Hindawi
Journal of Mathematics
Volume 2021, Article ID 5200979, 20 pages
https://doi.org/10.1155/2021/5200979

Research Article

Classical and Bayesian Inference of a Mixture of Bivariate Exponentiated Exponential Model

Refah Alotaibi [1], Mervat Khalifa [2,3], Ehab M. Almetwally [4,5], Indranil Ghosh [6], and Rezk. H. [3]

[1] *Mathematical Sciences Department, College of Science, Princess Nourah Bint Abdulrahman University, Riyadh 11671, Saudi Arabia*
[2] *Department of Statistics, King Abdulaziz University, Jeddah 21589, Saudi Arabia*
[3] *Department of Statistics, Al-Azhar University, Cairo, Egypt*
[4] *Department of Mathematical Statistical, Faculty of Graduate Studies for Statistical Research, Cairo University, Cairo, Egypt*
[5] *Department of Statistical, Faculty of Business Administration, Delta University for Science and Technology, Talkha, Egypt*
[6] *Department of Mathematics and Statistics, University of North Carolina, Wilmington, NC, USA*

Correspondence should be addressed to Rezk. H.; hodaragab2009@yahoo.com

Received 4 May 2021; Accepted 25 September 2021; Published 16 October 2021

Academic Editor: Markos Koutras

Exponentiated exponential (EE) model has been used effectively in reliability, engineering, biomedical, social sciences, and other applications. In this study, we introduce a new bivariate mixture EE model with two parameters assuming two cases, independent and dependent random variables. We develop a bivariate mixture starting from two EE models assuming two cases, two independent and two dependent EE models. We study some useful statistical properties of this distribution, such as marginals and conditional distributions and product moments and conditional moments. In addition, we study a dependent case, a new mixture of the bivariate model based on EE distribution marginal with two parameters and with a bivariate Gaussian copula. Different methods of estimation for the model parameters are used both under the classical and under the Bayesian paradigm. Some simulation studies are presented to verify the performance of the estimation methods of the proposed model. To illustrate the flexibility of the proposed model, a real dataset is reanalyzed.

1. Introduction

In the last two decades or so, a major point of interest for statisticians and practitioners was to study populations that exhibit similar behaviors with respect to some predetermined criteria. The earliest evidence regarding the study of heterogeneous populations was mostly due to Newcomb [1] and Pearson [2] who utilized/developed an approach, commonly known as finite mixture distributions. With modern days' stellar advancement on long data-related computation facilities, studies focusing on heterogeneous populations became more popular in the modern era; for some useful references, see the work of Titterington et al. [3], Everitt and Hand [4], McLachlan and Basford [5], and AL-Hussaini and Sultan [6] and the references cited therein. In recent times, there is a growing trend to study and explore the application of finite mixture models; for more details, see the work of Al-Hussaini and Sultan [6].

In several studies concerning heterogeneous population, the EE probability model appears to be really useful. Two-parameter EE distribution is a right skewed unimodal distribution. The behaviors of the probability density function and the hazard function of the EE distribution is quite close to the behavior of the pdf and the hazard function of the gamma or Weibull model. The two-parameter EE distribution has received an increasing amount of interest in recent times. The efficacy of the EE distribution in modeling lifetime data can be found in the works of Gupta and Kundu [7–12]. Several studies have demonstrated that, in specific real-life scenarios, the EE distribution provides a better fit (based on several well-known goodness-of-fit measures) as compared to the gamma or the Weibull model. Kundu and Gupta [13] introduced a bivariate generalized exponential (BGE) distribution and constructed BGE distribution which has three parameters. Some studies obtained a mixture of bivariate inverse Weibull and gamma models; for more details, see the work of Jones et al [14], Sarhan and Balakrishnan [15], Chen and Tan [16], Khosravi [17], and AL-Moisheer et al. [18] and the references cited therein.

The main objective of this paper is to develop and study the mixture of a new bivariate absolutely continuous distribution via a mixture of two independent two-parameter EE distributions. We call this new bivariate distribution as the bivariate mixture of exponentiated exponential distribution (henceforth, in short, the BMEE). The proposed model is constructed under two mechanisms. In the first case, let X and Y be two random variables where each variable is independent and distributed as EE distribution with parameters (λ_1, θ_1) and (λ_2, θ_2), respectively. In the second case, we construct the BMEE distribution via copula approach using the well-known bivariate Gaussian copula (see, for details, the work of Nelson [19]). Several useful mathematical properties of the proposed model are derived. Classical and Bayesian estimation methods are discussed. In addition, the performance of the suggested BMEE model is examined using simulation in estimating the model parameters and a real dataset. The rest of the paper is organized as follows. In Section 2, we introduce the BMEE (type I) distribution, discuss its construction via two mechanisms, and provide some contour plots. In Section 3, we provide some useful mathematical properties and obtain expressions for the bivariate survival function, hazard rate function, bivariate moment generating functions, conditional moments, joint moments, stochastic ordering, etc., for the BMEE distribution constructed starting from two independent EE distributions. In Section 4, we discuss the estimation strategy of the model parameters via EM algorithm. In Section 5, we discuss the estimation strategy for the BMEE (type II) distribution constructed via the bivariate Gaussian copula. In Section 6, we study and explore the estimation of the model parameters under the Bayesian paradigm. Simulation results are presented in Section 7. In Section 8, a well-known motor dataset has been reanalyzed to exhibit the efficacy of the proposed BMEE-type models. Finally, we conclude the paper by providing some final remarks in Section 9.

