Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 9(10): 29439–29473.
DOI:10.3934/math.20241427
Received: 20 June 2024
Revised: 16 September 2024
Accepted: 29 September 2024
Published: 17 October 2024

Research article

Bivariate exponentiated generalized inverted exponential distribution with applications on dependent competing risks data

Ammar M. Sarhan[1,2,*], Rabab S. Gomaa[2], Alia M. Magar[2] and Najwan Alsadat[3]

[1] Department of Mathematics and Statistics, Dalhousie University, Canada
[2] Department of Mathematics, Mansoura University, Mansoura 35516, Egypt
[3] Department of Quantitative Analysis, College of Business Administration, KSU, P.O. Box 71115, Riyadh 11587, Saudi Arabia

* **Correspondence:** Email: ammar.sarhan@dal.ca; Tel: +19022666423.

Abstract: This paper introduces a novel bivariate distribution derived from the univariate exponentiated generalized inverted exponential (EGIE) distribution, which we term the bivariate exponentiated generalized inverted exponential (BEGIE) distribution. The newly proposed distribution belongs to the Marshall-Olkin class. Several statistical attributes of the BEGIE distribution are explored. The utility of this distribution is examined through applications on both bivariate data and dependent competing risks data. Estimation processes for the model's parameters, using maximum likelihood and Bayesian methods, are outlined for scenarios involving both bivariate and dependent competing risks data. Due to the absence of closed-form solutions for these estimators, numerical optimization techniques are employed. Furthermore, the proposed distribution is illustrated and evaluated through the analysis of three real datasets: two involving bivariate data, and the other involving dependent competing risks data.

Keywords: reliability analysis; statistical inferences; Bayesian analysis; shock models
Mathematics Subject Classification: 62D05, 62G30, 62P99

1. Introduction

Failure and many causes of failure usually emerge in various application disciplines such as reliability, engineering, and lifetime study due to the complexity of internal structure and external environment, which is referred to as competing risks model in literature. In most literature, inference for competing risk models is studied under the assumption of independent causes of failure. Sarhan and Balakrishnan [28], Sarhan, El-Gohary, El-Bassiouny, and Balakrishnan [29], El-Gohary and Sarhan [10], Sarhan, Hamilton, Smith, and Kundu [30], Kundu, Sarhan, and Gupta [16], Sarhan,

Apaloo, and Kundu [27] and Sarhan [26] proposed bivariate/multivariate lifetime distributions using independent shock or competing risks models. This research mostly used bivariate distributions to fit bivariate data sets. However, the assumption of independence among the causes of failure may be inaccurate. For example, the classic two-unit series system failed due to shocks from three shock sources. The first and second shock sources affect the first and second units, respectively, and the third shock source hits both units at the same time. It is clear that the causes of failure in this system are dependent. In practice, the associated failure reasons may be independent or dependent. In a clinical medicine study, the causes of failure of colon cancer are cancer recurrence or death, and it is observed that such failure causes are also dependent in this situation. As a result, inference for dependent competing risks models will be more practical in reliability, engineering, and other related fields, and in the study of competing risks models involving dependent causes of failure, it is natural to consider a bivariate or multivariate distribution for the lifetimes.

Bivariate lifetime distributions are useful in modeling competing risks when there are two dependent reasons of failure. Wada, Sen, and Shimakura [33] and Wang and Ghosh [34] used bivariate exponential lifetime models to analyze data with two associated risk components. In situations where there is a positive probability of two failure modes occurring simultaneously, the Marshall-Olkin bivariate exponential (MOBE) distribution introduced by Marshall and Olkin [19] is probably the most extensively used model. Pena and Gupta [25], Kotz, Balakrishnan, and Johnson [13], Kundu and Dey [14], Li, Sun, and Song [18], Kundu and Gupta [15], Sarhan, Apaloo and Kundu [27], Alqallaf and Kundu [5], Sarhan and Kundu [31], and references therein are all worth reading.

The inverse exponential (IE) distribution is a life time model that can simulate real-world events, such as bathtub failure rates. Attempts to extend the flexibility of the IE distribution gave rise to the generalized inverted exponential (GIE) distribution by Abouammoh and Alshingiti [2]. Oguntunde et al. [24] presented a new lifetime distribution known as the exponentiated generalized inverted exponential (EGIE) distribution. The cumulative distribution function (cdf), probability density function (pdf), and hazard rate function (hrf) of EGIE are given, respectively, by

$$G(x;\lambda,\gamma,\alpha)=\left\{1-\left[1-\exp\left\{-\frac{\lambda}{x}\right\}\right]^{\gamma}\right\}^{\alpha}\ ,\ x\geq 0;\ \lambda,\gamma,\alpha>0, \tag{1.1}$$

$$g(x;\lambda,\gamma,\alpha)=\alpha\lambda\gamma\frac{1}{x^2}\exp\left\{-\frac{\lambda}{x}\right\}\left[1-\exp\left\{-\frac{\lambda}{x}\right\}\right]^{\gamma-1}\left\{1-\left[1-\exp\left\{-\frac{\lambda}{x}\right\}\right]^{\gamma}\right\}^{\alpha-1}\ , \tag{1.2}$$

and

$$h(x;\lambda,\gamma,\alpha)=\frac{\alpha\lambda\gamma\frac{1}{x^2}\exp\left\{-\frac{\lambda}{x}\right\}\left[1-\exp\left\{-\frac{\lambda}{x}\right\}\right]^{\gamma-1}\left\{1-\left[1-\exp\left\{-\frac{\lambda}{x}\right\}\right]^{\gamma}\right\}^{\alpha-1}}{1-\left\{1-\left[1-\exp\left\{-\frac{\lambda}{x}\right\}\right]^{\gamma}\right\}^{\alpha}}. \tag{1.3}$$

The EGIE distribution generalizes some well known lifetime distributions, and among them we list two special cases.

1) When $\gamma=1$, EGIE (λ,γ,α) reduces to the inverted exponential distribution, IE(λ,α), see Sanku [9].
2) When $\lambda=1$, EGIE (λ,γ,α) reduces to the exponentiated inverted exponential distribution, GIE(γ,α), see Kawsar and Ahmad [12].

The primary objective of this paper is to utilize the EGIE in proposing a novel bivariate distribution by employing a shock model with three independent shocks, each following a non-identical EGIE distribution. One advantage of the proposed distribution is its flexibility, as it generalizes several well-known bivariate distributions, such as the bivariate exponentiated inverse exponential, bivariate inverse exponential, and bivariate inverse Weibull distributions. This versatility allows it to fit a wide range of real data sets. However, a key disadvantage is its singularity, which complicates the theoretical calculations of certain statistical properties, such as expectations, variances, correlation coefficient, and entropy of its variables.

In some applications, we may encounter bivariate data where one variable is discrete and the other is continuous. In such cases, semiparametric approaches for bivariate modeling are highly recommended. Examples include Copula-Based semiparametric models (see Genest et al. [11]) and semiparametric transformation models (see Chen et al. [7]).

An alternative approach for proposing bivariate distributions involves the use of copulas. Several studies have employed this method, including a recent work by Al-Shomrani [4], which introduces a new bivariate family of distributions based on a copula function. As a special case of this family, Al-Shomrani proposed the bivariate Topp-Leone-Exponential-Exponential (BFGMTLEE) distribution. However, they did not apply the BFGMTLEE to fit any real data. In this manuscript, we will compare the performance of our newly proposed model with the BFGMTLEE distribution in fitting real data.

The rest of the paper is organized as follows. Section 2 introduces the BEGIE distribution and discusses some of its fundamental properties. Dependence measures, including correlation, joint entropy, and positively likelihood ratio dependent, are discussed in Section 3. Random samples generation from the proposed model is discussed in Section 4. Section 5 discusses the maximum likelihood method using bivariate data and dependent competing risks data. Section 6 discusses the Bayesian estimation of the model parameters. A simulation study is presented in Section 7. In Section 8, we analyze three real-world data sets using the new suggested distribution and compare it to the IE distribution and GIE distribution as sub-models and some other non-nested models. Section 9 concludes the paper.

2. BEGIE distribution

Now, we introduce the new bivariate exponentiated generalized inverted exponential (BEGIE) distribution as a mathematical model. Consider three independent random variables, denoted as U_j, $j = 1, 2, 3$. Each U_j follows an EGIE distribution with parameters $(\lambda, \gamma, \alpha_j)$ for $j = 1, 2, 3$. Define $X_i = \max(U_i, U_3)$ for $i = 1, 2$. In this scenario, the pair of random variables (X_1, X_2) follows the BEGIE distribution with parameters $(\lambda, \gamma, \alpha_1, \alpha_2, \alpha_3)$.

In a reliability context, this distribution can be understood as follows: Imagine a reliability system consisting of two units. These system units are exposed to three independent sources of shocks. In this scenario, a shock originating from the first source targets unit 1, a shock from the second source targets unit 2, and a shock from the third source affects both units. Unit 1 is considered destroyed when it receives shocks from both sources 1 and 3, while unit 2 is considered destroyed upon experiencing shocks from both sources 2 and 3. Let X_1 and X_2 represent the lifetimes of system units 1 and 2, respectively. In this context, we have $X_1 = \max(U_1, U_3)$ and $X_2 = \max(U_2, U_3)$, where U_1, U_2, and U_3 represent the times at which the shocks occur and are assumed to follow EGIE distributions, as

previously described. As a consequence of this assumption, the joint distribution of (X_1, X_2) will follow the BEGIE distribution.

The main characteristics of the BEGIE distribution are outlined in the following theorem. The proofs, which are straightforward, are omitted for brevity.

Theorem 1. *Suppose* $(X_1, X_2) \sim BEGIE(\lambda, \gamma, \alpha_1, \alpha_2, \alpha_3)$. *Then,*

1) The bivariate cumulative distribution function (cdf) of (X_1, X_2) *is*

$$F_{1,2}(x_1, x_2) = \begin{cases} G(x_1; \lambda, \gamma, \alpha_1 + \alpha_3)\, G(x_2; \lambda, \gamma, \alpha_2), & 0 < x_1 \leq x_2, \\ G(x_1; \lambda, \gamma, \alpha_1)\, G(x_2; \lambda, \gamma, \alpha_2 + \alpha_3), & 0 < x_2 < x_1. \end{cases} \tag{2.1}$$

2) The bivariate probability density function (pdf) of (X_1, X_2) *is*

$$f_{1,2}(x_1, x_2) = \begin{cases} f_1(x_1, x_2) = g(x_1; \lambda, \gamma, \alpha_1 + \alpha_3)\, g(x_2; \lambda, \gamma, \alpha_2), & 0 < x_1 < x_2\,, \\ f_2(x_1, x_2) = g(x_1; \lambda, \gamma, \alpha_1)\, g(x_2; \lambda, \gamma, \alpha_2 + \alpha_3), & 0 < x_2 < x_1\,, \\ f_3(x) = \dfrac{\alpha_3}{\alpha_1 + \alpha_2 + \alpha_3}\, g(x; \lambda, \gamma, \alpha_1 + \alpha_2 + \alpha_3), & 0 < x_1 = x_2 = x\,. \end{cases} \tag{2.2}$$

3) The marginal distribution of X_j *follows* $EGIE(\lambda, \gamma, \alpha_j + \alpha_3)$, $j = 1, 2$.

It is important to note that due to the relationship $P(X_1 = X_2) = \frac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3}$ and the set $A = \{(x_1, x_2) \mid x_1 = x_2 > 0\}$ having Lebesgue measure zero, the BEGIE distribution contains both absolutely continuous and singular components. As a result, the BEGIE distribution is continuous, but not absolutely continuous with respect to the standard Lebesgue measure on $(0, \infty) \times (0, \infty)$.

Substituting from (1.2) into (2.2), we can derive $f_1(x_1, x_2)$, $f_2(x_1, x_2)$, and $f_3(x)$ in the following forms:

$$\begin{aligned} f_1(x_1, x_2) &= \frac{(\alpha_1 + \alpha_3)\alpha_2 \lambda\gamma}{x_1^2\, x_2^2}\, e^{-\frac{\lambda}{x_1} - \frac{\lambda}{x_2}} \left[\left(1 - e^{-\frac{\lambda}{x_1}}\right)\left(1 - e^{-\frac{\lambda}{x_2}}\right)\right]^{\gamma-1} \times \\ &\quad \left[1 - \left(1 - e^{-\frac{\lambda}{x_1}}\right)^{\gamma}\right]^{(\alpha_1+\alpha_3)-1} \left[1 - \left(1 - e^{-\frac{\lambda}{x_2}}\right)^{\gamma}\right]^{\alpha_2-1}, \end{aligned} \tag{2.3}$$

$$\begin{aligned} f_2(x_1, x_2) &= \frac{(\alpha_2 + \alpha_3)\alpha_1 \lambda\gamma}{x_1^2\, x_2^2}\, e^{-\frac{\lambda}{x_1} - \frac{\lambda}{x_2}} \left[\left(1 - e^{-\frac{\lambda}{x_1}}\right)\left(1 - e^{-\frac{\lambda}{x_2}}\right)\right]^{\gamma-1} \times \\ &\quad \left[1 - \left(1 - e^{-\frac{\lambda}{x_1}}\right)^{\gamma}\right]^{\alpha_1-1} \left[1 - \left(1 - e^{-\frac{\lambda}{x_2}}\right)^{\gamma}\right]^{(\alpha_2+\alpha_3)-1}, \end{aligned} \tag{2.4}$$

and

$$f_3(x) = \frac{\alpha_3 \lambda\gamma}{x^2} e^{-\frac{\lambda}{x}} \left(1 - e^{-\frac{\lambda}{x}}\right)^{\gamma-1} \left[1 - \left(1 - e^{-\frac{\lambda}{x}}\right)^{\gamma}\right]^{(\alpha_1+\alpha_2+\alpha_3)-1}. \tag{2.5}$$

Figure 1 portrays the continuous components of the joint pdf (jpdf) for (X_1, X_2), accompanied by corresponding contour plots illustrating various parameter values of the model. Simultaneously, Figure 2 displays the pdf and hrf of the marginal distribution for X_1, while Figure 3 showcases the singular segments of the jpdf, utilizing three distinct sets of parameter values to emphasize diverse hazard function shapes. These figures were generated using MATLAB. It becomes evident from these figures that the distribution's support varies in response to the model parameters' values and the profiles of the pdf and hrf.

