An Absolute Continuous Bivariate Inverse Generalized Exponential Distribution: Properties, Inference and Extensions

Debasis Kundu *

Abstract

The aim of this paper is to introduce an absolutely continuous bivariate inverse generalized exponential (BIGE) distribution. The proposed distribution has been obtained by removing the singular component from the BIGE distribution similarly as the Block and Basu absolute continuous bivariate exponential distribution. This distribution has four parameters, and due to this the joint probability density function can take variety of shapes. This distribution can be used quite effectively if there are no ties in the bivariate data set and particularly if the marginals are from a heavy tailed distribution. We have developed different properties of this distribution and provide classical inference of the unknown parameters. The maximum likelihood (ML) estimators cannot be obtained in closed form and one needs to solve a four dimensional optimization problem to compute ML estimators in this case. To avoid that we propose to use expectation maximization (EM) algorithm to compute the ML estimators of the unknown parameters. The analysis of one data set has been performed to see the effectiveness of the proposed algorithm and we extend the results to the multivariate case also. Finally we conclude the paper with several open problems for future research.

Key Words and Phrases: Marshall-Olkin bivariate exponential distribution; Block and Basu bivariate distributions;maximum likelihood estimators; EM algorithm; competing risks.

AMS Subject Classifications: 62F10, 62F03, 62H12.

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

1 Introduction

Two-parameter generalized exponential (GE) has received a considerable amount of attention in the last two decades. It has been introduced by Gupta and Kundu [9] as a special case of the three-parameter exponentiated Weibull distribution originally proposed by Mudholkar and Srivastava [14]. It can also be obtained as a special case of the three-parameter generalized Gompertz-Verhulst family of distributions introduced by Ahuja and Nash [1], see also Verhulst [20] in this respect.

The two-parameter GE distribution has the following cumulative distribution function (CDF), probability density function (PDF) and hazard function (HF) for $x > 0$, $\alpha > 0$, $\lambda > 0$:

$$F_{GE}(x;\alpha,\lambda) = (1-e^{-\lambda x})^{\alpha} \quad (1)$$

$$f_{GE}(x;\alpha,\lambda) = \alpha\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\alpha-1}, \quad (2)$$

$$h_{GE}(x;\alpha,\lambda) = \frac{\alpha\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\alpha-1}}{1-(1-e^{-\lambda x})^{\alpha}}. \quad (3)$$

Here, α is the shape parameter and λ is the scale parameter. The PDF (2) of a GE distribution can be either a decreasing or an unimodal function depending on the values of α. If $\alpha \leq 1$, the PDF is a decreasing function, otherwise it becomes an unimodal function. When $\alpha = 1$, it coincides with the one-parameter exponential distribution. The hazard function of a GE distribution (3) can be increasing ($\alpha > 1$), decreasing ($\alpha < 1$) or constant ($\alpha = 1$). It is observed by Gupta and Kundu [9] that the GE distribution behaves very similarly as the two-parameter gamma distribution. But because of the explicit expression of the CDF it can be used very effectively for the censored data. Due to this reason, an extensive amount of work has been done in establishing different properties and also developing various inferential procedures of the unknown parameters of this model. A book length treatment can be found in Al-Hussaini and Ahsanullah [2], see also the review article by Nadarajah [16] and the references cited therein.

Although, the GE distribution is a very flexible distribution, it cannot have a non-monotone hazard function or it cannot be heavy tailed. Due to this reason, Oguntunde and Adejumo [17] introduced the inverted GE distribution (IGE) similar to the inverted Weibull distribution, see for example Murthy, Xie and Jiang [15]. The IGE distribution has the following survival function (SF), PDF and HF for $x > 0$, $\alpha > 0$ and $\lambda > 0$:

$$S_{IGE}(x;\alpha,\lambda) = (1 - e^{-\frac{\lambda}{x}})^{\alpha} \quad (4)$$

$$f_{IGE}(x;\alpha,\lambda) = \frac{\alpha\lambda}{x^2} e^{-\frac{\lambda}{x}} (1 - e^{-\frac{\lambda}{x}})^{\alpha-1} \quad (5)$$

$$h_{IGE}(x;\alpha,\lambda) = \frac{\alpha\lambda}{x^2(e^{-\frac{\lambda}{x}} - 1)}. \quad (6)$$

The PDF and HF of the IGE distribution are always unimodal for all values of $\alpha > 0$ and $\lambda > 0$. When $\lambda = 1$, the mode of an IGE is at 1/2 if $\alpha = 1$, for $\alpha > 1$, the mode is less than 1/2, and for $\alpha \leq 1$, the mode is greater than 1/2. Moreover, depending on the values of α it is a heavy tailed distribution. If $\alpha \leq 1$, the mean does not exist. If $1 < \alpha \leq 2$, the mean exists, but the variance does not exist. For $\alpha > 2$, the variance exist. The PDFs and hazard functions of an IGE distribution are always unimodal. From now on an absolutely continuous random variable with PDF (5) will be denoted by IGE(α, λ).

Recently, Alqallaf and Kundu [3] introduced bivariate IGE distribution similar to the Marshall-Olkin bivariate exponential (MOBE) distribution, see for example Marshall and Olkin [13], as follows: Suppose U_1 follows ($\sim$) IGE(α_1, λ), $U_2 \sim$ IGE(α_2, λ) and $U_0 \sim$ IGE(α_0, λ), and they are independently distributed. If $X = \min\{U_1, U_0\}$ and $Y = \min\{U_2, U_0\}$, then (X, Y) is said to have bivariate IGE (BIGE) distribution with parameters α_1, α_2, α_0 and λ. We denote this by BIGE$(\alpha_1, \alpha_2, \alpha_0, \lambda)$. It may be mentioned that this BIGE distribution has the same interpretation as the shock model similar to the MOBE model. Here the shock appears following an IGE distribution, which can be heavy tailed.

The BIGE introduced by Alqallaf and Kundu [3] has a singular component along $X = Y$, i.e. $P(X = Y) > 0$, similar to the MOBE distribution. Therefore, if there are no ties in the data, it may not be reasonable to use BIGE distribution in this case. It may be recalled that since MOBE has a singular component, it is not used when there are no ties in the data. Due to this reason, Block and Basu [5] introduced an absolutely continuous bivariate exponential distribution, from now on we call it as the Block and Basu bivariate exponential (BBBE) distribution, by removing the singular component from the MOBE distribution. Although, MOBE is a singular distribution, the BBBE distribution enjoys all the properties of an absolutely continuous distribution. It can be used quite effectively to analyze a bivariate data set, when there are no ties in the data.

The main aim of this paper is to introduce absolutely continuous BIGE (ABIGE) by removing the singular component of a BIGE distribution. Clearly, ABIGE is an absolutely continuous distribution with four parameters. We study different properties of the ABIGE distribution and its marginals. Due to presence of the four parameters, the joint PDF of an ABIGE can take variety of shapes. Moreover, the marginals of an ABIGE can be heavy tailed also. Hence, the ABIGE model can be used quite effectively for a bivariate data set when there are no ties. The ML estimators of the unknown parameters of a ABIGE cannot be obtained in explicit forms. They have to be obtained by solving four non-linear equations simultaneously. Therefore, some numerical algorithms like Newton-Raphson or Gauss-Newton method may be used to solve these non-linear equations. Hence, very accurate initial values are needed to start the iterative process, otherwise it may not converge or it may converge to some local optimum. To avoid that we have proposed a EM algorithm to compute the ML estimators. The proposed EM algorithm requires solving only one dimensional optimization problem at each 'E'-step of the EM algorithm. Hence, the implementation of the proposed EM algorithm is quite simple in practice. The analysis of one data set has been performed to see the performances of the proposed EM algorithm and the effectiveness of the model. Finally we have introduced the absolute continuous multivariate IGE (AMIGE) and showed how the EM algorithm can be developed for the multivariate model. We have indicated several open problems for further research.

The rest of the paper is organized as follows. In Section 2 we have defined the ABIGE model and discuss its different properties. The inference procedure has been developed in Section 3. In Section 4 we provide the analysis of one bivariate data set. In Section 5 we have discussed AMIGE distribution, and finally we conclude the paper and provide several open problems for future work, in Section 6.