2. Mixture Bivariate Independent EE Model

Here, we begin our discussion with two independent univariate EE distributions with parameters (λ_1, θ_1) and (λ_2, θ_2), respectively. The central idea of compounding is to consider that θ_1 and θ_2 are indeed random variables not constant, and the observed (marginal) distribution of X_{in} and Y_{in} can be obtained from the joint distribution of θ_1 and θ_2 which is as follows:

$$h_{x_{\text{in}},y_{\text{in}}}(x_{\text{in}}, y_{\text{in}}) = \iint h(x_{\text{in}}, y_{\text{in}}, \theta_1, \theta_2)\, d\theta_1\, d\theta_2. \tag{1}$$

Next, we construct a bivariate of EE mixture distribution by assuming two cases. In the first case, X_{in} and Y_{in} are independent EE distributions with the scale parameters having a generalized bivariate Bernoulli distribution. In the second case, X_d and Y_d are dependent. A random variable with an EE distribution has a cumulative distribution function (cdf) and a probability density function (pdf) for $X_{\text{in}} > 0$, given by

$$F(x_{\text{in}}, \theta, \lambda) = \left(1 - e^{-\lambda x_{\text{in}}}\right)^{\theta}, \quad x_{\text{in}} > 0, \tag{2}$$

$$f(x_{\text{in}}, \theta, \lambda) = \theta\lambda e^{-\lambda x_{\text{in}}}\left(1 - e^{-\lambda x_{\text{in}}}\right)^{\theta-1}, \quad x_{\text{in}} > 0, \tag{3}$$

where $\theta > 0$ and $\lambda > 0$ are the shape and scale parameters.

In the bivariate case, let X_{in} and Y_{in} be two random variables with parameters θ_1, and θ_2, respectively. For given fixed values of θ_1 and θ_2, X_{in} and Y_{in} are independent. The pdf of BMEE distribution is defined as

$$f(x_{in}, \theta,) = \sum_{i=1}^{2} p_i f_i\left((x_{in})_i, \theta_i\right), \tag{4}$$

where p_i are the mixing proportions which must satisfy $\sum_{i=1}^{2} p_i = 1$ and $p_i \geq 0$, and all parameters are unknowns. The pdf of the first component of EE is given by (2), with fixed shape parameter $\theta > 0$, and a random scale parameter $\lambda > 0$ that takes two distinct values λ_1 and λ_2. Likewise, for fixed shape parameter θ_2, let Y have an EE mixture density, and the pdf of second component (EE) is given by

$$g(y_{\text{in}}, \phi, \beta) = \phi\beta_i e^{-\beta_i y_{\text{in}}}\left(1 - e^{-\beta_i y_{\text{in}}}\right)^{\varphi-1}, \quad y_{\text{in}} > 0, \tag{5}$$

where β is a random scale parameter $(\beta > 0)$ that takes two distinct values β_1 and β_2. For given values of (λ, β), we assume that X_{in} and Y_{in} are independent, but λ and β are correlated through their generalized bivariate distribution with the following probability matrix:

$$P = \begin{array}{c} \\ \lambda_1 \\ \lambda_2 \end{array} \begin{array}{c} \beta_1 \quad \beta_2, \\ \begin{bmatrix} P_{\lambda_1\beta_1} & P_{\lambda_1\beta_2} \\ P_{\lambda_2\beta_1} & P_{\lambda_2\beta_2} \end{bmatrix}, \end{array} \tag{6}$$

where P is the mixture components and $P_{\lambda_1\beta_1} + P_{\lambda_1\beta_2} + P_{\lambda_2\beta_1} + P_{\lambda_2\beta_2} = 1$. Let $h_{x_{\text{in}},y_{\text{in}}}(x_{\text{in}}, y_{\text{in}})$ be the joint pdf of $(X_{\text{in}}, Y_{\text{in}})$; then,

$$\begin{aligned} h_{x_{\text{in}},y_{\text{in}}}(x_{\text{in}}, y_{\text{in}}) = {} & f(x_{\text{in}}|\theta,\lambda_1)g(y_{\text{in}}y|\phi,\beta_1)P_{\lambda_1\beta_1} + f(x_{\text{in}}|\theta,\lambda_1)g(y_{\text{in}}|\phi,\beta_2)P_{\lambda_1\beta_2} \\ & + f(x_{\text{in}}|\theta,\lambda_2)g(y_{\text{in}}|\phi,\beta_1)P_{\lambda_2\beta_1} + f(x_{\text{in}}|\theta,\lambda_2)g(y_{\text{in}}|\phi,\beta_2)P_{\lambda_2\beta_2}. \end{aligned} \tag{7}$$

