Multivariate Extension of Modified Sarhan-Balakrishnan Bivariate Distribution

Manuel Franco[1] & Debasis Kundu[2] & Juana-Maria Vivo[3]

Abstract

Recently Kundu and Gupta (2010, Modified Sarhan-Balakrishnan Singular Bivariate Distribution, Journal of Statistical Planning and Inference, 140, 526 - 538) introduced the modified Sarhan-Balakrishnan bivariate distribution and established its several properties. In this paper we provide a multivariate extension of the modified Sarhan-Balakrishnan bivariate distribution. It is a distribution with a singular part. Different ageing and dependence properties of the proposed multivariate distribution have been established. The moment generating function, the product moments can be obtained in terms of infinite series. The multivariate hazard rate has been obtained. We provide the EM algorithm to compute the maximum likelihood estimators and an illustrative example is performed to see the effectiveness of the proposed method.

Key Words and Phrases: Generalized exponential distribution; maximum likelihood estimator; multivariate failure rate; hazard gradient; EM algorithm; singular distribution.

[1] Department of Statistics and Operations Research, University of Murcia, 30100 Murcia, Spain

[2] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. Corresponding author. e-mail: kundu@iitk.ac.in

[3] Department of Quantitative Methods for Economy, University of Murcia, 30100 Murcia, Spain

1 Introduction

Gupta and Kundu [7] introduced the generalized exponential distribution (GE) as an alternative to the well known gamma or Weibull distribution. It has some desirable properties compared to Weibull and gamma distributions. Several interesting developments on GE distribution have taken place in the last few years. For a current account on the GE distribution, the readers are referred to the recent review article by Gupta and Kundu [8].

Recently, Sarhan and Balakrishnan [19] introduced a new class of bivariate distribution of Marshal-Olkin type based on GE and exponential distributions. From now on we call this as the bivariate Sarhan-Balakrishnan distribution. The bivariate Sarhan-Balakrishnan distribution is not an absolute continuous distribution, because it has a positive probability on the $x_1 = x_2$ axis. Several properties have been discussed by the authors, and recently Franco and Vivo [6] introduce multivariate Sarhan-Balakrishnan distribution, and discussed its ageing and dependence properties.

Very recently Kundu and Gupta [13] introduced the modified bivariate Sarhan-Balakrishnan distribution using similar approach as the bivariate Sarhan-Balakrishnan model. In the bivariate Sarhan-Balakrishnan model, Sarhan and Balakrishnan [19] used the minimization process between the GE and exponential distributions, where as in the modified bivariate Sarhan-Balakrishnan model Kundu and Gupta [13] used the minimization process between different GE distributions. Due to the presence of one extra shape parameter, the modified bivariate Sarhan-Balakrishnan model is more flexible than the bivariate Sarhan-Balakrishnan model. Different properties, and the estimation procedure have been developed in the same paper.

The aim of this paper is to develop the multivariate (p-variate) modified Sarhan-Balakrishnan model using the same minimization process, and from $(p+1)$ independent GE distributions.

We call this new distribution as the Marshall-Olkin multivariate GE (MOMGE) distribution. It introduces positive dependence among the variables. Any $q < p$ dimensional subset of p variate MOMGE model is a q-variate MOMGE model. The survival function of the proposed q-variate MOMGE distribution function can be written in a very convenient form. It is a distribution with a singular part, and the decomposition of the absolute continuous part and the singular part is clearly unique. We provide the explicit expression of the singular part and the absolute continuous part.

It may be mentioned that the importance of the ageing and dependence notions have been well established in the statistical literature, see for example Lai and Xie [15]. In many reliability and survival analysis applications, it has been observed that the components are often positively dependent in some stochastic sense. Hence, the derivations of ageing and dependence properties for the MOMGE model have their own importance. Similarly, the extreme order statistics, for example the minimum and maximum order statistics play an important role in several statistical applications, where the components are dependent, see for example Arnold, Balakrishnan and Nagaraja [1]. In this paper the distributions of both extreme order statistics of a random sample from a MOMGE model and their stochastic ageing are studied in details.

Estimation of the unknown parameters is an important problem in any statistical inference. The maximum likelihood estimators (MLEs), as expected, cannot be obtained in explicit forms. They have to be obtained by solving $(p+1)$ non-linear equations. We propose to use the EM algorithm to compute the MLEs. It is observed that in each 'E'-step, the corresponding 'M'-step can be performed by solving $(p + 1)$ - one dimensional non-linear equations. It definitely saves considerable amount of computational time. We perform one simulation example to illustrate the proposed method, and finally we conclude the paper.

Rest of the paper is organized as follows. In Section 2, we briefly discuss about the ageing and dependence properties. In Section 3, we introduce the MOMGE model and provide the survival function, marginals and conditional distribution. Different ageing and dependence properties are discussed in Section 4. EM algorithm is proposed in Section 5. Monte-Carlo simulations are presented in Section 6, and finally we conclude the paper in Section 7.

2 Preliminaries

2.1 Ageing

First we introduce the following notations: $I = (1, \cdots, p)$, $I_k = (i_1, \cdots, i_k) \subset I = (1, \cdots, p)$ where $1 \leq i_1 < i_2 < \cdots < i_k \leq p$. The distribution function of a random vector $\boldsymbol{X} = (X_1, \cdots, X_p)$ will be denoted by $F_I(\cdot)$ or $F_{\boldsymbol{X}}(\cdot)$, and for $(X_{i_1}, \cdots, X_{i_k})$ it will be denoted by $F_{I_k}(\cdot)$. The corresponding survival functions will be denoted by $S_I(\cdot)$ or $S_{\boldsymbol{X}}(\cdot)$, and $S_{I_k}(\cdot)$ respectively.

A p-dimensional random vector $\boldsymbol{X} = (X_1, \cdots, X_p)$, or its joint distribution function $F_I(\cdot)$ is said to have a multivariate increasing failure rate (MIFR) property, if

$$\frac{P(X_{i_1} > x_{i_1} + t, \cdots, X_{i_k} > x_{i_k} + t)}{P(X_{i_1} > x_{i_1}, \cdots, X_{i_k} > x_{i_k})} = \frac{S_{I_k}(x_{i_1} + t, \cdots, x_{i_k} + t)}{S_{I_k}(x_{i_1}, \cdots, x_{i_k})}, \tag{1}$$

decreases in $x_{i_1}, \cdots, x_{i_k}$, $\forall t > 0$, for each subset $I_k = (i_1, \cdots, i_k) \subseteq (1, \cdots, p)$.

The random vector $\boldsymbol{X}$ is said to have multivariate new better than used (MNBU) property, if $\forall t > 0$, and $\forall x_{i_j} \geq 0$

$$S_{I_k}(x_{i_1} + t, \cdots, x_{i_k} + t) \leq S_{I_k}(x_{i_1}, \cdots, x_{i_k}) S_{I_k}(t, \cdots, t). \tag{2}$$

If

$$\int_0^\infty \frac{S_{I_k}(x_{i_1} + t, \cdots, x_{i_k} + t)}{S_{I_k}(x_{i_1}, \cdots, x_{i_k})} dt \tag{3}$$

decreases in $x_{i_1}, \cdots, x_{i_k}$, then $\boldsymbol{X}$ us said to have a multivariate decreasing mean residual life (MDMRL).

Moreover, the random vector $\boldsymbol{X}$ is said to be new better than used in expectation (MNBUE) if for $\forall x_{ij} \geq 0$,

$$\int_0^\infty S_{I_k}(x_{i_1}+t,\cdots,x_{i_k}+t)dt \leq S_{I_k}(x_{i_1},\cdots,x_{i_k}) \int_0^\infty S_{I_k}(t,\cdots,t)dt. \tag{4}$$

Along the same line the dual classes also can be defined. Among the different ageing notions, the following relation holds;

$$\text{MNBUE} \;\Leftarrow\; \text{MNBU} \;\Leftarrow\; \text{MIFR} \;\Rightarrow \text{MDMRL} \;\Rightarrow \text{MNBUE}.$$

2.2 Multivariate Hazard Gradient

Johnson and Kotz [10] defined the multivariate hazard gradient of a p-variate random vector $\boldsymbol{X} = (X_1,\cdots,X_p)$ as follows. Suppose $X_1,\cdots,X_p$ are p absolutely continuous random variables, then the hazard gradient of $\boldsymbol{X}$ for $\mathbf{x} = (x_1,\cdots,x_p)$ is

$$h_X(\mathbf{x}) = \left(-\frac{\partial}{\partial x_1},\cdots,-\frac{\partial}{\partial x_p}\right) \ln P(X_1 > x_1,\cdots,X_p > x_p). \tag{5}$$

The meaning of the i-th component of (5) is similar to the univariate case. It is the punctual failure probability for the i-th component when all the components are working and have age x_i, $i = 1,\cdots,n$. Marshall and Olkin [17] showed that h_X defined above uniquely define the corresponding distribution function. See also, for example, Shanbhag and Kotz [20] or Marshall [16] in this connection. It may be noted that the multivariate extension of the hazard rate function is not unique. Basu [3] has also defined another multivariate hazard function, although it may not uniquely define the corresponding distribution function and it is not pursued here.

If for all values of $\mathbf{x}$, all components of $h_X(\mathbf{x})$ are increasing (decreasing) functions of the corresponding variables, then the distribution is called multivariate increasing (decreasing) hazard gradient.

2.3 Dependence

Several notions of positive and negative dependence for multivariate distributions of varying degree of strengths are available in the literature, see for example Colangelo, Hu and Shaked [4], Joe [9], Balakrishnan and Lai [2] and the references cited there.