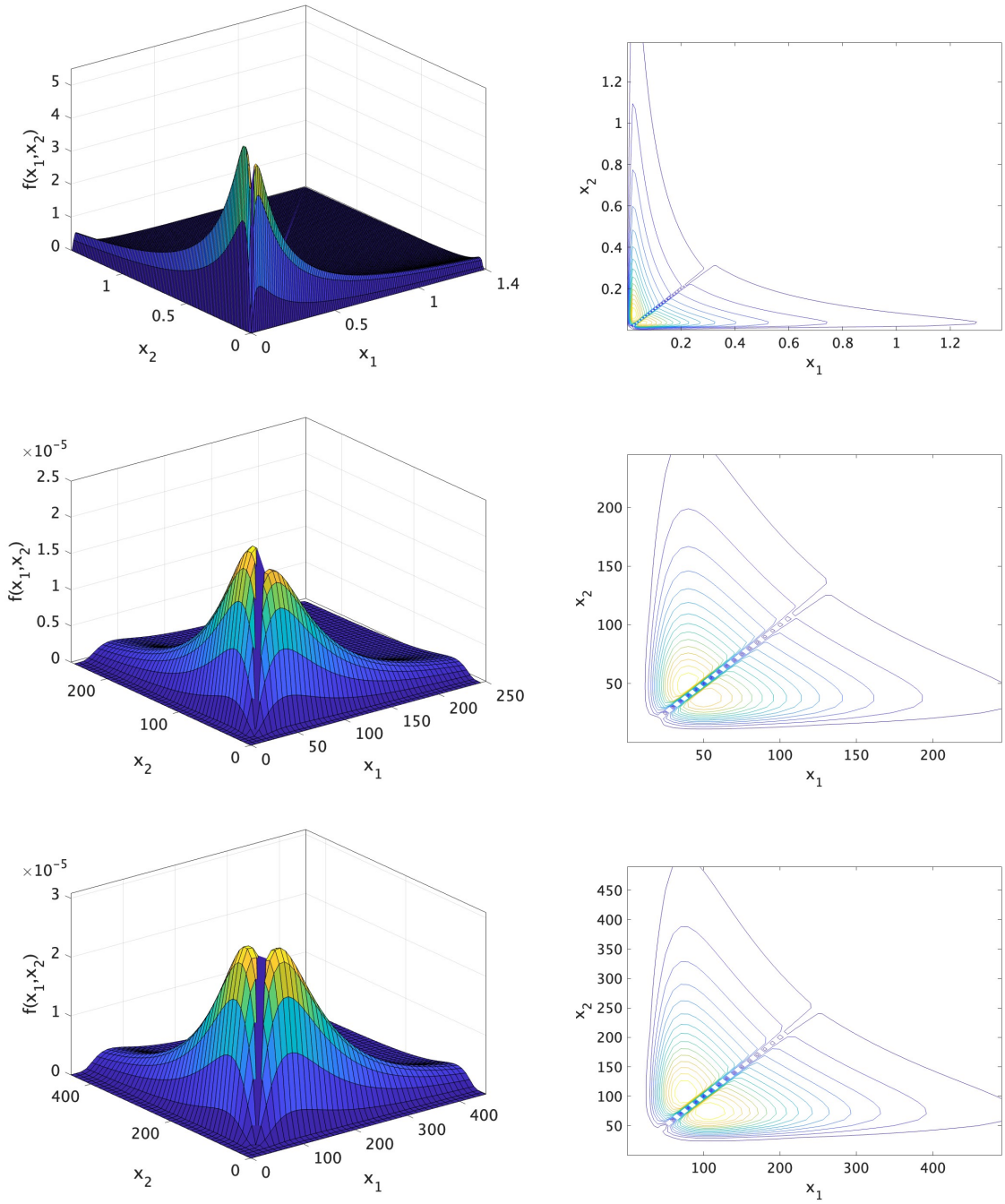

Figure 1. The absolutely continuous parts of the joint pdf and the corresponding contour plots of the BEGIE distribution at different values for the model parameters: (1) Set 1: $\lambda = 0.01, \gamma = 0.5, \alpha_1 = 2, \alpha_2 = 3, \alpha_3 = 2.5$; (2) Set 2: $\lambda = 69; \gamma = 2; a_1 = .4; a_2 = .3; a_3 = .5$; (3) Set 3: $\lambda = 69; \gamma = 2; a_1 = 2.4; a_2 = 2.3; a_3 = 2.5$ (from top to bottom).

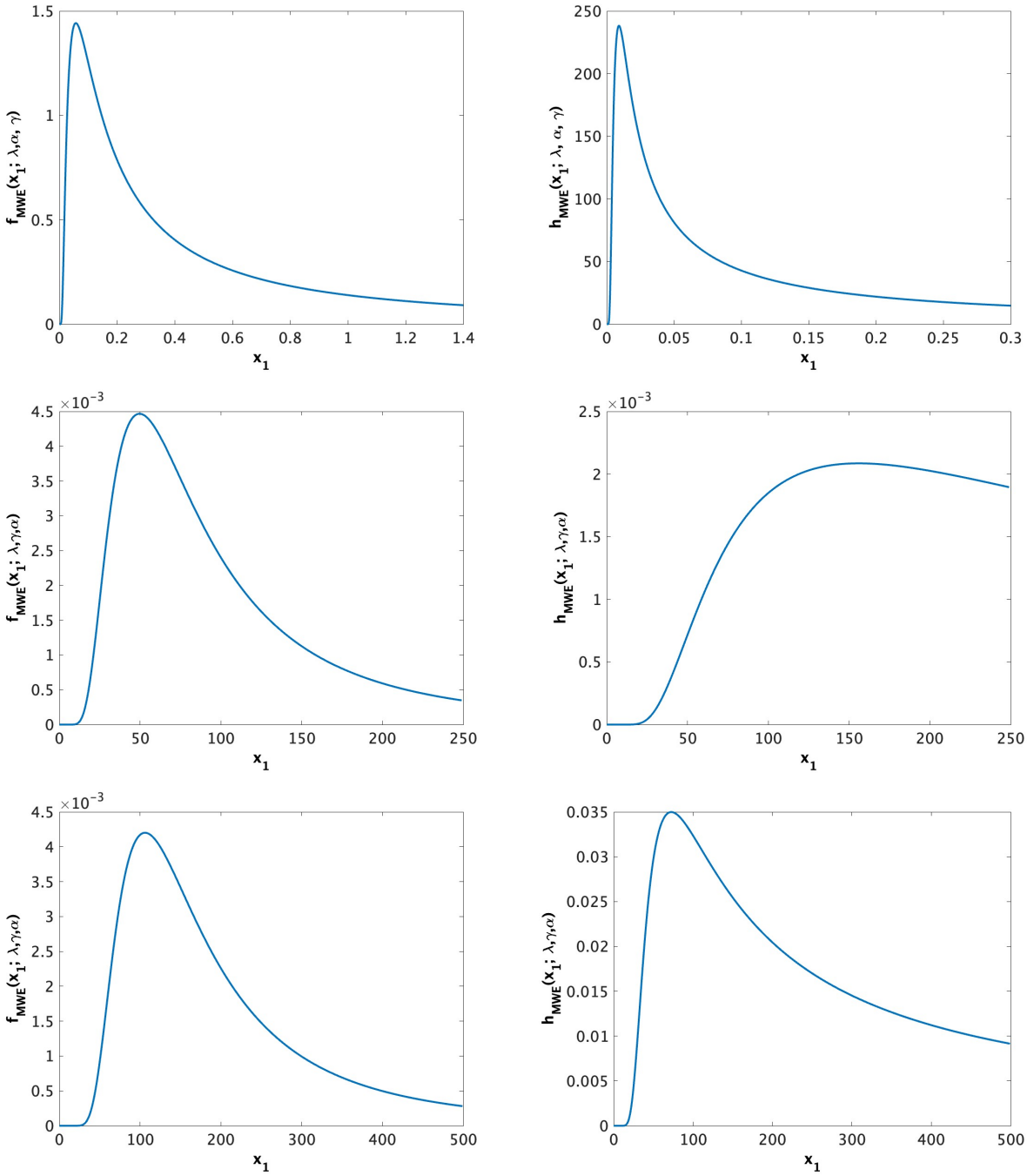

Figure 2. The pdf and the hrf of the marginal distribution of X_1 when: (1) Set 1: $\lambda = 0.01, \gamma = 0.5, \alpha_1 = 2, \alpha_2 = 3, \alpha_3 = 2.5$; (2) Set 2: $\lambda = 69; \gamma = 2; a_1 = .4; a_2 = .3; a_3 = .5$; (3) Set 3: $\lambda = 69; \gamma = 2; a_1 = 2.4; a_2 = 2.3; a_3 = 2.5$ (from top to bottom).

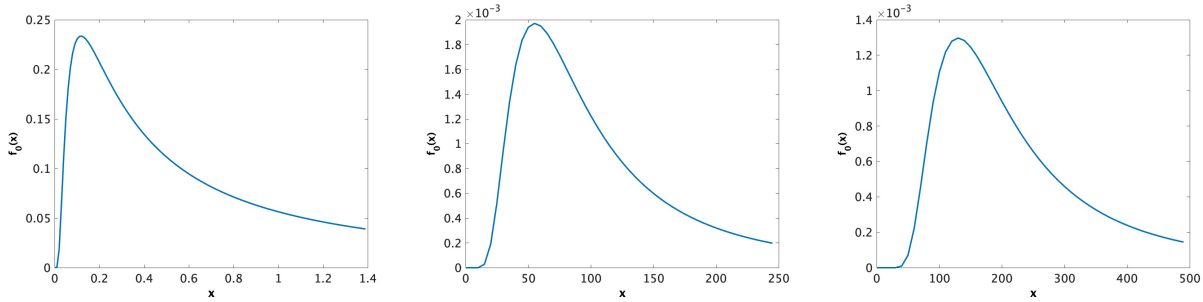

Figure 3. The singular part of the BEGIE$(\lambda, \gamma, \alpha_1, \alpha_2, \alpha_3)$ when: (1) Set 1: $\lambda = 0.01, \gamma = 0.5, \alpha_1 = 2, \alpha_2 = 3, \alpha_3 = 2.5$; (2) Set 2: $\lambda = 69; \gamma = 2; a_1 = .4; a_2 = .3; a_3 = .5$; (3) Set 3: $\lambda = 69; \gamma = 2; a_1 = 2.4; a_2 = 2.3; a_3 = 2.5$ (from left to right).

The BEGIE$(\lambda, \gamma, \alpha_1, \alpha_2, \alpha_3)$ distribution generalizes the following distributions:

1) Bivariate inverted exponential distribution, denoted by BIE$(\lambda, \alpha_1, \alpha_2, \alpha_3)$, by setting $\gamma = 1$;
2) Bivariate exponentiated inverted exponential distribution, denoted by BEIE$(\gamma, \alpha_1, \alpha_2, \alpha_3)$, by setting $\lambda = 1$.

Alqallaf and Kundu [5] applied a similar concept to propose a bivariate inverse generalized exponential (BIGE) distribution by combining three independent inverse generalized exponential (IGE) distributions. While the IGE distribution is a submodel of the EGIE distribution when $\alpha = 1$, the BIGE is not a submodel of the BEGIE.

Now, we discuss the Marshall–Olkin (MO) copula associated with the BEGIE distribution. The BEGIE distribution can be derived by applying the MO copula to EGIE distributions as the marginals. For any bivariate distribution function $F_{1,2}(x_1, x_2)$ with marginals $F_{X_1}(x_1)$ and $F_{X_2}(x_2)$, there exists a unique copula $C : [0,1]^2 \to [0,1]$ that has uniform margins, satisfying $F_{1,2}(x_1, x_2) = C(F_{X_1}(x_1), F_{X_2}(x_2))$ for all $(x_1, x_2) \in \mathbf{R}^2$, see Nelson [22]. The MO copula is given by

$$C_{\theta_1,\theta_2}(u_1, u_2) = u_1^{1-\theta_1} u_2^{1-\theta_2} \min\left(u_1^{\theta_1}, u_2^{\theta_2}\right),$$

for $0 < \theta_1, \theta_2 < 1$. By setting $u_j = F_{X_j}(x_j)$, $j = 1, 2$, where $X_j \sim$ EGIE$(\lambda, \gamma, \alpha_j + \alpha_3)$ and $\theta_j = \frac{\alpha_3}{\alpha_j + \alpha_3}$, we obtain the same joint cumulative distribution function $F_{1,2}(x_1, x_2)$ as in (2.1).

3. Dependence and information

This section discusses three mathematical properties of the BEGIE distribution. Correlation measures the linear relationship between two random variables, while joint entropy captures their combined uncertainty. Positively likelihood ratio dependent variables show a form of positive dependence.

Correlation coefficient: The correlation coefficient of X_1 and X_2, which has a joint pdf $f_{1,2}(x_1, x_2)$, is

$$\rho = \frac{\mathrm{Cov}(X_1, X_2)}{\sigma_{X_1}\, \sigma_{X_2}}, \tag{3.1}$$

where $\mathrm{Cov}(X_1, X_2) = E(X_1\, X_2) - E(X_1)\, E(X_2)$ and

$$E(X_i) \;=\; \int_0^\infty \int_0^\infty x_i f_{1,2}(x_1, x_2)\, dx_2 dx_1,\; i = 1, 2,$$

$$E(X_1\,X_2) \;=\; \int_0^\infty \int_0^\infty x_1\,x_2\,f_{1,2}(x_1,x_2)\,dx_2dx_1\,,$$

$$\sigma^2_{X_i} \;=\; \int_0^\infty \int_0^\infty (x_i - E(X_i))^2 f_{1,2}(x_1,x_2)dx_2dx_1,\; i=1,2.$$

Substituting from (2.2) into the above equations, we can get $E(X_i)$, $E(X_1X_2)$, and σ_{X_i}, and then we use (3.1) to get ρ. Unfortunately, those integrals cannot be computed analytically. Therefore, we should use numerical methods to get ρ.

Joint entropy: Joint entropy for bivariate distributions is a measure of the uncertainty or randomness associated with the joint distribution of two random variables. In information theory, entropy provides a quantitative way to capture the amount of information in a distribution. For bivariate distributions, this concept extends to joint entropy. The joint entropy $H(X_1, X_2)$ of two random variables X_1 and X_2 is a measure of the uncertainty associated with their joint distribution. It is defined as (see Cover and Thomas [6]):

$$H(X_1, X_2) = E\left[\log f_{1,2}(x_1, x_2)\right]\,. \tag{3.2}$$

For the underlying model, H does not have an analytical solution. Therefore, as for ρ, numerical methods must be used to compute H for specific parameter values.

We performed numerical calculations for the marginal means, standard deviations, covariance of (X_1, X_2), correlation, and joint entropy for two parameter sets: Set I ($\lambda = 69, \gamma = 2, \alpha_1 = .4, \alpha_2 = .3$) and Set II ($\lambda = 69, \gamma = 2, \alpha_1 = 2.4, \alpha_2 = 2.3$), with varying values of $\alpha_3 = 0.1, 0.2, \cdots, 3.5$. The results are presented in Tables 1 and 2. Additionally, Figure 4 illustrates the behavior of ρ and H as functions of α_3. From these results, we observe a positive correlation between the two variables that decreases with α_3. A decreasing correlation does not necessarily imply that there is no relationship between the variables, it simply means the linear relationship is weakening. We also observe that H increases and then decreases as α_3 increases. The "up then down" pattern in H indicates that the variables initially move towards independence (increasing uncertainty) but later exhibit stronger dependence, reducing uncertainty. This pattern often reflects dynamic interactions between variables as their relationship evolves under changing conditions.

Positively likelihood ratio dependent (PLRD): PLRD is a concept in probability and statistics that refers to a specific type of dependence between two random variables; for more details, see Lehmann [17]. It is a form of positive dependence that is stronger than simple positive correlation. Two random variables X_1 and X_2 are said to be PLRD if for any values $x_{11} < x_{12}$ and any value x_2, the likelihood ratio LR $= \frac{f_{2|1}(x_2|x_{12})}{f_{2|1}(x_2|x_{11})}$ is non-decreasing in x_2. Here, $f_{2|1}(x_2|x_1)$ is the conditional pdf of X_2, given $X_1 = x_1$. The following theorem summarizes PLRD in the BEGIE case. The proof is straightforward, so it is omitted for brevity.