2 Model Description and Properties

2.1 Model Description

We have already introduced BIGE in Section 1. If $(X, Y) \sim \text{BIGE}(\alpha_1, \alpha_2, \alpha_0, \lambda)$, then the joint SF of X and Y for $x > 0$ and $y > 0$ becomes

$$S_{X,Y}(x,y) = P(X > x, Y > y) = \begin{cases} (1 - e^{-\frac{\lambda}{x}})^{\alpha_1}(1 - e^{-\frac{\lambda}{y}})^{\alpha_2+\alpha_0} & \text{if } \; 0 < x < y < \infty \\ (1 - e^{-\frac{\lambda}{x}})^{\alpha_1+\alpha_0}(1 - e^{-\frac{\lambda}{y}})^{\alpha_2} & \text{if } \; 0 < y < x < \infty \\ (1 - e^{-\frac{\lambda}{x}})^{\alpha_1+\alpha_2+\alpha_0} & \text{if } \; 0 < x = y < \infty. \end{cases}$$

The joint SF of X and Y has the following unique decomposition:

$$S_{X,Y}(x,y) = \frac{\alpha_1 + \alpha_2}{\alpha_1 + \alpha_2 + \alpha_0} S_{ac}(x,y) + \frac{\alpha_0}{\alpha_1 + \alpha_2 + \alpha_0} S_{si}(x,y)$$

Here,

$$S_{si}(x,y) = \begin{cases} (1 - e^{-\frac{\lambda}{x}})^{\alpha_1+\alpha_2+\alpha_0} & \text{if } \; x = y, \\ 0 & \text{if } \; x \neq y, \end{cases}$$

and

$$S_{ac}(x,y) = \frac{\alpha_1 + \alpha_2 + \alpha_0}{\alpha_1 + \alpha_2}(1 - e^{-\frac{\lambda}{x}})^{\alpha_1}(1 - e^{-\frac{\lambda}{y}})^{\alpha_2}(1 - e^{-\frac{\lambda}{z}})^{\alpha_0} - \frac{\alpha_0}{\alpha_1 + \alpha_2}(1 - e^{-\frac{\lambda}{x}})^{\alpha_1+\alpha_2+\alpha_0},$$

here $z = \max\{x, y\}$. Note that here $S_{ac}(x, y)$ is the absolute continuous part and $S_{si}(x, y)$ is the singular part.

From the joint SF, the joint PDF of X and Y can be obtained as follows:

$$f_{X,Y}(x,y) = \frac{\alpha_1+\alpha_2}{\alpha_1+\alpha_2+\alpha_0} f_{ac}(x,y) + \frac{\alpha_0}{\alpha_1+\alpha_2+\alpha_0} f_{si}(u), \tag{7}$$

where

$$f_{ac}(x,y) = \frac{\alpha_1+\alpha_2+\alpha_0}{\alpha_1+\alpha_2} \times \begin{cases} f_{IGE}(x;\alpha_1,\lambda) f_{IGE}(y;\alpha_2+\alpha_0,\lambda) & \text{if } \; 0<x<y<\infty \\ f_{IGE}(x;\alpha_1+\alpha_0,\lambda) f_{IGE}(y;\alpha_2,\lambda) & \text{if } \; 0<y<x<\infty, \end{cases}$$

$u = x = y$, and

$$f_{si}(u) = f_{IGE}(u;\alpha_1+\alpha_2+\alpha_0,\lambda).$$

It should be mentioned that when we write the joint PDF of X and Y as in (7), it is understood that $f_{ac}(x,y)$ is a PDF with respect to two dimensional Lebesgue measure and $f_{si}(u)$ is a PDF with respect to one dimensional Lebesgue measure, see for example Bemis, Bain and Higgins [4].

Now we define ABIGE by removing the singular component from the BIGE, similar to the construction of BBBE from MOBE distribution, as follows: If the joint PDF of the random variables U and V is

$$f_{U,V}(u,v) = \frac{\alpha_1+\alpha_2+\alpha_0}{\alpha_1+\alpha_2} \times \begin{cases} f_{IGE}(u;\alpha_1,\lambda) f_{IGE}(v;\alpha_2+\alpha_0,\lambda) & \text{if } \; 0<u<v<\infty \\ f_{IGE}(u;\alpha_1+\alpha_0,\lambda) f_{IGE}(v;\alpha_2,\lambda) & \text{if } \; 0<v<u<\infty, \end{cases}$$

then (U,V) is said to have ABIGE distribution with parameters α_1, α_2, α_0 and λ and it will be denoted by ABIGE$(\alpha_1,\alpha_2,\alpha_0,\lambda)$.

The joint survival function of (U,V) becomes

$$S_{U,V}(u,v) = \frac{\alpha_1+\alpha_2+\alpha_0}{\alpha_1+\alpha_2}(1-e^{-\frac{\lambda}{u}})^{\alpha_1}(1-e^{-\frac{\lambda}{v}})^{\alpha_2}(1-e^{-\frac{\lambda}{w}})^{\alpha_0} - \frac{\alpha_0}{\alpha_1+\alpha_2}(1-e^{-\frac{\lambda}{u}})^{\alpha_1+\alpha_2+\alpha_0},$$

here $w = \max\{u,v\}$. The marginal survival functions of U and V become:

$$\begin{aligned} S_U(u) &= \frac{\alpha_1+\alpha_2+\alpha_0}{\alpha_1+\alpha_2}\left(1-e^{-\frac{\lambda}{u}}\right)^{\alpha_1+\alpha_0} - \frac{\alpha_0}{\alpha_1+\alpha_2}\left(1-e^{-\frac{\lambda}{u}}\right)^{\alpha_1+\alpha_2+\alpha_0} \\ S_V(v) &= \frac{\alpha_1+\alpha_2+\alpha_0}{\alpha_1+\alpha_2}\left(1-e^{-\frac{\lambda}{v}}\right)^{\alpha_2+\alpha_0} - \frac{\alpha_0}{\alpha_1+\alpha_2}\left(1-e^{-\frac{\lambda}{v}}\right)^{\alpha_1+\alpha_2+\alpha_0}, \end{aligned}$$

respectively. The marginal PDFs of U and V become

$$\begin{aligned} f_U(u) &= \frac{\alpha_1+\alpha_2+\alpha_0}{\alpha_1+\alpha_2} f_{IGE}(u;\alpha_1+\alpha_0,\lambda) - \frac{\alpha_0}{\alpha_1+\alpha_2} f_{IGE}(u;\alpha_1+\alpha_2+\alpha_0,\lambda) \\ f_V(v) &= \frac{\alpha_1+\alpha_2+\alpha_0}{\alpha_1+\alpha_2} f_{IGE}(v;\alpha_2+\alpha_0,\lambda) - \frac{\alpha_0}{\alpha_1+\alpha_2} f_{IGE}(v;\alpha_1+\alpha_2+\alpha_0,\lambda), \end{aligned}$$

respectively. It may be observed that the relation between the BIGE and ABIGE is the following.

$$(U,V) = (X,Y)|\{X \neq Y\}. \tag{8}$$

The above relation (8) can be used quite effectively to generate ABIGE distribution. The following algorithm can be used to generate $(U,V) \sim \text{ABIGE}(\alpha_1,\alpha_2,\alpha_0,\lambda)$ as follows:

Step 1: Generate $U_1 \sim \text{IGE}(\alpha_1,\lambda)$, $U_2 \sim \text{IGE}(\alpha_2,\lambda)$ and $U_0 \sim \text{IGE}(\alpha_0,\lambda)$.

Step 2: If $U_0 < \min\{U_1,U_2\}$, go to Step 1.

Step 3: $U = \min\{U_1,U_0\}$ and $V = \min\{U_2,U_0\}$.

2.2 Properties

The following result provides the shape of the joint PDF of ABIGE.

Theorem 2.1: Let $(U,V) \sim \text{ABIGE}(\alpha_1,\alpha_2,\alpha_0,\lambda)$. We use the following notations. $S_0 = \{(u,v); 0 < u = v < \infty\}$, $S_1 = \{(u,v); 0 < u < v < \infty\}$, $S_2 = \{(u,v); 0 < v < u < \infty\}$.

