Sankhyā : The Indian Journal of Statistics
2015, Volume 77-B, Part 2, pp. 175-206

Absolute Continuous Multivariate Generalized Exponential Distribution

Debasis Kundu and Ankush Kumar
Indian Institute of Technology, Kanpur, India
Arjun K. Gupta
Bowling Green State University, Bowling Green, USA

Abstract

Generalized exponential distribution has received some attention in the last few years. Recently, Kundu and Gupta (*Advances in Statistical Analysis*, 95, 169–185, 2011) and Shoaee and Khorram (*Journal of Statistical Planning and Inference*, **142**, 2203–2220, 2012) introduced an absolute continuous bivariate generalized exponential distribution. In this paper, we propose an absolute coinuous multivariate generalized exponential distribution. The proposed distribution is very flexible, and the joint probability density functions can take different shapes. We provide several properties of this model. Further, it is observed that the multivariate generalized exponential model can be obtained using multivariate Clayton copula. The maximum likelihood estimators are quite difficult to compute in practice. Due to this reason, we propose two step estimation procedure using the copula approach, which are quite easy to implement. Simulation experiments are performed to compare the performances of the two different estimators, and the performances are quite similar in nature particularly for large sample sizes. One multivariate bone mineral density data set has been analyzed for illustrative purposes, and it is observed that the proposed model provides a very good fit to the data set.

AMS (2000) *subject classification.* Primary 60E05, Secondary 62F10, 62P99.
Keywords and phrases. Generalized exponential distribution, Maximum likelihood estimators, Clayton copula, Fisher information matrix, Monte Carlo simulation, hazard function

1 Introduction

The generalized exponential (GE) distribution proposed by Gupta and Kundu (1999) has received some attention in the past few years. The two-parameter GE distribution has been used quite successfully to analyze lifetime data in place of two-parameter gamma or two-parameter Weibull distributions. The cumulative distribution function (CDF) of a two-parameter GE distribution has the following form;

$$F(x;\alpha,\lambda) = \left(1 - e^{-\lambda x}\right)^{\alpha}; \quad x > 0, \tag{1.1}$$

and 0, otherwise. The corresponding probability density function (PDF) becomes

$$f(x;\alpha,\lambda) = \alpha\lambda e^{-\lambda x}(1 - e^{-\lambda x})^{\alpha-1}; \quad x > 0, \tag{1.2}$$

and 0, otherwise. Here $\alpha > 0$, and $\lambda > 0$, are the shape and scale parameters, respectively. From now on it will be denoted by GE(α, λ). An extensive survey on the GE distribution can be obtained in Nadarajah (2011), see also Gupta and Kundu (2007).

Using the concept similar to Marshall and Olkin (1967), Kundu and Gupta (2009) introduced a bivariate GE distribution, which has a singular component. Following the approach of Block and Basu (1974), by removing the singular component, Shoaee and Khorram (2012) proposed an absolute continuous bivariate GE distribution, whose marginals are not the GE distributions. Using the one parameter bivariate exchangeable distribution of Mardia (1962), Kundu and Gupta (2011) introduced an absolute continuous bivariate generalized exponential distribution, whose marginals are the GE distributions. It is not very simple to generalize Shoaee and Khoram's bivariate generalized exponential distribution to its multivariate version. Kundu and Gupta's absolute continuous bivariate generalized exponential distribution can be obtained as a special case of the proposed multivariate distribution.

The main aim of this paper is to introduce an absolute continuous multivariate GE distribution. Takahasi (1965) introduced multivariate Burr distribution by compounding independent Weibull distributions with a gamma distribution, as a compounder. Using the same approach Crowder (1989) introduced multivariate Weibull distribution. Surles and Padgett (2005), using a suitable transformation of Takahasi (1965)'s multivariate Burr distribution proposed multivariate scaled Burr type X distribution. In this paper we introduce an absolute continuous multivariate GE (MVGE) distribution by making a suitable transformation from Takahasi's multivariate Burr distribution. The MVGE distribution has the GE marginals. Generation from the MVGE distribution has been addressed. We discuss several properties of the proposed distribution.

Estimation of the unknown parameters is an important problem. The maximum likelihood estimators (MLEs) cannot be obtained in closed form.

The MLEs can be obtained by solving multidimensional optimization problem. We propose alternative estimators based on copula which can be obtained quite conveniently. Monte Carlo simulations are performed to compare the performances of the different estimators and it is observed that the performance of the MLEs and the estimators based on copula are quite similar. One data analysis has been performed to show how the proposed model and the method work in a real life situation.

Rest of the paper is organized as follows. In Section 2, we provide the necessary preliminaries. The MVGE has been introduced in Section 3. In Section 4, we discuss several properties and in Section 5, we study inferential issues. Monte Carlo simulation results and the analysis of a data set are provided in Section 6. In Section 7, we conclude the paper.

2 Preliminaries

2.1. Dependence and Stochastic Order. Several notions of positive or negative dependence for multivariate distributions for varying degree of strengths are available in the literature, see for example Boland et al. (1996), Colangelo et al. (2005, 2008), and the references therein. The notions of positive dependence were introduced in the literature to model the fact that large values of a component of multivariate random vector are probabilistically associated with large values of the others. Similarly, the notion of negative dependence captures the fact that large or small values of a component of a random vector are probabilistically associated with small or large values of the others. In this paper we will consider three such positive dependence concepts.

A random vector $\boldsymbol{X} = (X_1, \ldots, X_p)^T$, is said to be *positively lower orthant dependent* (PLOD) if $F_{\boldsymbol{X}}(\cdot)$, the joint cumulative distribution function of $\boldsymbol{X}$ satisfies the following property:

$$F_{\boldsymbol{X}}(x_1, \ldots, x_p) \geq \prod_{i=1}^{p} F_{X_i}(x_i), \qquad \forall \boldsymbol{x} = (x_1, \ldots, x_p)^T, \tag{2.1}$$

here $F_{X_i}(\cdot)$ is the marginal distribution function of X_i for $i = 1, \ldots, p$.

Further, we will be using the following notations. For $\boldsymbol{x} \in \mathbb{R}^p$, a phrase such that 'non-decreasing' in $\boldsymbol{x}$ means non-decreasing in each component x_i, for $i = 1, \ldots, p$. If A is a subset of the set $\{1, \ldots, p\}$, then $\boldsymbol{X}_A$ denotes the vector $(X_i | i \in A)$, similarly, $\boldsymbol{x}_A$ is also defined. The following definitions are known in the statistical literature, see for example Joe (1997).

A p-dimensional random vector $\boldsymbol{X}$ is said to be *left tail decreasing* (LTD) if

$$P\left[\boldsymbol{X}_B \leq \boldsymbol{x}_B | \boldsymbol{X}_A \leq \boldsymbol{x}_A\right]$$

is non-increasing in $\boldsymbol{x}_A$ for all $\boldsymbol{x}_B$. Here the sets A and B are disjoint partition of the set $\{1, \ldots, p\}$.

Another multivariate dependence notion is the multivariate left corner set decreasing property. A random vector $\boldsymbol{X}$ is said to have *left corner set decreasing* property, if

$$P[X_1 \leq x_1, \ldots, X_P \leq x_p | X_1 \leq x_1', \ldots, X_P \leq x_p'] \tag{2.2}$$

is non-increasing in $\boldsymbol{x}'$, for every choice of $\boldsymbol{x} = (x_1, \cdots, x_p)^T$. Equivalently, (2.2) can be written as

$$\frac{F_{\boldsymbol{X}}(\boldsymbol{x} \wedge \boldsymbol{x}')}{F_{\boldsymbol{X}}(\boldsymbol{x}')} \qquad \text{non-increases in } \boldsymbol{x}',$$

where $\boldsymbol{x}' = (x_1', \ldots, x_p')^T$ and $\boldsymbol{x} \wedge \boldsymbol{x}' = (\min\{x_1, x_1'\}, \ldots, \min\{x_p, x_p'\})^T$.

Now we will define the following stochastic ordering for a multivariate distribution. It is a natural generalization from a univariate distribution to a multivariate distribution. Let us recall that for two random variables X and Y, X is said to be stochastically smaller than Y ($X \leq_{st} Y$) if $P(X \geq a) \leq P(Y \geq a)$, for all $a \in \mathbb{R}$. The concept can be generalized to p dimensional random vectors also as follows.

Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two p-dimensional random vectors such that

$$P(\boldsymbol{X} \in U) \leq P(\boldsymbol{Y} \in U) \qquad \text{for all upper sets } U \subset \mathbb{R}^p,$$

then $\boldsymbol{X}$ is said to be smaller than $\boldsymbol{Y}$ in *stochastic order*, and it will be denoted by $\boldsymbol{X} \leq_{st} \boldsymbol{Y}$.

2.2. Copula. The dependence among the random variables $X_1, \ldots, X_p$, is completely described by the joint distribution function $F_{\boldsymbol{X}}(x_1, \ldots, x_p)$. The idea of separating $F_{\boldsymbol{X}}(x_1, \ldots, x_p)$ in two parts - the one which describes the dependence structure, and the other one which describes the marginal behavior, leads to the concept of copula. To every p-variate distribution function $F_{\boldsymbol{X}}(\cdot)$, with continuous marginals $F_{X_1}, \ldots, F_{X_p}$, corresponds a unique function $C : [0,1]^p \to [0,1]$, called a copula function such that

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = C\left[F_{X_1}(x_1), \ldots, F_{X_p}(x_p)\right]; \qquad \text{for} \quad \boldsymbol{x} = (x_1, \ldots, x_p)^T \in \mathbb{R}^p.$$

We have the following relation between the joint PDF of $X_1, \ldots, X_p$, and the copula density function;

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = c\left[F_{X_1}(x_1), \ldots, F_{X_p}(x_p)\right] f_{X_1}(x_1) \ldots f_{X_p}(x_p);$$

$$\text{for } \boldsymbol{x} = (x_1, \ldots, x_p)^T \in \mathbb{R}^p. \tag{2.3}$$

Here $f_{\boldsymbol{X}}(\boldsymbol{x})$ is the joint PDF of $X_1, \ldots, X_p$, $f_{X_j}(x_j)$ is the PDF of X_j, for $j = 1, \ldots, p$, and $c(u_1, \ldots, u_p)$ is the copula density function of $C(u_1, \ldots, u_p)$. Moreover, from Sklar's theorem (see Nelsen: 2006, page 18), it follows that if $F_{\boldsymbol{X}}(\boldsymbol{x})$ is a joint distribution function with continuous marginals $F_{X_1}(\cdot), \ldots, F_{X_p}(\cdot)$, and if $F_{X_1}^{-1}(\cdot), \ldots, F_{X_p}^{-1}(\cdot)$ are the inverse functions of $F_{X_1}, \ldots, F_{X_p}$, respectively, then there exists a unique copula C in $[0,1]^p$, such that

$$C(u_1, \ldots, u_p) = F_{\boldsymbol{X}}\left(F_{X_1}^{-1}(u_1), \ldots, F_{X_p}^{-1}(u_p)\right) \text{ for } \boldsymbol{u} = (u_1, \ldots, u_p)^T \in [0,1]^p.$$

It is well known that many dependence properties of a multivariate distribution are copula properties, and therefore, can be obtained by studying the corresponding copula.