Theorem 2. *Let (X_1, X_2) follow BEGIE$(\lambda, \gamma, \alpha_1, \alpha_2, \alpha_3)$. Then, X_1 and X_2 are PLRD if $\alpha_2 \le \alpha_1 + \alpha_3$.*

It is important to note that PLRD is a stronger condition than positive correlation. While positive correlation implies that higher values of one variable are associated with higher values of another, PLRD imposes a stricter requirement on how their joint distribution behaves.

Table 1. The marginal mean, variance, covariance, and correlation of X_1 and X_2 using Set I of parameters' values: $\lambda = 69, \gamma = 2, \alpha_1 = .4, \alpha_2 = .3$, and different values of α_3.

α_3	$E(X_1)$	$E(X_2)$	$E(X_1X_2)$	$\mathrm{Cov}(X_1, X_2)$	σ_{X_1}	σ_{X_2}	ρ	H
0.1	12.5	11.6	238.2	93.2	14.2	12.8	0.524	−9.343
0.2	18	16.3	397.5	103.6	17.1	15.8	0.389	−8.906
0.3	21.5	19.4	520	102.5	19.4	17.8	0.302	−8.702
0.4	24.2	21.6	618.7	97	21.2	19.4	0.239	−8.914
0.5	26.2	23.3	702.3	91.8	22.6	21	0.196	−9.256
0.6	27.9	24.6	774	87.4	23.8	22.1	0.166	−9.652
0.7	29.3	25.7	832.4	80.8	24.9	23.1	0.142	−10.076
0.8	30.5	26.7	888.4	76.6	26.1	24.1	0.123	−10.496
0.9	31.4	27.4	934.9	73	26.9	25	0.108	−10.932
1	32.4	28.1	978.5	67	28.1	25.8	0.093	−11.366
1.1	33.2	28.8	1018.8	64.4	28.7	26.6	0.084	−11.794
1.2	33.9	29.3	1054.7	60.8	29.8	27.4	0.075	−12.22
1.3	34.5	29.9	1089.5	58.5	30.3	28.2	0.069	−12.643
1.4	35.1	30.3	1119.3	56	31	28.9	0.063	−13.084
1.5	35.7	30.7	1152.3	56.3	32.1	29.4	0.06	−13.514
1.6	36.1	31.1	1174.4	50.6	32.2	30	0.052	−13.921
1.7	36.6	31.5	1196.9	46.3	33	30.5	0.046	−14.343
1.8	37	31.9	1225.5	48	33.3	31.2	0.046	−14.778
1.9	37.3	32.1	1243.7	45.3	34.1	31.6	0.042	−15.207
2	37.7	32.3	1264.7	44.2	34.8	31.9	0.04	−15.627
2.1	38	32.6	1283.6	43.2	35.1	32.8	0.038	−16.046
2.2	38.4	32.9	1303.8	42.1	35.5	33.2	0.035	−16.46
2.3	38.8	33.1	1321.2	39.8	36.4	33.2	0.032	−16.891
2.4	39	33.3	1337.2	37.3	36.7	33.8	0.03	−17.309
2.5	39.2	33.6	1350.9	36.4	36.9	34.2	0.028	−17.729
2.6	39.5	33.7	1367.6	35.5	37.7	34.5	0.027	−18.173
2.7	39.7	33.9	1381	33.9	37.9	35.2	0.026	−18.593
2.8	40	34.1	1394	31.9	38.5	35.1	0.025	−19.004
2.9	40.2	34.2	1409.7	34.2	38.6	35.7	0.024	−19.426
3	40.4	34.4	1422.6	31.7	39.1	35.8	0.023	−19.85
3.1	40.5	34.6	1428.8	28.4	39.2	36.4	0.022	−20.294
3.2	40.8	34.7	1446.2	30.3	39.7	36.7	0.021	−20.703
3.3	41	34.8	1460.1	32.2	40.2	37.1	0.021	−21.126
3.4	41.1	35	1468.1	29	40.5	37.2	0.019	−21.532
3.5	41.3	35.1	1477.3	28	40.6	37.8	0.018	−21.964

Table 2. The marginal mean, variance, covariance, and correlation of X_1 and X_2 using Set II of parameters' values: $\lambda = 69, \gamma = 2, \alpha_1 = 2.4, \alpha_2 = 2.3$, and different values of α_3.

α_3	$E(X_1)$	$E(X_2)$	$E(X_1X_2)$	$\mathrm{Cov}(X_1, X_2)$	σ_{X_1}	σ_{X_2}	ρ	H
0.1	17	16.9	740.8	453.1	23.7	23.5	0.819	−35.213
0.2	26.3	26.2	1360.4	671	30	29.6	0.763	−27.834
0.3	33.3	33.2	1912.1	805.7	33.8	33.5	0.718	−20.501
0.4	39	38.7	2399.5	887.7	36.9	36.3	0.674	−16.933
0.5	43.8	43.4	2860.7	955.9	39	38.5	0.643	−14.875
0.6	48	47.5	3263	984.8	40.6	39.9	0.612	−13.581
0.7	51.5	51.1	3649.7	1015.3	42.4	41.9	0.579	−12.705
0.8	54.8	54.2	3999.6	1027.8	43.7	43.1	0.552	−12.085
0.9	57.7	57.1	4323.3	1027	44.9	44.4	0.522	−11.648
1	60.3	59.7	4638.6	1035.6	45.9	45.2	0.505	−11.328
1.1	62.7	62.1	4932.8	1035.2	47.1	46.2	0.482	−11.101
1.2	65.1	64.4	5222.9	1029	48.1	47.3	0.456	−10.922
1.3	67.2	66.5	5498	1026.4	49	48.6	0.438	−10.801
1.4	69	68.3	5731.6	1014.9	49.6	49.2	0.42	−10.716
1.5	71.1	70.1	5999.1	1012.6	50.7	49.8	0.405	−10.654
1.6	72.7	71.8	6222.1	1000.9	51.6	50.9	0.385	−10.62
1.7	74.3	73.3	6433.5	984	52.3	51.6	0.37	−10.593
1.8	75.9	74.9	6668.2	983.8	53.2	52.2	0.358	−10.596
1.9	77.3	76.4	6888.5	978.5	53.7	53.5	0.344	−10.602
2	78.6	77.6	7059.3	961	54.7	53.6	0.332	−10.617
2.1	80	78.9	7260.8	945	55.2	54.8	0.316	−10.638
2.2	81.2	80.1	7430	927.2	55.6	55.2	0.306	−10.668
2.3	82.2	81.2	7593.8	916.2	56.3	55.5	0.296	−10.706
2.4	83.4	82.2	7752.4	897.6	57.2	56.4	0.281	−10.751
2.5	84.4	83.3	7926.9	893.5	57.3	57.1	0.276	−10.796
2.6	85.5	84.3	8097.6	886.3	58.3	57.3	0.268	−10.839
2.7	86.6	85.3	8252.2	864.9	59	58.1	0.255	−10.892
2.8	87.5	86.2	8396.1	851.3	59.7	59.2	0.244	−10.949
2.9	88.4	87.2	8548	842.4	60.1	59.5	0.238	−11.002
3	89.3	88	8707.3	847.6	61	60.1	0.234	−11.063
3.1	90.1	88.8	8823.5	824.5	61.2	60.4	0.225	−11.115
3.2	90.8	89.7	8946.3	804.5	61.2	61.3	0.217	−11.176
3.3	91.7	90.4	9108.4	809.2	62.3	61.5	0.213	−11.232
3.4	92.4	91.1	9220.6	798.3	62.8	62	0.206	−11.295
3.5	93.3	91.8	9345.3	786	63.4	62.4	0.2	−11.357

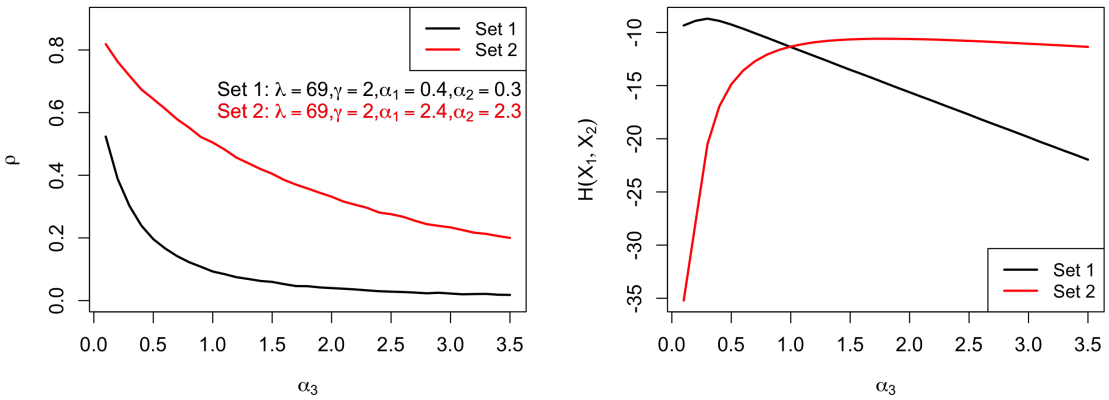

Figure 4. The correlation coefficient of X_1 and X_2 (left) and entropy (right) as functions of α_3 at different values of the other four parameters.

4. Random samples generation

When generating random numbers from the EGIE$(\lambda, \gamma, \alpha)$ distribution, if X follows EGIE$(\lambda, \gamma, \alpha)$, the subsequent equation provides the qth quantile of X:

$$x = \left[-\frac{1}{\lambda}\log\left(1-\left(1-q^{\frac{1}{\alpha}}\right)^{\frac{1}{\gamma}}\right)\right]^{-1}, \ q \in (0,1). \tag{4.1}$$

The following algorithm can be followed to generate bivariate data from the BEGIE$(\lambda, \gamma, \alpha_1, \alpha_2, \alpha_3)$ distribution.

1) Specify the sample size n;
2) Specify the values of the model parameters $\lambda, \gamma, \alpha_1, \alpha_2, \alpha_3$;
3) Generate the samples, each with size n, from EGIE$(\lambda, \gamma, \alpha_j)$, $j = 1, 2, 3$, according to Eq (4.1). Use U_j, $j = 1, 2, 3$, to denote these three generated samples respectively. That is, for $j = 1, 2, 3$, the jth element in vector U_j, is

$$U_{ji} = \left[-\frac{1}{\lambda}\log\left(1-\left(1-q_{ji}^{\frac{1}{\alpha_j}}\right)^{\frac{1}{\gamma}}\right)\right]^{-1}, i = 1, 2, \cdots, n;$$

   where q_{ji} is randomly generated from a uniform distribution on the $(0, 1)$ interval.
4) Generate a sample with size n of (X_1, X_2) by applying the following relationships:

$$X_{1i} = \max(U_{1i}, U_{3i}), \text{ and } X_{2i} = \max(U_{2i}, U_{3i}), \ i = 1, 2, \cdots, n.$$

The random sample (X_{1i}, X_{2i}) obtained from this algorithm follow a BEGIE$(\lambda, \gamma, \alpha_1, \alpha_2, \alpha_3)$ distribution. This bivaraite sample of (X_1, X_2) can be used to generate a dependent competing risks

sample, say (T_i, δ_i), $i = 1, 2, \cdots, n$, from the underlying distribution by setting $T_i = \min(X_{1i}, X_{2i})$ and

$$\delta_i = \begin{cases} 1 & \text{if } X_{1i} < X_{2i} \text{ (risk 1 causes the failure);} \\ 2 & \text{if } X_{2i} < X_{1i} \text{ (risk 2 causes the failure);} \\ 3 & \text{if } X_{1i} = X_{2i} \text{ (both risks 1 \& 2 cause the failure).} \end{cases}$$

5. Maximum likelihood estimation

In this section, we estimate the model parameters $\Theta = (\alpha_1, \alpha_2, \alpha_3, \lambda, \gamma)$ using two different types of data. Type I data consists of simple random samples of the bivariate vector (X_1, X_2) that follows the new bivariate distribution. Type II data consists of dependent competing risks data (T, δ), with $T = \min(X_1, X_2)$, where (X_1, X_2) follows a bivariate lifetime distribution and $\delta \in \{1, 2, 3\}$ represents the cause of failure. Theorem 1 will be helpful to estimate the model parameters using data of Type I. The following theorem is needed for Type II data.

Theorem 3. *The pdf of $T = \min(X_1, X_2)$, for a given cause of failure $\delta \in \{1, 2, 3\}$, is*

$$f^*(t;\delta) = \begin{cases} f_1^*(t) & \text{if } \delta = 1, \\ f_2^*(t) & \text{if } \delta = 2, \\ f_3^*(t) & \text{if } \delta = 3, \end{cases} \tag{5.1}$$

where

$$\begin{aligned} f_j^*(t) \;=\; & \frac{\alpha_j \lambda \gamma}{t^2} e^{-\frac{\lambda}{t}} \left(1 - e^{-\frac{\lambda}{t}}\right)^{\gamma-1} \left[1 - \left(1 - e^{-\frac{\lambda}{t}}\right)^{\gamma}\right]^{\alpha_j - 1} \\ & \left\{1 - \left[1 - \left(1 - e^{-\frac{\lambda}{t}}\right)^{\gamma}\right]^{\alpha_{3-j}}\right\} \left\{1 - \left[1 - \left(1 - e^{-\frac{\lambda}{t}}\right)^{\gamma}\right]^{\alpha_3}\right\}, \; j = 1, 2, \end{aligned} \tag{5.2}$$

and

$$f_3^*(t) = \frac{\alpha_j \lambda \gamma}{t^2} e^{-\frac{\lambda}{t}} \left(1 - e^{-\frac{\lambda}{t}}\right)^{\gamma-1} \left[1 - \left(1 - e^{-\frac{\lambda}{t}}\right)^{\gamma}\right]^{\alpha_1 + \alpha_2 + \alpha_3 - 1}. \tag{5.3}$$

Proof. Based on (2.1), one can show that the pdf of $T = \min(X_1, X_2)$ is

$$f^*(t;\delta) = \begin{cases} f_1^*(t) = g(t;\lambda,\gamma,\alpha_1)\,\overline{G}(t;\lambda,\gamma,\alpha_2)\,\overline{G}(t;\lambda,\gamma,\alpha_3) & \text{if } \delta = 1, \\ f_2^*(t) = g(t;\lambda,\gamma,\alpha_2)\,\overline{G}(t;\lambda,\gamma,\alpha_1)\,\overline{G}(t;\lambda,\gamma,\alpha_3) & \text{if } \delta = 2, \\ f_3^*(t) = \frac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3}\, g(t;\lambda,\gamma,\alpha_1+\alpha_2+\alpha_3) & \text{if } \delta = 3, \end{cases} \tag{5.4}$$

where $\overline{G}(t;\lambda,\gamma,\alpha_j) = 1 - G(t;\lambda,\gamma,\alpha_j)$ is the survival function of U_j, $j = 1, 2, 3$. Substituting from (1.2) and (1.1) into (5.4), we complete the proof.

The maximum likelihood method is utilized for estimating model parameters, employing both types of data. Additionally, we also apply the Bayesian method, specifically for bivariate data. Although the Bayesian method could similarly be applied to dependent competing risks data, it was omitted from the paper due to length constraints. For more details on estimation methods, we refer to Wang et al. [8] and Albert (2009) [3].