(a) If $\alpha_1 = \alpha_2 = \alpha$, then $f_{U,V}(u,v)$ is continuous on $S_0 \cup S_1 \cup S_2 = \mathbb{R}^2$, $f_{U,V}(u,v)$ is unimodal and the mode is at $(x_0,x_0) \in S_0$, where x_0 is the unique solution of the non-linear equation

$$2(e^{\frac{1}{x}}-1)(1-2x) = (2\alpha+\alpha_0-2). \tag{9}$$

(b) If $\alpha_2+\alpha_0 < 1 < \alpha_1$, then $f_{U,V}(u,v)$ is continuous on $S_1 \cup S_2$, $f_{U,V}(u,v)$ is unimodal and the mode is at $(x_1,x_2) \in S_1$, where x_1 and x_2 are unique solutions of the non-linear equations

$$(e^{\frac{1}{x}}-1)(1-2x) = (\alpha_1-1) \tag{10}$$

$$(e^{\frac{1}{x}} - 1)(1 - 2x) = (\alpha_2 + \alpha_0 - 1). \tag{11}$$

(c) If $\alpha_1 + \alpha_0 < 1 < \alpha_2$, then $f_{U,V}(u, v)$ is continuous on $S_1 \cup S_2$, $f_{U,V}(u, v)$ is unimodal and the mode is at $(x_1, x_2) \in S_2$, where x_1 and x_2 are unique solutions of the non-linear equations

$$\begin{aligned}(e^{\frac{1}{x}} - 1)(1 - 2x) &= (\alpha_2 - 1) \qquad (12)\\ (e^{\frac{1}{x}} - 1)(1 - 2x) &= (\alpha_1 + \alpha_0 - 1). \qquad (13)\end{aligned}$$

Proof: See in the Appendix.

In Figure 1, we have provided the surface plots of the of the joint PDF of ABIGE$(\alpha_1, \alpha_2, \alpha_0, \lambda)$ for different values of α_1, α_2 and α_0 keeping $\lambda = 1$. It is observed that for all values of α_1, α_2 and α_0, the joint PDF is an unimodal function. The following results are useful for data analysis purposes, they might have some independent interest also.

3 Inference

In this section we derive the ML estimators of the unknown parameters of a ABIGE$(\alpha_1, \alpha_2, \alpha_0, \lambda)$ based on a random sample of size n and it is as follows:

$$Data = \{(u_1, v_1), \ldots, (u_n, v_n)\}. \tag{14}$$

We use the following notations $I_1 = \{i : u_i < v_i\}$, $I_2 = \{i : u_i > v_i\}$ and $\mathbf{\Theta} = (\alpha_1, \alpha_2, \alpha_0, \lambda)^\top$. Moreover, $n_1 = |I_1|$ = number of elements in I_1 and similarly, $n_2 = |I_2|$. Based on the sample (14), the log-likelihood function can be written as

$$\begin{aligned} l(\mathbf{\Theta}|Data) &= n\ln(\alpha_1 + \alpha_2 + \alpha_0) - n\ln(\alpha_1 + \alpha_2) \\ &\quad + \sum_{i \in I_1} \ln f_{IGE}(u_i; \alpha_1, \lambda) + \sum_{i \in I_1} \ln f_{IGE}(v_i; \alpha_2 + \alpha_0, \lambda) \end{aligned}$$

$$
\begin{aligned}
& + \sum_{i \in I_2} \ln f_{IGE}(u_i; \alpha_1 + \alpha_0, \lambda) + \sum_{i \in I_2} \ln f_{IGE}(v_i; \alpha_2, \lambda) \\
= \;& C + n \ln(\alpha_1 + \alpha_2 + \alpha_0) - n \ln(\alpha_1 + \alpha_2) - A\lambda + 2n \ln \lambda + \\
& n_1(\ln \alpha_1 + \ln(\alpha_2 + \alpha_0)) + n_2(\ln(\alpha_1 + \alpha_0) + \ln \alpha_2) + \\
& (\alpha_1 - 1) \sum_{i \in I_1} \ln(1 - e^{-\frac{\lambda}{u_i}}) + (\alpha_2 + \alpha_0 - 1) \sum_{i \in I_1} \ln(1 - e^{-\frac{\lambda}{v_i}}) \\
& + (\alpha_1 + \alpha_0 - 1) \sum_{i \in I_2} \ln(1 - e^{-\frac{\lambda}{u_i}}) + (\alpha_2 - 1) \sum_{i \in I_2} \ln(1 - e^{-\frac{\lambda}{v_i}}). \qquad (15)
\end{aligned}
$$

Here C is a constant does not depend on the parameters and $A = \sum_{i=1}^{n} (u_i^{-1} + v_i^{-1})$. The ML estimate of $\boldsymbol{\Theta}$ can be obtained by maximizing (15) with respect to the unknown parameters. It is immediate that it cannot be obtained in explicit form. It has to be obtained by solving a four dimensional optimization problem. Therefore, one needs to use some iterative algorithms like Newton-Raphson or Gauss-Newton method to compute the ML estimates. Any iterative algorithm needs a very good initial guesses, which may not be a trivial issue in four dimension. Moreover, it may converge to a local maximum rather than the global maximum.

To avoid that we treat this problem as a missing value problem, and we will show that if we have the complete data set, then the ML estimates of $\boldsymbol{\Theta}$ can be obtained by solving only one non-linear equation. Hence, we propose to use a very simple EM algorithm, where at each 'E'-step the corresponding 'M'-step can be performed by solving only one non-linear equation, and that is the main motivation of the proposed EM algorithm.

Suppose instead of observing only (U, V), we also observe the indicator vector (Δ_1, Δ_2) associate with the corresponding U_1, U_2, U_0 defined before, as follows:

Case 1: $(U < V) : \Delta_1 = 1$ and

$$
\Delta_2 = \begin{cases} 2 & \text{if} \quad V = U_2 \\ 3 & \text{if} \quad V = U_0. \end{cases}
$$

Case 2: $(V < U) : \Delta_2 = 2$ and

$$\Delta_1 = \begin{cases} 1 & \text{if} \quad U = U_1 \\ 3 & \text{if} \quad U = U_0. \end{cases}$$

Now we will provide the log-likelihood contribution of a typical data point $(u, v, \delta_1, \delta_2)$ for different cases.

Case 1: $u < v$, $\delta_1 = 1$ and $\delta_2 = 2$. The log-likelihood contribution becomes:

$$\ln f_{IGE}(u; \alpha_1, \lambda) + \ln f_{IGE}(v; \alpha_2, \lambda) + \ln S_{IGE}(v; \alpha_0, \lambda).$$

Case 2: $u < v$, $\delta_1 = 1$ and $\delta_2 = 3$. The log-likelihood contribution becomes:

$$\ln f_{IGE}(u; \alpha_1, \lambda) + \ln f_{IGE}(v; \alpha_0, \lambda) + \ln S_{IGE}(v; \alpha_2, \lambda).$$

Case 3: $v < u$, $\delta_1 = 1$ and $\delta_2 = 2$. The log-likelihood contribution becomes:

$$\ln f_{IGE}(u; \alpha_1, \lambda) + \ln f_{IGE}(v; \alpha_2, \lambda) + \ln S_{IGE}(u; \alpha_0, \lambda).$$

Case 4: $u < v$, $\delta_1 = 3$ and $\delta_2 = 2$. The log-likelihood contribution becomes:

$$\ln f_{IGE}(u; \alpha_0, \lambda) + \ln f_{IGE}(v; \alpha_2, \lambda) + \ln S_{IGE}(u; \alpha_1, \lambda).$$

In this case it can be easily shown that for the complete data set namely $\{(u_i, v_i, \delta_{1i}, \delta_{2i}); i = 1, \ldots, n\}$, the ML estimates of α_1, α_2 and α_0 can be obtained in explicit forms if λ is known. Hence, the ML estimate of λ can be obtained by maximizing the profile log-likelihood function in one dimension only. The following Table 1 will be useful for further development of the EM algorithm.

Now following the idea of Dinse [7] for each incomplete data (u, v), we form 'pseudo observations' by fractioning (u, v) to two partially complete 'pseudo observation' of the form $\{(u, v, w_1), (u, v, 1 - w_1)\}$ and $\{(u, v, w_2), (u, v, 1 - w_2)\}$ depending on whether $u > v$ or

Table 1: Possible configuration of U_1, U_2, U_0 and the associated probabilities.