A random vector $\boldsymbol{X}$ is said to be positive upper orthant dependent (PUOD) if

$$S_I(\mathbf{x}) = P(X_1 > x_1, \cdots, X_p > x_p) \geq \prod_{i=1}^{p} P(X_i > x_i) = \prod_{i=1}^{p} S_i(x_i). \tag{6}$$

Now we will define the right tail increasing (RTI) property of a random vector, and for that we need the following notations. For any set $A \subseteq \{1, \cdots, p\}$, say $A = \{i_1, \cdots, i_q\}$, $\boldsymbol{X}_A = (X_{i_1}, \cdots, X_{i_q})$, similarly, $\mathbf{x}_A$ is also defined. The random vector $\boldsymbol{X}$ is said to have RTI property, if

$$P(X_B > \mathbf{x}_B | X_A > \mathbf{x}_A) \tag{7}$$

is a non-decreasing in $\mathbf{x}_A$ for all $\mathbf{x}_B$. Here the sets A and B are partitions of $\{1, \cdots, p\}$, and non-decreasing in $\mathbf{x}_A$ ($\mathbf{x}_B$) means for each components $\mathbf{x}_A$ ($\mathbf{x}_B$).

Another multivariate dependence notion is the multivariate right corner set increasing (RCSI). A random vector $\boldsymbol{X}$ is said to have multivariate RCSI property, if

$$P(X_1 > x_1, \cdots, X_p > x_p | X_1 > x_1', \cdots, X_p > x_p') \tag{8}$$

increases in $x_1', \cdots, x_p'$ for every choice of $(x_1, \cdots, x_p)$.

3 Marshall-Olkin Multivariate GE Distribution

The univariate generalized exponential distribution has the following cumulative distribution function (CDF) and the probability density function (PDF) respectively, for $x > 0$ are

$$F(x; \alpha, \lambda) = \left(1 - e^{-\lambda x}\right)^{\alpha}, \quad \text{and} \quad f(x; \alpha, \lambda) = \alpha \lambda e^{-\lambda x} \left(1 - e^{-\lambda x}\right)^{\alpha - 1}. \tag{9}$$

Here $\alpha > 0$, $\lambda > 0$ are the shape and scale parameters respectively. It is immediate that when $\alpha = 1$, it coincides with the exponential distribution. From now on, a generalized exponential (GE) distribution with the shape and scale parameters as α and λ respectively will be denoted by GE(α, λ). The PDF, CDF, and the survival function (SF) of GE$(\alpha, 1)$ will be denoted by $f(\cdot; \alpha)$, $F(\cdot; \alpha)$ and $S(\cdot; \alpha)$ respectively.

Now we are in a position to define MOMGE distribution, along the same line as the MOMGE distribution, as proposed by Kundu and Gupta [13]. From now on unless otherwise mentioned, it is assumed that $\alpha_1 > 0, \cdots, \alpha_{p+1} > 0, \lambda > 0$. Suppose U_1 follows ($\sim$) GE(α_1, λ), $\cdots, U_p \sim$ GE(α_p, λ), $V \sim$ GE(α_{p+1}, λ), and they are independently distributed. Now define

$$X_1 = \min\{U_1, V\}, \quad \cdots, \quad X_p = \min\{U_p, V\}. \tag{10}$$

Then we say, $\boldsymbol{X} = (X_1, \cdots, X_p)$ has the MOMGE distribution of order p, and it will be denoted by MOMGE$(\boldsymbol{\alpha}, \lambda, p)$, where $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_{p+1})$.

From now on unless otherwise mentioned, it is assumed that $\lambda = 1$, and in that case it will be denoted by MOMGE$(\boldsymbol{\alpha}, p)$. It is clear that similarly as MOMGE model, the proposed MOMGE model also can be used quite effectively as a competing risks model or a shock model.

The following result provides the joint survival function (JSF) of the MOMGE model.

THEOREM 3.1: If $\boldsymbol{X} = (X_1, \cdots, X_p) \sim$ MOMGE$(\boldsymbol{\alpha}, p)$, then the JSF is

$$S_X(\mathbf{x}) = P(X_1 > x_1, \cdots, X_p > x_p) = \prod_{i=1}^{p} S(x_i; \alpha_i) S(z; \alpha_{p+1}) \tag{11}$$

here $z = \max\{x_1, \cdots, x_p\}$, and $S(x; \alpha) = 1 - \left(1 - e^{-x}\right)^{\alpha}$

PROOF: It simply follows from the definition, and the details are omitted. ∎

It is clear that for $n > 1$, the MOMGE is not an absolutely continuous distribution. In this case MOMGE has an absolutely continuous part and a singular part. The MOMGE distribution function can be written as

$$F_X(\mathbf{x}) = \beta F_a(\mathbf{x}) + (1-\beta)F_s(\mathbf{x}).$$

Here $0 < \beta < 1$, $F_a(\cdot)$ and $F_s(\cdot)$ denote the absolute continuous part and singular part of $F_X(\cdot)$ respectively. The corresponding PDF of $\boldsymbol{X}$ can be written as

$$f_{\boldsymbol{X}}(\mathbf{x}) = \beta f_a(\mathbf{x}) + (1-\beta) f_s(\mathbf{x}) \quad (12)$$

In writing (12) it is understood that $f_a(\cdot)$ is a PDF with respect to p-dimensional Lebesgue measure, and the singular part $f_s(\cdot)$ may be considered as a collection of PDFs with respect to $1, ..., (p-1)$-dimensional Lebesgue measures. Although their expressions are not required here, the decomposition of $f_X(\cdot)$ into PDFs with respect to $(k+1)$-dimensional Lebesgue measures, $k = 0, ..., (p-1)$, is discussed in Appendix A.

Now we provide the distribution functions of the marginals, the conditionals and the extreme order statistics of the MOMGE model. We will be using the following notation. For any two vectors, $\mathbf{a} = (a_1, \cdots, a_m)$ and $\mathbf{b} = (b_1, \cdots, b_m)$ of same dimension, $\mathbf{a} > \mathbf{b}$ means $a_i > b_i$ for $i = 1, \cdots, m$.

THEOREM 3.2: If $(X_1, \cdots, X_p) \sim$ MOMGE$(\boldsymbol{\alpha}, p)$, then

(a) The marginal PDF of X_j for $x > 0$ is given by

$$f_{X_j}(x) = f(x;\alpha_j) + f(x;\alpha_{p+1}) - f(x;\alpha_j + \alpha_{p+1}). \quad (13)$$

(b) For any k-dimensional marginal $\mathbf{X}_{I_k} = (X_{i_1}, ..., X_{i_k}) \sim MOMGE(\alpha_{i_1}, ..., \alpha_{i_k}, \alpha_{p+1}, k)$.

(c) The conditional survival function of $(\mathbf{X}_{I_k} \mid \mathbf{X}_{I-I_k} > \mathbf{x}_{I-I_k})$ where $I - I_k = \{i \in I : i \notin I_k\}$, is an absolute continuous survival function as follows;

$$P(\mathbf{X}_{I_k} > \mathbf{x}_{I_k} \mid \mathbf{X}_{I-I_k} > \mathbf{x}_{I-I_k}) = \begin{cases} \prod_{i \in I_k} S(x_i;\alpha_i) & \text{if} \quad z = v \\ \prod_{i \in I_k} S(x_i;\alpha_i) \frac{S(z;\alpha_{p+1})}{S(v;\alpha_{p+1})} & \text{if} \quad z > v, \end{cases} \quad (14)$$

here $z = \max\{x_1, \cdots, x_p\}$ and $v = \max\{x_i : i \in I - I_k\}$.

(d) If $X_{(1)} = \min\{X_1, \cdots, X_p\}$, then

$$S_{X_{(1)}}(t) = P(X_{(1)} > t) = \prod_{i=1}^{p+1} S(t; \alpha_i).$$

(e) If $X_{(p)} = \max\{X_1, \cdots, X_p\}$, then

$$S_{X_{(p)}}(t) = P(X_{(p)} > t) = S(t; \alpha_{p+1})\ S(t, \bar{\alpha}),$$

here $\bar{\alpha} = \alpha_1 + \cdots + \alpha_p$.

(f) If $X_{(r)}$ denotes the r-th order statistic $X_{(r)}$ from $(X_1, \cdots, X_p)$, then its survival function is given by

$$S_{X_{(r)}}(t) = P(X_{(r)} > t) = S(t; \alpha_{p+1}) \sum_{j=0}^{r-1} \sum_{P_j} \left(1 - S(t; \alpha_{P_j})\right) \prod_{k=j+1}^{p} S(t; \alpha_{i_k})$$

where $\alpha_{P_j} = \sum_{k=1}^{j} \alpha_{i_k}$, and P_j is the set of all permutations $(i_1, ..., i_p)$ of $(1, 2, ..., p)$ such that $i_1 < ... < i_j$ and $i_{j+1} < ... < i_p$.

Proof: (a) - (d) are quite straight forward, and they can be easily obtained. The proof (e) can be obtained along the same line as the proof of part (4) of proposition 3.2 of Franco and Vivo [6]. The proof of (f) can be obtained along the same line as (e).

Comments: It is cleat that the survival functions of X_j and $X_{(p)}$ are product of two GE survival functions. Therefore, from Proposition 5.3 of Franco and Vivo [6], X_j and $X_{(p)}$ are classified in according to their parameters α_j, α_{p+1} and $\overline{\alpha}$, respectively.