2.3. Burr Type Distribution. Burr (1942) introduced twelve different distribution functions for modelling data. Among these twelve distribution functions, Burr Type X and Burr Type XII received the maximum attention. Thorough analysis of Burr Type XII distribution in Rodriguez (1977) (see also Wingo, 1993). A random variable Z is said to have a Burr Type XII distribution if the CDF of Z for $\delta > 0$, $\theta > 0$ and $\beta > 0$ is

$$F_Z(z) = 1 - \frac{1}{\left(1 + \theta z^{\beta}\right)^{\delta}}; \quad \text{for} \quad z > 0, \tag{2.4}$$

and 0 otherwise. The associated PDF becomes

$$f_Z(z) = \frac{\beta \theta z^{\beta - 1}}{\left(1 + \theta z^{\beta}\right)^{\delta + 1}}; \quad \text{for} \quad z > 0, \tag{2.5}$$

and 0 otherwise.

The following connection between a Burr Type XII and a GE distribution can be easily established, and it will be used later to construct a multivariate GE distribution. Consider a random variable

$$X = -\frac{1}{\lambda} \ln\left(1 - \left(1 + \theta Z^{\beta}\right)^{-1/\tau}\right), \tag{2.6}$$

where Z is a non-negative random variable with the CDF given in (2.4) and $\tau > 0$. Then, X follows ($\sim$) GE(α, λ) with $\alpha = \delta\tau$.

3 Multivariate GE Distribution

Takahasi (1965) introduced multivariate Burr (Type XII) distribution which can be described as follows. A p-variate random vector $\boldsymbol{Z} = (Z_1, \ldots, Z_p)^T$ is said to have a multivariate Burr (Type XII) distribution if it has the joint PDF

$$f_{\boldsymbol{Z}}(z_1, \ldots, z_p) = \frac{\Gamma(\theta+p)}{\Gamma(\theta)} \times \frac{\prod_{i=1}^{p} \beta_i \theta_i z_i^{\beta_i - 1}}{\left(1 + \sum_{i=1}^{p} \theta_i z_i^{\beta_i}\right)^{\theta+p}}; \quad \text{for} \quad z_1 > 0, \ldots, z_p > 0. \tag{3.1}$$

Here $\beta_1 > 0, \ldots, \beta_p > 0$, $\theta_1 > 0, \ldots, \theta_p > 0$ and $\theta > 0$. It has been shown by Takahasi (1965) that if $\boldsymbol{Z} = (Z_1, \ldots, Z_p)^T$ is a p-variate Burr distribution then for any $q < p$, $(Z_{i_1}, \ldots, Z_{i_q})$, for $1 \le i_1 < i_2 < \ldots < i_q \le p$, is a q-variate Burr distribution. Now for the random vector $(Z_1, \ldots, Z_p)^T$ with joint PDF (3.1), consider the following random vector $\boldsymbol{X} = (X_1, \ldots, X_p)^T$, where for $i = 1, \ldots, p$,

$$X_i = -\frac{1}{\lambda_i} \ln\left(1 - \left(1 + \theta_i Z_i^{\beta_i}\right)^{-1/\alpha_i}\right), \tag{3.2}$$

$\lambda_1 > 0, \ldots, \lambda_p > 0$, $\beta_1 > 0, \ldots, \beta_p > 0$, $\alpha_1 > 0, \ldots, \alpha_p > 0$. The joint PDF of $(X_1, \ldots, X_p)^T$ is

$$f_{\boldsymbol{X}}(x_1, \ldots, x_p) = \frac{c e^{-\sum_{i=1}^{p} \lambda_i x_i} \prod_{i=1}^{p} \left(1 - e^{-\lambda_i x_i}\right)^{-\alpha_i - 1}}{\left[\sum_{i=1}^{p} \left(1 - e^{-\lambda_i x_i}\right)^{-\alpha_i} - (p-1)\right]^{\theta+p}}. \tag{3.3}$$

Here the normalizing constant $c = \prod_{i=1}^{p} \alpha_i \lambda_i (\theta + i - 1)$. From now on a p-variate random vector with the joint PDF (5.1) will be called a multivariate generalized exponential (MVGE) distribution, and it will be denoted by the $\text{MVGE}_p(\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_p, \theta)$.

The proposed MVGE can be obtained in many other ways also. Consider the following one-parameter exchangeable p-variate distribution of Mardia (1962) defined on $(0, \infty)^p$. Mardia (1962) defined the following p-variate random vector $\boldsymbol{V} = (V_1, \ldots, V_p)^T$, for $\theta > 0$, with the joint PDF

$$f_{\boldsymbol{V}}(v_1, \ldots, v_p) = \frac{\theta(\theta+1)\ldots(\theta+p-1)}{(1 + v_1 + v_2 + \ldots + v_p)^{\theta+p}}, \tag{3.4}$$

where $v_1 > 0, \ldots, v_p > 0$. He showed that if $\boldsymbol{V}$ is a random vector with the joint PDF (3.4), then the marginals, the joint CDF, and the joint survival function can be obtained in explicit forms. Now consider the following p-variate random vector $\boldsymbol{X} = (X_1, \ldots, X_p)^T$ as follows

$$V_i = \left(1 - e^{-\lambda_i X_i}\right)^{-\alpha_i} - 1; \quad i = 1, \ldots, p. \tag{3.5}$$

Then the random vector $\boldsymbol{X}$ has the joint PDF (3.3).

We show that the MVGE distribution can be obtained from the multivariate Clayton copula also. Consider the one-parameter ($\theta > 0$), p-variate Clayton copula (see for example Nelsen, 2006),

$$C_\theta(u_1, \ldots, u_p) = \frac{1}{\left(u_1^{-1/\theta} + u_2^{-1/\theta} + \ldots + u_p^{-1/\theta} - (p-1)\right)^\theta}, \tag{3.6}$$

for $u_1 > 0, \ldots, u_p > 0$. Now we construct a new multivariate CDF for $x_1 > 0, \ldots, x_p > 0$, using the Clayton copula (3.6) as follows;

$$F_{\boldsymbol{X}}(x_1, \ldots, x_p) = C_\theta(F_{X_1}(x_1), \ldots, F_{X_p}(x_p)), \tag{3.7}$$

where for $i = 1, \ldots, p$,

$$F_{X_i}(x_i) = \left(1 - e^{-\lambda_i x_i}\right)^{\alpha_i \theta}.$$

Clearly, for $i = 1, \ldots, p$, $X_i \sim \text{GE}(\theta\alpha_i, \lambda_i)$, and the joint CDF of $\boldsymbol{X}$ becomes

$$F_{\boldsymbol{X}}(x_1, \ldots, x_p) = \frac{1}{\left[\sum_{i=1}^p \left(1 - e^{-\lambda_i x_i}\right)^{-\alpha_i} - (p-1)\right]^\theta}. \tag{3.8}$$

Now using the copula density function $c_\theta(u_1, \ldots, u_p)$, where

$$c_\theta(u_1, \ldots, u_p) = \frac{\theta(\theta+1)\ldots(\theta+p-1)}{\theta^p} \times \frac{\prod_{j=1}^p u_j^{-\left(\frac{1}{\theta}+1\right)}}{\left(u_1^{-1/\theta} + \ldots + u_p^{-1/\theta} - (p-1)\right)^{(\theta+p)}},$$

it is immediate that

$$\frac{\partial^p}{\partial x_1 \ldots \partial x_p} F_{\boldsymbol{X}}(x_1, \ldots, x_p) = f_{\boldsymbol{X}}(x_1, \ldots, x_p).$$

Here $f_{\boldsymbol{X}}(x_1, \ldots, x_p)$ is same as defined in (3.3). Hence, if $\boldsymbol{X} \sim \text{MVGE}_p(\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_p, \theta)$, then it has the CDF (3.8).

The following result will be useful to generate random samples from the MVGE distribution.

THEOREM 3.1. *Suppose the p-variate random vector* $\boldsymbol{V} = (V_1, \ldots, V_p)^T$ *has the joint PDF (3.4), then the marginal PDF of* V_1 *and conditional PDFs of*

$$\{V_2|V_1 = v_1\}, \quad \{V_3|V_1 = v_1, V_2 = v_2\}, \ldots, \quad \{V_p|V_1 = v_1, \ldots, V_{p-1} = v_{p-1}\},$$

for $v_1 > 0, \ldots, v_p > 0$, *are, respectively,*

$$\begin{aligned}
f_{V_1}(v_1) &= \frac{\theta}{(1+v_1)^{\theta+1}}; \\
f_{V_2|V_1=v_1}(v_2) &= \frac{(\theta+1)(1+v_1)^{\theta+1}}{(1+v_1+v_2)^{\theta+2}}; \\
f_{V_3|V_1=v_1,V_2=v_2}(v_3) &= \frac{(\theta+2)(1+v_1+v_2)^{\theta+2}}{(1+v_1+v_2+v_3)^{\theta+3}}; \\
&\vdots \\
f_{V_p|V_1=v_1,\ldots,V_{p-1}=v_{p-1}}(v_p) &= \frac{(\theta+p-1)(1+v_1+\ldots+v_{p-1})^{\theta+p-1}}{(1+v_1+\ldots+v_p)^{\theta+p}}.
\end{aligned}$$

The corresponding CDFs are,

$$\begin{aligned}
F_{V_1}(v_1) &= 1 - \frac{1}{(1+v_1)^{\theta}}; \\
F_{V_2|V_1=v_1}(v_2) &= 1 - \left(\frac{1+v_1}{1+v_1+v_2}\right)^{\theta+1}; \\
F_{V_3|V_1=v_1,V_2=v_2}(v_3) &= 1 - \left(\frac{1+v_1+v_2}{1+v_1+v_2+v_3}\right)^{\theta+2}; \\
&\vdots \\
F_{V_p|V_1=v_1,\ldots,V_{p-1}=v_{p-1}}(v_p) &= 1 - \left(\frac{1+v_1+\ldots+v_{p-1}}{1+v_1+\ldots+v_p}\right)^{\theta+p-1},
\end{aligned}$$

respectively.

PROOF. The proof of Theorem 3.1 can be obtained in a routine manner, and it is avoided.

Using Theorem 3.1, a random vector $\boldsymbol{V} = (V_1, \ldots, V_p)^T$ can be easily generated, sequentially. First generate V_1 using the inverse transformation, then V_2 and so on. Once V_i's are generated, X_i's can be easily obtained using the transformation (3.5).

4 Properties

In this section we provide several properties of the MVGE distribution. First we provide the distribution functions of the marginals, conditionals, and the extreme order statistics of the MVGE distribution.