Set	Relation Between U and V	Possible Configuration of U, V, W	Observed Variable	Conditional Probability	(Δ_1, Δ_2)
I_1	$U < V$	$U_1 < U_2 < U_0$	$U = U_1,\ V = U_2$	$\frac{\alpha_2}{(\alpha_2 + \alpha_0)}$	(1,2)
		$U_1 < U_0 < U_2$	$U = U_1,\ V = U_0$	$\frac{\alpha_0}{(\alpha_2 + \alpha_0)}$	(1,3)
I_2	$V < U$	$U_2 < U_1 < U_0$	$U = U_1,\ V = U_2$	$\frac{\alpha_1}{(\alpha_1 + \alpha_0)}$	(1,2)
		$U_2 < U_0 < U_1$	$U = U_0,\ V = U_2$	$\frac{\alpha_0}{(\alpha_1 + \alpha_0)}$	(3,2)

$u < v$, respectively. Here w_1 is the conditional probability that $U_1 < U_0$ given that $V < U$. Similarly, w_2 is the conditional probability that $U_2 < U_0$, given that $U < V$. From Table 1

$$w_1 = \frac{\alpha_1}{\alpha_1 + \alpha_0} \quad \text{and} \quad w_2 = \frac{\alpha_2}{\alpha_2 + \alpha_0}.$$

Now let us denote by $\Theta^{(k)} = (\alpha_1^{(k)}, \alpha_2^{(k)}, \alpha_0^{(k)}, \lambda^{(k)})^\top$ as the estimates of the parameters at the k-th stage of the EM algorithm. Similarly, let us denote $w_1^{(k)}$ and $w_2^{(k)}$ as the estimates of w_1 and w_2, respectively, at the k-st stage. At the $k - th$ stage the 'pseudo-log-likelihood' function can be written as follows:

$$\begin{aligned}
l(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(k)}) &= w_2^{(k)} \sum_{i \in I_1} (\ln f_{IGE}(u_i; \alpha_1, \lambda) + \ln f_{IGE}(v_i; \alpha_2, \lambda) + \ln S_{IGE}(v_i; \alpha_0, \lambda)) + \\
&\quad (1 - w_2^{(k)}) \sum_{i \in I_1} (\ln f_{IGE}(u_i; \alpha_1, \lambda) + \ln f_{IGE}(v_i; \alpha_0, \lambda) + \ln S_{IGE}(v_i; \alpha_2, \lambda)) + \\
&\quad w_1^{(k)} \sum_{i \in I_2} (\ln f_{IGE}(u_i; \alpha_1, \lambda) + \ln f_{IGE}(v_i; \alpha_2, \lambda) + \ln S_{IGE}(u_i; \alpha_0, \lambda)) + \\
&\quad (1 - w_1^{(k)}) \sum_{i \in I_2} (\ln f_{IGE}(u_i; \alpha_0, \lambda) + \ln f_{IGE}(v_i; \alpha_2, \lambda) + \ln S_{IGE}(u_i; \alpha_1, \lambda)) \\
&= \sum_{i \in I_1} \ln f_{IGE}(u_i; \alpha_1, \lambda) + \sum_{i \in I_2} \left\{ w_1^{(k)} \ln f_{IGE}(u_i; \alpha_1, \lambda) + (1 - w_1^{(k)}) \ln S_{IGE}(u_i; \alpha_1, \lambda) \right\} + \\
&\quad \sum_{i \in I_2} \ln f_{IGE}(v_i; \alpha_2, \lambda) + \sum_{i \in I_1} \left\{ w_2^{(k)} \ln f_{IGE}(v_i; \alpha_2, \lambda) + (1 - w_2^{(k)}) \ln S_{IGE}(v_i; \alpha_2, \lambda) \right\} +
\end{aligned}$$

$$
\begin{aligned}
& w_2^{(k)} \sum_{i \in I_1} \ln S_{IGE}(v_i; \alpha_0, \lambda) + (1 - w_2^{(k)}) \sum_{i \in I_1} \ln f_{IGE}(v_i; \alpha_0, \lambda) + \\
& w_1^{(k)} \sum_{i \in I_2} \ln S_{IGE}(u_i; \alpha_0, \lambda) + (1 - w_1^{(k)}) \sum_{i \in I_2} \ln f_{IGE}(u_i; \alpha_0, \lambda). \qquad (16)
\end{aligned}
$$

Therefore, $\mathbf{\Theta}^{(k+1)}$ can be obtained from $\mathbf{\Theta}^{(k)}$ by maximizing (16) with respect to $\mathbf{\Theta}$. For a fixed λ

$$
\begin{aligned}
\alpha_1^{(k+1)}(\lambda) &= -\frac{n_1 + w_1^{(k)} n_2}{\sum_{i \in I} \ln(1 - e^{-\frac{\lambda}{u_i}})} \\
\alpha_2^{(k+1)}(\lambda) &= -\frac{n_2 + w_2^{(k)} n_1}{\sum_{i \in I} \ln(1 - e^{-\frac{\lambda}{v_i}})} \\
\alpha_0^{(k+1)}(\lambda) &= -\frac{n_1(1 - w_2^{(k)}) + n_2(1 - w_1^{(k)})}{\sum_{i \in I_1} \ln(1 - e^{-\frac{\lambda}{v_i}}) + \sum_{i \in I_2} \ln(1 - e^{-\frac{\lambda}{u_i}})},
\end{aligned}
$$

maximize (16). Hence, $\lambda^{(k+1)}$ which maximizes (16) can be obtained by maximizing the profile log-likelihood function, i.e. $\lambda^{(k+1)} = \text{argmax } g(\lambda)$, where

$$
\begin{aligned}
g(\lambda) &= 2(n_1 + n_2) \ln \lambda + (n_1 + w_1^{(k)} n_2) \ln \alpha_1^{(k+1)}(\lambda) + (n_2 + w_2^{(k)} n_1) \ln \alpha_2^{(k+1)}(\lambda) + \\
& \quad (n_1(1 - w_2^{(k)}) + n_2(1 - w_1^{(k)})) \ln \alpha_0^{(k+1)}(\lambda) - \lambda \left(\sum_{i \in I_1 \cup I_2} \frac{1}{u_i} + \frac{1}{v_i} \right) \\
& \quad - \sum_{i \in I_1 \cup I_2} \left(\ln(1 - e^{-\frac{\lambda}{u_i}}) + \ln(1 - e^{-\frac{\lambda}{v_i}}) \right).
\end{aligned}
$$

Therefore

$$
\alpha_1^{(k+1)} = \alpha_1^{(k+1)}(\lambda^{(k+1)}), \quad \alpha_2^{(k+1)} = \alpha_2^{(k+1)}(\lambda^{(k+1)}), \quad \alpha_0^{(k+1)} = \alpha_0^{(k+1)}(\lambda^{(k+1)}).
$$

Once the ML estimates of α_1, α_2, α_3 and λ are obtained the associated confidence intervals can be obtained from the observed Fisher information matrix as suggested by Louis [12]. In the Appendix we have provided the observed Fisher information matrix.

4 Data Analysis

In this section we provide the analysis of a real data set to show how the proposed EM algorithm can be implemented in practice. This data set is obtained from Johnson and Wichern

[10] and it represents the cholesterol level of 23 adults at two different times. It is presented below for easy reference: (317,275), (186,190), (377,368), (229,282), (276,306), (272,250), (219,236), (260,264), (284,241), (365,294), (298,341), (274,262), (232,244), (367,358), (253,247), (230,245), (190,212), (290,291), (337,383), (283,277), (325,288), (266,253), (338,307).

Before progressing further we have subtracted 165 and divided by 10 for each data point, mainly for computational purposes. It is not going to affect the inference procedure. One natural question arises whether ABIGE can be used to analyze this data set or not. First of all there are no ties in the data set. We have fitted the IGE to the minimum of the two cholesterol levels, and the ML estimates of the shape and scale parameters are 4.7691 and 17.5489, respectively. The Kolmogorov-Smirnov distance between the fitted and empirical distribution is 0.1451 and the associated p value is 0.7178. Therefore, it is clear that IGE distribution fits the minimum quite well. Hence, we have fitted ABIGE distribution to the above bivariate data set.

We have used the EM algorithm to compute the ML estimates of the unknown parameters. The following initial values have been used: $\alpha_1 = \alpha_2 = \alpha_0 = 1$ and $\lambda = 17.0$. We have started the EM algorithm with these initial values and the EM algorithm stops when the relative difference between the two consecutive log-likelihood values is less than 10^{-6}. The iteration stops after 20 steps. The ML estimates and the associated 95% confidence intervals based on the observed Fisher information matrix are provided below:

$$\widehat{\alpha}_1 = 3.2683(\mp 0.9876), \quad \widehat{\alpha}_2 = 3.5433(\mp 1.0145), \quad \widehat{\alpha}_0 = 2.4060(\mp 0.7655)$$

$$\widehat{\lambda} = 19.6523(\mp 4.1276).$$

We have tried the above EM algorithm with different initial guesses, it converges to the same point, although the number of iterations are different. Another natural question is whether the EM algorithm converges to the global maximum or not. To verify that we have tried to find the maximum of the log-likelihood function using the grid search method. We have taken the range of α_0, α_1 and α_2 as $(0, 5)$ and the range of λ as $(0, 30)$ with grid size 0.0001 for each parameter. It gives the maximum at the same point although it took more than three hours to execute, whereas in the same machine the EM algorithm converges in few seconds.