4 Properties

First we provide the following multivariate stochastic ordering results.

Theorem 4.1: Let $\boldsymbol{X} = (X_1, \cdots, X_p) \sim \text{MOMGE}(\boldsymbol{\alpha}, p)$, and $\boldsymbol{Y} = (Y_1, \cdots, Y_p) \sim \text{MOMGE}(\boldsymbol{\beta}, p)$, here $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_{p+1})$, and $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_{p+1})$. If $\alpha_i \le \beta_i$, for $i = 1, \cdots, p+1$, then $\boldsymbol{X} <_{st} \boldsymbol{Y}$, *i.e.* $\boldsymbol{X}$ is smaller than $\boldsymbol{Y}$ in the usual stochastic order.

Proof: Note that $\boldsymbol{X} <_{st} \boldsymbol{Y}$, if and only if for all $u_1 > 0, \cdots, u_p > 0$,

$$P(X_1 > u_1, \cdots, X_p > u_p) \le P(Y_1 > u_1, \cdots, Y_p > u_p).$$

Since $S(u; \alpha_i) \le S(u, \beta_i)$, for $u > 0$, and for $i = 1, \cdots, p$, the result follows. ∎

Theorem 4.2: Let $\boldsymbol{X} \sim \text{MOMGE}(\boldsymbol{\alpha}, p)$, then $S_{\boldsymbol{X}}(\mathbf{x}) = P(\boldsymbol{X} \ge \mathbf{x})$ has a multivariate total positivity of order (MTP_2) property.

Proof: Recall that $S_{\boldsymbol{X}}(\mathbf{x})$ has MTP_2 property, if and only if

$$\frac{S_{\boldsymbol{X}}(\mathbf{x})S_{\boldsymbol{X}}(\mathbf{y})}{S_{\boldsymbol{X}}(\mathbf{x} \vee \mathbf{y})S_{\boldsymbol{X}}(\mathbf{x} \wedge \mathbf{y})} \le 1. \tag{15}$$

Here $\mathbf{x} = (x_1, \cdots, x_p)$, $\mathbf{y} = (y_1, \cdots, y_p)$, $\mathbf{x} \vee \mathbf{y} = \{x_1 \vee y_1, \cdots, x_p \vee y_p\}$, $\mathbf{x} \wedge \mathbf{y} = \{x_1 \wedge y_1, \cdots, x_p \wedge y_p\}$, where $c \vee d = \max\{c, d\}$, $c \wedge d = \min\{c, d\}$.

We will use the following notations:

$$u = \max\{x_1, \cdots, x_p\}, \quad v = \max\{y_1, \cdots, y_p\},$$

$$a = \max\{x_1 \vee y_1, \cdots, x_p \vee y_p\}, \quad b = \max\{x_1 \wedge y_1, \cdots, x_p \wedge y_p\}.$$

Therefore, observe that

$$b \le \min\{u, v\} \le \max\{u, v\} = a. \tag{16}$$

First consider the case when $u \le v$, therefore,

$$b \le u \le v = a.$$

Now the left hand side of (15) can be written as

$$\frac{S_{\boldsymbol{X}}(\mathbf{x})S_{\boldsymbol{X}}(\mathbf{y})}{S_{\boldsymbol{X}}(\mathbf{x} \vee \mathbf{y})S_{\boldsymbol{X}}(\mathbf{x} \wedge \mathbf{y})} = \frac{1 - (1 - e^{-u})^{\alpha_{p+1}}}{1 - (1 - e^{-b})^{\alpha_{p+1}}}. \tag{17}$$

Since $b \leq u$, the right hand side of (17) is less than or equal to 1. ∎

THEOREM 4.3: Let $\boldsymbol{X} \sim \text{MOMGE}(\boldsymbol{\alpha}, p)$. If $\alpha_i \geq 1$ for $i = 1, \cdots, p+1$, then $\boldsymbol{X}$ is MIFR, and if $\alpha_i \leq 1$, for $i = 1, \cdots, p+1$, then $\boldsymbol{X}$ is MDFR. Otherwise $\boldsymbol{X}$ can be neither MIFR nor MDFR.

PROOF: The proof can be obtained along the same line as the proof of Theorem 4.3 of Franco and Vivo [6], and therefore it is avoided. ∎

THEOREM 4.4: Let $\boldsymbol{X} \sim \text{MOMGE}(\boldsymbol{\alpha}, p)$. If $\alpha_i \geq 1$ for $i = 1, \cdots, p+1$, then $\boldsymbol{X}$ is MIHG, and if $\alpha_i \leq 1$, for $i = 1, \cdots, p+1$, then $\boldsymbol{X}$ is MDHG. Otherwise $\boldsymbol{X}$ can be neither MIHG nor MDHG.

PROOF: The i-th component of the hazard gradient of the random vector $\boldsymbol{X}$ can be written as

$$\begin{aligned} h_i(\mathbf{x}) &= -\frac{\partial}{\partial x_i} \ln S_{\boldsymbol{X}}(\mathbf{x}) \\ &= \begin{cases} h(x_i, \alpha_i) & \text{if} \quad x_i < \max\{x_1, \cdots, x_p\} \\ h(x_i; \alpha_i) + h(x_i; \alpha_{p+1}) & \text{if} \quad x_i = \max\{x_1, \cdots, x_p\} \end{cases} \end{aligned}$$

Here $h(\cdot; \alpha)$ denotes the hazard function of GE(α). Since GE(α) has increasing (decreasing) hazard rate for $\alpha > (<)1$, the result immediately follows. ∎

THEOREM 4.5: Let $\boldsymbol{X} \sim \text{MOMGE}(\boldsymbol{\alpha}, p)$. Then $\boldsymbol{X}$ is positive upper orthant dependent.

PROOF: The random vector $\boldsymbol{X}$ is PUOD if and only if its distribution function satisfies (6). To prove this we use the following notation for $i = 1, \cdots, p$;

$$a_i = \left(1 - e^{-x_i}\right)^{\alpha_i}, \quad b_i = \left(1 - e^{-x_i}\right)^{\alpha_{p+1}},$$

and $b = \left(1 - e^{-z}\right)^{\alpha_{p+1}}$, where z is same as defined before. Therefore, the left hand side and right hand side of (6) can be written as

$$\prod_{i=1}^{p}(1-a_i)(1-b) \quad \text{and} \quad \prod_{i=1}^{p}(1-a_i)(1-b_i)$$

respectively. Suppose $x_j = \max\{x_1, \cdots, x_p\}$. Since

$$\prod_{1 \le i \le p, i \ne j} (1-b_i) \le 1,$$

the result immediately follows. ∎

THEOREM 4.6: Let $\boldsymbol{X} \sim \text{MOMGE}(\boldsymbol{\alpha}, p)$. Then $\boldsymbol{X}$ has RTI property and also it has the RCSI property.

PROOF: Using the same notation as of part (c) of Theorem 3.2, since $v \le z$ and $S(z;\alpha_{p+1}) \le S(v;\alpha_{p+1})$, the result immediately follows. Now to prove the second part, we need to show that (8) is an increasing function in x_i', when $x_1, \cdots, x_p$ and $x_1', \cdots, x_{i-1}', x_{i+1}', \cdots, x_p'$ are kept fixed. Therefore, it is enough to prove that the following function

$$\frac{S(x_i \vee x_i'; \alpha_i) S(z''; \alpha_{p+1})}{S(x_i'; \alpha_i) S(z'; \alpha_{p+1})} \tag{18}$$

is an increasing function of x_i, here $z'' = \max\{x_1, \cdots, x_p, x_1', \cdots, x_p')$ and $z' = \max\{x_1', \cdots, x_p')$. Now considering all the cases *i.e.* $x_i' < x_i$, $x_i \le x_i' \le z'$, $z' \le x_i \le z''$, $z'' \le x_i$ and using the fact that $S(x;\alpha)$ is a decreasing function of x, the result follows. ∎

COMMENT: It may be noted that when $\alpha_1 = \cdots = \alpha_{p+1}$ most of the properties will follow using Marshall-Olkin copula properties, see for example Nelsen [18], although for arbitrary α_i's it is not true.

5 EM Algorithm

In this section we propose the EM algorithm to compute the maximum likelihood estimators, similarly as in Karlis [11] or Kundu and Dey [12]. It may be mentioned that the computation of the MLEs using the direct maximization of the likelihood function involves a $(p+1)$ dimensional optimization process. In this section we will show that the computation of the MLEs using EM algorithm involves a one dimensional optimization process at each 'E' step, which is definitely much easier to solve compared to the direct maximization.

First we will mention the possible available data. In general, for all $\mathbf{x} = (x_1, ..., x_p) \in I\!R^p$, there exists a permutation $\mathcal{P}_{k+1} = (i_1, ..., i_p)$ of $I = (1, ..., p)$ such that $x_{i_1} < ... < x_{i_{k+1}}$ are the $(k+1)$ different components of $\mathbf{x}$ with $0 \leq k \leq p-1$. Here, $k = p-1$ implies all the x_{i_j}'s are different, and $k = 0$ implies all the x_{i_j}'s are equal. So, we can consider a partition $J_{i_1}, ..., J_{i_{k+1}}$ of I, $\cup_{j=1}^{k+1} J_{i_j} = I$ and $J_{i_r} \cap J_{i_s} = \varnothing$ for $r \neq s \in \{1, ..., k+1\}$, where $J_{i_j} = \{i_l \in \mathcal{P}_{k+1} : x_{i_l} = x_{i_j}\}$ and $m_{i_j} =\mid J_{i_j} \mid$ for $j = 1, ..., k+1$, with $m_{i_1} + ... + m_{i_{k+1}} = p$.