For simplicity in the independent case, we use $x_{\text{in}} = x$ and $y_{\text{in}} = y$ as follows:

$$\begin{aligned} h(x,y) &= \theta\lambda_1 e^{-\lambda_1 x}\left(1-e^{-\lambda_1 x}\right)^{\theta-1}\phi\beta_1 e^{-\beta_1 y}\left(1-e^{-\beta_1 y}\right)^{\varphi-1}P_{\lambda_1\beta_1} + \theta\lambda_1 e^{-\lambda_1 x}\left(1-e^{-\lambda_1 x}\right)^{\theta-1}\phi\beta_2 e^{-\beta_2 y}\left(1-e^{-\beta_2 y}\right)^{\phi-1}P_{\lambda_1\beta_2} \\ &\quad + \theta\lambda_2 e^{-\lambda_2 x}\left(1-e^{-\lambda_2 x}\right)^{\theta-1}\phi\beta_1 e^{-\beta_1 y}\left(1-e^{-\beta_1 y}\right)^{\varphi-1}P_{\lambda_2\beta_1} + \theta\lambda_2 e^{-\lambda_2 x}\left(1-e^{-\lambda_2 x}\right)^{\theta-1}\phi\beta_2 e^{-\beta_2 y}\left(1-e^{-\beta_2 y}\right)^{\varphi-1}P_{\lambda_2\beta_2}. \end{aligned} \tag{8}$$

For simplification, let $a = p_{\lambda_1\beta_1}$, $b = p_{\lambda_1\beta_2}$, $c = p_{\lambda_2\beta_1}$, and $d = p_{\lambda_2\beta_2}$.

From Figures 1 and 2, it is evident that the joint pdf in equation (4) can produce various shapes corresponding to several parameter choices. The joint pdf mixture of four univariate EE mixture distributions involves a total of 9 parameters for its specification. In application to real-life datasets, as one can imagine, not all four components might be necessary. Consequently, one may put some restrictions, such as $b=c=0$, $a=d=0$, or $a=b=c=0$. These restrictions result in correlation values (among scale parameters) of +1, −1, and 0, respectively.

The marginal densities of X and Y, respectively, are given as follows:

$$h_x(x) = \pi_1 f_1(x) + (1-\pi_1) f_2(x), \tag{9}$$

where $\pi_1 = a + b$.

$$h_y(Y) = \pi_2 g_1(Y) + (1-\pi_2) g_2(Y), \tag{10}$$

where $\pi_2 = a + c$.

The joint cdf will be

$$\begin{aligned} F(x,y) &= \int_0^y\int_0^x h(t,s)dt\ ds \\ &= \int_0^y\int_0^x\Big[a\theta\lambda_1 e^{-\lambda_1 x}\left(1-e^{-\lambda_1 x}\right)^{\theta-1}\phi\beta_1 e^{-\beta_1 y}\left(1-e^{-\beta_1 y}\right)^{\varphi-1} + b\theta\lambda_1 e^{-\lambda_1 x}\left(1-e^{-\lambda_1 x}\right)^{\theta-1}\phi\beta_2 e^{-\beta_2 y}\left(1-e^{-\beta_2 y}\right)^{\phi-1} \\ &\quad + c\theta\lambda_2 e^{-\lambda_2 x}\left(1-e^{-\lambda_2 x}\right)^{\theta-1}\phi\beta_1 e^{-\beta_1 y}\left(1-e^{-\beta_1 y}\right)^{\varphi-1} + d\theta\,\lambda_2 e^{-\lambda_2 x}\left(1-e^{-\lambda_2 x}\right)^{\theta-1}\phi\beta_2 e^{-\beta_2 y}\left(1-e^{-\beta_2 y}\right)^{\varphi-1}\Big]dt\ ds \\ &= a\left(1-e^{-\lambda_1 x}\right)^{\theta}\left(1-e^{-\beta_1 y}\right)^{\varphi} + b\left(1-e^{-\lambda_1 x}\right)^{\theta}\left(1-e^{-\beta_2 y}\right)^{\varphi} + c\left(1-e^{-\lambda_2 x}\right)^{\theta}\left(1-e^{-\beta_1 y}\right)^{\varphi} + d\left(1-e^{-\lambda_2 x}\right)^{\theta}\left(1-e^{-\beta_2 y}\right)^{\varphi}. \end{aligned} \tag{11}$$

The associated survival function of BMEE distribution will be

$$\begin{aligned} R(x,y) &= P(X>x, Y>y) = 1 - P(X<x) - P(Y<y) + P(X<x, Y<y) \\ &= 1 - F_X(x) - F_Y(y) + F_{X,Y}(x,y), \end{aligned} \tag{12}$$

where $F_X(x)$ and $F_Y(y)$ are the marginal density functions of X and Y, respectively. The hazard rate function (hrf) is given as follows:

$$\text{hrf}(x,y) = \frac{A}{B}, \tag{13}$$

where A is given in (8) and B is given in (12). The conditional pdf of X, given Y, for each fixed $Y=y$ will be

$$h(x|y) = \frac{h(x,y)}{h_y(y)}. \tag{14}$$

The conditional pdf of Y, given X, will be

$$h(y|x) = \frac{h(x,y)}{h_x(x)}. \tag{15}$$

3. Structural Properties

Here, we derive some properties of the BMEE distribution which are as follows.