Theorem 4.1. *If* $\boldsymbol{X} = (X_1, \ldots, X_P)^T \sim MVGE_p(\alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_p, \theta)$*, then*

(a) $X_1 \sim GE(\alpha_1\theta, \lambda_1)$*, …,* $X_p \sim GE(\alpha_p\theta, \lambda_p)$*.*

(b) For any non-empty subset $I_q = (i_1, \ldots, i_q) \subset (1, \ldots, p)$*, the* q*-dimensional marginal* $\boldsymbol{X}_{I_q} = (X_{i_1}, \ldots, X_{i_q})^T \sim MVGE_q(\alpha_{i_1}, \ldots, \alpha_{i_q}, \lambda_{i_1}, \ldots, \lambda_{i_q}, \theta)$*.*

(c) The conditional distribution function of $(\boldsymbol{X}_{i_q} | \boldsymbol{X}_{I-I_q} \leq \boldsymbol{x}_{I-I_q})$*, where the set* $I - I_q = \{i \in I, i \neq i_1, \ldots, i_q\}$*, is*

$$P(\boldsymbol{X}_{i_q} \leq \boldsymbol{x}_{I_q} | \boldsymbol{X}_{I-I_q} \leq \boldsymbol{x}_{I-I_q}) = \left[\frac{\sum_{i\in I-I_q}(1-e^{-\lambda_i x_i})^{-\alpha_i} - (p-q-1)}{\sum_{i\in I}(1-e^{-\lambda_i x_i})^{-\alpha_i} - (p-1)}\right]^\theta. \tag{4.1}$$

(d) The survival function of $\boldsymbol{X} = (X_1, \ldots, X_p)^T$ *is*

$$\begin{aligned} S_{\boldsymbol{X}}(\boldsymbol{x}) &= 1 - \sum_{i=1}^{p}(1-e^{-\lambda_i x_i})^{\theta\alpha_i} \\ &\quad + \sum_{1\leq i<j\leq p} \frac{1}{\left[(1-e^{-\lambda_i x_i})^{-\alpha_i} + (1-e^{-\lambda_j x_j})^{-\alpha_j} - 1\right]^\theta} \\ &\quad + \ldots + (-1)^{p+1}\frac{1}{\left[\sum_{i=1}^{p}(1-e^{-\lambda_i x_i})^{-\alpha_i} - (p-1)\right]^\theta}. \end{aligned} \tag{4.2}$$

Proof. The proofs of (a), (b) and (c) can be obtained using (3.8). The proof of (d) can be obtained using the following relation:

$$\begin{aligned} P(\boldsymbol{X} > \boldsymbol{x}) &= 1 - P((\boldsymbol{X} > \boldsymbol{x})^c) \\ &= 1 - P(\{X_1 \leq x_1\} \cup \{X_2 \leq x_2\} \cup \ldots \cup \{X_p \leq x_p\}) \\ &= 1 - \sum_{i=1}^{p} P(X_i \leq x_i) + \sum_{1\leq i<j\leq p} P(X_i \leq x_i, X_j \leq x_j) \\ &\quad + \ldots + \ldots (-1)^{p+1} P\left(\prod_{i=1}^{p} X_i \leq x_i\right). \end{aligned}$$

It is clear that if $X_{(1)} = \min\{X_1, \dots, X_p\}$ and $X_{(p)} = \max\{X_1, \dots, X_p\}$, then their distributions can be easily obtained from the expressions (4.2) and (3.8), respectively.

THEOREM 4.2. *If* $\boldsymbol{X} = (X_1, \dots, X_P)^T \sim MVGE_p(\alpha_1, \dots, \alpha_p, \lambda_1, \dots, \lambda_p, \theta)$, *then* $\boldsymbol{X}$ *is*

(a) PLOD, positively lower orthant dependent.

(b) LTD, left tail decreasing.

(c) LCSD, left corner set decreasing.

PROOF.

(a) Note that a random vector $\boldsymbol{X}$ is PLOD if and only if it satisfies (2.1). Since, PLOD property is a copula property, Nelsen (2006), note that to prove (2.1) in case of the MVGE, it is equivalent to show

$$C_\theta(u_1, \dots, u_p) = \frac{1}{\left(u_1^{-1/\theta} + u_2^{-1/\theta} + \dots + u_p^{-1/\theta} - (p-1)\right)^\theta} \geq u_1 \dots u_p, \tag{4.3}$$

for all $0 < u_1, \dots, u_p < 1$. Make the following transformation $v_1 = u_1^{1/\theta_1}, \dots, v_p = u_p^{1/\theta_p}$. Hence proving (4.3) is equivalent to prove

$$0 \leq v_1 v_2 \dots v_p \left[\frac{1}{v_1} + \dots + \frac{1}{v_p} - (p-1)\right] \leq 1, \tag{4.4}$$

for all $0 < v_1, \dots, v_p < 1$. Now we will prove (4.4) by induction on p. Clearly the result is true for $p = 1$. It is assumed that the result is true for $p = k$, and we will prove that the result is true for $p = k + 1$. For $0 \leq v_1, \dots, v_{k+1} \leq 1$, let us write

$$D = v_1 v_2 \dots v_{k+1} \left[\frac{1}{v_1} + \dots + \frac{1}{v_{k+1}} - k\right] = A + B.$$

Here, due to induction hypothesis

$$A = v_1 v_2 \dots v_k \left[\frac{1}{v_1} + \dots + \frac{1}{v_k} - (k-1)\right] v_{k+1} \leq v_{k+1}$$

and

$$B = v_1 v_2 \dots v_k v_{k+1} \left[\frac{1}{v_{k+1}} - 1\right] = v_1 v_2 \dots v_k (1 - v_{k+1}) \leq (1 - v_{k+1}).$$

Hence $D = A + B \leq v_{k+1} + (1 - v_{k+1}) = 1$.

(b) To prove (b), without loss of generality, let us take $A = \{1, \ldots, q\}$ and $B = \{q+1, \ldots, p\}$. If $\boldsymbol{x} = (x_1, \ldots, x_p)^T$, $x_i \geq 0$, for $i = 1, \ldots, p$, then

$$P(\boldsymbol{X}_B \leq \boldsymbol{x}_B | \boldsymbol{X}_A \leq \boldsymbol{x}_A) = \left[\frac{\sum_{i=1}^{q}(1-e^{-\lambda_i x_i})^{-\alpha_i} - (q-1)}{\sum_{i=1}^{p}(1-e^{-\lambda_i x_i})^{-\alpha_i} - (p-1)}\right]^{\theta}. \quad (4.5)$$

The right hand side of (4.5) can be written as

$$\left[\frac{1}{1+[\sum_{i=q+1}^{p}(1-e^{-\lambda_i x_i})^{-\alpha_i} - (p-q)][\sum_{i=1}^{q}(1-e^{-\lambda_i x_i})^{-\alpha_i} - (q-1)]^{-1}}\right]^{\theta}. \quad (4.6)$$

Since for fixed $x_{q+1}, \ldots, x_p$, the function defined in (4.6), is a non-increasing function of $x_1, \ldots, x_q$, the result follows.

(c) In order to prove the part (c), let us consider

$$\frac{F_{\boldsymbol{X}}(\boldsymbol{x} \wedge \boldsymbol{x}')}{F_{\boldsymbol{X}}(\boldsymbol{x}')} = \left[\frac{\sum_{i=1}^{p}(1-e^{-\lambda_i x_i'})^{-\alpha_i} - (p-1)}{\sum_{i=1}^{p}(1-e^{-\lambda_i \min\{x_i, x_i'\}})^{-\alpha_i} - (p-1)}\right]^{\theta}. \quad (4.7)$$

We will show that the function defined in (4.7) is a non-increasing function of x_1', when $x_2', \ldots, x_p'$ are kept fixed, and that will prove the result. Consider two cases separately. Suppose $x_1' \leq x_1$, in this case the right hand side of (4.7) as a function of x_1', can be written as $\left[g(x_1')\right]^{\theta}$, where

$$g(x_1') = \frac{u(x_1') + c_1}{u(x_1') + c_2},$$

$$u(x_1') = (1 - e^{-\lambda_1 x_1'})^{-\alpha_1},$$

$$c_1 = \sum_{i=2}^{p}(1-e^{-\lambda_i x_i'})^{-\alpha_i} - (p-1) \text{ and } c_2 = \sum_{i=2}^{p}(1-e^{-\lambda_i \min\{x_i, x_i'\}})^{-\alpha_i} - (p-1).$$

Clearly, $c_2 \geq c_1 \geq 0$. Since $u(x_1')$ is a non-decreasing function of x_1', it follows that $g(x_1')$, hence $\left[g(x_1')\right]^{\theta}$, is a non-increasing function of x_1'. Now for $x_1' > x_1$, the right hand side of (4.7) as a function of x_1', can be written as $c(u(x_1') + c_1)$ for $c > 0$. Here, c_1 is same as before, and

$$c = \left[\sum_{i=1}^{p}(1-e^{-\lambda_i \min\{x_i, x_i'\}})^{-\alpha_i} - (p-1)\right]^{-\theta}.$$

Since $u(x_1')$ is a non-decreasing function of x_1', the result immediately follows.

THEOREM 4.3. *If* $\boldsymbol{X} = (X_1, \ldots, X_P)^T \sim MVGE_p(\alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_p, \theta)$*, then* $\boldsymbol{X}$ *has a multivariate total positivity of order two (MTP*$_2$*) property.*

PROOF. Recall that $F_{\boldsymbol{X}}(\boldsymbol{x})$ has the MTP$_2$ property, if and only if

$$\frac{F_{\boldsymbol{X}}(\boldsymbol{x})F_{\boldsymbol{Y}}(\boldsymbol{y})}{F_{\boldsymbol{X}}(\boldsymbol{x} \vee \boldsymbol{y})F_{\boldsymbol{Y}}(\boldsymbol{x} \wedge \boldsymbol{y})} \leq 1. \tag{4.8}$$

Here $\boldsymbol{x} = (x_1, \ldots, x_p)^T$, $\boldsymbol{y} = (y_1, \ldots, y_p)^T$, $\boldsymbol{x} \vee \boldsymbol{y} = (x_1 \vee y_1, \ldots, x_p \vee y_p)^T$, $\boldsymbol{x} \wedge \boldsymbol{y} = (x_1 \wedge y_1, \ldots, x_p \wedge y_p)^T$, where $c \vee d = \max\{c, d\}$ and $c \wedge d = \min\{c, d\}$. We will use the following notations for $i = 1, \ldots, p$;

$$a_i = (1 - e^{-\lambda_i x_i})^{-\alpha_i} \quad \text{and} \quad b_i = (1 - e^{-\lambda_i y_i})^{-\alpha_i}.$$

Therefore, proving (4.8) is equivalent to proving

$$\begin{aligned}\left(\sum_{i=1}^{p} \min\{a_i, b_i\} - (p-1)\right)\left(\sum_{i=1}^{p} \max\{a_i, b_i\} - (p-1)\right) \\ \leq \left(\sum_{i=1}^{p} a_i - (p-1)\right)\left(\sum_{i=1}^{p} b_i - (p-1)\right).\end{aligned} \tag{4.9}$$