Now, taking into account that $X \sim MOMGE(\alpha, p)$, for $i \neq j \neq k+1$, we have that $(X_i = X_j < X_{k+1})$ has null probability, since

$$\begin{aligned} P\left(X_i = X_j < X_{k+1}\right) &= P\left(\min\left(U_i, V\right) = \min\left(U_j, V\right) < \min\left(U_{k+1}, V\right)\right) \\ &= P\left(U_i = U_j < \min\left(U_{k+1}, V\right)\right) \\ &= P\left(U_j = U_i, U_{k+1} > U_i, V > U_i\right) = 0, \end{aligned} \tag{19}$$

and consequently, there are not possible ties $m_{i_j} = 1$ for $j < k+1$.

Therefore, if $\mathcal{P}_{k+1}$ is same as defined at the beginning of this section, the possible available data will be of the form;

$$\{x_{i_1} < \cdots < x_{i_k} < x^* = x_{i_{k+1}} = \cdots = x_{i_p}\}, \tag{20}$$

where $k = 0, \cdots, p-1$. Note that for $k = 0, \cdots, (p-2)$, we observe the data (20) if

$$U_{i_1} < \cdots < U_{i_k} < V < \min\{U_{i_{k+1}}, \cdots, U_{i_p}\}, \tag{21}$$

happens, and for $k = p-1$, we observe the data (20) if

$$U_{i_1} < \cdots < U_{i_k} < U_{i_{k+1}} < \cdots < U_{i_p} < V \quad \text{or} \quad U_{i_1} < \cdots < U_{i_k} < U_{i_{k+1}} < \cdots < V < U_{i_p}, \tag{22}$$

happens.

We treat this problem as a missing value problem. Here the complete observation denotes the complete information about $U_1, \cdots, U_p$ and V. First we will show that if we know $U_1, \cdots, U_p$ and V, then the MLEs of $\alpha_1, \cdots, \alpha_{p+1}$ can be obtained by a one dimensional optimization process. The log-likelihood contribution of the observed $\{u_1, \cdots, u_p, v\}$ is

$$\sum_{j=1}^{p} \log f(u_j; \alpha_j, \lambda) + \log f(v; \alpha_{p+1}, \lambda). \tag{23}$$

Note that in writing the log-likelihood function (23), it is assumed that the scale parameter λ is also present in $U_i, i = 1, \cdots, p$ and in V. Therefore, based on the complete observations (CO), say

$$\{u_{i1}, \cdots, u_{ip}, v_i\}, \qquad i = 1, \cdots, n, \tag{24}$$

the log-likelihood function is;

$$l(\alpha_1, \cdots, \alpha_{p+1}, \lambda | CO) = \sum_{i=1}^{n} \sum_{j=1}^{p} \log f(u_{ij}; \alpha_j, \lambda) + \sum_{i=1}^{n} \log f(v_i; \alpha_{p+1}, \lambda). \tag{25}$$

For fixed λ, the MLEs of α's can be obtained as

$$\widehat{\alpha}_1(\lambda) = -\frac{n}{\sum_{i=1}^{n} \ln\left(1 - e^{-\lambda u_{i1}}\right)}, \cdots, \widehat{\alpha}_p(\lambda) = -\frac{n}{\sum_{i=1}^{n} \ln\left(1 - e^{-\lambda u_{ip}}\right)},$$
$$\widehat{\alpha}_{p+1}(\lambda) = -\frac{n}{\sum_{i=1}^{n} \ln\left(1 - e^{-\lambda v_i}\right)}, \tag{26}$$

and the MLE of λ can be obtained by maximizing the profile log-likelihood function

$$l(\widehat{\alpha}_1(\lambda), \cdots, \widehat{\alpha}_{p+1}(\lambda), \lambda | CO) \tag{27}$$

with respect to λ. Therefore, it is clear that if we have the information about all the U_i's and V, then the MLEs of $(p+2)$ unknown parameters can be obtained by solving one non-linear equation only.

Now we are in a position to provide the EM algorithm. First we provide the 'E'-step of the proposed algorithm. In case of (20) it is clear that $u_{i_1}, \cdots, u_{i_k}$, v are observable, and $u_{i_{k+1}}, \cdots, u_{i_p}$ are missing. In writing the 'pseudo' log-likelihood function, we replace the missing values by their expected values. We need the following result for further development.

LEMMA 5.1: If $X \sim \text{GE}(\alpha, \lambda)$, then

$$A(c; \alpha, \lambda) = E(X|X > c) = \frac{1}{\lambda S(c; \alpha, \lambda)} \int_{\left(1-e^{-\lambda c}\right)^{\alpha}}^{1} -\log\left(1 - u^{\frac{1}{\alpha}}\right) du. \tag{28}$$

PROOF: Since

$$E(X|X > c) = \frac{1}{S(c; \alpha, \lambda)} \int_c^{\infty} x f(x; \alpha, \lambda) dx,$$

the result follows by using the change of variable. ∎

First we will discuss the 'pseudo' log-likelihood contribution of the observed data (20), for $k = 0, \cdots, p-2$. In this case the 'pseudo' log-likelihood function can be obtained by replacing u_{i_j} with its expectation, namely $u^*_{i_j} = A(v; \alpha_{i_j}, \lambda)$, for $j = k+1, \cdots, p$. For $k = p-1$, all the x_{i_j}'s are distinct, and in this case the original configuration of U_i's and V given $x_{i_1} < \cdots < x_{i_p}$, is

$$U_{i_1} < \cdots < U_{i_k} < U_{i_{k+1}} < \cdots < U_{i_p} < V \quad \text{or} \quad U_{i_1} < \cdots < U_{i_k} < U_{i_{k+1}} < \cdots < V < U_{i_p},$$

with probability $P(U_{i_p} < V) = \dfrac{\alpha_{i_p}}{\alpha_{p+1} + \alpha_{i_p}} = C_{i_p}$ (say) and $P(V < U_{i_p}) = \dfrac{\alpha_{p+1}}{\alpha_{p+1} + \alpha_{i_p}} = D_{i_p}$ (say), respectively. Therefore, using similar idea as in Dinse [5] or Kundu [14], the 'pseudo' log-likelihood contribution of $x_{i_1} < \cdots < x_{i_p}$, is

$$\begin{aligned}
&C_{i_p}\left(\sum_{j=1}^{p} \log f(x_{i_j}; \alpha_{i_j}, \lambda) + \log f(v^*; \alpha_{p+1}, \lambda)\right) + \\
&D_{i_p}\left(\sum_{j=1}^{p-1} \log f(x_{i_j}; \alpha_{i_j}, \lambda) + \log f(x_{i_p}; \alpha_{p+1}, \lambda) + \log f(x^*_{i_p}; \alpha_{i_p}, \lambda)\right) = \\
&\quad \sum_{j=1}^{p-1} \log f(x_{i_j}; \alpha_{i_j}, \lambda) + C_{i_p} \log f(x_{i_p}; \alpha_{i_p}, \lambda) + C_{i_p} \log f(v^*; \alpha_{p+1}, \lambda) + D_{i_p} \log f(x_{i_p}; \alpha_{p+1}, \lambda) + \\
&\qquad D_{i_p} \log f(x^*_{i_p}; \alpha_{i_p}, \lambda),
\end{aligned} \tag{29}$$

here

$$v^* = A(x_{i_p}; \alpha_{p+1}, \lambda), \quad \text{and} \quad x^*_{i_p} = A(x_{i_p}; \alpha_{i_p}, \lambda).$$

In the 'M'-step the maximization of the 'pseudo' log-likelihood function can be performed by first maximizing with respect to λ, as a one dimensional optimization process, and then obtain the estimates of α_j's using (26), by replacing missing u_{ij}'s and v_i's by the corresponding expected values. The process should be continued unless convergence occurs. For better understanding we present the EM algorithm explicitly when $p = 3$ in the Appendix B, which is applied in the following section.

6 Illustrative Example

In this section we present one illustrative example to show how the proposed EM algorithm can be used in practice. One simulated data set has been used for this purpose. We have used FORTRAN-77 to compute the EM algorithm. The code can be obtained from the corresponding author on request.

We have generated a sample of size 30 from a trivariate modified Sarhan-Balakrishnan model with $\alpha_1 = 1.5$, $\alpha_2 = 1.5$, α_3 1.5, $\alpha_4 = 1.5$ and $\lambda = 1.0$. The data set is presented below.

In this case $I_0 = \{7, 18, 23\}$, $I_{10} = \{30\}$, $I_{20} = \{10, 14, 17, 25\}$, $I_{30} = \{19\}$, $I_{123} = \{28, 29\}$, $I_{132} = \{6, 8, 12, 20\}$, $I_{213} = \{2, 16, 21, 22\}$, $I_{231} = \{9, 13, 26, 27\}$, $I_{312} = \{3, 4, 5, 11, 15\}$, $I_{321} = \{1, 24\}$, as defined in Appendix B. To start the EM algorithm, we need some initial guesses of α_1, α_2, α_3 and α_4, although we do not need any initial guess for λ. We use the following procedure to get initial guesses of α's. From the sets I_0, I_{10}, I_{20} and I_{30}, we have observations from the random variable V. From these 9 observations, we get MLE of α_4, using the method proposed by Gupta and Kundu [7], as $\widetilde{\alpha}_4 = 1.54$. That will be used as the

Table 1: Simulated Data Set.