Proposition 1. *Let $(X, Y) \sim BMEE(a,b,c,d,\lambda_1,\lambda_2,\beta_1,\beta_2,\theta,\phi)$. Then, (X,Y) has a total positivity of order 2 positive association (TP_2 property).*

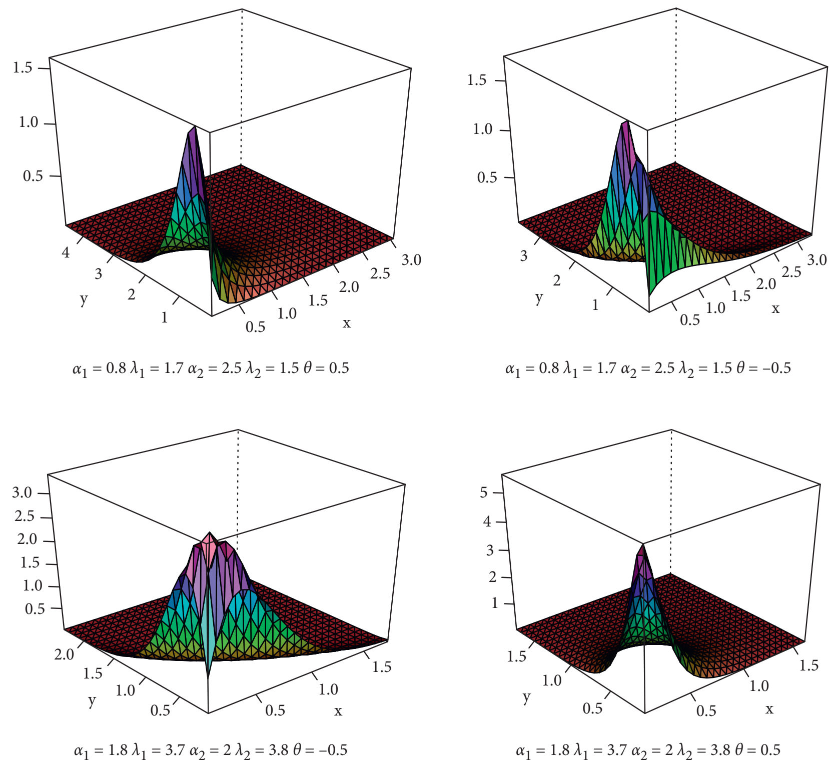

Figure 1: The plot of the BMEE pdf model for varying parameter choices in equation (4).

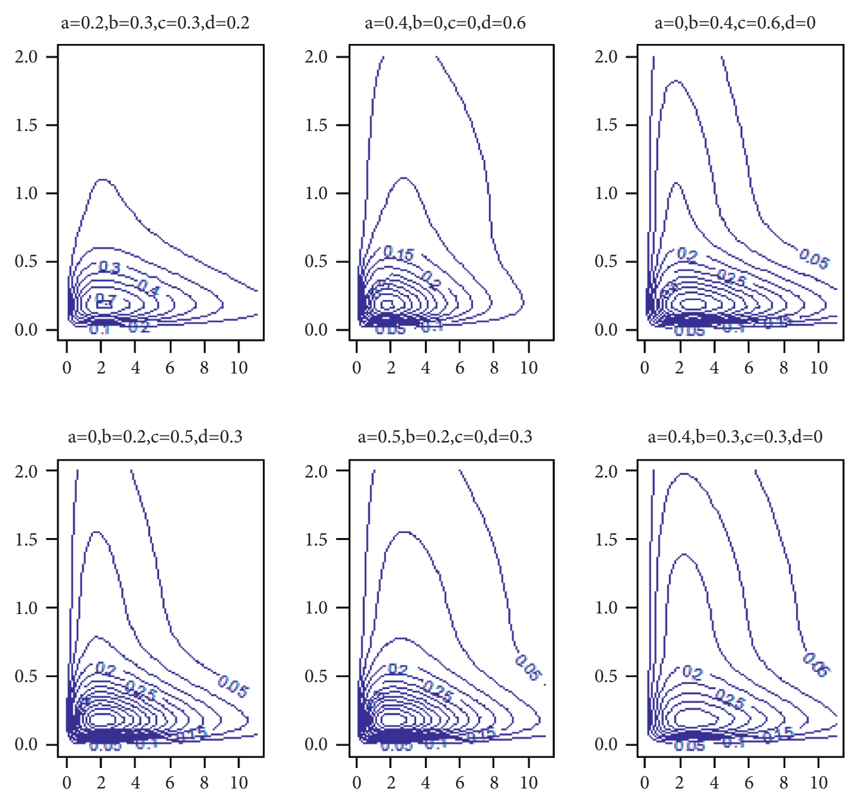

Figure 2: The contour plots of BMEE distribution for varying parameter choices in equation (4).