5.1. Bivariate data

Let us assume that $(X_{11}, X_{21}), (X_{12}, X_{22}), \cdots, (X_{1n}, X_{2n})$ is an independent and identically distributed random sample of $(X_1,\ X_2)$ that follows BEGIE$(\alpha_1,\ \alpha_2,\ \alpha_3,\ \lambda,\ \gamma)$. For simplicity, let us introduce the indicator variables δ_i, $i = 1, 2, ..., n$, where

$$\delta_i = \begin{cases} 1 & \text{if } X_{1i} < X_{2i}, \\ 2 & \text{if } X_{1i} > X_{2i}, \\ 3 & \text{if } X_{1i} = X_{2i}. \end{cases}$$

Using the above bivariate sample, the likelihood function can be expressed as

$$L_1(\Theta) = \prod_{i=1}^{n} [f_1(x_{1i}, x_{2i})]^{I[\delta_i=1]} [f_2(x_{1i}, x_{2i})]^{I[\delta_i=2]} [f_3(x_{1i}]^{I[\delta_i=3]}. \tag{5.5}$$

Substituting from (2.3)–(2.5) into (5.5), we get the log-likelihood function as

$$\begin{aligned}
\mathcal{L}_1 \;=\; & \sum_{j=1}^{3} n_j \log \alpha_j + (2n - n_3) \log(\lambda\, \gamma) + (n_1 + n_2) \log(\alpha_j + \alpha_3) \\
& -2 \sum_{i=1}^{n} \left[I[\delta_i = 3] \log x_{1i} + \sum_{j=1}^{2} I[\delta_i = j] \log(x_{1i} x_{2i}) \right] \\
& + \sum_{i=1}^{n} I[\delta_i = 3] \left(\frac{-\lambda}{x_{1i}} + (\gamma - 1) \log(1 - e^{-\frac{\lambda}{x_{1i}}}) \right) \\
& + \sum_{i=1}^{n} \sum_{j=1}^{2} I[\delta_i = j] \left[\sum_{\ell=1}^{2} \frac{-\lambda}{x_{\ell i}} + (\gamma - 1) \log(1 - e^{-\frac{\lambda}{x_{\ell i}}}) \right] \\
& + (\alpha_2 - 1) \sum_{i=1}^{n} I[\delta_i = 1] \log \left(1 - (1 - e^{-\frac{\lambda}{x_{2i}}})^{\gamma} \right) \\
& + \sum_{i=1}^{n} \sum_{j=1}^{2} (\alpha_j + \alpha_3 - 1) I[\delta_i = j] \log(1 - (1 - e^{-\frac{\lambda}{x_{ji}}})^{\gamma}) \\
& + \sum_{i=1}^{n} \left(I[\delta_i = 3] \left(\sum_{j=1}^{3} \alpha_j - 1 \right) + I[\delta_i = 2](\alpha_1 - 1) \right) \log \left(1 - \left(1 - e^{-\frac{\lambda}{x_{1i}}} \right)^{\gamma} \right),
\end{aligned} \tag{5.6}$$

where $n_j = \sum_{i=1}^{n} I[\delta_i = j],\ j = 1, 2, 3$ with $I[A] = 1$ if A is true and 0 otherwise.

Consequently, the maximum likelihood estimates (MLEs) of the unknown parameters can be determined by maximizing the log-likelihood function pertaining to the given data in relation to these parameters. Similarly, we can formulate the likelihood equations, which result from equating the first partial derivatives of the log-likelihood function concerning the unknown parameters to zero. The MLEs correspond to the solutions of these derived likelihood equations, at which point the Fisher information matrix should exhibit positive definiteness. This Fisher information matrix encompasses the second partial derivatives of the log-likelihood function with respect to the parameters.

In the case of the distribution under consideration, the likelihood equations are not amenable to an analytical solution. Consequently, we have resorted to employing numerical techniques within the R software framework to solve the system of five non-linear equations involving five unknown parameters.

5.2. Dependent competing risks data

In this type of data, we observe a pair of quantities: T, the system time to failure, and δ, an indicator of the cause of failure. This means that the observations in this case are $(T_i,\ \delta_i), i = 1, 2, ..., n$. The lifetime experiment that produces this type of data can be illustrated as follows: (1) we put n independent and identical devices (system/objects) on the life test; (2) each system is under attack from two dependent competing risks which occur at times X_1 and X_2, and the system will fail once it receives one of the two attacks; (3) (X_1, X_2) follows the BEGIE distribution; (4) we observe $(T_i, \delta_i), i = 1, 2, ..., n$, where $T_i = \min(X_{1i}, X_{2i})$, and δ_i is defined as $\delta_i = 1$ if risk 1 causes the failure $(X_{1i} < X_{2i})$, $\delta_i = 2$ if risk 2 causes the failure $(X_{2i} < X_{1i})$, and $\delta_i = 3$ if both risks cause the failure $(X_{1i} = X_{2i})$. The likelihood function using such dependent competing risks data is

$$L_2(\Theta) = \prod_{i=1}^{n} \left\{ \prod_{j=1}^{3} [f_j^*(t_i)]^{I[\delta_i = j]} \right\}, \tag{5.7}$$

where $f_j^*(t)$ are given in Theorem 3. Substituting (5.2) and (5.3) into (5.7), we can get the log-likelihood function as

$$\begin{aligned} \mathcal{L}_2 \;=\; & \sum_{j=1}^{3} n_j \log \alpha_j + n \log(\lambda\gamma) - 2\sum_{i=1}^{n} \log t_i - \sum_{i=1}^{n} \frac{\lambda}{t_i} + (\gamma - 1) \sum_{i=1}^{n} \log(1 - e^{-\frac{\lambda}{t_i}}) \\ & + \sum_{j=1}^{2} \left\{ (\alpha_j - 1) \sum_{i=1}^{n} I[\delta_i = j]\ \log\left[1 - \left(1 - e^{-\frac{\lambda}{t_i}}\right)^{\gamma}\right] \right\} \\ & + (\alpha_1 + \alpha_2 + \alpha_3 - 1) \sum_{i=1}^{n} I[\delta_i = 3] \log\left[1 - \left(1 - e^{-\frac{\lambda}{t_i}}\right)^{\gamma}\right], \end{aligned} \tag{5.8}$$

where $n_j = \sum_{i=1}^{n} I[\delta_i = j],\ j = 1, 2, 3$.

To get the MLE of the model parameters, using dependent competing risks data, we set the first partial derivatives of $\mathcal{L}_2$ with respect to the five parameters $\alpha_1, \alpha_2, \alpha_3, \lambda, \gamma$ equal to zero, and we get a system of five non-linear equations in five unknowns. Solving the first three equations in α_j, $j = 1, 2, 3$, we get α_j as a function of (λ, γ):

$$\alpha_j(\lambda, \gamma) = \frac{-n_j}{\sum_{i=1}^{n} (I[\delta_i = j] + I[\delta_i = 3]) \log(1 - (1 - e^{-\frac{\lambda}{t_i}})^{\gamma})},\ j = 1, 2. \tag{5.9}$$

and

$$\alpha_3(\lambda, \gamma) = \frac{-n_3}{\sum_{i=1}^{n} I[\delta_i = 3]\ \log(1 - (1 - e^{-\frac{\lambda}{t_i}})^{\gamma})}. \tag{5.10}$$

Substituting (5.9) and (5.10) into (5.8), we can express the log-likelihood function as a function of two parameters (λ, γ) , say $\mathcal{L}_2^*(\lambda, \gamma)$. This will make the optimization of the log-likelihood function much

easier since we will deal only with two parameters instead of five. Once we get the MLE of λ and γ, we can use (5.9) and (5.10) to get the MLE of α_j, $j = 1, 2, 3$.

Since the MLE of the parameters using either type of data are not obtained in closed form, we cannot get the explicit sampling distributions of the MLEs of these parameters. Therefore, we cannot obtain exact confidence intervals for the model parameters. Alternatively, we can use the large sample distribution of the MLE and the corresponding observed Fisher information matrix to derive approximate confidence intervals for the model parameters.

6. Bayesian estimation

In this section, we implement a Bayesian approach to estimate the model parameters. We will demonstrate this method using the two types of data previously discussed. To derive Bayesian estimations, we assume independence among the five model parameters: λ, γ, α_1, α_2, and α_3, considering them as independent random variables following gamma prior distributions. These gamma prior distributions have the hyperparameters (a_i, b_i), where $i = 1, 2, \cdots, 5$, respectively. This implies that the logarithm of the joint prior density of these five parameters, referred to as $g(\theta)$, is up to a normalized constant, given by

$$\log g(\theta) \propto \sum_{j=1}^{5} \left\{ (a_j - 1) \log \theta_j - b_j \theta_j \right\} \tag{6.1}$$

where, θ_j is the jth component in the vector of unknown parameter θ. Combining the logarithm of the joint prior density function with the logarithm of likelihood function allows us to formulate the logarithm of the joint posterior density function of θ, given data, up to a normalized constant as

$$\log g(\theta|\text{data}) = \log g(\theta) + \mathcal{L}(\text{data}, \theta). \tag{6.2}$$

Substituting (6.1) and (5.6) into (6.2), we can obtain the $\log g(\theta|\text{data})$ for the bivariate data, as

$$\log g_I(\theta|\text{data}) = \log g(\theta) + \mathcal{L}_I(\text{data}, \theta). \tag{6.3}$$

While substituting (6.1) and (5.8) into (6.2), we can obtain the $\log g(\theta|\text{data})$ for the dependent competing risks data, as

$$\log g_{II}(\theta|\text{data}) = \log g(\theta) + \mathcal{L}_{II}(\text{data}, \theta). \tag{6.4}$$

The logarithm of the posterior density functions, $\log g_I(\theta|\text{data})$ and $\log g_{II}(\theta|\text{data})$ does not take the form of well-known multivariate distributions, making it challenging to derive analytic solutions for the Bayesian analysis of the model parameters. Consequently, we will need to resort to numerical computational methods. Specifically, in this paper, we will utilize the Markov chain Monte Carlo (MCMC) and sampling importance resampling (SIR) methods to approximate Bayesian estimations for the model parameters. Within MCMC, our primary task involves simulating samples from a proposal distribution that closely resembles the posterior distribution. For more comprehensive insights, readers are directed to Albert [3]. Typically, multivariate normal and t distributions serve as common proposal distributions. However, as the underlying parameters are all positive, whereas these distributions have

real-valued support, we need to re-parameterize θ into $\phi = \log(\theta) \in \mathbf{R}_5$. Through this logarithmic transformation of θ, we can derive the logarithm of the posterior density functions, of ϕ, as

$$\log g_I^T(\phi|\text{data}) = \log g_I(\theta = e^{\phi}|\text{data}) + \sum_{j=1}^{5} \phi_j \,, \tag{6.5}$$

and

$$\log g_{II}^T(\phi|\text{data}) = \log g_{II}(\theta = e^{\phi}|\text{data}) + \sum_{j=1}^{5} \phi_j \,. \tag{6.6}$$

We will also consider the normal prior distribution of $\phi_j = \log \theta_j \sim N(\mu_j, \sigma_j)$. In this case, $\log g(\theta)$, given in (6.1), will be replaced with

$$\log g(\phi) \propto -\sum_{j=1}^{5} (\phi_j - \mu_j)^2 \,.$$

7. Simulation study

To demonstrate the robustness of the model and its practical implications under various theoretical conditions, we conduct a simulation study of the bivariate model using two different sets of model parameters. The study follows the steps outlined below:

1) **Specify the parameter values:** Define the sample size n and the model parameters $\lambda, \gamma, \alpha_1, \alpha_2$, and α_3.
2) **Generate random samples:** Using the random sample generation method outlined in Section 3, generate M random samples of (X_1, X_2).
3) **Estimate parameters:** For each sample generated in Step 2, compute the MLE, Bayes estimate (BE) of each parameter, and the corresponding asymptotic 95% confidence interval and 95% credible interval.
4) **Compute performance metrics:** Using the M results from Step 3, calculate the following:
   (a) The average of the point estimation (APE), using the MLE and BE.
   (b) The mean squared error (MSE), using the MLE and BE.
   (c) The average length of the confidence/credible interval (ALCI).
   (d) The coverage probability (CP) for both maximum likelihood and Bayes methods.

   The mean squared error (MSE) and average of the point estimations (APE) are calculated as

$$\text{APE} = \frac{\sum_{i=1}^{M} \hat{\theta}^{(i)}}{M}, \quad \text{MSE} = \frac{\sum_{i=1}^{M} \left(\theta - \hat{\theta}^{(i)}\right)^2}{M} \,,$$

   where $\hat{\theta}^{(i)}$ is the either the MLE or BE of the parameter θ, $\theta = \lambda, \gamma, \alpha_1, \alpha_2, \alpha_3$, for the i-th sample, $i = 1, 2, \cdots, M$.

Table 3. The estimation results for two sets of model parameters: Set I: $\lambda = 69, \gamma = 2, \alpha_1 = .4, \alpha_2 = .3, \alpha_3 = .5$ and Set II: $\lambda = 69, \gamma = 2, \alpha_1 = 2.4, \alpha_2 = 2.3, \alpha_3 = 2.5$.

n	Par	AMLE	MSE	ALCI	CP	AMLE	MSE	ALCI	CP
		Set I				Set II			
50	λ	70.4797	108.0786	39.0713	94.78	70.2358	107.3271	38.936	94.84
		70.3798	107.1782	38.9473	94.4	70.1365	106.4778	38.8127	94.82
	γ	2.0446	0.08677	1.1335	95.2	2.0435	0.08892	1.1328	94.96
		2.0446	0.08675	1.1315	94.94	2.0435	0.08891	1.1308	94.88
	α_1	0.4081	0.00356	0.2262	95.54	2.4474	0.1279	1.3567	95.12
		0.4081	0.00356	0.2258	95.08	2.4473	0.12787	1.3543	95.1
	α_2	0.3062	0.00198	0.1698	95.46	2.3489	0.11669	1.3022	94.98
		0.3062	0.00198	0.1695	95.2	2.3489	0.11667	1.2998	95.16
	α_3	0.5096	0.00547	0.2825	95.56	2.5462	0.13861	1.4115	94.94
		0.5096	0.00547	0.282	95.44	2.5461	0.13858	1.409	95
75	λ	69.9846	69.88682	31.6775	94.92	69.9421	66.7908	31.6582	95.58
		69.9194	69.50135	31.6108	94.68	69.877	66.42289	31.5916	95.42
	γ	2.0285	0.0574	0.9182	94.86	2.0307	0.05747	0.9192	95.38
		2.0285	0.05739	0.9171	94.58	2.0307	0.05747	0.9181	95.04
	α_1	0.4046	0.00221	0.1832	95.2	2.4363	0.08479	1.1028	94.48
		0.4046	0.00221	0.1829	95.24	2.4363	0.08478	1.1015	94.68
	α_2	0.3045	0.00133	0.1378	94.88	2.3315	0.07415	1.0553	95.16
		0.3045	0.00133	0.1377	94.94	2.3315	0.07414	1.0541	95.1
	α_3	0.5062	0.00355	0.2291	94.98	2.5338	0.09091	1.1469	95.36
		0.5062	0.00355	0.2289	95	2.5338	0.09089	1.1455	95.06
100	λ	69.7553	49.43844	27.3436	94.92	69.7013	49.31247	27.3224	95.06
		69.7068	49.2296	27.3005	95.1	69.6529	49.10912	27.2794	95.16
	γ	2.0224	0.04295	0.7928	94.96	2.0241	0.04206	0.7934	95.5
		2.0224	0.04295	0.7921	94.74	2.0241	0.04206	0.7927	95.32
	α_1	0.4046	0.00171	0.1586	95.08	2.421	0.05836	0.949	95.64
		0.4046	0.00171	0.1585	95.28	2.421	0.05835	0.9482	95.72
	α_2	0.3029	0.00091	0.1187	95.6	2.3216	0.05465	0.9101	95.28
		0.3029	0.00091	0.1186	95.48	2.3216	0.05464	0.9092	95.34
	α_3	0.5053	0.00275	0.1981	94.82	2.5276	0.06571	0.9908	95.26
		0.5053	0.00275	0.1979	94.8	2.5275	0.0657	0.9899	95.18
150	λ	69.4208	32.23925	22.2189	95.36	69.3671	31.94475	22.2017	94.88
		69.3889	32.15379	22.1956	95.38	69.3353	31.86328	22.1785	94.62
	γ	2.0114	0.0268	0.6438	95.18	2.0146	0.02736	0.6448	95.08
		2.0114	0.0268	0.6434	95.34	2.0146	0.02736	0.6444	95
	α_1	0.4022	0.00113	0.1287	94.48	2.4182	0.03949	0.774	94.96
		0.4022	0.00113	0.1286	94.58	2.4182	0.03949	0.7735	95.1
	α_2	0.3017	0.00062	0.0965	94.86	2.3117	0.03587	0.7399	95.04
		0.3017	0.00062	0.0965	94.84	2.3117	0.03586	0.7395	94.94
	α_3	0.5032	0.0017	0.1611	95.26	2.5145	0.04243	0.8048	94.94
		0.5032	0.0017	0.161	95.36	2.5145	0.04243	0.8043	94.92