Sl. No.	X_1	X_2	X_3	Sl. No.	X_1	X_2	X_3	Sl. No.	X_1	X_2	X_3
1.	3.095	1.893	0.262	11.	0.955	2.283	0.481	21.	0.269	0.258	0.418
2.	1.132	0.138	1.227	12.	0.694	2.122	1.654	22.	1.131	0.699	1.426
3.	0.461	1.111	0.425	13.	0.855	0.157	0.387	23.	0.056	0.056	0.056
4.	0.884	1.105	0.354	14.	0.224	0.097	0.224	24.	2.387	1.502	0.868
5.	1.334	1.475	0.161	15.	1.496	1.684	0.497	25.	0.300	0.221	0.300
6.	0.462	1.350	0.907	16.	0.927	0.101	2.189	26.	1.811	0.311	0.331
7.	0.430	0.430	0.430	17.	0.390	0.236	0.390	27.	0.951	0.420	0.785
8.	0.053	0.292	0.199	18.	0.247	0.247	0.247	28.	0.327	0.937	1.424
9.	1.312	0.477	0.673	19.	0.467	0.467	0.265	29.	0.317	0.431	0.850
10.	1.133	0.121	1.133	20.	0.208	0.636	0.240	30.	1.469	1.669	1.669

guess value of α_4. Similarly, from the sets I_{10}, I_{123}, I_{132}, I_{213}, I_{312}, we have the observations from U_1, and using those observations we obtain an guess value of α_1 as $\widetilde{\alpha}_1 = 1.63$. Finally from the sets I_{20}, I_{213}, I_{123}, I_{231}, I_{321}, and I_{30}, I_{312}, I_{321}, I_{132}, I_{231}, we obtain the initial estimates of α_2 and α_3 as $\widetilde{\alpha}_2 = 1.48$ and $\widetilde{\alpha}_3 = 1.46$ respectively.

We start the EM algorithm with these initial guesses. We provide the profile 'pseudo' log-likelihood in Figure 1. It shows that it is an unimodal function. In fact it is observed that the profile 'pseudo' log-likelihood function in each iterate is unimodal. Therefore, the maximization becomes quite simple in each case. We have stopped iteration procedure if the relative absolute difference between two consecutive log-likelihood values is less than $\epsilon = 10^{-6}$. The iteration stops after 6 steps and we provide the estimates and the associated log-likelihood values at each iteration step. We obtain the MLEs of the unknown parameters at the last step. We computed the parametric bootstrap confidence intervals (95%) for all the parameters, and they are (0.8089, 1.3046), (1.1183, 2.4977), (0.6949, 1.7891), (1.0799, 2.0796), (1.2631, 2.4859) for λ, α_1, α_2, α_3 and α_4 respectively.

Moreover, the suitability of the estimated model, MOMGE($\widehat{\alpha}_1 = 1.810310$, $\widehat{\alpha}_2 = 1.254555$,

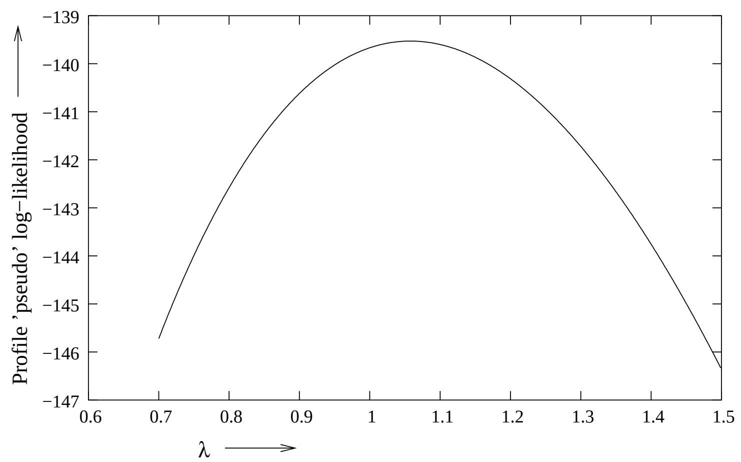

Figure 1: Profile 'pseudo' log-likelihood function at the 1st iterate.

Table 2: Parameter estimates and the associated log-likelihood value, at different EM steps.

Iter. No.	λ	α_1	α_2	α_3	α_4	log-likelihood value
1	1.057051	1.808038	1.242048	1.514702	1.916045	-140.01872
2	1.056851	1.811170	1.255755	1.520792	1.870050	-139.92827
3	1.056851	1.810386	1.254531	1.519807	1.875120	-139.90946
4	1.056151	1.809353	1.254055	1.519041	1.873439	-139.71152
5	1.056751	1.810310	1.254555	1.519783	1.874466	-139.71124
6	1.056751	1.810310	1.254555	1.519783	1.874458	-139.71124

$\widehat{\alpha}_3 = 1.519783$, $\widehat{\alpha}_4 = 1.874458$, $\widehat{\lambda} = 1.056751$), is judged from Kolmogorov-Smirnov goodness-of-fit test. The Kolmogorov-Smirnov statistic $D = 0.18219$ and its p-value$= 0.272261$ indicated that the estimated model could not be rejected as an acceptable fit to the simulated data set.

7 Conclusions

In this paper we have considered multivariate version of the modified Sarhan-Balakrishnan bivariate distribution. We have discussed several ageing, dependence and ordering properties of the proposed multivariate distribution. We have provided the EM algorithm which can be used to compute the MLEs of the unknown parameters quite effectively. This proposed distribution is a multivariate distribution with a singular part. Since not too many multivariate distributions with singular parts, are available in the literature, this model can be used quite effectively for analyzing multivariate data with singular components.

Acknowledgements

The authors would like to thank the referees and the editor-in-chief Professor N. Balakrishnan for many valuable comments. This work was partially supported by Fundacion Seneca of the Regional Government of Murcia (Spain) under Grant 11886/PHCS/09. The work of the second author has been supported by a grant from the Department of Science and Technology, Government of India.

Appendix A

In this appendix, we provide the decomposition of (12) into PDFs with respect to $(k+1)$-dimensional Lebesgue measures, $0 \leq k \leq p-1$. First, we present the explicit expressions of β and $f_a(\mathbf{x})$ of (12). The absolute continuous part $f_a(\mathbf{x})$ and β can be obtained from $\dfrac{\partial^p S(x_1, \cdots, x_p)}{\partial x_1 \cdots \partial x_p}$. It is immediate that $\mathbf{x} = (x_1, \cdots, x_p)$ belongs to the set where $S_{\mathbf{X}}(\cdot)$ is absolutely continuous if and only if all the x_i's are different, *i.e.* $k = p-1$. For a given $(\mathbf{x})$, so that all the x_i's are different, there exists a permutation $\mathcal{P}_p = (i_1, \cdots, i_p)$, so that $x_{i_1} < x_{i_2} < \cdots < x_{i_p}$. Let us define the following for $x_{i_1} < \cdots < x_{i_p}$

$$\begin{aligned} f_{\mathcal{P}_p}(\mathbf{x}) &= f(x_{i_1};\alpha_{i_1}) \cdots f(x_{i_{p-1}};\alpha_{i_{p-1}}) \left(f(x_{i_p};\alpha_{i_p}) + f(x_{i_p};\alpha_{p+1}) - f(x_{i_p};\alpha_{i_p}+\alpha_{p+1}) \right) \\ &= \left(\prod_{j=1}^{p-1} f(x_{i_j};\alpha_{i_j}) \right) \left(f(x_{i_p};\alpha_{i_p}) + f(x_{i_p};\alpha_{p+1}) - f(x_{i_p};\alpha_{i_p}+\alpha_{p+1}) \right) \end{aligned} \quad (30)$$

Then from (12) we obtain for $x_{i_1} < \cdots < x_{i_p}$

$$(-1)^p \frac{\partial^p S(x_1, \cdots, x_p)}{\partial x_1 \cdots \partial x_p} = \beta f_a(x_1, \cdots, x_p) = f_{\mathcal{P}_p}(x_1, \cdots, x_p). \tag{31}$$

From (31) we have the following relation;

$$\begin{aligned}\beta = \beta \int_{R^p} f_a(x_1, \cdots, x_p) dx_1 \cdots dx_p &= \sum_{\mathcal{P}_p} \int_{x_{i_p}=0}^{\infty} \int_{x_{i_{p-1}}=0}^{x_{i_p}} \cdots \int_{x_{i_1}=0}^{x_{i_2}} f_{\mathcal{P}_p}(x_1, \cdots, x_p) dx_1 \cdots dx_p \\ &= \sum_{\mathcal{P}_p} J_{\mathcal{P}_p}, \quad \text{(say)}. \end{aligned} \tag{32}$$

Since

$$\int_{x_{i_1}=0}^{x_{i_2}} f(x_{i_1}; \alpha_{i_1}) dx_{i_1} = F(x_{i_2}; \alpha_{i_1});$$

$$\begin{aligned}\int_{x_{i_{p-1}}=0}^{x_{i_p}} \cdots \int_{x_{i_1}=0}^{x_{i_2}} \prod_{j=1}^{p-1} f(x_{i_j}; \alpha_{i_j}) dx_{i_1} ... dx_{i_{p-1}} &= \left(\prod_{j=2}^{p-1} \frac{\alpha_{i_j}}{\alpha_{i_1} + ... + \alpha_{i_j}} \right) F\left(x_{i_p}; \alpha_{i_1} + ... + \alpha_{i_{p-1}}\right) \\ &= \left(\prod_{j=1}^{p-1} \frac{\alpha_{i_j}}{\sum_{l=1}^{j} \alpha_{i_l}} \right) F\left(x_{i_p}; \alpha_{i_1} + ... + \alpha_{i_{p-1}}\right). \end{aligned} \tag{33}$$