Here $a_i \geq 1$ and $b_i \geq 1$ for $i = 1, \ldots, p$. Note that (4.9) can be established if we can show for $c_i \geq 0$ and $d_i \geq 0$,

$$\left(\sum_{i=1}^{p} \min\{c_i, d_i\}\right)\left(\sum_{i=1}^{p} \max\{c_i, d_i\}\right) \leq \left(\sum_{i=1}^{p} c_i\right)\left(\sum_{i=1}^{p} d_i\right). \tag{4.10}$$

If $c_i \geq d_i$ or $c_i \leq d_i$ for $i = 1, \ldots, p$, (4.10) easily follows. Suppose there exists a $1 < q < p$, and without loss of generality we assume that $c_1 \leq d_1, \ldots, c_q \leq d_q, c_{q+1} > d_{q+1}, \ldots, c_p > d_p$. Therefore

$$\begin{aligned}&\left(\sum_{i=1}^{p} c_i\right)\left(\sum_{i=1}^{p} d_i\right) - \left(\sum_{i=1}^{p} \min\{c_i, d_i\}\right)\left(\sum_{i=1}^{p} \max\{c_i, d_i\}\right) \\ &= \left(\sum_{i=q+1}^{p} (c_i - d_i)\right)\left(\sum_{i=1}^{q} (d_i - c_i)\right) \geq 0.\end{aligned}$$

THEOREM 4.4. *Suppose* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *are p-variate random vectors, such that*

$$\boldsymbol{X} \sim MVGE_p(\alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_p, \theta) \ \ \textit{and} \ \ \boldsymbol{Y} \sim MVGE_p(\beta_1, \ldots, \beta_p, \lambda_1, \ldots, \lambda_p, \theta).$$

If $\alpha_i \le \beta_i$ *for* $i = 1, \ldots, p$, *then* $\boldsymbol{X} \le_{st} \boldsymbol{Y}$.

Proof: *Since* $\alpha_i \le \beta_i$, *it follows* $X_i \le_{st} Y_i$, *for* $i = 1, \ldots, p$. *since* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *have same copula, the result follows using the Theorem 6.B.1 of Joe (1997).*

Theorem 4.5. *Suppose* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *are p-variate random vectors, such that*

$$\boldsymbol{X} \sim MVGE_p(\alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_p, \theta), \quad \boldsymbol{Y} \sim MVGE_p(\alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_p, \delta),$$

and they are independently distributed. Then

$$P(\boldsymbol{Y} \le \boldsymbol{X}) = P(Y_1 \le X_1, \ldots, Y_p \le X_p) = \prod_{i=1}^{p} \frac{\theta + i - 1}{\theta + \delta + i - 1}.$$

Proof. For $c = \prod_{i=1}^{p} \alpha_i \lambda_i (\theta + i - 1)$,

$$\begin{aligned} P(\boldsymbol{Y} \le \boldsymbol{X}) = & \; P(Y_1 \le X_1, \ldots, Y_p \le X_p) \\ = & \int_0^{\infty} \cdots \int_0^{\infty} f_{\boldsymbol{X}}(x_1, \ldots, x_p) F_{\boldsymbol{Y}}(x_1, \ldots, x_p) dx_1 \ldots dx_p \\ = & \int_0^{\infty} \cdots \int_0^{\infty} \frac{c e^{-\sum_{i=1}^{p} \lambda_i x_i} \prod_{i=1}^{p} \left(1 - e^{-\lambda_i x_i}\right)^{-\alpha_i - 1}}{\left[\sum_{i=1}^{p} \left(1 - e^{-\lambda_i x_i}\right)^{-\alpha_i} - (p-1)\right]^{\theta + \delta + p}} dx_1 \ldots dx_p \\ = & \prod_{i=1}^{p} \frac{\theta + i - 1}{\theta + \delta + i - 1}. \end{aligned}$$

Interestingly $P(\boldsymbol{Y} \le \boldsymbol{X})$ does not depend on α_i's and λ_i's, it just depends on the copula parameters.

5 Different Estimators

5.1. Maximum Likelihood Estimators. Suppose $\{(x_{i1}, \ldots, x_{ip}); i = 1, \ldots, n\}$ is a random sample of size n from a MVGE($\boldsymbol{\Theta}$), where $\boldsymbol{\Theta} = (\alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_p, \theta)$, and consider the maximum likelihood estimation of the $2p + 1$ unknown parameters of $\boldsymbol{\Theta}$. The likelihood function of the unknown parameters can be written as

$$L(\boldsymbol{\Theta}) = \left(\prod_{i=1}^{p} \alpha_i \lambda_i (\theta + i - 1)\right)^n \frac{e^{-\sum_{i=1}^{n} \sum_{j=1}^{p} \lambda_j x_{ij}} \prod_{i=1}^{n} \prod_{j=1}^{p} \left(1 - e^{-\lambda_j x_{ij}}\right)^{-\alpha_j - 1}}{\prod_{i=1}^{n} \left[\sum_{j=1}^{p} \left(1 - e^{-\lambda_j x_{ij}}\right)^{-\alpha_j} - (p-1)\right]^{\theta + p}}.$$

Hence, the log-likelihood function becomes

$$\begin{aligned} l(\mathbf{\Theta}) \;=\;& n\sum_{j=1}^{p}\ln\alpha_j + n\sum_{j=1}^{p}\ln\lambda_j - \sum_{j=1}^{p}\lambda_j\sum_{i=1}^{n}x_{ij} \\ & -\sum_{j=1}^{p}(\alpha_j+1)\sum_{i=1}^{n}\ln(1-e^{-\lambda_j x_{ij}}) + n\sum_{j=1}^{p}\ln(\theta-j+1) \\ & -\sum_{j=1}^{p}\sum_{i=1}^{n}(\theta+p)\ln\left\{\sum_{j=1}^{p}(1-e^{-\lambda_j x_{ij}})^{-\alpha_j}-(p-1)\right\}. \end{aligned} \quad (5.1)$$

The MLEs of the unknown parameters can be obtained by maximizing the log-likelihood function (5.1) with respect to $2p+1$ unknown parameters. The explicit solutions are not available, hence we need to obtain numerical solutions. It involves solving simultaneously $2p+1$ non-linear equations. Newton-Raphson method may be used to solve the $2p+1$ non-linear equations, but for large p, it is quite difficult to implement.

Due to this reason, we propose to use the MBP (maximization by parts) method proposed by Song et al. (2005). The MBP method can be used quite effectively to compute the MLEs in this case, and it involves solving only $p+1$ non-linear equations separately, in each iteration. Implementation of the proposed algorithm is quite simple, and it can be used quite effectively for large p also. Moreover, in this case it can be proved that it produces consistent estimates of the unknown parameters at each iteration. We make the following one to one transformation of the parameters:

$$(\alpha_1,\ldots,\alpha_p,\lambda_1,\ldots,\lambda_p,\theta) \Leftrightarrow (\beta_1,\ldots,\beta_p,\lambda_1,\ldots,\lambda_p,\theta),$$

where $\beta_i = \alpha_i\theta$, for $i = 1,\ldots,p$. For convenience we denote $\mathbf{\Theta} = (\beta_1,\ldots,\beta_p,\lambda_1,\ldots,\lambda_p,\theta)$. We re-write the log-likelihood function (5.1) , in terms of the copula density as follows:

$$l(\mathbf{\Theta}) = l_1(\mathbf{\Theta}_1) + l_2(\mathbf{\Theta}_1,\theta), \quad (5.2)$$

where $\mathbf{\Theta}_1 = (\beta_1,\ldots,\beta_p,\lambda_1,\ldots,\lambda_p)$, $l_1(\mathbf{\Theta}_1) = \sum_{j=1}^{p} l_{1j}(\beta_j,\lambda_j)$,

$$l_{1j}(\beta_j,\lambda_j) = n\ln\beta_j + n\ln\lambda_j - \sum_{i=1}^{n}\lambda_j x_{ij} + \sum_{i=1}^{n}(\beta_j-1)\ln(1-e^{-\lambda_j x_{ij}}), \quad (5.3)$$

and for $u_{ij}(\beta_j, \lambda_j) = (1 - e^{-\lambda_j x_{ij}})^{\beta_j}$, $i = 1, \ldots, n$, $j = 1, \ldots, p$,

$$\begin{aligned} l_2(\boldsymbol{\Theta}_1, \theta) &= -np \ln \theta + n \sum_{j=1}^{p} \ln(\theta + j - 1) - \left(\frac{1}{\theta} + 1\right) \sum_{i=1}^{n} \sum_{j=1}^{p} \ln u_{ij}(\beta_j, \lambda_j) \\ &\quad -(\theta + p) \sum_{i=1}^{n} \ln \Big((u_{i1}(\beta_1, \lambda_1))^{-1/\theta} + \ldots + (u_{ip}(\beta_p, \lambda_p))^{-1/\theta} \\ &\quad -(p-1) \Big). \end{aligned} \tag{5.4}$$

Now we propose the following MBP algorithm to compute the MLEs of $\boldsymbol{\Theta}$. We use the following notations:

$$\dot{l}_1(\boldsymbol{\Theta}_1) = \frac{\partial}{\partial \boldsymbol{\Theta}_1} l_1(\boldsymbol{\Theta}_1), \quad \dot{l}_{21}(\boldsymbol{\Theta}_1, \theta) = \frac{\partial}{\partial \boldsymbol{\Theta}_1} l_2(\boldsymbol{\Theta}_1, \theta), \quad \dot{l}_{22} = \frac{\partial}{\partial \theta} l_2(\boldsymbol{\Theta}_1, \theta)$$

Algorithm 1.

Step 1: Find $\boldsymbol{\Theta}_1^{(1)}$ by solving $\dot{l}_1(\boldsymbol{\Theta}_1) = 0$.

Step 2: Find $\theta^{(1)}$, by solving $\dot{l}_{22}(\boldsymbol{\Theta}_1^{(1)}, \theta) = 0$.

Step 3: At the k-th iteration, for $k > 0$, find $\boldsymbol{\Theta}_1^{(k+1)}$, by solving $\dot{l}_1(\boldsymbol{\Theta}_1) = -\dot{l}_{21}(\boldsymbol{\Theta}_1^{(k)}, \theta^{(k)})$.

Step 4: Find $\theta^{(k+1)}$, by solving $\dot{l}_{22}(\boldsymbol{\Theta}^{(k+1)}, \theta) = 0$.

The implementation details are provided in Appendix A.

Since the MVGE satisfies all the conditions for the consistency and asymptotic normality of the MLEs to hold, we have the following result. If $\widehat{\boldsymbol{\Theta}}$ is the MLE of $\boldsymbol{\Theta}$, then

$$\sqrt{n}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}) \rightarrow N_{2p+1}(\mathbf{0}, \boldsymbol{I}^{-1}). \tag{5.5}$$

Here $\boldsymbol{I}$ is the expected Fisher information matrix. Note that it is not difficult to compute the expected Fisher information matrix which can be expressed in p dimensional integration. The observed information matrix can be used in construction of asymptotic confidence intervals of the unknown parameters. The observed Fisher information matrix can be obtained by taking second derivatives of the log-likelihood function. They are provided in Appendix B for easy reference.