Table 3 presents the results of the simulation study for sample sizes $n = 50, 75, 100$, and 150, with $M = 5000$, using two sets of model parameter values. Based on the results in Table 3, we observe the following trends for each parameter:

1. **APE approaches actual value:** As the sample size increases, both the average MLE and Bayesian estimate (APE) converge toward the true parameter value.
2. **MSE decreases:** The mean squared error (MSE) associated with the point estimates, using both MLE and Bayesian estimates, decreases with increasing sample size, indicating improved estimation accuracy.
3. **ALCI decreases:** The average length of the confidence/credible interval (ALCI) decreases as the sample size increases, reflecting greater precision in the intervals.
4. **CP aligns with nominal confidence level:** The coverage probability (CP), using either method, remains very close to the nominal confidence level, confirming the reliability of the intervals.
5. **No significant difference between methods:** There is no significant difference between the results obtained from the maximum likelihood and Bayesian methods.

8. Applications

To demonstrate the applicability of the proposed model, this section focuses on the analysis of three distinct real data sets: two bivariate data and one dependent competing risks data. The primary objective is to illustrate the practical application and effectiveness of the proposed model in a real-world context. By examining these data sets, we aim to demonstrate how the proposed model can be applied to address complex scenarios and provide valuable insights in practical settings.

8.1. Bivariate data

First bivariate data set (UEFA Champions League): The UEFA Champions League data, sourced from Meintanis [21] and displayed in Table 4, pertains to soccer matches. This data set captures instances where the home team scores at least one goal, as well as instances where a goal is scored directly from a penalty kick, foul kick, or any other direct kick (referred to as kick goals) by any team under consideration. In this data set, the variable X_1 represents the time in minutes when the first kick goal is scored by any team, while X_2 represents the time in minutes when the first goal of any type is scored by the home team.

The initial step in analyzing this dataset involved constructing a non-parametric and parametric scaled TTT-transform plot, using the EGIE distribution, for the marginal variables X_1 and X_2, as illustrated in Figure 5.

The visual analysis of the plot in Figure 5 leads to the inference that (for illustration see Aarset [1]) the hazard rate initially exhibits an upward trend and subsequently experiences a decline as time progresses or aging occurs. Consequently, considering the discerned pattern in the hazard rate, it is reasonable to anticipate that the EGIE distribution would be well-suited for modeling the marginal distribution of the data.

Table 4. UEFA Champion's League data.

2005-2006	X_1	X_2	2004-2005	X_1	X_2
Lyon-Real Madrid	26	20	Internazionale-Bremen	34	34
Milan-Fenerbahce	63	18	Real Madrid-Roma	53	39
Chelsea-Anderlecht	19	19	Man. United-Fenerbahce	54	7
Club Brugge-Juventus	66	85	Bayern-Ajax	51	28
Fenerbahce-PSV	40	40	Moscow-PSG	76	64
Internazionale-Rangers	49	49	Barcelona-Shakhtar	64	15
Panathinaikos-Bremen	8	8	Leverkusen-Roma	26	48
Ajax-Arsenal	69	71	Arsenal-Panathinaikos	16	16
Man. United-Benfica	39	39	Dynamo Kyiv-Real Madrid	44	13
Real Madrid-Rosenborg	82	48	Man. United-Sparta	25	14
Villarreal-Benfica	72	72	Bayern-M. Tel Aviv	55	11
Juventus-Bayern	66	62	Bremen-Internazionale	49	49
Club Brugge-Rapid	25	9	Anderlecht-Valencia	24	24
Olympiacos-Lyon	41	3	Panathinaikos-PSV	44	30
Internazionale-Porto	16	75	Arsenal-Rosenborg	42	3
Schalke-PSV	18	18	Liverpool-Olympiacos	27	47
Barcelona-Bremen	22	14	M. Tel Aviv-Juventus	28	28
Milan-Schalke	42	42	Bremen-Panathinaikos	2	2
Rapid–Juventus	36	52			

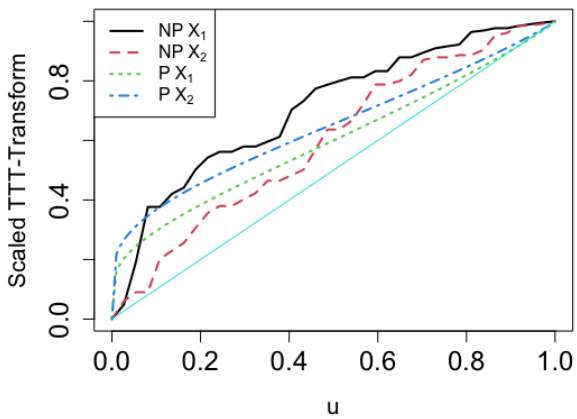

Figure 5. The scaled TTT-transform plot for the marginal data X_1 and X_2 and the fitted marginal distributions using UEFA data. NP means non-parametric, while P means parametric.

To facilitate a thorough comparison between the EGIE distribution and its sub-model distributions, to fit the marginal data, we conducted the following two hypotheses:

1) X_1 and/or X_2 follows IE vs they follow EGIE, by testing $H_0 : \gamma = 1$ vs $H_1 : \gamma \neq 1$.
2) X_1 and/or X_2 follows GIE vs they follow EGIE, by testing $H_0 : \lambda = 1$ vs $H_1 : \lambda \neq 1$.

To accomplish this objective, the likelihood ratio test statistic, $\Lambda = -2(\mathcal{L}_0 - \mathcal{L}_1)$, is used. Table 5 displays the model, null hypothesis, MLE, negative log-likelihood value, the likelihood ratio statistic value, degrees of freedom (df), and p-value using marginal data.

Table 5. Statistical analysis of the marginal data values X_1 and X_2.

Model	H_0	MLE	$-\mathcal{L}$	Λ	df	p-value
		X_1				
IE	$\gamma = 1$	$\hat{\lambda} = 10.913$, $\hat{\alpha} = 2.0783$	183.184	9.367	1	0.00221
GIE	$\lambda = 1$	$\hat{\gamma} = 1.017$, $\hat{\alpha} = 23.773$	183.173	9.344	1	0.00224
EGIE	–	$\hat{\lambda} = 72.663$, $\hat{\gamma} = 2.999$, $\hat{\alpha} = 0.4372$	178.501	–	–	–
		X_2				
IE	$\gamma = 1$	$\hat{\lambda} = 10.033$, $\hat{\alpha} = 1.374$	173.897	3.730219	1	0.05343
GIE	$\lambda = 1$	$\hat{\gamma} = 0.958$, $\hat{\alpha} = 12.508$	173.838	3.61268	1	0.05734
EGIE	–	$\hat{\lambda} = 71.641$, $\hat{\gamma} = 2.5116$, $\hat{\alpha} = 0.229$	172.032		–	–

From Table 5, it can be concluded that for both X_1 and X_2, at any significance level greater than or equal to 0.00224(0.0573), the IE and GIE distributions are rejected in favor of the EGIE distribution as a better fit for the marginal data. This suggests that the BEGIE distribution may be more appropriate for fitting the current bivariate data set. To further investigate, the BGIE, BIE, and BEGIE distributions were applied, with the results summarized in Table 6. Based on these findings, there is significant evidence to reject both the BIE and BGIE distributions in favor of the BEGIE distribution for fitting the current dataset.

Table 6. Statistical results for the UEFA Champion's League data.

Model	H_0	MLE	$-\mathcal{L}$	Λ	df	p-value
BIE	$\gamma = 1$	$\hat{\lambda} = 70.211$, $\hat{\alpha}_1 = 0.177$, $\hat{\alpha}_2 = 0.057$, $\hat{\alpha}_3 = 0.148$	318.3554	6.429	1	0.0112
BGIE	$\lambda = 1$	$\hat{\gamma} = 0.938$, $\hat{\alpha}_1 = 10.453$, $\hat{\alpha}_2 = 3.443$, $\hat{\alpha}_3 = 9.079$	318.1551	6.028	1	0.0141
BEGIE	–	$\hat{\lambda} = 70.311$ $\hat{\gamma} = 2.099$, $\hat{\alpha}_1 = 0.223$, $\hat{\alpha}_2 = 0.066$, $\hat{\alpha}_3 = 0.169$	315.1409	–	–	–

Moreover, we compare the goodness of fit of the BEGIE model with the BFGMTLEE and BIGE distributions, whose joint PDFs are provided in Appendix A. This comparison is conducted using the

Akaike information criterion (AIC). Table 7 displays the MLE, $-\mathcal{L}$, and AIC values for BFGMTLEE and BIGE. The AIC for the BEGIE is 640.282. Based on the AIC values, we can conclude that the BEGIE model provides a better fit for the current data set than both the BFGMTLEE and BIGE models.

Table 7. MLE, negative log-likelihood value, and AIC for the BFGMTLEE and BIGE models, using the UEFA Champion's League data.

Model	MLE	$-\mathcal{L}$	AIC
BFGMTLE	$\hat{\alpha}_1 = 3.27, \hat{\alpha}_2 = 1.85, \hat{\lambda}_1 = 0.026, \hat{\lambda}_2 = 0.022, \hat{\delta} = 0.99$	372.397	754.794
BIGE	$\hat{\lambda} = 2.1943, \hat{\alpha}_1 = 0.3582, \hat{\alpha}_2 = 0.8402, \hat{\alpha}_3 = 0.8902$	419.006	846.012

We also used the Bayesian approach to estimate the five parameters of the BEGIE distribution by adopting two sets of the hyperparameters for both the two types of prior distributions (gamma on the original parameters and normal on the log-transformed ones), as shown in Table 8.

Table 8. Sets of hyperparameter values of gamma and normal prior distributions.

Set	$\theta_i \sim \text{Gamma}(a_i, b_i)$	$\phi_i = \log\theta_i \sim N(\mu_i, \sigma_i), i = 1, 2, \cdots, 5$	$E(\theta_i)$	$\text{Var}(\theta_i)$
I	$a_i = b_i = 0.001$	$\mu_i = -.0445, \sigma_i = 0.77$	1	1000
II	$a_1 = 20973, b_1 = 289$	$\phi_1 = \log\lambda \sim N(4.284, 0.012)$	72.57	0.251
	$a_2 = 16, b_2 = 7$	$\phi_2 = \log\gamma \sim N(0.9196, 0.020)$	2.286	0.327
	$a_3 = 13, b_3 = 60$	$\phi_3 = \log\alpha_1 \sim N(-1.641, 0.225)$	0.217	0.004
	$a_4 = 8, b_4 = 121$	$\phi_2 = \log\alpha_2 \sim N(-2.629, 0.401)$	0.066	0.001
	$a_5 = 21, b_5 = 123$	$\phi_3 = \log\alpha_3 \sim N(-1.817, 0.216)$	0.171	0.001

Under the above two sets of hyperparameters and the two types of prior distributions, we conducted 50,000 random draws from the joint posterior distribution using the MCMC and SIR algorithms. For the MCMC, we used the multivariate normal distribution with mean vector equal to the MLE of the logarithm of the vector of the five unknown parameters and variance-covariance matrix equal to 2 times the Fisher information matrix associated with the MLE. Subsequently, we approximated the posterior mean, median, and 95% credible interval "the 2.5th and 97.5th percentiles" for each parameter. The results are shown in Table 9.

As diagnostic tests for the random draws using the MCMC and SIR methods from the posterior distribution, we present trace plots for 2,000 draws out of the 50,000 generated draws using two sets of hyperparameters, as shown in Figures 6 and 7. These figures demonstrate well-mixed draws, indicating that the samples are randomly generated from the posterior distribution. Additionally, the trace plots suggest that the acceptance rate is higher for Set II of hyperparameters compared to Set I.

Utilizing the draws generated through the MCMC and SIR algorithms, we present histograms that illustrate the marginal posterior distributions of the model parameters as shown in Figures 8–11. The figures suggest that the normal prior distribution, particularly in scenarios with limited prior information about the parameters, does not perform as effectively as when more substantial prior information is available. However, this discrepancy is not observed when using the gamma prior distribution.

Table 9. Bayesian results of the model parameters using different sets of the hyperparameters for the gamma and normal prior distributions, using the MCMC and SIR algorithms.