Thus

$$\begin{aligned} J_{\mathcal{P}_p} &= \prod_{j=1}^{p-1} \frac{\alpha_{i_j}}{\sum_{l=1}^{j} \alpha_{i_l}} \int_{x_{i_p}=0}^{\infty} \left(\frac{\alpha_{i_p}}{\sum_{j=1}^{p} \alpha_{i_j}} f\left(x_{i_p}; \sum_{j=1}^{p} \alpha_{i_j}\right) \right. \\ &\quad + \frac{\alpha_{p+1}}{\sum_{j=1}^{p-1} \alpha_{i_j} + \alpha_{p+1}} f\left(x_{i_p}; \sum_{j=1}^{p-1} \alpha_{i_j} + \alpha_{p+1}\right) \\ &\quad \left. - \frac{\alpha_{i_p} + \alpha_{p+1}}{\sum_{j=1}^{p} \alpha_{i_j} + \alpha_{p+1}} f\left(x_{i_p}; \sum_{j=1}^{p} \alpha_{i_j} + \alpha_{p+1}\right) \right) dx_{i_p} \\ &= \prod_{j=1}^{p-1} \frac{\alpha_{i_j}}{\sum_{l=1}^{j} \alpha_{i_l}} \left(\frac{\alpha_{i_p}}{\sum_{j=1}^{p} \alpha_{i_j}} + \frac{\alpha_{p+1}}{\sum_{j=1}^{p-1} \alpha_{i_j} + \alpha_{p+1}} - \frac{\alpha_{i_p} + \alpha_{p+1}}{\sum_{j=1}^{p} \alpha_{i_j} + \alpha_{p+1}} \right). \end{aligned}$$

Therefore,

$$\beta = \sum_{\mathcal{P}_p} \left(\prod_{j=1}^{p-1} \frac{\alpha_{i_j}}{\sum_{l=1}^{j} \alpha_{i_l}} \right) \left(\frac{\alpha_{i_p}}{\sum_{j=1}^{p} \alpha_{i_j}} + \frac{\alpha_{p+1}}{\sum_{j=1}^{p-1} \alpha_{i_j} + \alpha_{p+1}} - \frac{\alpha_{i_p} + \alpha_{p+1}}{\sum_{j=1}^{p} \alpha_{i_j} + \alpha_{p+1}} \right), \tag{34}$$

and for $x_{i_1} < \cdots < x_{i_p}$

$$f_a(\mathbf{x}) = \frac{1}{\beta} f_{\mathcal{P}_p}(\mathbf{x}), \tag{35}$$

where $f_{\mathcal{P}_p}$ is same as defined in (30).

Now we delve into the decomposition of $f_X(\cdot)$ taking into account that (12) can be rewritten as

$$f_X(\mathbf{x}) = \beta f_a(\mathbf{x}) + \sum_{k=2}^{p} \sum_{I_k \subset I} \beta_{I_k} f_{I_k}(\mathbf{x}) \tag{36}$$

where $I_k = (i_1, ..., i_k) \subset I = (1, ..., p)$ such that $i_1 < ... < i_k$. Here, it is understood that each $f_{I_k}(\mathbf{x})$ is a PDF with respect to $(p-k+1)$-dimensional Lebesgue measure on the hyperplane $A_{I_k} = \{\mathbf{x} \in I\!R^p : x_{i_1} = ... = x_{i_k}\}$. The exact meaning of $f_X(\mathbf{x})$ is as follows. For any Borel measurable set $B \subset I\!R^p$

$$P(\mathbf{X} \in B) = \beta \int_B f_a(\mathbf{x}) + \sum_{k=2}^{p} \sum_{I_k \subset I} \beta_{I_k} \int_{B_{I_k}} f_{I_k}(\mathbf{x})$$

where $B_{I_k} = B \cap A_{I_k}$ denotes the projection of the set B onto the $(p-k+1)$-dimensional hyperplane A_{I_k}.

Now we provide the explicit expressions of β_{I_k} and $f_{I_k}(\cdot)$.

For a given $\mathbf{x} \in I\!R^p$, we define a function g_{I_k} from the $(p-k+1)$-dimensional hyperplane A_{I_k} to $I\!R$ as follows, taking into account that when $\mathbf{x} \in A_{I_k}$ then it is of the form $\mathbf{x} = (x_1, ..., x_{i_1-1}, x^*, x_{i-1+1}, ..., x_{i_k-1}, x^*, x_{i_k+1}, ..., x_p)$

$$g_{I_k}(\mathbf{x}) = f(x^*; \alpha_{p+1}) \prod_{i \in I_k} S(x^*; \alpha_i) \prod_{i \in I - I_k} f(x_i; \alpha_i) \tag{37}$$

if $x_i < x^*$ for $i \in I - I_k$ and zero otherwise, where $\prod_{i \in \varnothing} = 1$ when $k = p$. Hence, from (37), we have that

$$\begin{aligned} \int_{A_{I_k}} g_{I_k}(\mathbf{x}) d\mathbf{x} &= \sum_{\mathcal{P}_{I-I_k}} \int_{x^*=0}^{\infty} \int_{x_{j_{p-k}}=0}^{x^*} \int_{x_{j_{p-k-1}}=0}^{x_{j_{p-k}}} ... \int_{x_{j_1}=0}^{x_{j_2}} g_{I_k}(\mathbf{x}) dx_{j_1} ... dx_{j_{p-k}} dx^* \\ &= \sum_{\mathcal{P}_{I-I_k}} J_{\mathcal{P}_{I-I_k}} \end{aligned} \tag{38}$$

where $\mathcal{P}_{I-I_k} = (j_1, ..., j_{p-k})$ denotes a permutation of $I - I_k$, so that $x_{j_1} < ... < x_{j_{p-k}}$. Then, it can be shown along the same line as before that

$$\int_{x_{j_{p-k}}=0}^{x^*}\int_{x_{j_{p-k-1}}=0}^{x_{j_{p-k}}}\cdots\int_{x_{j_1}=0}^{x_{j_2}}\prod_{i\in I-I_k} f(x_i;\alpha_i)dx_{j_1}...dx_{j_{p-k}} = \left(\prod_{r=1}^{p-k}\frac{\alpha_{j_r}}{\sum_{s=1}^{r}\alpha_{j_s}}\right)F(x^*;\gamma_{I_k}-\alpha_{p+1})$$

where $\gamma_{I_k} = \sum_{i=1}^{p}\alpha_i - \sum_{i\in I_k}\alpha_i$ (i.e. $\gamma_{I_k} = \alpha_{p+1} + \sum_{i\in I-I_k}\alpha_i$), and so

$$\begin{aligned} J_{\mathcal{P}_{I-I_k}} &= \left(\prod_{r=1}^{p-k}\frac{\alpha_{j_r}}{\sum_{s=1}^{r}\alpha_{j_s}}\right)\frac{\alpha_{p+1}}{\gamma_{I_k}}\int_{x^*=0}^{\infty} f(x^*;\gamma_{I_k})\prod_{i\in I_k} S(x^*;\alpha_i)dx^* \\ &= \left(\prod_{r=1}^{p-k}\frac{\alpha_{j_r}}{\sum_{s=1}^{r}\alpha_{j_s}}\right)\frac{\alpha_{p+1}}{\gamma_{I_k}}\int_{x^*=0}^{\infty}\sum_{r=0}^{k}(-1)^r\sum_{(I_k)_r\subset I_k} f(x^*;\gamma_{I_k})F(x^*;\textstyle\sum_{m\in(I_k)_r}\alpha_{i_m})dx^* \\ &= \left(\prod_{r=1}^{p-k}\frac{\alpha_{j_r}}{\sum_{s=1}^{r}\alpha_{j_s}}\right)\frac{\alpha_{p+1}}{\gamma_{I_k}}\sum_{r=0}^{k}(-1)^r\sum_{(I_k)_r\subset I_k}\frac{\gamma_{I_k}}{\gamma_{I_k}+\sum_{m\in(I_k)_r}\alpha_{i_m}} \end{aligned}$$

and consequently, from (38), we have

$$\beta_{I_k} = \sum_{\mathcal{P}_{I-I_k}}\left(\prod_{r=1}^{p-k}\frac{\alpha_{j_r}}{\sum_{s=1}^{r}\alpha_{j_s}}\right)\frac{\alpha_{p+1}}{\gamma_{I_k}}\sum_{r=0}^{k}(-1)^r\sum_{(I_k)_r\subset I_k}\frac{\gamma_{I_k}}{\gamma_{I_k}+\sum_{m\in(I_k)_r}\alpha_{i_m}}$$

where $(I_k)_r$ denotes any subset of I_k with r different components, $r = 0, 1, ..., k$, and

$$f_{I_k}(\mathbf{x}) = \frac{1}{\beta_{I_k}}g_{I_k}(\mathbf{x}). \tag{39}$$

Note that, the above decomposition of $f_X(\cdot)$ given by (36) for $p = 2$ coincides with the one given by Kundu and Gupta [13].