5 Absolute Continuous Multivariate IGE Distribution

In this section we define absolute continuous multivariate IGE (AMIGE) distribution along the same way as the multivariate Block and Basu absolutely continuous exponential distribution, see for example Pradhan and Kundu [18]. The basic idea is the same. First we define the multivariate IGE distribution as follows. Suppose $U_0, \ldots, U_p$ are independent IGE distributions, and $U_i \sim \text{IGE}(\alpha_i, \lambda)$, for $i = 0, \ldots, p$. Now define $X_i = \min\{U_i, U_0\}$, for $i = 0, \ldots, p$. Then $(X_1, \ldots, X_p)^\top$ is called the MIGE distribution with parameters $\alpha_0, \ldots, \alpha_p, \lambda$ and it is denoted by $\text{MIGE}(\alpha_1, \ldots, \alpha_p, \alpha_0, \lambda)$. Now, an AMIGE distribution can be constructed from a MIGE distribution by removing the singular components. We give the formal definition of a AMIGE distribution.

Definition: A random vector $(Y_1, \ldots, Y_p)^\top$ is said to have a p-variate AMIGE distribution with parameters $\alpha_0, \alpha_1, \ldots, \alpha_p$ and λ, if the joint PDF of $(Y_1, \ldots, Y_p)^\top$ is of the form:

$$f_{Y_1,\ldots,Y_p}(y_1, \ldots, y_p) = c f_{IGE}(y_{i_1}; \alpha_{i_1}, \lambda) \times \ldots \times f_{IGE}(y_{i_{p-1}}; \alpha_{i_{p-1}}, \lambda) \times f_{IGE}(y_{i_p}; \alpha_{i_p} + \alpha_0, \lambda), \quad (17)$$

here c is the normalizing constant and $\{i_1, \ldots, i_p\}$ is a permutation of $\{1, \ldots, p\}$, where $y_{i_1} < \ldots < y_{i_p}$. From now on it will be denoted by $\text{AMIGE}(\alpha_1, \ldots, \alpha_p, \alpha_0, \lambda)$.

The normalizing constant c is such that

$$\int_{\mathbb{R}^p} f_{Y_1,\ldots,Y_p}(y_1, \ldots, y_p) dy_1 \ldots dy_p = 1.$$

It can be seen (see Appendix C) from the simple multiple integration that

$$c^{-1} = \sum_{\mathcal{P}} \frac{\alpha_{i_1}}{\alpha_{i_1} + \ldots + \alpha_{i_p} + \alpha_0} \times \ldots \times \frac{\alpha_{i_{p-1}}}{\alpha_{i_{p-1}} + \alpha_{i_p} + \alpha_0}. \tag{18}$$

Here $\mathcal{P}$ denotes the set of all permutations of $\{1, \ldots, p\}$. Note that when $p = 2$, $c = \frac{\alpha_1 + \alpha_2 + \alpha_0}{\alpha_1 + \alpha_2}$. The relation between a MIGE and AMIGE can be described as follows:

$$(Y_1, \ldots, Y_p)^\top = (X_1, \ldots, X_p)^\top | \{X_{i_1} \neq X_{i_j}, 1 \leq i_1, i_j \leq p\}. \tag{19}$$

The above relation (19) can be easily used to generate random samples from a AMIGE distribution. The following results can be easily obtained, the proofs are avoided.

Theorem 5.1: Let $(Y_1, \ldots, Y_p)^\top \sim \text{AMIGE}(\alpha_1, \ldots, \alpha_p, \alpha_0, \lambda)$.

(a) If $q < p$, then $(Y_1, \ldots, Y_q)^\top \sim \text{AMIGE}(\alpha_1, \ldots, \alpha_q, \alpha_0, \lambda)$.

(b) If $\alpha_1 = \ldots = \alpha_p = 1$, then $f_{Y_1,\ldots,Y_p}(y_1, \ldots, y_p)$ is continuous on $\mathbb{R}^p$, $f_{Y_1,\ldots,Y_p}(y_1, \ldots, y_p)$ is unimodal and the mode is at $(x_0, \ldots, x_0)$, where x_0 is the unique solution of the non-linear equation

$$n(e^{\frac{1}{x}} - 1)(1 - 2x) = n(\alpha - 1) + \alpha_0.$$

(c) $Z = \min\{Y_1, \ldots, Y_p\} \sim \text{IGE}(\alpha_1 + \ldots + \alpha_p + \alpha_0, \lambda)$.

(d) $P(Y_i < Y_j) = \dfrac{\alpha_i}{\alpha_i + \alpha_j}$.

(e) $Y_i | \{Y_i < Y_j\} \sim \text{IGE}(\alpha_i + \alpha_j + \alpha_0, \lambda)$.

Proof: See Appendix D.

Now we consider the estimation of the unknown parameters based on a random sample from a AMIGE distribution. For notational simplicity we illustrate the procedure for $p = 3$, although the result can be easily obtained for a general p also. It is assumed that we have a random sample of size n from AMIGE$(\alpha_1, \alpha_2, \alpha_3, \alpha_0, \lambda)$ as follows:

$$Data = \{(y_{1i}, y_{2i}, y_{3i}); i = 1, \ldots, n\}. \tag{20}$$

In this case also we use the same notations as before, i.e. $\boldsymbol{\Theta} = (\alpha_1, \alpha_2, \alpha_3, \alpha_0, \lambda)^\top$. It is clear that the MLE of $\boldsymbol{\Theta}$ cannot be obtained in explicit forms, and we use the EM algorithm as before. We use the following notations $I_{jkm} = \{i : y_{1j} < y_{2k} < y_{3m}\}$, here $\{jkm\}$ belongs to the class of all permutations of $\{1,2,3\}$. Moreover,

$$w_1 = \frac{\alpha_1}{\alpha_1 + \alpha_0}, \quad w_2 = \frac{\alpha_2}{\alpha_2 + \alpha_0}, \quad w_3 = \frac{\alpha_3}{\alpha_3 + \alpha_0}.$$

Let us denote $n_1 = |I_{123}| + |I_{213}| + |I_{312}| + |I_{132}|$, $n_2 = |I_{123}| + |I_{213}| + |I_{321}| + |I_{231}|$, $n_3 = |I_{312}| + |I_{321}| + |I_{231}| + |I_{132}|$. The 'pseudo-log-likelihood' function at the k-th stage of the EM algorithm can be written as follows:

$$\begin{aligned} l(\boldsymbol{\Theta}|\boldsymbol{\Theta}_k) &= \sum_{i \in I_{123} \cup I_{132} \cup I_{213} \cup I_{312}} \ln f_{IGE}(y_{1i}; \alpha_1, \lambda) + \\ & \sum_{i \in I_{231} \cup I_{321}} \left\{ w_1^{(k)} \ln f_{IGE}(y_{1i}; \alpha_1, \lambda) + (1 - w_1^{(k)}) \ln S_{IGE}(y_{1i}; \alpha_1, \lambda) \right\} + \\ & \sum_{i \in I_{123} \cup I_{213} \cup I_{321} \cup I_{231}} \ln f_{IGE}(y_{2i}; \alpha_2, \lambda) + \\ & \sum_{i \in I_{132} \cup I_{312}} \left\{ w_2^{(k)} \ln f_{IGE}(y_{2i}; \alpha_2, \lambda) + (1 - w_2^{(k)}) \ln S_{IGE}(y_{2i}; \alpha_2, \lambda) \right\} + \\ & \sum_{i \in I_{312} \cup I_{132} \cup I_{321} \cup I_{231}} \ln f_{IGE}(y_{3i}; \alpha_3, \lambda) + \\ & \sum_{i \in I_{123} \cup I_{213}} \left\{ w_3^{(k)} \ln f_{IGE}(y_{3i}; \alpha_3, \lambda) + (1 - w_3^{(k)}) \ln S_{IGE}(y_{3i}; \alpha_3, \lambda) \right\} + \\ & \sum_{i \in I_{231} \cup I_{321}} \left\{ w_1^{(k)} \ln S_{IGE}(y_{1i}; \alpha_0, \lambda) + (1 - w_1^{(k)}) \ln f_{IGE}(y_{1i}; \alpha_0, \lambda) \right\} + \\ & \sum_{i \in I_{132} \cup I_{312}} \left\{ w_2^{(k)} \ln S_{IGE}(y_{2i}; \alpha_0, \lambda) + (1 - w_2^{(k)}) \ln f_{IGE}(y_{2i}; \alpha_0, \lambda) \right\} + \\ & \sum_{i \in I_{123} \cup I_{213}} \left\{ w_3^{(k)} \ln S_{IGE}(y_{3i}; \alpha_0, \lambda) + (1 - w_3^{(k)}) \ln f_{IGE}(y_{3i}; \alpha_0, \lambda) \right\}. \end{aligned} \quad (21)$$