5.2. Two Stage Copula Estimators. Since the MVGE has a very convenient copula structure, the method of Joe (2005) can be immediately used to provide two stage copula estimators of the unknown parameters. In this case the estimators can be obtained by solving $p + 1$ non-linear equations separately, hence computationally two-stage copula estimators can be obtained more conveniently than the MLEs. The two-stage copula estimators can be obtained very easily using the structure of the log-likelihood function (5.2). The following algorithm can be used for that purpose:

ALGORITHM 2.

Step 1: Maximize $l_1(\boldsymbol{\Theta}_1)$ with respect to $\boldsymbol{\Theta}_1$ to compute the estimate of $\boldsymbol{\Theta}_1$, say $\widetilde{\boldsymbol{\Theta}}_1$.

Step 2: Maximize $l_2(\widetilde{\boldsymbol{\Theta}}_1, \theta)$ with respect to θ, to get an estimate of θ, say $\widetilde{\theta}$.

Since Step 1 can be obtained by solving p non-linear equations separately, it is immediate that $\widetilde{\boldsymbol{\Theta}}$ can be obtained by solving $p+1$ non-linear equations separately. We have the following result. If $\widetilde{\boldsymbol{\Theta}}$ is the two-stage estimator of $\boldsymbol{\Theta}$, then

$$\sqrt{n}(\widetilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}) \rightarrow N_{2p+1}(\mathbf{0}, \boldsymbol{W}). \tag{5.6}$$

The exact expressions of the different elements of $\boldsymbol{W}$ can be obtained using the method provided by Joe (2005) and they are provided in Appendix C.

6 Simulations and Data Analysis

6.1. Simulations. In this section we present some simulation results mainly to compare the performances of the MLEs and the two-stage estimators. We have taken different p ($p = 3$ and 5) values, different sample sizes ($n = 20$, 40, 60, 80 and 100) and different θ ($\theta = 1$ and 2) values. The samples have been generated using the method proposed in Section 3. For each data set, we have calculated the estimators of the unknown parameters using two-stage estimators, and also based on MLEs. In calculating the MLEs we have used the two-stage estimators as initial guesses. We report the average estimates and the square root of the mean squared errors based on 1000 replications. The results are reported in Tables 1, 2, 3, 4, 5, 6, 7 and 8. In each box the upper figures indicate the average estimate and the associated square root of the mean squared errors (MSE) is reported below.

It is clear from the simulation results that as the sample size increases, as expected, the biases and the MSEs decrease in both cases. Both methods perform quite satisfactorily. The performance of the MLEs are slightly better

Table 1: The average estimates based on two-stage copula method and the associated square root mean squared errors (reported below) for different parameters when $\alpha_1 = \lambda_1 = \alpha_2 = \lambda_2 = \alpha_3 = \lambda_3 = 1$ and $\theta = 1.0$

n	α_1	λ_1	α_2	λ_2	α_3	λ_3	θ
20	1.1216	1.1459	1.0900	1.1112	1.1060	1.1375	1.2021
	0.4505	0.3775	0.4136	0.3518	0.4596	0.3864	0.6974
40	1.0542	1.0694	1.0391	1.0501	1.0429	1.0597	1.0874
	0.2751	0.2407	0.2617	0.2341	0.2716	0.2375	0.3745
60	1.0391	1.0383	1.0341	1.0303	1.0324	1.0303	1.0467
	0.2146	0.1808	0.2132	0.1823	0.2148	0.1743	0.2746
80	1.0314	1.0285	1.0262	1.0227	1.0228	1.0205	1.0327
	0.1905	0.1538	0.1824	0.1553	0.1824	0.1522	0.2295
100	1.0264	1.0223	1.0188	1.0144	1.0160	1.0118	1.0239
	0.1647	0.1377	0.1568	0.1345	0.1592	0.1305	0.2006

Table 2: The average estimates and the associated square root mean squared errors (reported below) of the MLEs for different parameters when $\alpha_1 = \lambda_1 = \alpha_2 = \lambda_2 = \alpha_3 = \lambda_3 = 1$ and $\theta = 1.0$

n	α_1	λ_1	α_2	λ_2	α_3	λ_3	θ
20	1.1142	1.1301	1.0698	1.1096	1.0978	1.1167	1.1723
	0.4221	0.3552	0.4013	0.3498	0.4327	0.3689	0.6767
40	1.0481	1.0456	1.0227	1.0448	1.0352	1.0448	1.0765
	0.2522	0.2317	0.2445	0.2289	0.2598	0.2267	0.3592
60	1.0267	1.0228	1.0167	1.0234	1.0291	1.0299	1.0401
	0.2089	0.1756	0.2078	0.1778	0.2098	0.1705	0.2611
80	1.0229	1.0199	1.0198	1.0202	1.0177	1.0196	1.0298
	0.1889	0.1501	0.1801	0.1498	0.1801	0.1503	0.2109
100	1.0199	1.0201	1.0187	1.0098	1.0098	1.0101	1.0178
	0.1589	0.1354	0.1498	0.1311	0.1497	0.1298	0.1996

Table 3: The average estimates based on two-stage copula method and the associated square root mean squared errors (reported below) for different parameters when $\alpha_1 = \lambda_1 = \alpha_2 = \lambda_2 = \alpha_3 = \lambda_3 = 1$ and $\theta = 2.0$

n	α_1	λ_1	α_2	λ_2	α_3	λ_3	θ
20	1.2628	1.1172	1.2132	1.0873	1.2401	1.1071	2.4206
	0.7666	0.3108	0.7771	0.2929	0.7872	0.3164	1.3776
40	1.1098	1.0553	1.0864	1.0388	1.0939	1.0462	2.2607
	0.4405	0.2001	0.4119	0.1972	0.4367	0.1984	1.0397
60	1.0701	1.0299	1.0651	1.0250	1.0652	1.0258	2.1613
	0.3341	0.1522	0.3314	0.1554	0.3387	0.1500	0.8182
80	1.0570	1.0223	1.0529	1.0197	1.0479	1.0168	2.1182
	0.2935	0.1295	0.2856	0.1357	0.2902	0.1289	0.7127
100	1.0542	1.0176	1.0418	1.0123	1.0378	1.0094	2.0806
	0.2547	0.1171	0.2470	0.1154	0.2529	0.1104	0.6089

Table 4: The average estimates and the associated square root mean squared errors (reported below) of the MLEs for different parameters when $\alpha_1 = \lambda_1 = \alpha_2 = \lambda_2 = \alpha_3 = \lambda_3 = 1$ and $\theta = 2.0$

n	α_1	λ_1	α_2	λ_2	α_3	λ_3	θ
20	1.2427	1.0988	1.1943	1.0678	1.1756	1.0896	2.3452
	0.7332	0.2934	0.7215	0.2786	0.7567	0.2998	1.3561
40	1.0799	1.0336	1.0643	1.0265	1.0653	1.0227	2.2225
	0.4228	0.1889	0.3991	0.1801	0.4118	0.1768	1.0222
60	1.0565	1.0111	1.0338	1.0114	1.0399	1.0234	2.1556
	0.3139	0.1447	0.3098	0.1410	0.3089	0.1392	0.7789
80	1.0338	1.0210	1.0234	1.0102	1.0198	1.0119	2.0798
	0.2815	0.1198	0.2789	0.1210	0.2817	0.1219	0.7070
100	1.0220	1.0165	1.0228	1.0011	1.0229	1.0057	2.0182
	0.2489	0.1143	0.2389	0.1099	0.2498	0.1101	0.5988

Table 5: The average estimates based on two-stage copula method and the associated square root mean squared errors (reported below) for different parameters when $\alpha_1 = \lambda_1 = \alpha_2 = \lambda_2 = \alpha_3 = \lambda_3 = \alpha_4 = \lambda_4 = \alpha_5 = \lambda_5 = 1$ and $\theta = 1.0$

n	α_1	λ_1	α_2	λ_2	α_3	λ_3	α_4	λ_4	α_5	λ_5	θ
20	1.0876	1.1195	1.0788	1.1273	1.0753	1.1246	1.0700	1.1192	1.0750	1.1141	1.1478
	0.3419	0.3722	0.3491	0.3736	0.3467	0.3749	0.3429	0.3656	0.3443	0.3537	0.5226
40	1.0273	1.0594	1.0214	1.0542	1.0211	1.0542	1.0264	1.0592	1.0195	1.0550	1.0777
	0.2028	0.2327	0.2075	0.2189	0.2149	0.2195	0.2118	0.2345	0.2058	0.2307	0.2859
60	1.0149	1.0340	1.0167	1.0389	1.0129	1.0335	1.0148	1.0314	1.0150	1.0347	1.0454
	0.1606	0.1811	0.1629	0.1851	0.1639	0.1795	0.1668	0.1814	0.1608	0.1775	0.2157
80	1.0066	1.0242	1.0111	1.0287	1.0046	1.0182	1.0097	1.0267	1.0082	1.0260	1.0332
	0.1350	0.1517	0.1376	0.1501	0.1345	0.1484	0.1426	0.1578	0.1374	0.1529	0.1762
100	1.0064	1.0174	1.0094	1.0205	1.0028	1.0126	1.0099	1.0209	1.0087	1.0219	1.0259
	0.1214	0.1373	0.1236	0.1338	0.1225	0.1337	0.1283	0.1382	0.1209	0.1362	0.1561

Table 6: The average estimates and the associated square root mean squared errors (reported below) of the MLEs for different parameters when $\alpha_1 = \lambda_1 = \alpha_2 = \lambda_2 = \alpha_3 = \lambda_3 = \alpha_4 = \lambda_4 = \alpha_5 = \lambda_5 = 1$ and $\theta = 1.0$

n	α_1	λ_1	α_2	λ_2	α_3	λ_3	α_4	λ_4	α_5	λ_5	θ
20	1.0523	1.1098	1.0498	1.1156	1.0439	1.1110	1.0232	1.0101	1.0543	1.0909	1.1167
	0.3118	0.3434	0.3098	0.3423	0.3115	0.3389	0.3119	0.3432	0.3114	0.3427	0.4876
40	1.0195	1.0276	1.0114	1.0345	1.0134	1.0345	1.0087	1.0329	1.0101	1.0367	1.0435
	0.1818	0.2029	0.1834	0.2019	0.1919	0.2001	0.1889	0.1999	0.1876	0.2101	0.2675
60	1.0121	1.0178	1.0098	1.0141	1.0108	1.0076	1.0111	1.0198	1.0008	1.0187	1.0213
	0.1497	0.1634	0.1423	0.1671	0.1501	0.1598	0.1428	0.1699	0.1501	0.1701	0.1987
80	1.0022	1.0098	1.0043	1.0098	1.0025	1.0043	1.0065	1.0054	1.0031	1.0098	1.0114
	0.1298	0.1489	0.1301	0.1467	0.1289	0.1402	0.1388	0.1498	0.1287	0.1491	0.1698
100	1.0022	1.0008	1.0021	1.0039	1.0009	1.0001	1.0001	1.0101	1.0025	1.0111	1.0119
	0.1198	0.1332	0.1205	0.1313	0.1189	0.1315	0.1197	0.1312	0.1189	0.1334	0.1523