Proof. Observe that an absolute continuous bivariate random vector, say (U_1, U_2), has TP_2 property if and only if, for any $u_{11}, u_{12}, u_{21}, u_{22}$, whenever $u_{11} < u_{12}$ and $u_{21} > u_{22}$, we have

$$f_{U_1,U_2}(u_{11}, u_{21}) f_{U_1,U_2}(u_{11}, u_{12}) - f_{U_1,U_2}(u_{12}, u_{21}) f_{U_1,U_2}(u_{11}, u_{22}) \geq 0, \tag{16}$$

where $f_{U_1,U_2}(,)$ is the joint PDF of (U_1, U_2). Observe that, by taking different ordered $u_{11}, u_{12}, u_{21}, u_{22}$, such that $u_{11} < u_{12}$, and $u_{21} > u_{22}$, our result immediately follows.

As a consequence, the positive quadrant dependence property (alternatively, the TP_2 property) will indicate several other nonincreasing properties related to conditional survival, conditional cdf of X, given Y, and Y, given X, including this result: $g_1(.)$ and $g_2(.)$, $Cov(g_1(X), g_2(Y)) \geq 0$. □

Proposition 2 (moments). *The product moments, $\mu'_{x,y}$, about zero is*

$$\begin{aligned}
\mu'_{X,Y} &= \int_0^\infty \int_0^\infty x\, y\, h(x,y) dx\ dy = \int_0^\infty \int_0^\infty x\, y \Big[a\theta\lambda_1 e^{-\lambda_1 x}\left(1 - e^{-\lambda_1 x}\right)^{\theta-1} \phi\beta_1 e^{-\beta_1 y}\left(1 - e^{-\beta_1 y}\right)^{\varphi-1} \\
&\quad + b\theta\lambda_1 e^{-\lambda_1 x}\left(1 - e^{-\lambda_1 x}\right)^{\theta-1} \phi\beta_2 e^{-\beta_2 y}\left(1 - e^{-\beta_2 y}\right)^{\phi-1} + c\theta\lambda_2 e^{-\lambda_2 x}\left(1 - e^{-\lambda_2 x}\right)^{\theta-1} \phi\beta_1 e^{-\beta_1 y}\left(1 - e^{-\beta_1 y}\right)^{\varphi-1} \\
&\quad + d\theta\,\lambda_2 e^{-\lambda_2 x}\left(1 - e^{-\lambda_2 x}\right)^{\theta-1} \phi\beta_2 e^{-\beta_2 y}\left(1 - e^{-\beta_2 y}\right)^{\varphi-1} \Big] dx\ dy \\
&= \text{Harmonic number}[\theta]\text{Harmonic number}[\varphi]\left[\frac{a}{\lambda_1\beta_1} + \frac{b}{\lambda_1\beta_2} + \frac{c}{\lambda_2\beta_1} + \frac{d}{\lambda_2\beta_2}\right].
\end{aligned} \tag{17}$$

where Harmonic number[.] is the sum of the reciprocals of the first n natural numbers.

The expected value of X and Y are, respectively, given by

$$\begin{aligned}
E(X) &= \int_0^\infty x\ h_x(x) dx = \text{Harmonic number}[\varphi]\theta\left\{\left[\frac{a\lambda_1}{\beta_1} e^{-\lambda_1 x}\left(1 - e^{-\lambda_1 x}\right)^{\theta-1}\right] + \left[\frac{b\lambda_2}{\beta_1} e^{-\lambda_2 x}\left(1 - e^{-\lambda_2 x}\right)^{\theta-1}\right]\right. \\
&\quad \left. + \left[\frac{\lambda_1}{\beta_2} e^{-\lambda_1 x}\left(1 - e^{-\lambda_1 x}\right)^{\theta-1}\right] + \left[\frac{d\lambda_2}{\beta_2} e^{-\lambda_2 x}\left(1 - e^{-\lambda_2 x}\right)^{\theta-1}\right]\right\},
\end{aligned} \tag{18}$$

$$\begin{aligned}
E(Y) &= \int_0^\infty y\ h_x(y) dy = \text{Harmonic number}[\theta]\varphi\left\{\left[\frac{a\beta_1}{\lambda_1} e^{-\beta_1 y}\left(1 - e^{-\beta_1 y}\right)^{\phi-1}\right] + \left[\frac{b\beta_2}{\lambda_1} e^{-\beta_2 y}\left(1 - e^{-\beta_2 y}\right)^{\phi-1}\right]\right. \\
&\quad \left. + \left[\frac{c\beta_1}{\lambda_2} e^{-\beta_1 y}\left(1 - e^{-\beta_1 y}\right)^{\phi-1}\right] + \left[\frac{d\beta_2}{\lambda_2} e^{-\beta_2 y}\left(1 - e^{-\beta_2 y}\right)^{\phi-1}\right]\right\}.
\end{aligned} \tag{19}$$

A Mathematica 11.2 is used to obtain the integral in (17)–(19).