It is clear that for a given λ

$$\begin{aligned} \widehat{\alpha}_1^{(k+1)}(\lambda) &= \frac{n_1 + w_1^{(k)}(n - n_1)}{\sum_{i \in I} \ln(1 - e^{-\frac{\lambda}{y_{1i}}})}, \\ \widehat{\alpha}_2^{(k+1)}(\lambda) &= \frac{n_2 + w_2^{(k)}(n - n_2)}{\sum_{i \in I} \ln(1 - e^{-\frac{\lambda}{y_{2i}}})}, \end{aligned}$$

$$\widehat{\alpha}_3^{(k+1)}(\lambda) = \frac{n_3 + w_3^{(k)}(n-n_3)}{\sum_{i\in I}\ln(1-e^{-\frac{\lambda}{y_{3i}}})},$$

$$\widehat{\alpha}_0^{(k+1)}(\lambda) = \frac{(n-n_1)(1-w_1^{(k)}) + (n-n_2)(1-w_2^{(k)}) + (n-n_3)(1-w_3^{(k)})}{\sum_{i\in I}\ln(1-e^{-\frac{\lambda}{y_{1i}}}) + \sum_{i\in I}\ln(1-e^{-\frac{\lambda}{y_{2i}}}) + \sum_{i\in I}\ln(1-e^{-\frac{\lambda}{y_{3i}}})},$$

maximize (21). Moreover, $\lambda^{(k+1)}$ can be obtained by maximizing the profile ''pseudo-log-likelihood' function of λ. Hence, 'M'-step can be performed by solving only one one-dimensional optimization problem.

6 Conclusions and Some Open Problems

In this paper we have introduced a new absolutely continuous bivariate distribution by removing the singular component of the singular bivariate inverse generalized exponential distribution. The marginals of the present bivariate distribution can be heavy tailed and have non-monotone hazard function. Due to presence of four parameters, the proposed distribution is very flexible. We have developed an EM algorithm which can be used very conveniently to compute the ML estimators of the unknown parameters. The method has been extended to the multivariate case also.

As it has been mentioned before that GE distribution can be obtained as a special case of the generalized Gompertz-Verhulst (GGV) distribution as introduced by Ahuja and Nash [1]. The CDF of a GGV distribution with parameters $p > 0$, $\sigma > 0$ and $\theta > 0$ takes the following form:

$$F_{GGV}(x;\theta,\sigma,p) = \begin{cases} 0 & \text{if} \quad x \le \sigma \ln p \\ \left(1 - pe^{-\frac{x}{\sigma}}\right)^{\theta} & \text{if} \quad x > \sigma \ln p. \end{cases} \tag{22}$$

Therefore, if we reparameterize as $\lambda = \sigma^{-1}$ and $\mu = \sigma \ln p$, then (22) can be written as

$$F_{GGV}(x;\theta,\lambda,\mu) = \begin{cases} 0 & \text{if} \quad x \le \mu \\ \left(1 - e^{-\lambda(x-\mu)}\right)^{\theta} & \text{if} \quad x > \mu. \end{cases} \tag{23}$$

Hence, the GGV distribution is same as the three-parameter (location shift) GE distribution. Three-parameter IGE can be analogously defined from the three-parameter GE distribution. Similarly, five-parameter ABIGE can be defined with one location, one scale and three shape parameters and all the properties also remain same. If the common location parameter is known then the proposed EM algorithm can be used to compute the MLEs of the unknown parameters. All the results can be easily generalized to the multivariate case also. But if the location parameter is unknown, then it is no more a regular family and MLEs may not exist always. It will be interesting to develop proper inference procedure in this case. More work is needed in this direction.

Recently, Feizjavadian and Hashemi [8], Cai, Shi and Liu [6] and Shen and Xu [19] developed dependent competing risks model based Marshall-Olkin bivariate Weibull distribution. The main assumption in developing the model based on Marshall-Olkin bivariate Weibull distribution is that an experimental unit can fail at a particular time due to two competing causes simultaneously. But it may not be true in many cases. It seems this distribution can be used to develop dependent competing risks data when there are no ties on the cause of failure. The detailed inference procedure needs to be developed. The work is in progress, it will be reported later.

Acknowledgements:

The author would like to thank two unknown reviewers and Professor Barry C. Arnold for providing some constructive suggestions which had helped to improve the manuscript significantly. Part of the work has been funded by a grant from the Science and Engineering Research Board, Government of India, Grant no. SERB MTR/2018/000179.

Appendix A: Proofs

Proof of Theorem 2.1:

(a) It is clear that $f_{U,V}(u,v)$ is continuous in $S_1 \cup S_2$. Since $f_{U,V}(x,x) = \lim_{u,v\to x} f_{U,V}(u,v)$, it follows that $f_{U,V}(u,v)$ is continuous in $S_0 \cup S_1 \cup S_2$. Since for all $0 < u, v < \infty$,

$$f_{U,V}(0,0) = f_{U,V}(\infty,\infty) = f_{U,V}(u,0) = f_{U,V}(u,\infty) = f_{U,V}(0,v) = f_{U,V}(\infty,v) = 0,$$

that $f_{U,V}(u,v)$ has a local maximum. It can be easily checked by taking derivatives of $\ln f_{U,V}(u,v)$ that $f_{U,V}(u,v)$ does not have any critical point in the region $S_1 \cup S_2$, hence $f_{U,V}(u,v)$ does not have any critical point in the region $S_1 \cup S_2$, hence it does not have any local maximum in $S_1 \cup S_2$. Therefore, in this case the local maximum will be at S_0. By taking derivative with respect to x of $\ln f_{U,V}(x,x)$ and equating it to zero, we can get one needs to solve the equation (9). It can be easily seen that the left hand side of (9) is a decreasing function of x, and it decreases from ∞ to -4. Hence, it has a unique solution.
(b) Note that since $\alpha_1 > 1$ and $\alpha_2+\alpha_0 < 1$, it can be easily seen by taking partial derivatives of $\ln f_{U,V}(u,v)$ that $f_{U,V}(u,v)$ has a critical point at (x_1,x_2), where x_1 and x_2 are solutions of the non-linear equations (10) and (11), respectively. Clearly $x_1 < 1/2$, since $\alpha_1 > 1$ and $x_2 < 1/2$, since $\alpha_2 + \alpha_0 < 1$. Hence, $(x_1,x_2) \in S_1$. Uniqueness follows using the same argument as in (a). It can be easily checked that $f_{U,V}(u,v)$ does not have a critical point in S_2.
(c) Follows similarly as in (b). ∎

Appendix B: Observed Fisher Information Matrix

Using the same notation as Louis [12], the observed Fisher information matrix can be written as

$$\boldsymbol{F}_{obs} = \boldsymbol{B} - \boldsymbol{S}\boldsymbol{S}^\top,$$

here $\boldsymbol{B}$ is the negative of the second derivative of the log-likelihood function and $\boldsymbol{S}$ is the derivative vector. We provide the elements of the matrix $\boldsymbol{B}$ and the vector $\boldsymbol{S}$. We will use the following notation for brevity.

$$a_{11} = \sum_{i \in I_1} \frac{1}{u_i^2(1 - e^{-\frac{\widehat{\lambda}}{u_i}})^2}, \quad a_{12} = \sum_{i \in I_2} \frac{1}{u_i^2(1 - e^{-\frac{\widehat{\lambda}}{u_i}})^2},$$

$$a_{22} = \sum_{i \in I_2} \frac{1}{v_i^2(1 - e^{-\frac{\widehat{\lambda}}{v_i}})^2}, \quad a_{21} = \sum_{i \in I_1} \frac{1}{v_i^2(1 - e^{-\frac{\widehat{\lambda}}{v_i}})^2}$$

$$b_{11} = \sum_{i \in I_1} \frac{1}{u_i(1 - e^{-\frac{\widehat{\lambda}}{u_i}})}, \quad b_{12} = \sum_{i \in I_2} \frac{1}{u_i(1 - e^{-\frac{\widehat{\lambda}}{u_i}})},$$

$$b_{22} = \sum_{i \in I_2} \frac{1}{v_i(1 - e^{-\frac{\widehat{\lambda}}{v_i}})}, \quad b_{21} = \sum_{i \in I_1} \frac{1}{v_i(1 - e^{-\frac{\widehat{\lambda}}{v_i}})}$$

$$c_{11} = \sum_{i \in I_1} \frac{1}{u_i}, \quad c_{12} = \sum_{i \in I_2} \frac{1}{u_i}, c_{22} = \sum_{i \in I_2} \frac{1}{v_i}, \quad c_{21} = \sum_{i \in I_1} \frac{1}{v_i},$$

$$d_{11} = \sum_{i \in I_1} \ln(1 - e^{-\frac{\lambda}{u_i}}), d_{12} = \sum_{i \in I_2} \ln(1 - e^{-\frac{\lambda}{u_i}})$$

$$d_{22} = \sum_{i \in I_2} \ln(1 - e^{-\frac{\lambda}{v_i}}), d_{21} = \sum_{i \in I_1} \ln(1 - e^{-\frac{\lambda}{v_i}}).$$