Table 7: The average estimates based on two-stage copula method and the associated square root mean squared errors (reported below) for different parameters when $\alpha_1 = \lambda_1 = \alpha_2 = \lambda_2 = \alpha_3 = \lambda_3 = \alpha_4 = \lambda_4 = \alpha_5 = \lambda_5 = 1$ and $\theta = 2.0$

n	α_1	λ_1	α_2	λ_2	α_3	λ_3	α_4	λ_4	α_5	λ_5	θ
20	1.1976	1.0928	1.1907	1.1050	1.1857	1.1011	1.1788	1.0929	1.1913	1.0961	2.2948
	0.5842	0.3046	0.5862	0.3133	0.5872	0.3104	0.5985	0.3026	0.6032	0.2931	1.1052
40	1.0645	1.0461	1.0606	1.0443	1.0617	1.0455	1.0633	1.0467	1.0542	1.0429	2.1925
	0.3205	0.1935	0.3414	0.1879	0.3399	0.1852	0.3257	0.1942	0.3212	0.1889	0.7529
60	1.0409	1.0266	1.0458	1.0323	1.0418	1.0284	1.0408	1.0239	1.0413	1.0278	2.1084
	0.2495	0.1531	0.2558	0.1564	0.2523	0.1492	0.2601	0.1517	0.2481	0.1483	0.5677
80	1.0236	1.0191	1.0318	1.0239	1.0217	1.0147	1.0284	1.0207	1.0257	1.0205	2.0786
	0.2047	0.1282	0.2105	0.1278	0.2012	0.1240	0.2168	0.1331	0.2060	0.1275	0.4518
100	1.0199	1.0133	1.0260	1.0175	1.0162	1.0100	1.0269	1.0172	1.0239	1.0179	2.0610
	0.1843	0.1163	0.1879	0.1148	0.1839	0.1111	0.1975	0.1174	0.1825	0.1141	0.3999

Table 8: The average estimates and the associated square root mean squared errors (reported below) of the MLEs for different parameters when $\alpha_1 = \lambda_1 = \alpha_2 = \lambda_2 = \alpha_3 = \lambda_3 = \alpha_4 = \lambda_4 = \alpha_5 = \lambda_5 = 1$ and $\theta = 2.0$

n	α_1	λ_1	α_2	λ_2	α_3	λ_3	α_4	λ_4	α_5	λ_5	θ
20	1.1623	1.0876	1.1543	1.0897	1.1659	1.0987	1.1683	1.0581	1.1459	1.0563	2.1778
	0.5325	0.2676	0.5410	0.2669	0.5289	0.2688	0.5411	0.2781	0.5386	0.2689	0.8998
40	1.0344	1.0312	1.0317	1.0117	1.0361	1.0009	1.0267	1.0056	1.0034	1.0101	2.1111
	0.2898	0.1551	0.2999	0.1523	0.2956	0.1489	0.3011	0.1423	0.2998	0.1501	0.5923
60	1.0319	1.0009	1.0219	1.0110	1.0278	1.0210	1.0391	1.0211	1.0313	1.0091	2.0198
	0.2398	0.1415	0.2367	0.1493	0.2381	0.1397	0.2391	0.1401	0.2369	0.1407	0.5197
80	1.0118	1.0023	1.0219	1.0110	1.0109	1.0054	1.0115	1.0107	1.0165	1.0167	2.0514
	0.1923	0.1175	0.2001	0.1198	0.1996	0.1210	0.1998	0.1218	0.1976	0.1212	0.4389
100	1.0019	1.0111	1.0213	1.0110	1.0009	1.0001	1.0191	1.0019	1.0154	1.0080	2.0098
	0.1818	0.1112	0.1825	0.1125	0.1801	0.1109	0.1889	0.1135	0.1822	0.1122	0.3818

than the estimators based on two-stage methods in terms of lower MSEs and biases particularly for small sample sizes. For large sample sizes both the methods behave quite similarly in terms of biases and MSEs. Although both methods perform quite satisfactorily, computationally the two-stage estimators are much easier to obtain than the MLEs.

6.2. Data Analysis. In this section we present the analysis of a data set for illustrative purposes mainly to show how the proposed model and the estimators work in practice. We analyze a multivariate data set obtained from Johnson and Wichern (1999, page 34), representing the bone mineral density (BMD) measured in g/cm^2 for 25 individuals. The data are presented in Table 9.

Table 9: The BMD data

Subject No.	Dominant radius	Radius	Dominant humerus	Humerus	Dominant ulna	Ulna
1	1.103	1.052	2.139	2.238	0.873	0.872
2	0.842	0.859	1.873	1.741	0.590	0.744
3	0.925	0.873	1.887	1.809	0.767	0.713
4	0.857	0.744	1.739	1.547	0.706	0.674
5	0.795	0.809	1.734	1.715	0.549	0.654
6	0.787	0.779	1.509	1.474	0.782	0.571
7	0.933	0.880	1.695	1.656	0.737	0.803
8	0.799	0.851	1.740	1.777	0.618	0.682
9	0.945	0.876	1.811	1.759	0.853	0.777
10	0.921	0.906	1.954	2.009	0.823	0.765
11	0.792	0.825	1.624	1.657	0.686	0.668
12	0.815	0.751	2.204	1.846	0.678	0.546
13	0.755	0.724	1.508	1.458	0.662	0.595
14	0.880	0.866	1.786	1.811	0.810	0.819
15	0.900	0.838	1.902	1.606	0.723	0.677
16	0.764	0.757	1.743	1.794	0.586	0.541
17	0.733	0.748	1.863	1.869	0.672	0.752
18	0.932	0.898	2.028	2.032	0.836	0.805
19	0.856	0.786	1.390	1.324	0.578	0.610
20	0.890	0.950	2.187	2.087	0.758	0.718
21	0.688	0.532	1.650	1.378	0.533	0.482
22	0.940	0.850	2.334	2.225	0.757	0.731
23	0.493	0.616	1.037	1.268	0.546	0.615
24	0.835	0.752	1.509	1.422	0.618	0.664
25	0.915	0.936	1.971	1.869	0.869	0.868

Preliminary data analysis suggests that marginals are coming from skewed distributions and they have increasing hazard functions. Therefore, the GE distribution can be used to fit the marginals. We have used the three-parameter GE for fitting the marginals, *i.e.* it has the PDF

$$f(x;\alpha,\lambda,\mu)=\alpha\lambda e^{-\lambda(x-\mu)}(1-e^{-\lambda(x-\mu)})^{\alpha-1},$$

and the location parameter μ is assumed to be known. We have fitted the three-parameter GE distributions to the marginals. In this case we first obtain the unbiased estimator of the location parameter using the method proposed by Hall and Wang (2005), and then compute the MLEs of the shape and scale parameters, and the results are presented in Table 10. We have subtracted the estimates of the location parameters from the corresponding marginals, and use this data set for fitting the model $\text{MVGE}_6(\alpha_1,\ldots,\alpha_6,\lambda_1,\ldots,\lambda_6,\theta)$.

We present below the two-stage copula estimators of the unknown parameters and the associated 95 % confidence interval in brackets based on the asymptotic distribution of the two-stage estimators as provided in (5.6).

$$\widetilde{\alpha}_1=23.8616(\mp 7.1513),\ \widetilde{\lambda}_1=7.5595(\mp 2.8776),$$

$$\widetilde{\alpha}_2=32.3204(\mp 9.0118),\ \widetilde{\lambda}_2=8.7521(\mp 3.1234),$$

$$\widetilde{\alpha}_3=10.8818(\mp 3.2567),\ \widetilde{\lambda}_3=3.4830(\mp 0.9117),$$

$$\widetilde{\alpha}_4=16.5540(\mp 4.9987),\ \widetilde{\lambda}_4=4.5939(\mp 1.7876),$$

$$\widetilde{\alpha}_5=33.7024(\mp 9.1165),\ \widetilde{\lambda}_5=11.0702(\mp 4.1113),$$

$$\widetilde{\alpha}_6=31.5931(\mp 8.5467),\widetilde{\lambda}_6=10.7384(\mp 4.0112),$$

$$\widetilde{\theta}=2.0912(\mp 0.7511).$$

We use these estimates as the initial estimates to compute the MLEs of the unknown parameters. The MLEs of the unknown parameters, and the

Table 10: Parameter estimates for marginal GE

Parameter	Dominant radius	Radius	Dominant humerus	Humerus	Dominant ulna	Ulna
α	49.9006	67.5898	22.7566	36.6185	70.4800	66.0690
λ	7.5595	8.7521	3.4830	4.5939	11.0702	10.7384
μ	0.2578	0.2762	0.7159	0.8185	0.2636	0.2483

associated 95% confidence intervals based on the Fisher information matrix presented within brackets are as follows:

$$\widehat{\alpha}_1 = 23.3272(\mp 6.1217),\ \widehat{\lambda}_1 = 7.6402(\mp 2.1521),$$

$$\widehat{\alpha}_2 = 31.3839(\mp 8.1176),\ \widehat{\lambda}_2 = 8.5622(\mp 2.3654),$$

$$\widehat{\alpha}_3 = 10.7247(\mp 2.7664),\ \widehat{\lambda}_3 = 3.2199(\mp 0.8675),$$

$$\widehat{\alpha}_4 = 15.3030(\mp 4.0145),\ \widehat{\lambda}_4 = 4.3723(\mp 1.1657),$$

$$\widehat{\alpha}_5 = 32.9878(\mp 8.2525),\ \widehat{\lambda}_5 = 10.7200(\mp 3.5641),$$

$$\widehat{\alpha}_6 = 30.9029(\mp 7.8876),\ \widehat{\lambda}_6 = 10.4922(\mp 3.3176),$$

$$\widehat{\theta} = 2.2257(\mp 0.7223).$$

Now the natural question is how good the model fits the data. Although, we have several satisfactory goodness of fit tests available for univariate data set, the same is not true in case of multivariate data set. Therefore, we test the marginals only. It is known that this is not sufficient, but necessary at least. We computed the Kolmogorov-Smirnov (KS) distances between the empirical marginals and the fitted marginals, and the associated p values (reported within brackets) for the six marginals to be; 0.1989 (0.2754), 0.1565 (0.5728), 0.1174 (0.8807), 0.1045 (0.9474), 0.1123 (0.9106) 0.0944 (0.9790), respectively. Since in all the cases the p values are quite high, it indicates that the proposed MVGE is indeed a good model for this BMD multivariate data set.

7 Conclusions

In this paper we have proposed a new multivariate absolute continuous distribution whose marginals are the GE distributions. The proposed model is a very flexible multivariate model, and it is observed that the proposed model can be obtained in three different ways. We have developed various properties of the model and discussed two different estimation procedures. It is observed that the MLEs are computationally quite difficult to compute but the two-stage estimators are very easy to implement in practice. Simulation results suggest that the performances of the two estimators are quite similar in nature mainly for large sample sizes. One multivariate bone mineral density data has been analyzed using this model, and it is observed that the proposed model provides a good fit to the above data set.