Algorithm	Hyper	Statistic	λ	γ	α_1	α_2	α_3
MCMC		Gamma prior for the parameters					
	Set I	2.5^{th} percentile	71.523	1.214	0.122	0.025	0.101
		Median	72.710	2.109	0.213	0.061	0.159
		Mean	72.675	2.243	0.215	0.063	0.163
		97.5^{th} percentile	73.791	3.530	0.329	0.116	0.240
	Set II	2.5^{th} percentile	71.619	1.501	0.148	0.036	0.120
		Median	72.800	2.134	0.215	0.065	0.161
		Mean	72.807	2.209	0.217	0.065	0.164
		97.5^{th} percentile	73.930	3.054	0.293	0.099	0.215
		Normal prior for the logarithm of parameters					
	Set I	2.5^{th} percentile	70.972	2.421	0.142	0.033	0.109
		Median	72.576	2.512	0.234	0.071	0.164
		Mean	72.538	2.509	0.237	0.078	0.166
		97.5^{th} percentile	74.020	2.586	0.350	0.141	0.248
	Set II	2.5^{th} percentile	72.218	2.485	0.167	0.036	0.122
		Median	72.541	2.508	0.235	0.072	0.167
		Mean	72.542	2.508	0.238	0.074	0.169
		97.5^{th} percentile	72.860	2.532	0.327	0.123	0.227
SIR		Gamma prior for the parameters					
	Set I	2.5^{th} percentile	71.148	1.182	0.125	0.024	0.102
		Median	72.853	2.245	0.218	0.061	0.166
		Mean	72.819	2.260	0.219	0.063	0.168
		97.5^{th} percentile	74.295	3.538	0.327	0.118	0.256
	Set II	2.5^{th} percentile	71.553	1.486	0.148	0.034	0.118
		Median	72.768	2.227	0.214	0.062	0.163
		Mean	72.744	2.250	0.217	0.063	0.165
		97.5^{th} percentile	73.909	3.137	0.298	0.100	0.217
		Normal prior for the logarithm of parameters					
	Set I	2.5^{th} percentile	67.040	2.401	0.175	0.027	0.118
		Median	70.540	2.661	0.313	0.049	0.196
		Mean	70.199	2.622	0.280	0.056	0.176
		97.5^{th} percentile	74.371	2.715	0.349	0.113	0.222
	Set II	2.5^{th} percentile	72.212	2.485	0.167	0.036	0.121
		Median	72.542	2.508	0.235	0.072	0.169
		Mean	72.536	2.508	0.237	0.075	0.170
		97.5^{th} percentile	72.847	2.533	0.319	0.121	0.225

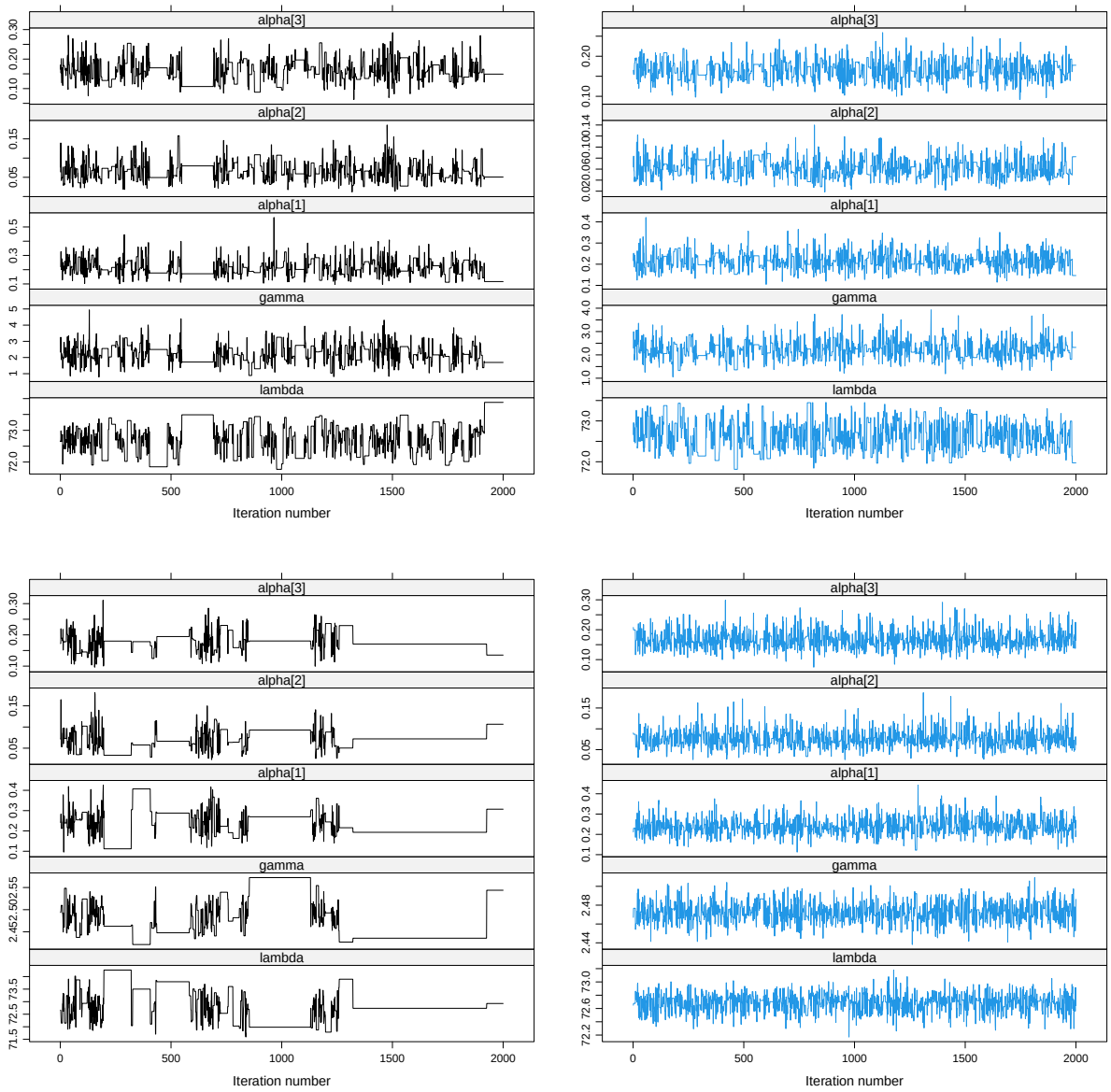

Figure 6. Trace plots for the early 2000 random draws from the posterior distribution using gamma prior (top row) and normal prior (bottom row) using Set I (left) and Set II (right) of the hyperparameters.

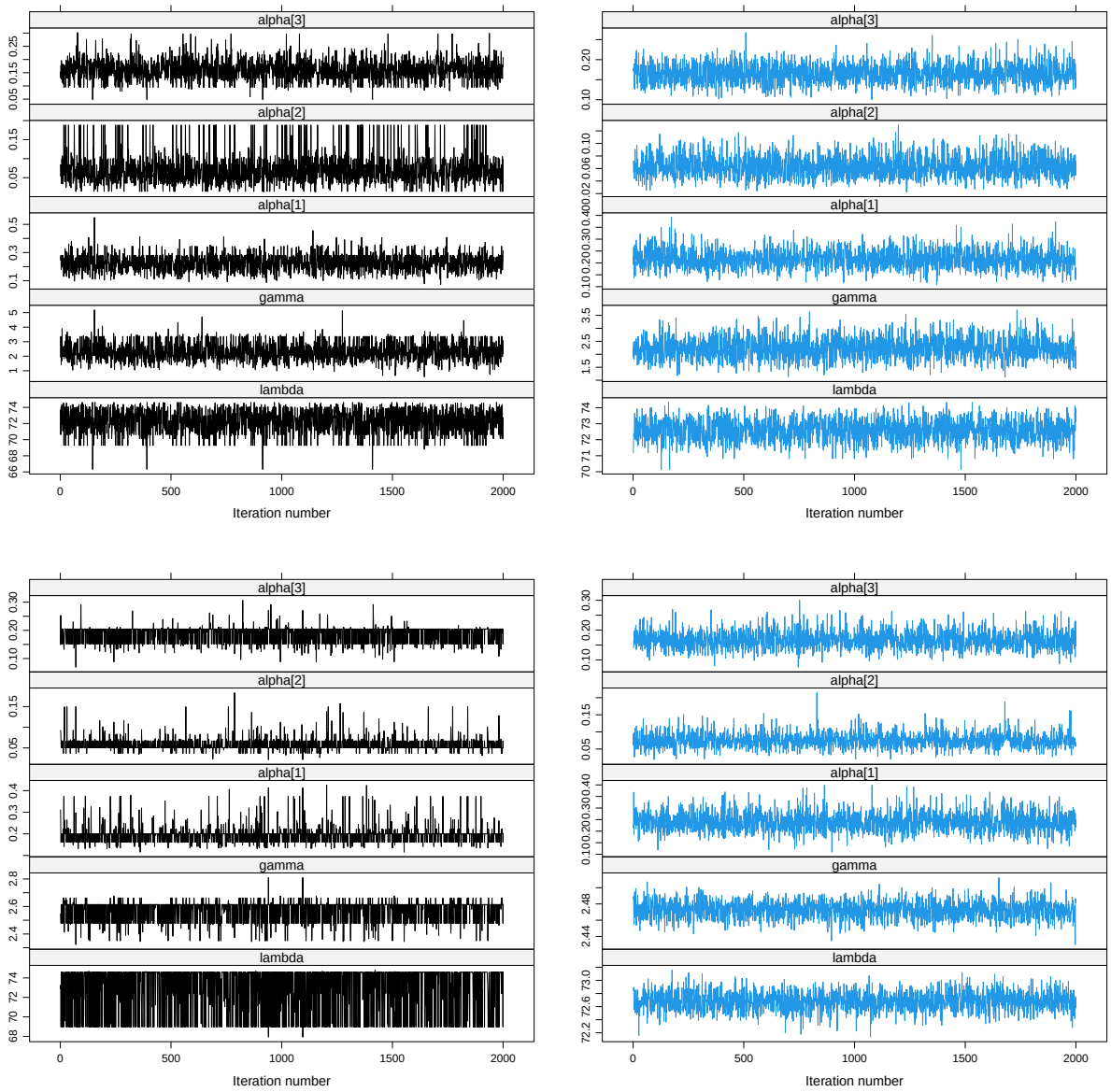

Figure 7. Trace plots for the early 2000 SIR random draws from the posterior distribution using gamma prior (top row) and normal prior (bottom row) using Set I (left) and Set II (right) of the hyperparameters.

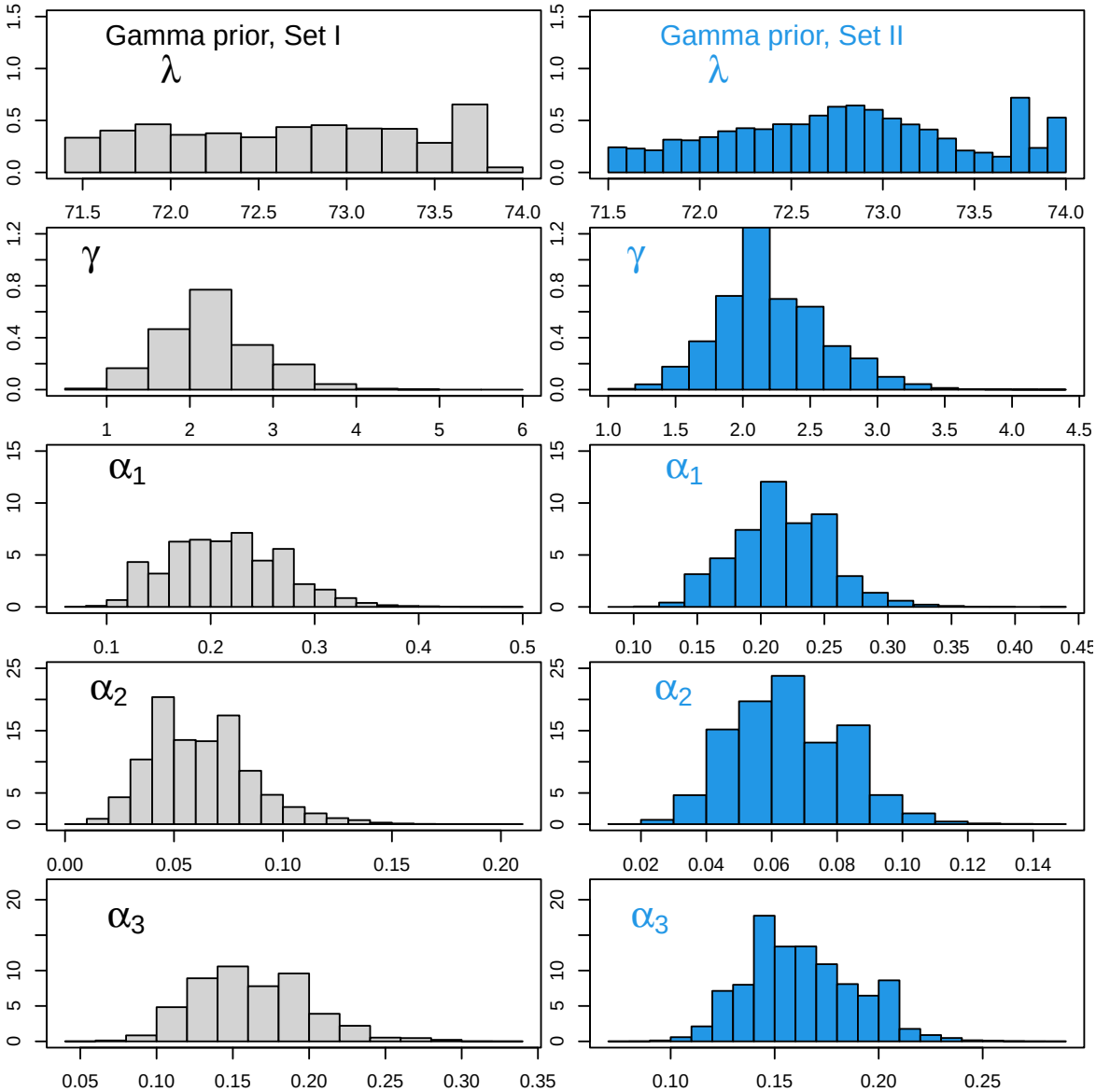

Figure 8. Histogram plots for the marginal distributions of the model parameters using the MCMC algorithm assuming gamma prior distribution of the parameters for two sets of hyperparameters, Set I (left) and Set II (right).

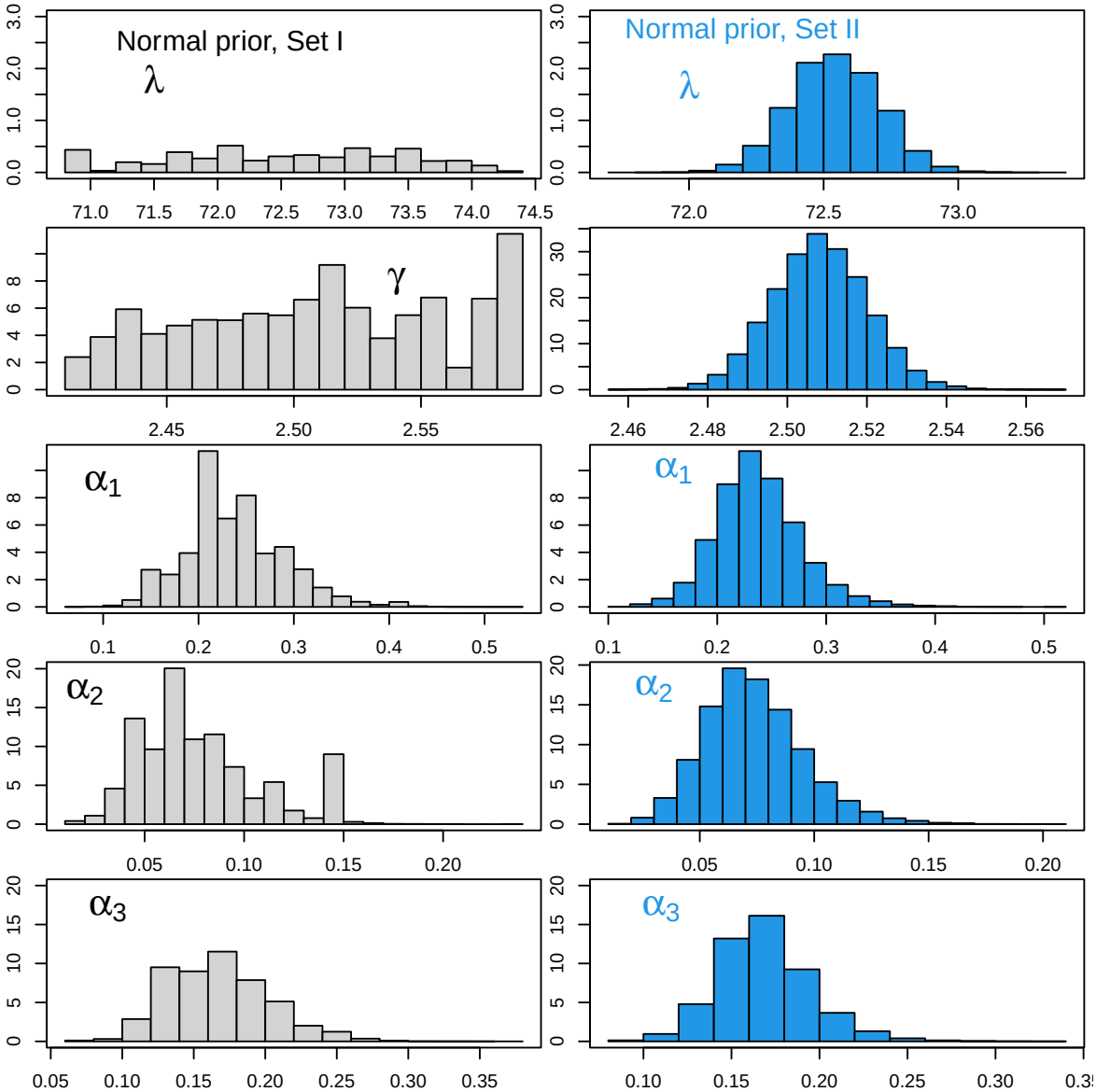

Figure 9. Histogram plots for the marginal distributions of the model parameters using the MCMC algorithm assuming normal prior distribution of the log parameters for two sets of hyperparameters, Set I (left) and Set II (right).