Proposition 3. *The joint moment generating function of BMEE will be*

$$\begin{aligned}M_{x,y}(t,s) &= \int_0^\infty \int_0^\infty e^{tx+sy}\ h(x,y)dx\ dy = \int_0^\infty \int_0^\infty e^{tx+sy}\Big[a\theta\lambda_1 e^{-\lambda_1 x}\left(1-e^{-\lambda_1 x}\right)^{\theta-1}\phi\beta_1 e^{-\beta_1 y}\left(1-e^{-\beta_1 y}\right)^{\varphi-1} \\ &\quad + b\theta\lambda_1 e^{-\lambda_1 x}\left(1-e^{-\lambda_1 x}\right)^{\theta-1}\phi\beta_2 e^{-\beta_2 y}\left(1-e^{-\beta_2 y}\right)^{\phi-1} + c\theta\lambda_2 e^{-\lambda_2 x}\left(1-e^{-\lambda_2 x}\right)^{\theta-1}\phi\beta_1 e^{-\beta_1 y}\left(1-e^{-\beta_1 y}\right)^{\varphi-1} \\ &\quad + d\theta\ \lambda_2 e^{-\lambda_2 x}\left(1-e^{-\lambda_2 x}\right)^{\theta-1}\phi\beta_2 e^{-\beta_2 y}\left(1-e^{-\beta_2 y}\right)^{\varphi-1}\Big]dx\ dy \\ &= \theta\varphi\Gamma\theta\Gamma\varphi\left[\begin{array}{c}\dfrac{a\Gamma(1-s/\beta_1)\Gamma(1-t/\lambda_1)}{\Gamma(1+\theta-t/\lambda_1)\Gamma(1+\varphi-s/\beta_1)} + \dfrac{b\Gamma(1-s/\beta_2)\Gamma(1-t/\lambda_2)}{\Gamma(1+\theta-t/\lambda_2)\Gamma(1+\varphi-s/\beta_2)} + \\ \dfrac{c\Gamma(1-s/\beta_1)\Gamma(1-t/\lambda_1)}{\Gamma(1+\theta-t/\lambda_1)\Gamma(1+\varphi-s/\beta_1)} + \dfrac{d\Gamma(1-s/\beta_2)\Gamma(1-t/\lambda_2)}{\Gamma(1+\theta-t/\lambda_2)\Gamma(1+\varphi-s/\beta_2)}\end{array}\right],\end{aligned} \tag{20}$$

provided $|1+\theta-t/\lambda_2|\le 1$ *and* $|1+\varphi-s/\beta_2|\le 1$.

Using the joint pdf of X and Y and/or using the MGF expression given above, the different product moments of the order $X^m X^n$, $(m,n)\ge 1$, can be obtained from the above.

Proposition 4 (shape of the distribution). *A critical point of a function with two variables is a point where the partial derivatives of first order are equal to zero. The most compelling two reasons to study the critical points for a bivariate distribution are as follows:*

(1) A real-life dataset(s) can have several different shapes. The flexibility of any proposed model can be well determined from such a study.

(2) In dealing with bivariate distributions, quite often, it is imperative to study the tails of the joint pdf as well as the point of inflection. Knowledge on critical point(s) will help to better understand these properties.

Let us now consider the shape of the BMEE distribution:

$$\begin{aligned}\frac{\partial f(x,y)}{\partial x\ x} &= \theta\phi\Bigg[\frac{-a\beta_1\lambda_1^2\left(e^{\lambda_1 x}-\theta\right)\left(1-e^{-\lambda_1 x}\right)^{\theta}\left(1-e^{-\beta_1 y}\right)^{\varphi}}{\left(e^{-\lambda_1 x}-1\right)^2\left(e^{\beta_1 y}-1\right)} \\ &\quad - \frac{\left(b\beta_2\lambda_1^2\left(e^{\lambda_1 x}-\theta\right)\left(1-e^{-\lambda_1 x}\right)^{\theta}\left(1-e^{-\beta_2 y}\right)^{\varphi}/\left(e^{-\lambda_1 x}-1\right)^2\right) + \left(\left(\lambda_2^2\left(e^{\lambda_2 x}-\theta\right)\left(1-e^{-\lambda_2 x}\right)^{\theta}\left(c\beta_1\left(e^{\beta_2 y}-1\right)\left(1-e^{-\beta_1 y}\right)^{\varphi}+d\beta_2\left(e^{\beta_1 y}-1\right)\left(1-e^{-\beta_2 y}\right)^{\varphi}\right)\right)/\left(e^{-\lambda_1 x}-1\right)^2\left(e^{\beta_1 y}-1\right)\right)}{\left(e^{\beta_2 y}-1\right)}\Bigg]. \\ \frac{\partial f(x,y)}{\partial y} &= \theta\phi\Bigg[\frac{-a\beta_1\lambda_1\left(e^{\beta_1 y}-\phi\right)\left(1-e^{-\lambda_1 x}\right)^{\theta}\left(1-e^{-\beta_1 y}\right)^{\varphi}}{\left(e^{\beta_1 y}-1\right)^2\left(e^{-\lambda_1 x}-1\right)} \\ &\quad - \left(\frac{\left(b\lambda_1\beta_2^2\left(e^{\beta_2 y}-\phi\right)\left(1-e^{-\lambda_1 x}\right)^{\theta}\left(1-e^{-\beta_2 y}\right)^{\varphi}/\left(e^{-\lambda_1 x}-1\right)\right) + \left(\lambda_2\left(1-e^{-\lambda_2 x}\right)^{\theta}\left(\begin{array}{c}c\beta_1^2\left(e^{\beta_2 y}-1\right)^2\left(e^{\beta_1 y}-\varphi\right)\left(1-e^{-\beta_1 y}\right)^{\varphi} \\ +d\beta_2^2\left(e^{\beta_1 y}-1\right)^2\left(1-e^{-\beta_2 y}\right)^{\varphi}\left(e^{\beta_2 y}-\varphi\right)\end{array}\right)/\left(e^{-\lambda_1 x}-1\right)^2\left(e^{\beta_1 y}-1\right)\right)}{\left(e^{\beta_2 y}-1\right)^{-2}}\right)\Bigg].\end{aligned} \tag{21}$$