Acknowledgements. The authors would like to thank the associate editor and two referees for their constructive suggestions which have helped to improve the paper significantly. Part of the work of the first author has been supported by a grant from the Mathematical Sciences Initiative, IIT Kanpur.

Appendix A

In Step 1, $\mathbf{\Theta}^{(1)} = (\beta_1^{(1)}, \ldots, \beta_p^{(1)}, \lambda_1^{(1)}, \ldots, \lambda_p^{(1)})$ can be obtained by solving

$$\frac{n}{\lambda_j} - \sum_{i=1}^{n} x_{ij} + (\beta_j - 1)\sum_{i=1}^{n} \frac{x_{ij}e^{-\lambda_j x_{ij}}}{1 - e^{-\lambda_j x_{ij}}} = 0, \tag{7.1}$$

$$\frac{n}{\beta_j} + \sum_{i=1}^{n} \ln(1 - e^{-\lambda_j x_{ij}}) = 0. \tag{7.2}$$

First obtain $\lambda_j^{(1)}$ by solving

$$\frac{n}{\lambda_j} - \sum_{i=1}^{n} x_{ij} + (\beta_j^{(1)}(\lambda_j) - 1)\sum_{i=1}^{n} \frac{x_{ij}e^{-\lambda_j x_{ij}}}{1 - e^{-\lambda_j x_{ij}}} = 0,$$

where

$$\beta_j^{(1)}(\lambda_j) = -\frac{n}{\sum_{i=1}^{n} \ln(1 - e^{-\lambda_j x_{ij}})}$$

and finally obtain

$$\beta_j^{(1)} = -\frac{n}{\sum_{i=1}^{n} \ln(1 - e^{-\lambda_j^{(1)} x_{ij}})}$$

In Step 2 and Step 4, $\theta^{(k)}$ for $k \geq 1$, can be obtained by solving the following non-linear equation on θ:

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{p} \ln(u_{ij}(\beta_j^{(k)}, \lambda_j^{(k)})) &= \sum_{j=1}^{p} \frac{n\theta(j-1)}{\theta + j - 1} + \theta^2 \sum_{i=1}^{n} \ln(v_i(\mathbf{\Theta}_1^{(k)}, \theta)) \\ &+(\theta + p)\sum_{i=1}^{n} \frac{\sum_{j=1}^{p}\left(\left(u_{ij}\left(\beta_j^{(k)}, \lambda_j^{(k)}\right)\right)^{-1/\theta} \ln\left(u_{ij}\left(\beta_j^{(k)}, \lambda_j^{(k)}\right)\right)\right)}{v_i\left(\mathbf{\Theta}_1^{(k)}, \theta\right)},\end{aligned}$$

where

$$v_i(\mathbf{\Theta}_1^{(k)}, \theta) = \sum_{j=1}^{p} u_{ij}(\beta_j^{(k)}, \lambda_j^{(k)})^{-1/\theta} - (p-1)$$

Step 3 can be performed as follows. Let us define the following notations for $j = 1, \ldots, p$.

$$c_{1j}^{(k)} = \beta_j^{(k)}\left(\frac{1}{\theta^{(k)}} + 1\right)\sum_{i=1}^{n} \frac{x_{ij}e^{-\lambda_j^{(k)} x_{ij}}}{(1 - e^{-\lambda_j^{(k)} x_{ij}})}$$

$$
\begin{aligned}
c_{2j}^{(k)} &= \left(\frac{1}{\theta^{(k)}}+1\right)\sum_{i=1}^{n}\ln(1-e^{-\lambda_j^{(k)}x_{ij}}) \\
d_{1j} &= -\frac{\beta_j^{(k)}(\theta^{(k)}+p)}{\theta^{(k)}}\sum_{i=1}^{n}\frac{(1-e^{-\lambda_j^{(k)}x_{ij}})^{-1-\beta_j^{(k)}/\theta^{(k)}}x_{ij}e^{-\lambda_j^{(k)}x_{ij}}}{v_i(\mathbf{\Theta}_1^{(k)},\theta^{(k)})} \\
d_{2j} &= -\frac{(\theta^{(k)}+p)}{\theta^{(k)}}\sum_{i=1}^{n}\frac{\ln(1-e^{-\lambda_j^{(k)}x_{ij}})\times(1-e^{-\lambda_j^{(k)}x_{ij}})^{-\beta_j^{(k)}/\theta^{(k)}}}{v_i(\mathbf{\Theta}_1^{(k)},\theta^{(k)})}.
\end{aligned}
$$

Therefore, $\lambda_j^{(k+1)}$ can be obtained by solving the following non-linear equation on λ_j

$$
\frac{n}{\lambda_j}-\sum_{i=1}^{n}x_{ij}+(\beta_j^{(k+1)}(\lambda_j)-1)\sum_{i=1}^{n}\frac{x_{ij}e^{-\lambda_jx_{ij}}}{1-e^{-\lambda_jx_{ij}}}=c_{1j}+d_{1j}
$$

where

$$
\beta_j^{(k+1)}(\lambda_j)=-\frac{n}{\sum_{i=1}^{n}\ln(1-e^{-\lambda_jx_{ij}})-c_{1j}-d_{1j}}
$$

and finally obtain

$$
\beta_j^{(k+1)}=\beta_j^{(k+1)}(\lambda_j^{(k+1)})
$$

Appendix B

In this section we provide the elements of the observed Fisher information matrix. We will provide the following elements for $j,k=1,\ldots,p$, $j\neq k$.

$$
\frac{\partial^2 l}{\partial\beta_j^2},\frac{\partial^2 l}{\partial\lambda_j^2},\frac{\partial^2 l}{\partial\theta^2},\frac{\partial^2 l}{\partial\beta_j\partial\beta_k},\frac{\partial^2 l}{\partial\beta_j\partial\lambda_j},\frac{\partial^2 l}{\partial\beta_j\partial\lambda_k},\frac{\partial^2 l}{\partial\beta_j\partial\theta},\frac{\partial^2 l}{\partial\lambda_j\partial\theta}.
$$

$$
\begin{aligned}
\frac{\partial^2 l}{\partial\beta_j^2} &= -\frac{n}{\beta_j^2}+\frac{(\theta+p)}{\theta^2\beta_j^2}\sum_{i=1}^{n}\frac{[\ln(u_{ij}(\beta_j,\lambda_j)]^2\times[(u_{ij}(\beta_j,\lambda_j))]^{-1/\theta}}{v_i(\mathbf{\Theta}_1,\theta)} \\
&\quad\times\left\{1-\frac{[(u_{ij}(\beta_j,\lambda_j))]^{-1/\theta}}{v_i(\mathbf{\Theta}_1,\theta)}\right\} \\
\frac{\partial^2 l}{\partial\lambda_j^2} &= -\frac{n}{\lambda_j^2}+\left(\frac{\beta_j}{\theta}-1\right)\sum_{i=1}^{n}\frac{x_{ij}^2e^{-\lambda_jx_{ij}}}{(1-e^{-\lambda_jx_{ij}})^2} \\
&\quad+\frac{(\theta+p)\beta_j}{\theta}\sum_{i=1}^{n}\frac{A_{ij}(\mathbf{\Theta}_1,\theta)x_{ij}^2w_{ij}(\lambda_j)}{(v_i(\mathbf{\Theta}_1,\theta))^2}
\end{aligned}
$$

where

$$A_{ij}(\mathbf{\Theta}_1,\theta) = \frac{1}{(u_{ij}(\beta_j,\lambda_j))^{1/\theta}} \times \left[\left\{\left(\frac{\beta_j}{\theta}+1\right) w_{ij}(\lambda_j)+1\right\} v_i(\mathbf{\Theta}_1,\theta) + w_{ij}(\lambda_j)(u_{ij}(\beta_j,\lambda_j))^{-1/\theta}\right]$$

$$\frac{\partial^2 l}{\partial\theta^2} = \frac{np}{\theta^2} - n\sum_{j=1}^{p}\frac{1}{(\theta+j-1)^2} - \frac{2}{\theta^3}\sum_{i=1}^{n}\sum_{j=1}^{p}\ln u_{ij}(\beta_j,\lambda_j) - \sum_{i=1}^{n}\frac{2B_i(\mathbf{\Theta}_1,\theta)}{v_i(\mathbf{\Theta}_1,\theta)} -(\theta+p)\sum_{i=1}^{n} c_i(\mathbf{\Theta}_1,\theta)$$

where

$$\begin{aligned} B_i(\mathbf{\Theta}_1,\theta) &= \sum_{j=1}^{p} B_{ij}(\mathbf{\Theta}_1,\theta) \\ B_{ij}(\mathbf{\Theta}_1,\theta) &= \frac{1}{\theta^2}(u_{ij}(\beta_j,\lambda_j))^{-1/\theta}\ln(u_{ij}(\beta_j,\lambda_j)) \\ C_i(\mathbf{\Theta}_1,\theta) &= -\frac{2B_i(\mathbf{\Theta}_1,\theta)}{\theta v_i(\mathbf{\Theta}_1,\theta)} + \frac{1}{\theta^4}\sum_{j=1}^{p}[u_{ij}(\beta_j,\lambda_j)]^{-1/\theta}[\ln(u_{ij}(\beta_j,\lambda_j))]^2 \\ &\quad + \frac{B_i(\mathbf{\Theta}_1,\theta)}{[v_i(\mathbf{\Theta}_1,\theta)]^2}. \end{aligned}$$