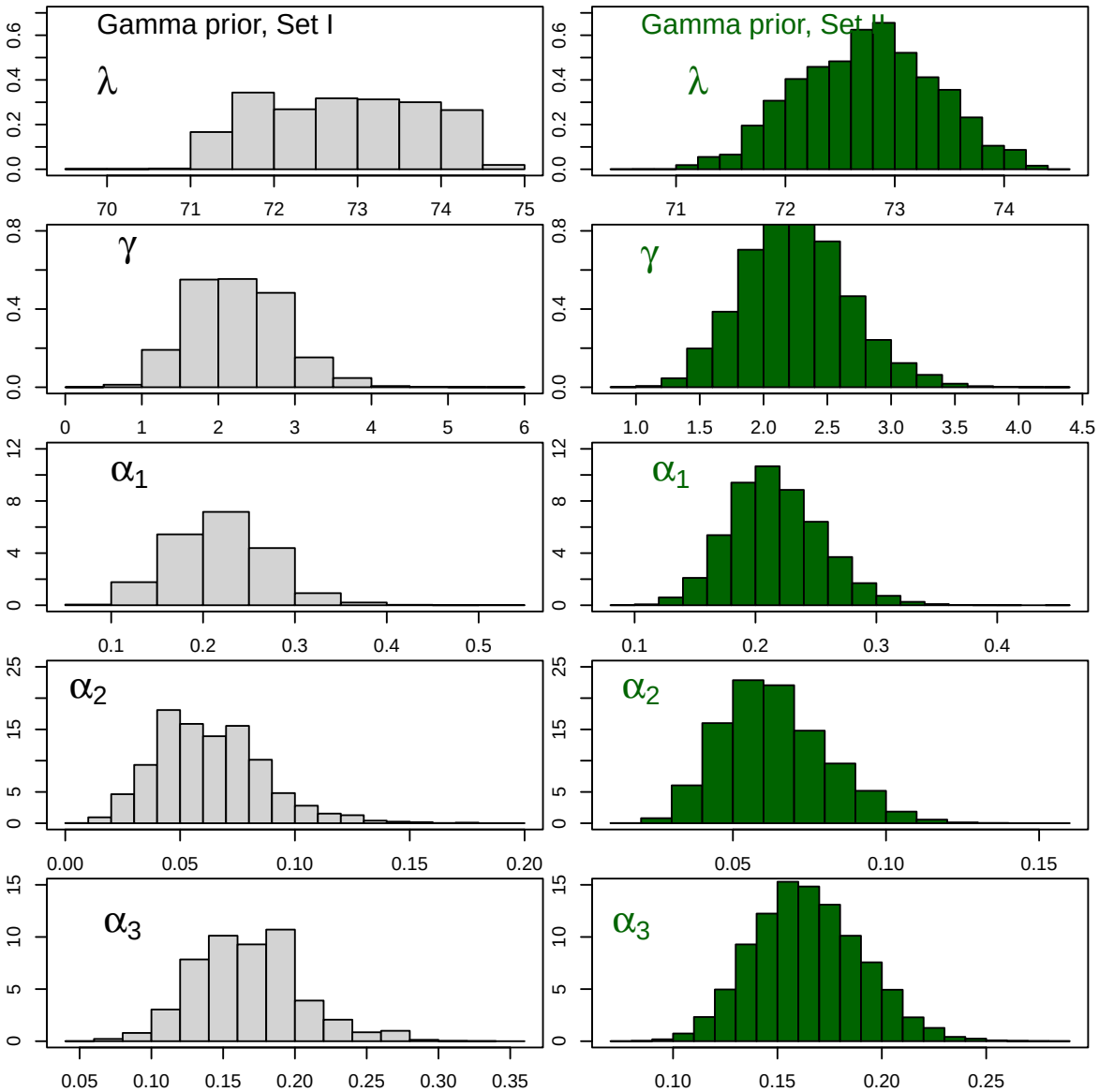

Figure 10. Histogram plots for the marginal distributions of the model parameters using the SIR algorithm under the assumption of gamma prior distribution of the parameters for two sets of hyperparameters, Set I (left) and Set II (right).

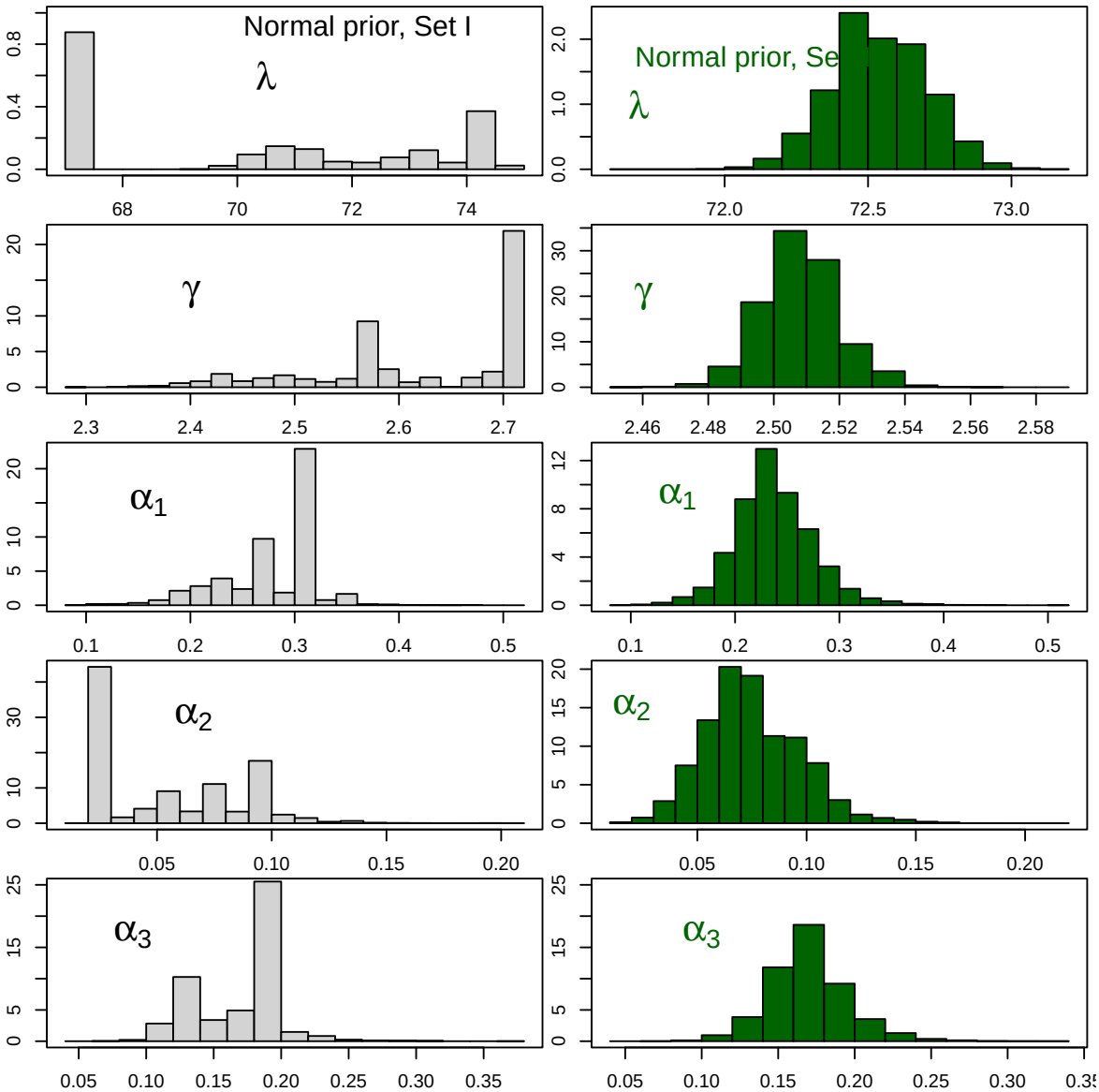

Figure 11. Histogram plots for the marginal distributions of the model parameters using the SIR algorithm under the assumption of normal prior distribution of the log parameters for two sets of hyperparameters, Set I (left) and Set II (right).

Second bivariate data set (cholesterol data set): This dataset includes cholesterol levels measured at 5 and 25 weeks after treatment in 30 patients. Shoaee [32] applied three bivariate distributions of the Marshall-Olkin type to model this data: the bivariate Weibull (BWG), bivariate Chen (BCHG), and bivariate Gompertz (BGG) distributions. For computational reasons, Shoaee [32] transformed the marginal data values using the formula $\frac{(X-150)}{100}$ before conducting the analysis. The transformed data are presented in Table 10.

Table 10. The transformed cholesterol levels of 30 patients after 5 and 25 weeks of treatment.

5th	25th	5th	25th	5th	25th	5th	25th	5th	25th	5th	25th
1.75	0.96	0.42	0.42	1.59	0.82	0.65	1.11	0.75	0.75	0.48	0.48
0.64	1.24	0.93	0.95	0.93	0.97	1.65	1.05	1.28	0.95	1.26	1.75
1.37	1.37	0.67	1.02	1.37	0.58	0.79	0.29	1.03	0.59	1.55	1.22
1.65	1.33	1.13	0.65	1.07	0.62	1.12	1.44	1.54	0.95	0.98	1.55
0.83	0.67	1.60	2.02	1.66	1.33	0.47	0.47	0.55	0.55	0.60	1.21

Based on AIC statistics, Shoaee [32] reported that the BCHG model is the best fit for this dataset. In the current work, we use the BEGIE model to reanalyze this data and compare its performance with the aforementioned three models, as well as the BFGMTLEE model by Al-Shomrani [4]. Table 11 presents the MLE, negative log-likelihood values, and AIC for the five models used to analyze this dataset. It is important to note that the results for the BWG, BCHG, and BGG models are sourced from Shoaee [32]. Based on the AIC criterion, the BEGIE model provides a better fit for this data than the other four models.

Table 11. MLE, negative log-likelihood values, and AIC based on the cholesterol data set.

Model	MLE	$-\mathcal{L}$	AIC
BFGMTLE	$\hat{\alpha}_1 = 10.783$, $\hat{\alpha}_2 = 8.876$, $\hat{\lambda}_1 = 1.441$, $\hat{\lambda}_2 = 1.443$, $\hat{\delta} = 0.990$	64.046	138.092
BWG	$\hat{\alpha} = 2.898$, $\hat{\lambda}_1 = 0.134$, $\hat{\lambda}_1 = 0.272$, $\hat{\lambda}_2 = 0.307$, $\hat{\theta} = 0.300$	47.345	104.689
BCHG	$\hat{\alpha} = 1.581$, $\hat{\lambda}_1 = 0.039$, $\hat{\lambda}_1 = 0.084$, $\hat{\lambda}_2 = 0.089$, $\hat{\theta} = 0.282$	47.322	104.665
BGG	$\hat{\alpha} = 2.034$, $\hat{\lambda}_1 = 0.019$, $\hat{\lambda}_1 = 0.041$, $\hat{\lambda}_2 = 0.040$, $\hat{\theta} = 0.683$	55.940	128.885
BEGIE	$\hat{\lambda} = 9.778$, $\hat{\gamma} = 334.075$, $\hat{\alpha}_1 = 0.130$, $\hat{\alpha}_2 = 0.093$, $\hat{\alpha}_3 = 0.067$	46.615	103.230

A Bayesian analysis can be conducted similarly to how we approached the first dataset. However, we have not included the Bayesian component for the second dataset due to the appropriate length considerations.

8.2. Diabetic retinopathy application

Diabetic retinopathy is a significant ocular condition capable of inducing vision loss and blindness in individuals with diabetes. It affects the blood vessels within the retina, which is the light-sensitive tissue layer at the back of the eye. The National Eye Institute in Bethesda, Maryland, conducted an experiment involving 71 patients to investigate the impact of laser treatment on reducing the risk of blindness associated with diabetic retinopathy. In this study, laser treatment was administered to a randomly selected eye for each patient. The study recorded the time to blindness and whether one or both eyes became blind. The recorded data are presented in Table 12. This data was analyzed by Sarhan, Apaloo, and Kundu [27] and Manshi, Sarhan, and Smith [20]. For more information about diabetic retinopathy, we refer the readers to [23].

Table 12. Diabetic retinopathy data.

i	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
t_i	266	91	154	285	583	547	79	622	707	469	93	1313
δ_i	1	2	2	0	1	2	1	0	2	2	1	2
i	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
t_i	805	344	790	125	777	306	415	307	637	577	178	517
δ_i	1	1	2	2	2	1	1	2	2	2	1	2
i	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**
t_i	272	1137	1484	315	287	1252	717	642	141	407	356	1653
δ_i	0	0	1	1	2	1	2	1	2	1	1	0
i	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**
t_i	427	699	36	667	588	471	126	350	350	663	567	966
δ_i	2	1	2	1	2	0	1	2	1	0	2	0
i	**49**	**50**	**51**	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**
t_i	203	84	392	1140	901	1247	448	904	276	520	485	248
δ_i	0	1	1	2	1	0	2	2	1	1	2	2
i	**61**	**62**	**63**	**64**	**65**	**66**	**67**	**68**	**69**	**70**	**71**	
t_i	503	423	285	315	727	210	409	584	355	1302	227	
δ_i	1	2	2	2	2	2	2	1	1	1	2	

The primary objective of this experiment was to assess whether the laser treatment could effectively delay the onset of blindness in individuals with diabetic retinopathy. The dataset derived from this study can be treated as one involving dependent competing risks data, with two potential causes of failure (blindness).