If the (i,j)-th element of the matrix $\boldsymbol{B}$ is $B(i,j)$, then $B(i,j) = B(j,i)$, for $1 \le i,j \le 4$, and for $1 \le i \le j \le 4$,

$$B(1,1) = \frac{n_1 + n_2 w_2}{\widehat{\alpha}_1^2}, \quad B(2,2) = \frac{n_2 + n_1 w_1}{\widehat{\alpha}_2^2}, \quad B(3,3) = \frac{n_1(1 - w_1) + n_2(1 - w_2)}{\widehat{\alpha}_0^2},$$

$$B(4,4) = \frac{2}{\widehat{\lambda}^2} + a_{11}[\widehat{\alpha}_1 - 1] + a_{22}[\widehat{\alpha}_2 - 1] + a_{12}[\widehat{\alpha}_1 + \widehat{\alpha}_0 - 1] + a_{21}[\widehat{\alpha}_2 + \widehat{\alpha}_0 - 1]$$

$$B(1,4) = -\frac{1}{\widehat{\lambda}^2} - (c_{11} + b_{11}), B(2,4) = -\frac{1}{\widehat{\lambda}^2} - (c_{22} + b_{22}), B(3,4) = b_{12} + b_{21},$$

$$B(1,3) = B(2,3) = B(1,2) = 0.$$

If $\boldsymbol{S} = (S(1), S(2), S(3), S(4))^\top$, then

$$S(1) = \frac{n_1 + w_2 n_2}{\widehat{\alpha}_1} + (d_{11} + d_{12}), S(2) = \frac{n_2 + w_1 n_1}{\widehat{\alpha}_2} + (d_{22} + d_{21}), S(3) = d_{12} + d_{21},$$

$$S(4) = \widehat{\lambda}(c_{11}+c_{12}+c_{21}+c_{22}) - \frac{2n}{\widehat{\lambda}} - \widehat{\alpha}_1(b_{11}+b_{12}) + \widehat{\alpha}_2(b_{22}+b_{21}) + \widehat{\alpha}_0(b_{21}+b_{12}) + (b_{11}+b_{12}+b_{21}+b_{22}).$$

Appendix C: Normalizing Constant c

In this section we show that the normalizing constant c satisfies (18). We will show the result for $p = 3$, the general result easily follows from there. If $(Y_1, Y_2, Y_3)^\top$ follows a AMIGE with parameters α_0, α_1, α_2, α_3 and λ, then

$$f_{Y_1,Y_2,Y_3}(y_1,y_2,y_3) = c \begin{cases} f_{IGE}(y_1;\alpha_1,\lambda) f_{IGE}(y_2;\alpha_2,\lambda) f_{IGE}(y_3;\alpha_0+\alpha_3,\lambda) & \text{if} \quad y_1 < y_2 < y_3 \\ f_{IGE}(y_1;\alpha_1,\lambda) f_{IGE}(y_3;\alpha_3,\lambda) f_{IGE}(y_2;\alpha_0+\alpha_2,\lambda) & \text{if} \quad y_1 < y_3 < y_2 \\ f_{IGE}(y_2;\alpha_2,\lambda) f_{IGE}(y_1;\alpha_1,\lambda) f_{IGE}(y_3;\alpha_0+\alpha_3,\lambda) & \text{if} \quad y_2 < y_1 < y_3 \\ f_{IGE}(y_2;\alpha_2,\lambda) f_{IGE}(y_3;\alpha_3,\lambda) f_{IGE}(y_1;\alpha_0+\alpha_1,\lambda) & \text{if} \quad y_2 < y_3 < y_1 \\ f_{IGE}(y_3;\alpha_3,\lambda) f_{IGE}(y_1;\alpha_1,\lambda) f_{IGE}(y_2;\alpha_0+\alpha_2,\lambda) & \text{if} \quad y_3 < y_1 < y_2 \\ f_{IGE}(y_3;\alpha_3,\lambda) f_{IGE}(y_2;\alpha_2,\lambda) f_{IGE}(y_1;\alpha_0+\alpha_1,\lambda) & \text{if} \quad y_3 < y_2 < y_1. \end{cases}$$

Now note that

$$\begin{aligned} &\int_0^\infty \int_{y_1}^\infty \int_{y_2}^\infty f_{IGE}(y_1;\alpha_1,\lambda) f_{IGE}(y_2;\alpha_2,\lambda) f_{IGE}(y_3;\alpha_0+\alpha_3,\lambda) dy_3 dy_2 dy_1 = \\ &\int_0^\infty \int_{y_1}^\infty f_{IGE}(y_1;\alpha_1,\lambda) f_{IGE}(y_2;\alpha_2,\lambda) S_{IGE}(y_2;\alpha_0+\alpha_3,\lambda) dy_3 dy_2 = \\ &\frac{\alpha_2}{\alpha_2+\alpha_3+\alpha_0} \int_0^\infty \int_{y_1}^\infty f_{IGE}(y_1;\alpha_1,\lambda) S_{IGE}(y_1;\alpha_2+\alpha_3+\alpha_0,\lambda) = \\ &\frac{\alpha_1}{\alpha_1+\alpha_2+\alpha_3+\alpha_0} \times \frac{\alpha_2}{\alpha_2+\alpha_3+\alpha_0}. \end{aligned}$$

Similarly, the other integrations also can be obtained. Hence

$$\begin{aligned} c^{-1} \quad = \quad & \frac{\alpha_1}{\alpha_1+\alpha_2+\alpha_3+\alpha_0} \times \frac{\alpha_2}{\alpha_2+\alpha_3+\alpha_0} + \frac{\alpha_1}{\alpha_1+\alpha_2+\alpha_3+\alpha_0} \times \frac{\alpha_3}{\alpha_2+\alpha_3+\alpha_0} + \\ & \frac{\alpha_2}{\alpha_1+\alpha_2+\alpha_3+\alpha_0} \times \frac{\alpha_1}{\alpha_1+\alpha_3+\alpha_0} + \frac{\alpha_2}{\alpha_1+\alpha_2+\alpha_3+\alpha_0} \times \frac{\alpha_3}{\alpha_1+\alpha_3+\alpha_0} + \\ & \frac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3+\alpha_0} \times \frac{\alpha_1}{\alpha_1+\alpha_2+\alpha_0} + \frac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3+\alpha_0} \times \frac{\alpha_2}{\alpha_1+\alpha_2+\alpha_0}. \end{aligned}$$

Appendix D: Proof of Theorem 5.1

(a) Follows from the definition.

(b) Proof follows along the same way as the proof of Part (a) of Theorem 2.1.

(c)

$$\begin{aligned} P(Z > z) &= P(U_1 > z, \ldots, U_p > z, U_0 > z) \\ &= S_{IGE}(z; \alpha_1, \lambda) \times \ldots \times S_{IGE}(z; \alpha_p, \lambda) S_{IGE}(z; \alpha_0, \lambda) \\ &= S_{IGE}(z; \alpha_1 + \ldots + \alpha_p + \alpha_0, \lambda). \end{aligned}$$

(d) Observe that $(Y_i, Y_j) \sim$ ABIGE$(\alpha_i, \alpha_j, \lambda)$. Hence,

$$\begin{aligned} P(Y_i < Y_j) &= \frac{\alpha_i + \alpha_j + \alpha_0}{\alpha_i + \alpha_j} \int_0^\infty \int_u^\infty f_{IGE}(u; \alpha_i, \lambda) f_{IGE}(v; \alpha_j + \alpha_0, \lambda) dv du \\ &= \frac{\alpha_i + \alpha_j + \alpha_0}{\alpha_i + \alpha_j} \int_0^\infty f_{IGE}(u; \alpha_i, \lambda) S_{IGE}(u; \alpha_j + \alpha_0, \lambda) du \\ &= \frac{\alpha_i}{\alpha_i + \alpha_j}. \end{aligned}$$

(e) Observe that $(Y_i, Y_j) \sim$ ABIGE$(\alpha_i, \alpha_j, \lambda)$. Hence,

$$\begin{aligned} P(Y_i > a | Y_i < Y_j) &= \frac{P(a < Y_i < Y_j)}{P(Y_i < Y_j)} \\ &= \frac{\alpha_i + \alpha_j + \alpha_0}{\alpha_i} \int_a^\infty \int_u^\infty f_{IGE}(u; \alpha_i, \lambda) f_{IGE}(v, \alpha_j + \alpha_0, \lambda) dv du \\ &= \frac{\alpha_i + \alpha_j + \alpha_0}{\alpha_i} \int_a^\infty f_{IGE}(u; \alpha_i, \lambda) S_{IGE}(u, \alpha_j + \alpha_0, \lambda) du \\ &= S_{IGE}(a; \alpha_i + \alpha_j + \alpha_0, \lambda). \end{aligned}$$

References

[1] Ahuja, J. C. and Nash, S. W. (1967), "The generalized Gompertz-Verhulst family of distributions", *Sankhyā Ser. A*, vol. 29, 141–156.