For $1 \le j \ne k \le p$

$$\begin{aligned} \frac{\partial^2 l}{\partial\beta_j\partial\beta_k} &= -\frac{\theta+p}{\theta}\sum_{i=1}^{n}\frac{1}{(v_i(\mathbf{\Theta}_1,\theta))^2} \times [u_{ij}(\beta_j,\lambda_j)]^{-1/\theta}\ln(1-e^{-\lambda_j x_{ij}}) \\ &\quad \times [u_{ik}(\beta_k,\lambda_k)]^{-1/\theta}\ln(1-e^{-\lambda_k x_{ik}}). \\ \frac{\partial^2 l}{\partial\beta_j\partial\theta} &= \frac{1}{\theta^2}\sum_{i=1}^{n}\sum_{j=1}^{p}\ln(1-e^{-\lambda_j x_{ij}}) + p\sum_{i=1}^{n}\frac{B_i(\mathbf{\Theta}_1,\theta)}{v_i(\mathbf{\Theta}_1,\theta)} \\ &\quad +\theta(\theta+p)\sum_{i=1}^{n}\left(\frac{B_i(\mathbf{\Theta}_1,\theta)}{v_i(\mathbf{\Theta}_1,\theta)}\right)^2 \\ &\quad -\left(\frac{p}{\theta}+1\right)\sum_{i=1}^{n}\sum_{j=1}^{p}\frac{B_{ij}(\mathbf{\Theta}_1,\theta)\ln(1-e^{-\lambda_j x_{ij}})}{v_i(\mathbf{\Theta}_1,\theta)} \\ \frac{\partial^2 l}{\partial\beta_j\partial\lambda_j} &= -\frac{1}{\theta}\sum_{i=1}^{n}\frac{x_{ij}e^{-\lambda_j x_{ij}}}{1-e^{-\lambda_j x_{ij}}} \end{aligned}$$

$$-(\theta+p)\sum_{i=1}^{n} D_{ij}(\beta_j,\lambda_j,\theta)\left(1-\frac{\beta_j}{\theta}\ln(1-e^{-\lambda_j x_{ij}})\right)$$
$$-\beta_j^2(\theta+p)\sum_{i=1}^{n}(D_{ij}(\beta_j,\lambda_j,\theta))^2$$

$$\begin{aligned}\frac{\partial^2 l}{\partial\beta_k\partial\lambda_j} &= -\frac{\beta_j(\theta+p)}{\theta}\sum_{i=1}^{n} D_{ij}(\beta_j,\lambda_j)\times\frac{1}{v_i(\boldsymbol{\Theta}_1,\theta)}\\ &\quad\times(u_{ik}(\beta_k,\lambda_k))^{-1/\theta}\times\ln(1-e^{-\lambda_k x_{ik}}),\end{aligned}$$

where

$$D_{ij}(\beta_j,\lambda_j,\theta)=\frac{(1-e^{-\lambda_j x_{ij}})^{-(\beta_j/\theta+1)}x_{ij}e^{-\lambda_j x_{ij}}}{\theta v_i(\boldsymbol{\Theta}_1,\theta)}.$$

$$\begin{aligned}\frac{\partial^2 l}{\partial\theta\partial\lambda_j} &= \frac{1}{\theta^2}\sum_{i=1}^{n}\frac{\beta_j x_{ij}e^{-\lambda_j x_{ij}}}{1-e^{-\lambda_j x_{ij}}}\\ &\quad+\frac{\beta_j}{\theta}\sum_{i=1}^{n} D_{ij}(\beta_j,\lambda_j,\theta)\left(p-\frac{\beta_j(\theta+p)}{\theta}\ln(1-e^{-\lambda_j x_{ij}})\right)\\ &\quad-\frac{\beta_j(\theta+p)}{\theta^2}\sum_{i=1}^{n}\sum_{m=1}^{p} D_{ij}(\beta_j,\lambda_j,\theta)\times\frac{1}{v_i(\boldsymbol{\Theta}_1,\theta)}\\ &\quad\times[u_{im}(\beta_m,\lambda_m)]^{-1/\theta}\ln u_{im}(\beta_m,\lambda_m).\end{aligned}$$

Appendix C

In this section we provide the elements of the matrix $\boldsymbol{W}$, which is a $2p+1\times 2p+1$ positive definite matrix. In the calculation of the different elements of the matrix $\boldsymbol{W}$, we need to compute the expected Fisher information matrix also. They are not presented here, as these elements can be easily obtained from the expressions provided in Appendix B, and after performing required one or two dimensional integration numerically. We will be using the following notations. The expected $2p+1\times 2p+1$ Fisher information matrix is denoted by

$$\boldsymbol{I}=\begin{bmatrix}\boldsymbol{I}_{11} & \dots & \boldsymbol{I}_{1p} & \boldsymbol{I}_{1d}\\ \vdots & \ddots & \vdots & \vdots\\ \boldsymbol{I}_{p1} & \dots & \boldsymbol{I}_{pp} & \boldsymbol{I}_{pd}\\ \boldsymbol{I}_{d1} & \dots & \boldsymbol{I}_{dp} & I_{dd}\end{bmatrix}.$$

Here each $\boldsymbol{I}_{jk}$, for $1 \le j,k \le p$ is a 2×2 matrix, each $\boldsymbol{I}_{jd}$, for $1 \le j \le p$ is a 2×1, vector, $\boldsymbol{I}_{jd} = \boldsymbol{I}_{dj}^T$, and I_{dd} is a real number. From Joe (2005), we have

$$\boldsymbol{W} = \left(-\boldsymbol{U}^{-1}\right) \boldsymbol{M} \left(-\boldsymbol{U}^{-1}\right)^T.$$

Here

$$-\boldsymbol{U} = \begin{bmatrix} \boldsymbol{J}_{11} & \dots & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{0} & \dots & \boldsymbol{J}_{pp} & \boldsymbol{0} \\ \boldsymbol{I}_{d1} & \dots & \boldsymbol{I}_{dp} & I_{dd} \end{bmatrix}, \quad \text{and} \quad -\boldsymbol{M} = \begin{bmatrix} \boldsymbol{J}_{11} & \dots & \boldsymbol{J}_{1p} & \boldsymbol{0} \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{J}_{p1} & \dots & \boldsymbol{J}_{pp} & \boldsymbol{0} \\ \boldsymbol{0} & \dots & \boldsymbol{0} & I_{dd} \end{bmatrix}.$$

Now we provide the elements of each of the 2×2 matrix $\boldsymbol{J}_{jk}$, for $1 \le j,k \le p$. If we denote for $1 \le j \ne k \le p$,

$$\boldsymbol{J}_{jj} = \begin{bmatrix} a_{11}^j & a_{12}^j \\ a_{21}^j & a_{22}^j \end{bmatrix} \quad \text{and} \quad \boldsymbol{J}_{jk} = \begin{bmatrix} b_{11}^{jk} & b_{12}^{jk} \\ b_{21}^{jk} & b_{22}^{jk} \end{bmatrix},$$

then for $(X_j, X_k) \sim \text{MVGE}_2(\alpha_j, \alpha_k, \beta_j, \beta_k, \theta)$,

$$a_{11}^j = \frac{1}{\beta^2}, \quad a_{12}^j = a_{21}^j = -E\left[\frac{X_j e^{-\lambda_j X_j}}{1 - e^{-\lambda_j X_j}}\right],$$

$$a_{22}^j = \frac{1}{\lambda_j^2} + (\beta_j - 1)E\left[\frac{X_j^2 e^{-\lambda_j X_j}}{(1 - e^{-\lambda_j X_j})^2}\right],$$

$$\begin{aligned}
b_{11}^{jk} &= \operatorname{cov}\left\{\ln\left(1 - e^{-\lambda_j X_j}\right), \ln\left(1 - e^{-\lambda_k X_k}\right)\right\} \\
b_{12}^{jk} &= \operatorname{cov}\left\{\ln\left(1 - e^{-\lambda_j X_j}\right), (\beta_k - 1)\left(\frac{X_k e^{-\lambda_k X_k}}{(1 - e^{-\lambda_k X_k})}\right)\right\} \\
b_{21}^{jk} &= \operatorname{cov}\left\{(\beta_j - 1)\left(\frac{X_j e^{-\lambda_j X_j}}{(1 - e^{-\lambda_j X_j})}\right), \ln\left(1 - e^{-\lambda_k X_k}\right)\right\} \\
b_{22}^{jk} &= \operatorname{cov}\left\{(\beta_j - 1)\left(\frac{X_j e^{-\lambda_j X_j}}{(1 - e^{-\lambda_j X_j})}\right), (\beta_k - 1)\left(\frac{X_k e^{-\lambda_k X_k}}{(1 - e^{-\lambda_k X_k})}\right)\right\}.
\end{aligned}$$

References

BLOCK, H.W. and BASU, A.P. (1974). A continuous bivariate exponential extension. *Journal of the American Statistical Association* **69**, 1031–1037.

BOLAND, P.J., HOLLANDER, M., JOAG-DEV, K. and KOCHAR, S. (1996). Bivariate dependence properties of order statistics. *Journal of Multivariate Analysis* **56**, 75–89.

BURR, I.W. (1942). Cumulative frequency functions. *Annals of Mathematical Statistics* **13**, 215–232.

COLANGELO, A., SCARSINI, M. and SHAKED, M. (2005). Some notions of multivariate positive dependence. *Insurance Mathematical Economics* **37**, 13–26.

COLANGELO, A., HU, T. and SHAKED, M. (2008). Conditional orderings and positive dependence. *Journal of Multivariate Analysis* **99**, 358–371.

CROWDER, M. (1989). A multivariate distribution with Weibull connections. *Journal of the Royal Statistical Society, Series B* **51**, 93–107.

GUPTA, R.D. and KUNDU, D. (1999). Generalized exponential distribution. *Australian New Zealand Journal of Statistics* **41**, 173–188.

GUPTA, R.D. and KUNDU, D. (2007). Generalized exponential distribution; existing methods and some recent developments. *Journal of Statistical Planning and Inference* **137**, 3537–3547.

HALL, P. and WANG, J.Z. (2005). Bayesian likelihood methods for estimating the end point of a distribution. *Journal of the Royal Statistical Society, Series B* **67**, 717–729.

JOHNSON, R.A. and WICHERN, D.W. (1999). *Applied multivariate analysis*, 4-th edition. Prentice and Hall, New Jersey.

JOE, H. (1997). *Multivariate models and dependence concepts*. Chapman and Hall, London.

JOE, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis* **94**, 401–419.

KUNDU, D. and GUPTA, R.D. (2009). Bivariate generalized exponential distribution. *Journal of Multivariate Analysis* **100**, 581–593.

KUNDU, D. and GUPTA, R.D. (2011). Absolute continuous bivariate generalized exponential distribution. *Advances in Statistical Analysis* **95**, 169–185.

MARDIA, K.V. (1962). Multivariate Pareto distributions. *Annals of Mathematical Statistics* **33**, 1008–1015.

MARSHALL, A.W. and OLKIN, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association* **62**, 30–44.

NADARAJAH, S. (2011). The exponentiated exponential distribution: a survey. *Advances in Statistical Analysis* **95**, 219–251.

NELSEN, R.B. (2006). *An introduction to copula*, 2nd edition. Springer, New York.

RODRIGUEZ, R.N. (1977). A guide to Burr Type XII distributions. *Biometrika* **64**, 129–134.

SHOAEE, S. and KHORRAM, E. (2012). A new absolute continuous bivariate generalized exponential distribution. *Journal of Statistical Planning and Inference* **142**, 2203–2220.

SONG, P.X.-K., FAN, Y. and KALBFLEISCH, J.D. (2005). Maximization by parts in likelihood inference, (with discussions). *Journal of the American Statistical Association* **100**, 1145–1167.

SURLES, J.G. and PADGETT, W.J. (2005). Some properties of a scaled Burr type X distribution. *Journal of Statistical Planning and Inference* **128**, 271–280.

TAKAHASI, K. (1965). Note on the multivariate Burr's distribution. *Annals of the Institute of Statistical Mathematics* **17**, 257–260.

WINGO, D.R. (1993). Maximum likelihood methods for fitting the Burr Type XII distribution to multiply (progressively) censored lifetest data. *Metrika* **40**, 203–210.

Debasis Kundu
Ankush Kumar
Department of Mathematics
and Statistics, Indian Institute
of Technology, Kanpur
208016, India
E-mail: kundu@iitk.ac.in

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

Paper received: 21 April 2014.