We utilize the BEGIE distribution along with its sub-models (BIE and BGIE) to perform fitting on this dataset, followed by a subsequent comparison. To facilitate this comparative analysis, Table 13 provides the MLEs, corresponding log-likelihood functions, and Akaike information criterion (AIC) for all the aforementioned models. Additionally, the likelihood ratio test (LRT) statistic and corresponding p-values for the sub-models are also included in the table. The outcomes presented in Table 13 lead us to the conclusion that the sub-models BIE and BGIE distributions exhibit statistical significance in their rejection, favoring the BEGIE distribution instead, as evidenced by their considerably small p-values.

Table 13. Statistical analysis of the diabetic retinopathy data.

Model	MLE	$\mathcal{L}$	AIC	Λ	P-value
BIE	$\hat{\lambda} = 5.278, \hat{a}_1 = 0.692, \hat{a}_2 = 0.562,$ $\hat{a}_3 = 1.774$	−590.49	1188.98	5.319	0.0211
BGIE	$\hat{\gamma} = 1.03 \times 10^{-5}$, $\hat{a}_1 = 5.969, \hat{a}_2 = 5.839,$ $\hat{a}_3 = 7.053$	−592.04	1188.98	8.416	0.0037
BEGIE	$\hat{\lambda} = 1307.93, \hat{\gamma} = 1.464, \hat{a}_1 = 0.330,$ $\hat{a}_2 = 0.283, \hat{a}_3 = 0.976$	−587.83	1185.66	–	–

To ascertain whether the laser treatment significantly contributes to delaying blindness among individuals with Diabetic Retinopathy, a straightforward hypothesis testing approach can be employed. Previous studies conducted by Sarhan, Apaloo, and Kundu [27] and Manshi, Sarhan, and Smith [20] concluded that the laser treatment does not yield a substantial impact in delaying blindness in these patients. In this study, we address the same question by utilizing the BEGIE distribution. To this end, we consider the following hypotheses for testing:

Null Hypothesis (H_0): The laser surgery does not exhibit a significant effect in delaying blindness ($\alpha_1 = \alpha_2$).

Alternative Hypothesis (H_1): The laser surgery has a meaningful impact in delaying blindness ($\alpha_1 \neq \alpha_2$).

We can employ the likelihood ratio test to assess the aforementioned hypotheses, defined as $\Lambda = -2(\hat{\mathcal{L}}_{02} - \hat{\mathcal{L}}_2)$, which conforms to a chi-square distribution with one degree of freedom. In this context, $\hat{\mathcal{L}}_{02}$ and $\hat{\mathcal{L}}_2$ signify the maximum values of the log-likelihood functions under H_0 and H_1, respectively.

Deriving the log-likelihood function under H_0 involves a special case of Eq (5.8), where we assume $\alpha_1 = \alpha_2 = \alpha$, thus yielding:

$$\begin{aligned}\mathcal{L}_{02}(\theta) &= (n_1 + n_2)\log\alpha + n_3\log\alpha_3 + n\log(\lambda\gamma) - 2\sum_{i=1}^{n}\log t_i - \sum_{i=1}^{n}\frac{\lambda}{t_i} \\ &\quad +(\gamma-1)\sum_{i=1}^{n}\log\left(1-e^{-\frac{\lambda}{t_i}}\right) + (2\alpha+\alpha_3-1)\sum_{i=1}^{n}\log\left[1-\left(1-e^{-\frac{\lambda}{t_i}}\right)^{\gamma}\right]. \qquad (8.1)\end{aligned}$$

Under the null hypothesis, the MLEs of the unknown parameters are as follows: $\hat{\lambda} = 1308.073$, $\hat{\gamma} = 1.382$, $\hat{\alpha}_1 = \hat{\alpha}_2 = 0.301$, and $\hat{\alpha}_3 = 0.982$. The corresponding log-likelihood value is calculated as -588.06. Consequently, the test statistic value is determined as $\Lambda = -2(-588.06 + 587.83) = 0.46$, and the resulting p-value is remarkably small at 0.498. This leads us to the conclusion that the null hypothesis (H_0) cannot be significantly rejected in favor of the alternative hypothesis (H_1). Consequently, we can confidently affirm that the laser treatment does not exhibit a significant effectiveness in delaying the onset of blindness.

9. Conclusions

In this study, we harnessed the fatal shock model concept to introduce a fresh bivariate distribution in the realm of the Marshall-Olkin type. This distribution incorporates three independent univariate random variables following the exponentiated generalized inverted exponential distribution. Termed the "bivariate exponentiated generalized inverted exponential distribution" (BEGIE), this novel proposal extends the scope of both the bivariate inverted exponential (BIE) and bivariate exponentiated inverted exponential (BGIE) distributions. We explore several statistical properties inherent to this newly introduced distribution.

The versatility of the BEGIE distribution is demonstrated through its application in modeling dependent competing risks data and bivariate data scenarios. Estimation of the unknown parameters within the BEGIE distribution was achieved via the maximum likelihood method, effectively applied to both bivariate and dependent competing risks datasets.

A Bayesian technique is applied on the dependent competing risks data. Additionally, the performance of the BEGIE distribution was evaluated using real datasets, comparing its effectiveness

against the BIE and BGIE distributions. The results prominently indicated the BEGIE distribution's superior fit for these datasets.

Looking ahead, we are considering the utilization of step-stress acceleration life test plans or progressively type-II censored data as part of our future research trajectory. These approaches will aid in estimating BEGIE distribution parameters, particularly in scenarios involving bivariate data or dependent competing risks data.

Author contributions

A.M.S.: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing-original draft preparation, Writing-review and editing, Visualization; R.S.G.: Methodology, Validation, Resources, Writing-original draft preparation, Writing-review and editing; N.A.: Formal analysis, Investigation; A.M.M.: Methodology, Writing-original draft preparation, Writing-review and editing. All authors have read and approved the final version of the manuscript for publication.

Acknowledgments

The authors would like to thank the five external experts and the editor for their valuable comments, which have significantly enhanced the overall quality and presentation of the manuscript. This research is supported by researchers supporting project number (RSPD2024R548), King Saud University, Riyadh, Saudi Arabia.

Conflict of interest

The authors declare no conflict of interest.

References

1. M. V. Aarset, How to identify a bathtub hazard rate, *IEEE Trans. Reliab.*, **R-36** (1987), 106–108. https://doi.org/10.1109/TR.1987.5222310
2. A. M. Abouammoh, A. M. Alshingiti, Reliability estimation of generalized inverted exponential distribution, *J. Stat. Comput. Simul.*, **79** (2009), 1301–1315. http://doi.org/10.1080/00949650802261095
3. J. Albert, *Bayesian computation with R*, 2 Eds., New York: Springer, 2009. https://doi.org/10.1007/978-0-387-92298-0
4. A. A. Al-Shomrani, New bivariate family of distributions based on any copula function: statistical properties, *Heliyon*, **9** (2023), e15160. https://doi.org/10.1016/j.heliyon.2023.e15160
5. F. A. Alqallaf, D. Kundu, A bivariate inverse generalized exponential distribution and its applications in dependent competing risks model, *Commun. Stat.-Simul. Comput.*, **51** (2022), 7019–7036. https://doi.org/10.1080/03610918.2020.1821888
6. T. M. Cover, J. A. Thomas, *Elements of information theory*, 2 Eds., Wiley, 2012. https://doi.org/10.1002/047174882X

7. K. Chen, Z. Jin, Z. Ying, Semiparametric analysis of transformation models with censored data, *Biometrika*, **89** (2002), 659–668. https://doi.org/10.1093/biomet/89.3.659

8. W. X. Chen, C. X. Long, R. Yang, D. S. Yao, Maximum likelihood estimator of the location parameter under moving extremes ranked set sampling design, *Acta Math. Appl. Sin. Engl. Ser.*, **37** (2021), 101–108. https://doi.org/10.1007/s10255-021-0998-8

9. S. Dey, Inverted exponential distribution as a life distribution model from a Bayesian viewpoint, *Data Science Journal*, **6** (2007), 107–113. https://doi.org/10.2481/dsj.6.107

10. A. El-Gohary, A. M. Sarhan, The Distributions of sums, products, ratios and differences for Marshall-Olkin Bivariate exponential distribution, *International Journal of Applied Mathematics*, **20** (2007), 351.

11. C. Genest, K. Ghoudi, L.-P. Rivest, A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, *Biometrika*, **82** (1995), 543–552. https://doi.org/10.1093/biomet/82.3.543

12. F. Kawsar, S. P. Ahmad, The exponentiated inverted exponential distribution, *Journal of Applied Information Science*, **5** (2017), 35–41.

13. S. Kotz, N. Balakrishnan, N. L. Johnson, *Continuous multivariate distributions: models and applications*, 2 Eds., New York: Wiley, 2000. https://doi.org/10.1002/0471722065

14. D. Kundu, A. K. Dey, Estimating the parameters of the Marshall Olkin bivariate Weibull distribution by EM algorithm, *Comput. Stat. Data Anal.*, **53** (2009), 956–965. https://doi.org/10.1016/j.csda.2008.11.009

15. D. Kundu, A. K. Gupta, Bayes estimation for the Marshall-Olkin bivariate Weibull distribution, *Comput. Stat. Data Anal.*, **57** (2013), 271–281. https://doi.org/10.1016/j.csda.2012.06.002

16. D. Kundu, A. M. Sarhan, R. D. Gupta, On Sarhan-Balakrishnan bivariate distribution, *Journal of Statistics Applications & Probability*, **1** (2012), 163–170. https://doi.org/10.12785/jsap/010302

17. E. L. Lehmann, Some concepts of dependence, *Ann. Math. Statist.*, **37** (1966), 1137–1153. https://doi.org/10.1214/aoms/1177699260

18. Y. Li, J. Sun, S. Song, Statistical analysis of bivariate failure time data with Marshall-Olkin Weibull models, *Comput. Stat. Data Anal.*, **56** (2012), 2041–2050. https://doi.org/10.1016/j.csda.2011.12.010

19. A. W. Marshall, I. Olkin, A multivariate exponential distribution, *J. Amer. Stat. Assoc.*, **62** (1967), 30–44. https://doi.org/10.1080/01621459.1967.10482885

20. T. Manshi, A. M. Sarhan, B. Smith, A new bivariate distribution with applications on dependent competing risks data, *International Journal of Statistics and Probability*, **12** (2023), 27–39. https://doi.org/10.5539/ijsp.v12n3p27

21. S. G. Meintanis, Test of fit for marshall-olkin distributions with applications, *J. Stat. Plan. Infer.*, **137** (2007), 3954–3963. https://doi.org/10.1016/j.jspi.2007.04.013

22. R. B. Nelsen, *An introduction to copulas*, 2 Eds., New York: Springer, 2006. https://doi.org/10.1007/0-387-28678-0

23. National Eye Institute, Diabetic retinopathy, 2024. Available from: https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/diabetic-retinopathy.

24. P. E. Oguntunde, A. O. Adejumo, O. S. Balogun, Statistical properties of the exponentiated generalized inverted exponential distribution, *Applied Mathematics*, **4** (2014), 47–55. https://doi.org/10.5923/j.am.20140402.02

25. E. A. Pena, A. K. Gupta, Bayes estimation for the Marshall-Olkin exponential distribution, *J. Roy. Stat. Soc. B*, **52** (1990), 379–389. https://doi.org/10.1111/j.2517-6161.1990.tb01794.x

26. A. M. Sarhan, The bivariate generalized Rayleigh distribution, *Journal of Mathematical Sciences and Modelling*, **2** (2019), 99–111. https://doi.org/10.33187/jmsm.439873

27. A. M. Sarhan, J. Apaloo, D. Kundu, A new bivariate lifetime distribution: properties, estimations and its extension, *Commun. Stat.-Simul. Comput.*, **53** (2024), 879–896. https://doi.org/10.1080/03610918.2022.2034866

28. A. M. Sarhan, N. Balakrishnan, A new class of bivariate distributions and its mixture, *J. Multivariate Anal.*, **98** (2007), 1508–1527. https://doi.org/10.1016/j.jmva.2006.07.007

29. A. M. Sarhan, A. El-Gohary, A. H. El-Bassiouny, N. Balakrishnan, A class of bivariate linear failure rate distributions and their mixtures, *International Journal of Reliability and Applications*, **10** (2009), 63–79.

30. A. M. Sarhan, D. C. Hamilton, B. Smith, D. Kundu, The bivariate generalized linear failure rate distribution and its multivariate extension, *Comput. Stat. Data Anal.*, **55** (2011), 644–654. https://doi.org/10.1016/j.csda.2010.06.006

31. A. M. Sarhan, D. Kundu, Generalized linear failure rate distribution, *Commun. Stat.-Theory Meth.*, **38** (2009), 642–660. https://doi.org/10.1080/03610920802272414

32. S. Shoaee, On a new class of bivariate survival distributions based on the model of dependent lives and its generalization, *Applications and Applied Mathematics*, **15** (2020), 801–829.

33. C. Y. Wada, P. K. Sen, S. E. Shimakura, A bivariate exponential model with covariates in competing risk data, *Calcutta Statistical Association Bulletin*, **46** (1996), 197–210. https://doi.org/10.1177/0008068319960305

34. C.-P. Wang, M. Ghosh, Bayesian analysis of bivariate competing risks models with covariates, *J. Stat. Plan. Infer.*, **115** (2003), 441–459. https://doi.org/10.1016/S0378-3758(02)00177-5

Appendix

A. The joint pdf of non-nested models

The joint probability density function (pdf) of the BFGMTLEE, as introduced by Al-Shomrani [4]:

$$\begin{aligned} f(x,y) &= 4\alpha_1\alpha_2\lambda_1\lambda_2 e^{-2\lambda_1 x-2\lambda_2 y}\left[1-e^{-2\lambda_1 x}\right]^{\alpha_1-1}\left[1-e^{-2\lambda_2 y}\right]^{\alpha_2-1} \\ &\quad \left(1+\delta\left(1-2\left[1-e^{-2\lambda_1 x}\right]^{\alpha_1}\right)\left(1-2\left[1-e^{-2\lambda_2 y}\right]^{\alpha_2}\right)\right), x,y\geq 0, \end{aligned}$$

where $\alpha_1, \alpha_2, \lambda_1, \lambda_2 > 0$, and $-1 \leq \delta \leq 1$.

The joint pdf of the BIGE, as given by Alqallaf and Kundu [5], is

$$f_{X,Y}(x,y) = \begin{cases} f_{\text{IGE}}(x;\alpha_1,\lambda)\, f_{\text{IGE}}(y;\alpha_2+\alpha_3,\lambda), & 0 < x < y < \infty \\ f_{\text{IGE}}(x;\alpha_1+\alpha_2,\lambda)\, f_{\text{IGE}}(y;\alpha_2,\lambda), & 0 < y < x < \infty \\ \frac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3}\, f_{\text{IGE}}(x;\alpha_1+\alpha_2+\alpha_3,\lambda), & 0 < x = y < \infty \end{cases}$$

where

$$f_{\text{IGE}}(x;\alpha,\lambda) = \frac{\alpha\lambda}{x^2}\, e^{-\frac{\lambda}{x}} \left(1 - e^{-\frac{\lambda}{x}}\right)^{\alpha-1}, \; x \geq 0, \; \alpha, \lambda > 0.$$

Note: $f_{\text{IGE}}(x;\alpha,\lambda)$ can be obtained from (1.2) by setting $\gamma = \alpha$ and $\alpha = 1$.

AIMS Press

© 2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)