Appendix B

In this appendix we present the explicit EM algorithm when $p = 3$. For $p = 3$, we have the following unknown parameters $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \lambda)$, and the following available data $\{(x_{1i}, x_{2i}, x_{3i}),\ i = 1, \cdots, n\}$. We use the following notation: $I_0 = \{i; x_{1i} = x_{2i} = x_{3i} = x_i\}$, $I_{10} = \{i; x_{1i} < x_{2i} = x_{3i} = x_{10i}\}$, $I_{20} = \{i; x_{2i} < x_{1i} = x_{3i} = x_{20i}\}$, $I_{30} = \{i; x_{3i} < x_{2i} = x_{1i} = x_{30i}\}$, $I_{123} = \{i; x_{1i} < x_{2i} < x_{3i}\}$, $I_{132} = \{i; x_{1i} < x_{3i} < x_{2i}\}$, $I_{213} = \{i; x_{2i} < x_{1i} < x_{3i}\}$, $I_{231} = \{i; x_{2i} < x_{3i} < x_{1i}\}$, $I_{312} = \{i; x_{3i} < x_{1i} < x_{2i}\}$, $I_{321} = \{i; x_{3i} < x_{2i} < x_{1i}\}$.

Clearly, in the set I_0, $V = x_i$ and $U_1 > x_i, U_2 > x_i, U_3 > x_i$. Similarly, in the set I_{10}, $U_1 = x_{1i}$, $V = x_{10i}$, $U_2 > x_{10i}, U_3 > x_{10i}$ and so on. Let us assume that at the j-th stage of the EM algorithm we have the estimates of α's and λ are $\alpha_1^{(j)}$, $\alpha_2^{(j)}$, $\alpha_3^{(j)}$, $\alpha_4^{(j)}$ and $\lambda^{(j)}$ respectively. By maximizing the 'pseudo' log-likelihood function obtained at the j-th stage we will obtain $\alpha_1^{(j+1)}$, $\alpha_2^{(j+1)}$, $\alpha_3^{(j+1)}$, $\alpha_4^{(j+1)}$ and $\lambda^{(j+1)}$.

We first present the 'pseudo' log-likelihood function at the j-stage. The 'pseudo' log-likelihood contributions from the different sets are presented below:

From I_0:

$$\begin{aligned} & n_0(\ln\alpha_1 + \ln\alpha_2 + \ln\alpha_3 + \ln\alpha_4 + 4\ln\lambda) - \lambda\sum_{i\in I_0}(x_i + x_i(1) + x_i(2) + x_i(3)) + \\ & (\alpha_1 - 1)\sum_{i\in I_0}\ln(1 - e^{-\lambda x_i(1)}) + (\alpha_2 - 1)\sum_{i\in I_0}\ln(1 - e^{-\lambda x_i(2)}) + \\ & (\alpha_3 - 1)\sum_{i\in I_0}\ln(1 - e^{-\lambda x_i(3)}) + (\alpha_4 - 1)\sum_{i\in I_0}\ln(1 - e^{-\lambda x_i}). \end{aligned} \tag{40}$$

Here $x_i(1) = A(x_i; \alpha_1^{(j)}, \lambda^{(j)})$, $x_i(2) = A(x_i; \alpha_2^{(j)}, \lambda^{(j)})$, $x_i(3) = A(x_i; \alpha_3^{(j)}, \lambda^{(j)})$, and they depend on j, but we do not make it explicit for brevity.

From I_{10}

$$\begin{aligned} & n_{10}(\ln\alpha_1 + \ln\alpha_2 + \ln\alpha_3 + \ln\alpha_4 + 4\ln\lambda) - \lambda\sum_{i\in I_{10}}(x_{1i} + x_{10i}(2) + x_{10i}(3) + x_{10i}) + \\ & (\alpha_1 - 1)\sum_{i\in I_{10}}\ln(1 - e^{-\lambda x_{1i}}) + (\alpha_2 - 1)\sum_{i\in I_{10}}\ln(1 - e^{-\lambda x_{10i}(2)}) + \\ & (\alpha_3 - 1)\sum_{i\in I_{10}}\ln(1 - e^{-\lambda x_{10i}(3)}) + (\alpha_4 - 1)\sum_{i\in I_{10}}\ln(1 - e^{-\lambda x_{10i}}). \end{aligned} \tag{41}$$

Here $x_{10i}(2) = A(x_{10i}; \alpha_2^{(j)}, \lambda^{(j)})$, $x_{10i}(3) = A(x_{10i}; \alpha_3^{(j)}, \lambda^{(j)})$.

From I_{20}

$$n_{20}(\ln\alpha_1 + \ln\alpha_2 + \ln\alpha_3 + \ln\alpha_4 + 4\ln\lambda) - \lambda\sum_{i\in I_{20}}(x_{20i}(1) + x_{2i} + x_{20i}(3) + x_{20i}) +$$

$$(\alpha_3 - 1)\sum_{i\in I_{20}} \ln(1-e^{-\lambda x_{20i}(1)}) + (\alpha_2 - 1)\sum_{i\in I_{20}} \ln(1-e^{-\lambda x_{2i}}) +$$

$$(\alpha_3 - 1)\sum_{i\in I_{20}} \ln(1-e^{-\lambda x_{20i}(3)}) + (\alpha_4 - 1)\sum_{i\in I_{20}} \ln(1-e^{-\lambda x_{20i}}). \tag{42}$$

Here $x_{20i}(1) = A(x_{20i}; \alpha_1^{(j)}, \lambda^{(j)})$, $x_{20i}(3) = A(x_{20i}; \alpha_3^{(j)}, \lambda^{(j)})$.

From I_{30}

$$n_{30}(\ln\alpha_1 + \ln\alpha_2 + \ln\alpha_3 + \ln\alpha_4 + 4\ln\lambda) - \lambda\sum_{i\in I_{30}}(x_{30i}(1) + x_{30i}(2) + x_{3i} + x_{30i}) +$$

$$(\alpha_1 - 1)\sum_{i\in I_{30}} \ln(1-e^{-\lambda x_{30i}(1)}) + (\alpha_2 - 1)\sum_{i\in I_{30}} \ln(1-e^{-\lambda x_{30i}(2)}) +$$

$$(\alpha_3 - 1)\sum_{i\in I_{30}} \ln(1-e^{-\lambda x_{3i}}) + (\alpha_4 - 1)\sum_{i\in I_{30}} \ln(1-e^{-\lambda x_{30i}}). \tag{43}$$

Here $x_{30i}(1) = A(x_{30i}; \alpha_1^{(j)}, \lambda^{(j)})$, $x_{30i}(2) = A(x_{30i}; \alpha_2^{(j)}, \lambda^{(j)})$.

From I_{123}

$$n_{123}(\ln\alpha_1 + \ln\alpha_2 + \ln\alpha_3 + \ln\alpha_4 + 4\ln\lambda) -$$

$$\lambda\sum_{i\in I_{123}}(x_{1i} + x_{2i} + x_{3i}) - \lambda\sum_{i\in I_{123}}(C_3 x_{123i}(4) + D_3 x_{123i}(3)) +$$

$$(\alpha_1 - 1)\sum_{i\in I_{123}} \ln(1-e^{-\lambda x_{1i}}) + (\alpha_2 - 1)\sum_{i\in I_{123}} \ln(1-e^{-\lambda x_{2i}}) +$$

$$(\alpha_3 - 1)\left(C_3\sum_{i\in I_{123}} \ln(1-e^{-\lambda x_{3i}}) + D_3\sum_{i\in I_{123}} \ln(1-e^{-\lambda x_{123i}(3)})\right) +$$

$$(\alpha_4 - 1)\left(C_3\sum_{i\in I_{123}} \ln(1-e^{-\lambda x_{123i}(4)}) + D_3\sum_{i\in I_{123}} \ln(1-e^{-\lambda x_{3i}})\right). \tag{44}$$

Here $x_{123i}(3) = A(x_{3i}; \alpha_3^{(j)}, \lambda^{(j)})$, $x_{123i}(4) = A(x_{3i}; \alpha_4^{(j)}, \lambda^{(j)})$. Similarly, the contribution from the set I_{213} can be obtained by replacing $\{123\}$ with $\{213\}$ everywhere in (44).

From I_{132}

$$n_{132}(\ln\alpha_1 + \ln\alpha_2 + \ln\alpha_3 + \ln\alpha_4 + 4\ln\lambda) -$$

$$\lambda\sum_{i\in I_{132}}(x_{1i} + x_{2i} + x_{3i}) - \lambda\sum_{i\in I_{132}}(C_2 x_{132i}(4) + D_2 x_{132i}(2))$$

$$(\alpha_1 - 1)\sum_{i\in I_{132}} \ln(1-e^{-\lambda x_{1i}}) + (\alpha_3 - 1)\sum_{i\in I_{132}} \ln(1-e^{-\lambda x_{3i}}) +$$
$$(\alpha_2 - 1)\left(C_2\sum_{i\in I_{132}} \ln(1-e^{-\lambda x_{2i}}) + D_2\sum_{i\in I_{132}} \ln(1-e^{-\lambda x_{132i}(2)})\right) +$$
$$(\alpha_4 - 1)\left(C_2\sum_{i\in I_{132}} \ln(1-e^{-\lambda x_{132i}(4)}) + D_2\sum_{i\in I_{132}} \ln(1-e^{-\lambda x_{2i}})\right). \qquad (45)$$

Here $x_{132i}(2) = A(x_{2i};\alpha_2^{(j)},\lambda^{(j)})$, $x_{132i}(4) = A(x_{2i};\alpha_4^{(j)},\lambda^{(j)})$. Similarly, the contribution from the set I_{312} can be obtained by replacing $\{132\}$ with $\{312\}$ everywhere in (45),

