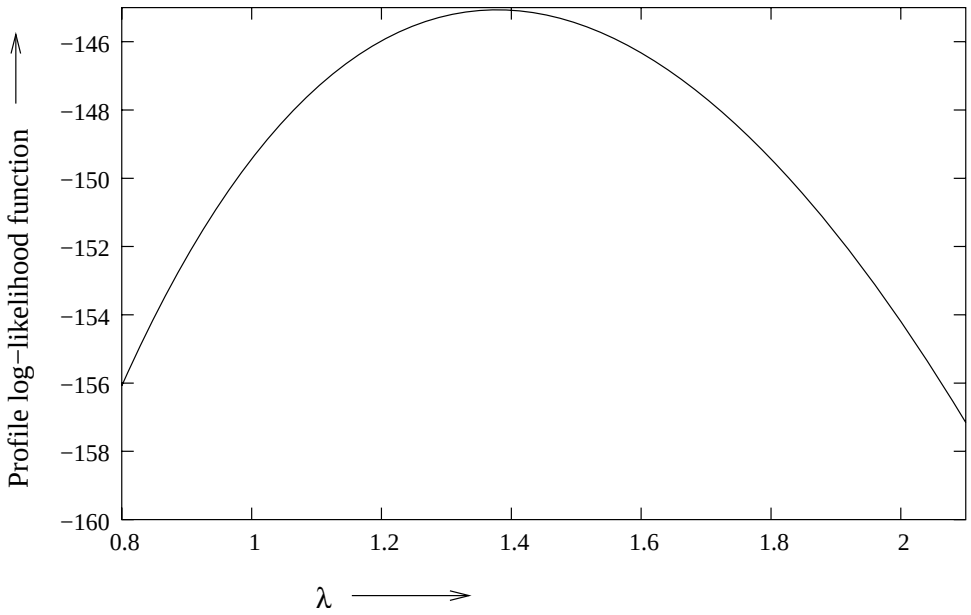

Figure 2 Profile pseudo log-likelihood function of λ as given by (21).

For comparison purposes, we have also fitted the MOBE model to this data set. The estimates of the three parameters are 0.0715, 0.0456 and 0.0030 respectively and the corresponding log-likelihood value is -231.4609. Since MOBE model is not a sub-model of SBBV, we can not use the chi-square test directly. Although, based on AIC or BIC values, we can say that SBBV model is preferable than MOBE model to analyze this data set.

5. Conclusions

In this paper we have mainly consider the maximum likelihood estimation of the unknown parameters of the Sarhan-Balakrishnan bivariate distribution. With the presence of the scale parameter, it becomes a four-parameter model. It is observed that the calculation of the maximum likelihood estimates is not a trivial issue. We suggested to use the EM algorithm to compute the maximum likelihood estimators of the unknown parameters. It may be mentioned that our method can be used to compute the maximum likelihood estimators of the unknown parameters of the MOBE model also.

Acknowledgment

D. Kundu acknowledges the support from DST, Govt. of India. Rameshwar D. Gupta acknowledges the support from the Natural Sciences and Engineering Research Council.

APPENDIX A

In this appendix we provide different expressions of p_{ijk}.

$$p_{012} = P(U_0 < U_1 < U_2) = \int_0^\infty \theta e^{-\theta y} \left[1 - (1-e^{-y})^{\alpha_1} - \frac{\alpha_1}{\alpha_1+\alpha_2}\left(1-(1-e^{-y})^{\alpha_1+\alpha_2}\right)\right] dy$$

$$p_{021} = P(U_0 < U_2 < U_1) = \int_0^\infty \theta e^{-\theta y} \left[1 - (1-e^{-y})^{\alpha_2} - \frac{\alpha_2}{\alpha_1+\alpha_2}\left(1-(1-e^{-y})^{\alpha_1+\alpha_2}\right)\right] dy$$

$$p_{102} = P(U_1 < U_0 < U_2) = \int_0^\infty \theta e^{-\theta y}(1-e^{-y})^{\alpha_1}\left(1-(1-e^{-y})^{\alpha_2}\right) dy$$

$$p_{201} = P(U_1 < U_0 < U_2) = \int_0^\infty \theta e^{-\theta y}(1-e^{-y})^{\alpha_2}\left(1-(1-e^{-y})^{\alpha_1}\right)dy$$

$$p_{120} = P(U_1 < U_2 < U_0) = \frac{\alpha_2}{\alpha_1+\alpha_2}\int_0^\infty \theta e^{-\theta y}(1-e^{-y})^{\alpha_1+\alpha_2}dy$$

$$p_{210} = P(U_2 < U_1 < U_0) = \frac{\alpha_1}{\alpha_1+\alpha_2}\int_0^\infty \theta e^{-\theta y}(1-e^{-y})^{\alpha_1+\alpha_2}dy$$

References

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

[2] Bemis, B., Bain, L.J. and Higgins, J.J. (1972), "Estimation and hypothesis testing for the parameters of a bivariate exponential distribution", *Journal of the American Statistical Association*, vol. 67, 927 - 929.

[3] Csorgo, S. and Welsh, A.H. (1989), "Testing for exponential and Marshall-Olkin distribution", *Journal of Statistical Planning and Inference*, vol. 23, 287 - 300.

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

[5] Gupta, R. D. and Kundu, D. (2007), "Generalized exponential distributions: existing results and some recent developments", *Journal of Statistical Planning and Inference*, vol. 137, 3525 - 3536.

[6] Marshall, A.W. and Olkin, I. (1967), "A multivariate exponential distribution", *Journal of the American Statistical Association*, vol. 62, 30 - 44.

[7] Sarhan, A. and Balakrishnan, N. (2007), "A new class of bivariate distribution and its mixture", *Journal of Multivariate Analysis*, vol. 98, 1508 - 1527.