[2] Al-Hussaini, E.K. and Ahsanullah, M. (2015), *Exponentiated Distributions*, Atlantis Press, Paris, France.

[3] Alqallaf, F.A. and Kundu, D. (2020), "A Bivariate Inverse Generalized Exponential Distribution and its Applications in Dependent Competing Risks Model", submitted for publication.

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

[5] Block, H. and Basu, A.P. (1974), "A continuous bivariate exponential extension", *Journal of the American Statistical Association*, vol. 69, 1031–1037.

[6] Cai, J., Shi, Y., Liu, B. (2017), "Analysis of incomplete data in the presence of dependent competing risks from Marshall-Olkin bivariate Weibull distribution under progressive hybrid censoring", *Communications in Statistics - Theory and Methods*, vol. 46, 6497 – 6511.

[7] Dinse, G. E. (1982). Non-parametric estimation of partially incomplete time and types of failure data.Biometrics38, 417–431

[8] Feizjavadian, S.H. and Hashemi, R. (2015), "Analysis of dependent competing risks in presence of progressive hybrid censoring using Marshall-Olkin bivariate Weibull distribution", *Computational Statistics and Data Analysis*, vol. 82, 19 – 34.

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

[10] Johnson, R.A. and Wichern, D.W. (1999), *Applied Multivariate Analysis*, 4th edition, Prentice-Hall, New Jersey.

[11] Kundu, D. and Gupta, R.D. (2009), "Bivariate generalized exponential distribution", *Journal of Multivariate Analysis*, vol. 100, 581 – 593.

[12] Louis, T. A. (1982), "Finding the observed information matrix when using the EM algorithm", *Journal of the Royal Statistical Society*, Series B 44, 226 – 233.

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

[14] Mudholkar, G.S. and Srivastava, D.K. (1993), "Exponentiated Weibull family for analyzing bathtub failure data", *IEEE Transactions on Reliability*, vol. 42, 299–302.

[15] Murthy, D.N.P., Xie, M. and Jiang, R. (2004), *Weibull Models*, Wiley, New-York.

[16] Nadarajah, S. (2011), "The exponentiated exponential distribution; a survey", *Advances in Statistical Analysis*, vol. 95, 219 - 251

[17] Oguntunde, P.E. and Adejumo, A.O. (2015), "The generalized inverted generalized exponential distribution with an application to a censored Data", *Journal of Statistics Applications and Probability*, vol. 4, 223 - 230.

[18] Pradhan, B. and Kundu, D. (2016), "Bayes estimation for the Block and Basu bivariate and multivariate Weibull distributions", *Journal of Statistical Computation and Simulation*, vol. 86, no. 1, 170 – 182.

[19] Shen, Y. and Xu, A. (2018), "On the dependent competing risks using Marshall-Olkin bivariate Weibull model: parameter estimation with different methods", *Communications in Statistics - Theory and Methods*, vol. 47, 5558 – 5572.

[20] Verhulst,P.F. (1945) Recherches mathematique sur la loi d'-accroissement do la population. Nouvelles Memoires de l'Academie Royale des Sciences et Belles-Lettres de Bruxelles [i.e. Mémoires, Series 2], vol. 18, 38 pp.

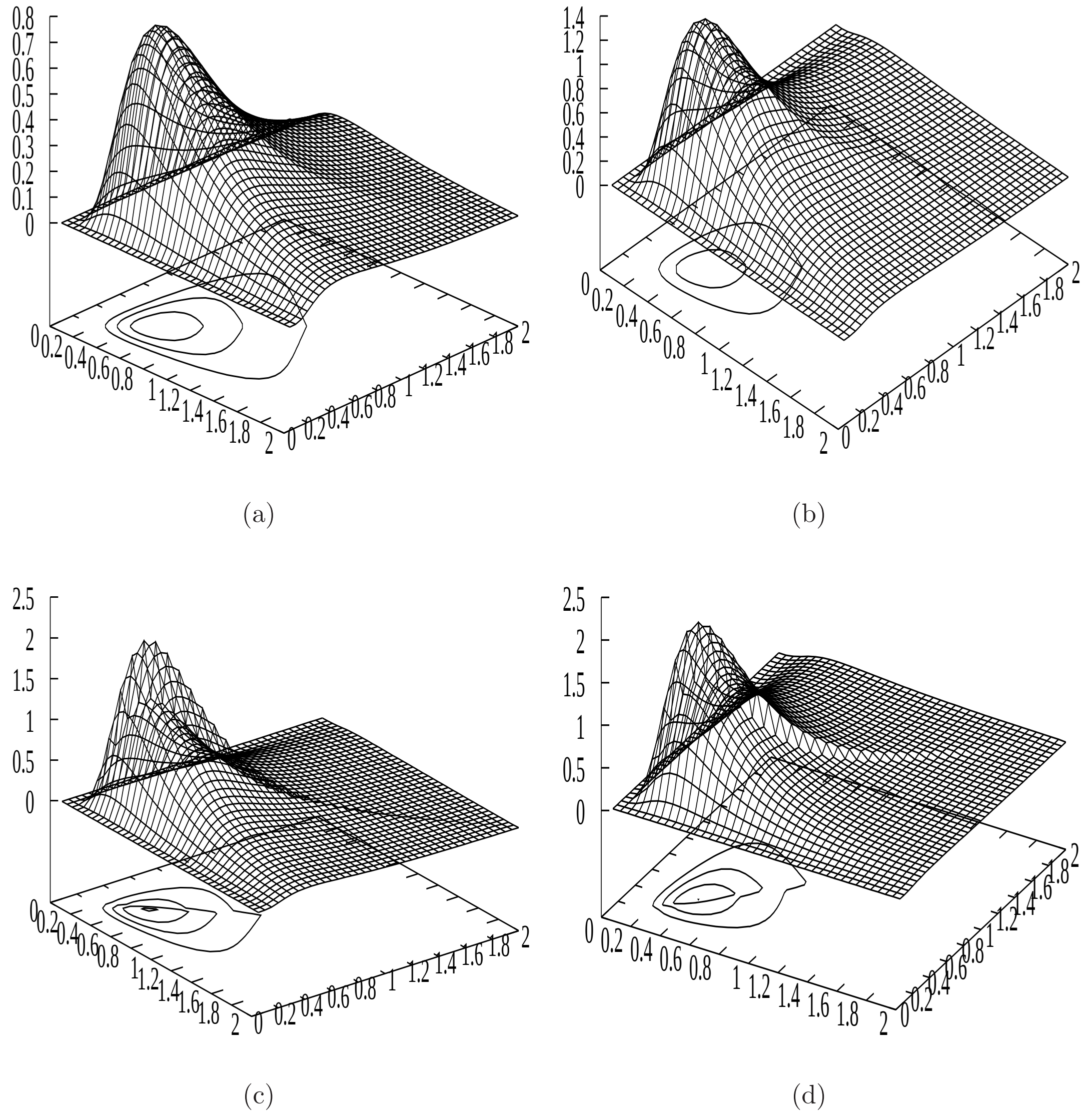

Figure 1: Bivariate surface plots of BIGE$(\alpha_1, \alpha_2, \alpha_0, \lambda)$ distribution for different α_1, α_2, α_0, λ values: (a) (1.0,1.0,1.0,1.0) (b) (1.0,1.0,2.0,1.0) (c) (1.0,2.0,1.0,1.0) (d) (2.0,1.0,2.0,1.0).