Consequently, for BMEE distribution, there may be several critical points. For specific choices of the model parameters, a numerical study can be made here.

4. Estimation (Independence Case Scenario)

In this section, we discuss the estimation of the model parameters of the BMEE distribution assuming the independence of X and Y under the EM algorithm.

4.1. EM Algorithm. Mclachlan and Krishnan [20] introduced expectation-maximization (EM) algorithm which is an iterative method to find the maximum likelihood estimator of a parameter Φ of a parametric probability distribution. To invoke the idea of EM algorithm, we augment the data $((x_k,\ y_k),\ k=1,\dots,n$, with the group membership variables $\Phi_k=(a_k,b_k,c_k)$, $k=1,\dots,\ n$, where a_k is an indicator variable, and one if the k^{th} observation is in $f(x,\lambda_1,\beta_1)$, and zero otherwise.

Similarly, for b_k, c_k, we have four groups G_{ij}, $i, j = 1, 2$, for which the densities are

$$f_{ij}(x, y) = f_i(x) f_j(y) = \theta\lambda_1 e^{-\lambda_1 x}\left(1 - e^{-\lambda_1 x}\right)^{\theta-1} \phi\beta_1 e^{-\beta_1 y}\left(1 - e^{-\beta_1 y}\right)^{\varphi-1}. \tag{22}$$

The corresponding mixing proportions are $P(G_{11}) = a$, $P(G_{12}) = b$, $P(G_{21}) = c$, and $P(G_{22}) = 1 - a - b - c$.

The corresponding log-likelihood function to the complete sample, $\ell_{ij}(x, y) = \log f_{ij}(x, y)$, is represented by

$$\ell = \sum_{k=1}^{n} a_k \ell_{11}(x_k, y_k) + = \sum_{k=1}^{n} b_k \ell_{12}(x_k, y_k) + \sum_{k=1}^{n} c_k \ell_{21}(x_k, y_k) + \sum_{k=1}^{n} (1 - a_k - b_k - c_k)\ell_{22}(x_k, y_k). \tag{23}$$

Each iteration of the EM algorithm involves two steps: step E (expectation) and step M (maximization), defined by (16). This is linear in the group membership variables Φ_k, so in the E-step, we replace in (9) the associated expected values, given the current estimates $((\widehat{\theta}, \widehat{\varphi}, \widehat{\lambda}_1, \widehat{\lambda}_2, \widehat{\beta}_1, \widehat{\beta}_2, \widehat{a}, \widehat{b}, \widehat{c})$ of the parameters are calculated as

$$\widehat{a}_k = \frac{\widehat{a} f_{11}(x_k, y_k)}{\widehat{a} f_{11}(x_k, y_k) + \widehat{b} f_{12}(x_k, y_k) + \widehat{c} f_{21}(x_k, y_k) + (1 - \widehat{a} - \widehat{b} - \widehat{c}) f_{22}(x_k, y_k)}, \tag{24}$$

and follow the same procedure similarly for b_k and c_k.