From I_{231}

$$n_{231}(\ln\alpha_1 + \ln\alpha_2 + \ln\alpha_3 + \ln\alpha_4 + 4\ln\lambda) -$$
$$\lambda\sum_{i\in I_{231}} (x_{1i} + x_{2i} + x_{3i}) - \lambda\sum_{i\in I_{231}} (C_1 x_{231i}(4) + D_1 x_{231i}(1))$$
$$(\alpha_2 - 1)\sum_{i\in I_{231}} \ln(1-e^{-\lambda x_{2i}}) + (\alpha_3 - 1)\sum_{i\in I_{231}} \ln(1-e^{-\lambda x_{3i}}) +$$
$$(\alpha_1 - 1)\left(C_1\sum_{i\in I_{231}} \ln(1-e^{-\lambda x_{1i}}) + D_1\sum_{i\in I_{231}} \ln(1-e^{-\lambda x_{231i}(1)})\right) +$$
$$(\alpha_4 - 1)\left(C_1\sum_{i\in I_{231}} \ln(1-e^{-\lambda x_{231i}(4)}) + D_1\sum_{i\in I_{231}} \ln(1-e^{-\lambda x_{1i}})\right). \qquad (46)$$

Here $x_{231i}(1) = A(x_{1i};\alpha_1^{(j)},\lambda^{(j)})$, $x_{231i}(4) = A(x_{1i};\alpha_4^{(j)},\lambda^{(j)})$. Similarly, the contribution from the set I_{321} can be obtained by replacing $\{231\}$ with $\{321\}$ everywhere in (46),

Therefore, it is clear that for fixed λ, $\widehat{\alpha}_1^{(j+1)}(\lambda)$, $\widehat{\alpha}_2^{(j+1)}(\lambda)$, $\widehat{\alpha}_3^{(j+1)}(\lambda)$, and $\widehat{\alpha}_4^{(j+1)}(\lambda)$ maximize the 'pseudo' log-likelihood function. When

$$\widehat{\alpha}_1^{(j+1)}(\lambda) = -\frac{n}{A_1},\quad \widehat{\alpha}_2^{(j+1)}(\lambda) = -\frac{n}{A_2},\quad \widehat{\alpha}_3^{(j+1)}(\lambda) = -\frac{n}{A_3},\quad \widehat{\alpha}_4^{(j+1)}(\lambda) = -\frac{n}{A_4}, \qquad (47)$$

and

$$\begin{aligned} A_1 \;=\; & \sum_{i\in I_0} \ln(1-e^{-\lambda x_i(1)}) + \sum_{i\in I_{10}} \ln(1-e^{-\lambda x_{1i}}) + \sum_{i\in I_{20}} \ln(1-e^{-\lambda x_{20i}(1)}) + \sum_{i\in I_{30}} \ln(1-e^{-\lambda x_{30i}(1)}) \\ & + \sum_{i\in I_{123}\cup I_{213}\cup I_{132}\cup I_{312}} \ln(1-e^{-\lambda x_{1i}}) + C_1\sum_{i\in I_{231}\cup I_{321}} \ln(1-e^{-\lambda x_{1i}}) \end{aligned}$$

$$
\begin{aligned}
& +D_1 \sum_{i \in I_{231}} \ln(1-e^{-\lambda x_{231i}(1)}) + D_1 \sum_{i \in 321} \ln(1-e^{-\lambda x_{321i}(1)}) \\
A_2 \;=\; & \sum_{i \in I_0} \ln(1-e^{-\lambda x_i(2)}) + \sum_{i \in I_{10}} \ln(1-e^{-\lambda x_{10i}(2)}) + \sum_{i \in I_{20}} \ln(1-e^{-\lambda x_{2i}}) + \sum_{i \in I_{30}} \ln(1-e^{-\lambda x_{30i}(2)}) \\
& + \sum_{i \in I_{123} \cup I_{213} \cup I_{231} \cup I_{321}} \ln(1-e^{-\lambda x_{2i}}) + C_2 \sum_{i \in I_{132} \cup I_{312}} \ln(1-e^{-\lambda x_{2i}}) \\
& +D_2 \sum_{i \in I_{132}} \ln(1-e^{-\lambda x_{132i}(2)}) + D_2 \sum_{i \in I_{312}} \ln(1-e^{-\lambda x_{312i}(2)}) \\
A_3 \;=\; & \sum_{i \in I_0} \ln(1-e^{-\lambda x_i(3)}) + \sum_{i \in I_{10}} \ln(1-e^{-\lambda x_{10i}(3)}) + \sum_{i \in I_{20}} \ln(1-e^{-\lambda x_{20i}(3)}) + \sum_{i \in I_{30}} \ln(1-e^{-\lambda x_{3i}}) \\
& + \sum_{i \in I_{132} \cup I_{312} \cup I_{231} \cup I_{321}} \ln(1-e^{-\lambda x_{3i}}) + C_3 \sum_{i \in I_{123} \cup I_{213}} \ln(1-e^{-\lambda x_{3i}}) \\
& +D_3 \sum_{i \in I_{123}} \ln(1-e^{-\lambda x_{123i}(3)}) + D_3 \sum_{i \in I_{213}} \ln(1-e^{-\lambda x_{213i}(3)}) \\
A_4 \;=\; & \sum_{i \in I_0} \ln(1-e^{-\lambda x_i}) + \sum_{i \in I_{10}} \ln(1-e^{-\lambda x_{10i}}) + \sum_{i \in I_{20}} \ln(1-e^{-\lambda x_{20i}}) + \sum_{i \in I_{30}} \ln(1-e^{-\lambda x_{30i}}) \\
& +C_3 \sum_{i \in I_{123}} \ln(1-e^{-\lambda x_{123i}(4)}) + D_3 \sum_{i \in I_{123}} \ln(1-e^{-\lambda x_{3i}}) \\
& +C_3 \sum_{i \in I_{213}} \ln(1-e^{-\lambda x_{213i}(4)}) + D_3 \sum_{i \in I_{213}} \ln(1-e^{-\lambda x_{3i}}) \\
& +C_2 \sum_{i \in I_{132}} \ln(1-e^{-\lambda x_{132i}(4)}) + D_2 \sum_{i \in I_{132}} \ln(1-e^{-\lambda x_{2i}}) \\
& +C_2 \sum_{i \in I_{312}} \ln(1-e^{-\lambda x_{312i}(4)}) + D_2 \sum_{i \in I_{312}} \ln(1-e^{-\lambda x_{2i}}) \\
& +C_1 \sum_{i \in I_{231}} \ln(1-e^{-\lambda x_{231i}(4)}) + D_1 \sum_{i \in I_{231}} \ln(1-e^{-\lambda x_{1i}}) \\
& +C_1 \sum_{i \in I_{321}} \ln(1-e^{-\lambda x_{321i}(4)}) + D_1 \sum_{i \in I_{321}} \ln(1-e^{-\lambda x_{1i}}).
\end{aligned}
$$

Note that $\lambda^{(k+1)}$ can be obtained by maximizing the 'pseudo' profile log-likelihood function with respect to λ. The 'pseudo' profile log-likelihood function is obtained by adding all the 'pseudo' log-likelihood contribution from different sets as provided in (40) - (46), and replacing α_k by $\alpha_k^{(j+1)}(\lambda)$, for $k = 1, \cdots, 4$. The iteration process continues until convergence takes place.

References

[1] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1992), *A first course in order statistics*, Wiley, New York.

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

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

[4] Colangelo, A., Hu, T. and Shaked, M. (2008), "Conditional ordering and positive dependence", *Journal of Multivariate Analysis*, vol. 99, 358 - 371.

[5] Dinse, G.E. (1982), "Non-parametric estimation of partially incomplete time and type of failure data", *Biometrics*, vol. 38, 417 - 431.

[6] 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.

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

[8] Gupta, R.D. and Kundu, D. (1999), "Generalized exponential distribution: existing results and some recent developments", *Journal of Statistical Planning and Inference*, vol. 137, 3525 - 3536.

[9] Joe, H. (1997), *Multivariate model and dependence concepts*, Chapman and Hall, London.

[10] Johnson, N.L. and Kotz, S. (1975), "A vector multivariate hazard rate",*Journal of Multivariate Analysis*, vol. 5, 53 - 66.

[11] Karlis, D. (2003), "ML estimation for multivariate shock models via an EM algorithm", *Annals of the Institute of Statistical Mathematics*, vol. 55, 817 - 830.

[12] Kundu, D. and Dey, A.K. (2009), "Estimating the parameters of the Marshall Olkin bivariate Weibull distribution by EM algorithm", *Computational Statistics and Data Analysis*, vol. 53, 956-965.

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

[14] Kundu, D. (2004), "Parameter estimation of partially incomplete time and type of failure data", *Biometrical Journal*, vol. 46, 165 - 179.

[15] Lai, C.D. and Xie, M. (2006), *Stochastic ageing and dependence for reliability*, Springer, New York.

[16] Marshall, A.W. (1975), "Some comments on the hazard gradient", *Stochastic Processes and their Applications*, vol. 3, 293-300.

[17] Marshall, A.W. and Olkin, I. (1979), *Inequalities: Theory of majorization and its applications*, Academic Press, San Diego, California.

[18] Nelsen, R. B (2006), *Introduction to copulas*, second edition, Springer, New York.

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

[20] Shanbhag, D.N. and Kotz, S. (1987), "Some new approaches to multivariate probability distributions", *Journal of Multivariate Analysis*, vol. 22, 89 - 211.