Next, in the M-step, we need to maximize (23) over $(\theta, \varphi, \beta_1, \beta_2, \lambda_1, \lambda_2)$ for fixed values of Φ_k. It is achieved by the conditional independence of X and Y, given the group membership. Differentiating (23), we obtain

$$\frac{\partial \iota}{\partial \lambda_1} = \frac{n}{\lambda_1} \sum_{k=1}^{n} (a_k + b_k) - \sum_{k=1}^{n} (a_k + b_k) x_k + (\theta - 1) \sum_{k=1}^{n} \frac{x_k (a_k + b_k) e^{-\lambda_1 x_k}}{\left(1 - e^{-\lambda_1 x_k}\right)}, \tag{25}$$

$$\frac{\partial \iota}{\partial \lambda_2} = \frac{n}{\lambda_2} \sum_{k=1}^{n} (c_k + d_k) - \sum_{k=1}^{n} (c_k + d_k) x_k + (\theta - 1) \sum_{k=1}^{n} \frac{x_k (c_k + d_k) e^{-\lambda_2 x_k}}{\left(1 - e^{-\lambda_2 x_k}\right)}, \tag{26}$$

$$\frac{\partial \iota}{\partial \beta_1} = \frac{n}{\beta_1} \sum_{k=1}^{n} (a_k + c_k) - \sum_{k=1}^{n} (a_k + c_k) y_k + (\varphi - 1) \sum_{k=1}^{n} \frac{y_k (a_k + c_k) e^{-\beta_1 y_k}}{\left(1 - e^{-\beta_1 y_k}\right)}, \tag{27}$$

$$\frac{\partial \iota}{\partial \beta_2} = \frac{n}{\beta_2} \sum_{k=1}^{n} (b_k + d_k) - \sum_{k=1}^{n} (b_k + d_k) y_k + (\varphi - 1) \sum_{k=1}^{n} \frac{y_k (b_k + d_k) e^{-\beta_2 y_k}}{\left(1 - e^{-\beta_2 y_k}\right)}, \tag{28}$$

$$\frac{\partial \iota}{\partial \theta} = \frac{n}{\theta} \sum_{k=1}^{n} (a_k + b_k + c_k + d_k) + \sum_{k=1}^{n} (a_k + c_k)\left(1 - e^{-\lambda_1 x_k}\right)^{\theta-1} \log\left(1 - e^{-\lambda_1 x_k}\right) + \sum_{k=1}^{n} (b_k + d_k)\left(1 - e^{-\lambda_2 x_k}\right)^{\theta-1} \log\left(1 - e^{-\lambda_2 x_k}\right), \tag{29}$$

$$\frac{\partial \iota}{\partial \varphi} = \frac{n}{\varphi} \sum_{k=1}^{n} (a_k + b_k + c_k + d_k) + \sum_{k=1}^{n} (a_k + c_k)\left(1 - e^{-\beta_1 y_k}\right)^{\varphi-1} \log\left(1 - e^{-\beta_1 y_k}\right) + \sum_{k=1}^{n} (b_k + d_k)\left(1 - e^{-\beta_2 y_k}\right)^{\varphi-1} \log\left(1 - e^{-\beta_2 y_k}\right). \tag{30}$$

Next, the M-step is completed by setting

$$\widehat{a} = \frac{1}{n}\sum_{k=1}^{n} a_k, \text{etc.} \tag{31}$$

This algorithm requires an initial value of the model parameters, designated by $\Phi_{(0)}$. A judicious choice of these initial values requires special attention on the fact that the rate of convergence of the assumed EM algorithm may not become quite slow. Another point of concern is that the maximum likelihood equation may have multiple solutions corresponding to local maxima; therefore, the selection of the starting values is indeed very important. A comparative study of various strategies in the choice of initial values can be found in Karlis and Xekalaki [21]. We use the copula R package to solve these equations numerically. After the maximum likelihood estimators for $\theta, \lambda_1, \beta_1, \lambda_2, \beta_2, \delta, \rho_{11}, \rho_{12}, \rho_{21}$, and ρ_{22} are obtained, we substitute these estimates in (a_k, b_k, c_k). We complete the M-step by setting $\widehat{a} = 1/n\sum_{k=1}^{n} a_k$, etc.

Initial values for the mixing proportions are obtained by the moment's method of the marginal univariate EE parameter separately. Next, we take the resulting estimates of the BEE parameters as starting values for the EM algorithm. After that, we merge the moment estimators of the marginal mixing parameters to obtain initial values for the bivariate mixing parameters, assuming that the independence between two variables X and Y. We apply this method in application as mentioned in Section 8; later on, specifically, in Tables 1–4, we equate (25)–(30) to zero, to obtain estimates of the parameter of the distribution.

5. BMEE Model (Type II) Distribution Using Gaussian Copula

The BMEE model proposed in this paper involves EE marginals, with greatest flexibility with its marginals as well as in the correlation structure. On the contrary, this proposed distribution has several fields of applicability. Usually, in the dependence study, copulas play a vital role. Copulas are a general tool to construct bivariate and multivariate distributions and to study dependence structure between random variables. Several bivariate and multivariate lifetime distributions are suggested using several methods of constructing bivariate and multivariate distributions, and copula functions have been proposed by Nelsen [22], Trivedi and Zimmer [23], Adham and Walker [24], Kundu et al. [25], Kundu and Gupta [26], Kundu [27], El-Morshedy et al. [28], and Alotaibi et al. [29].

5.1. BMEE Distribution Based on Gaussian Copula. The concept of copula, suggested and derived by Sklar [30], states that any multivariate distribution can be disintegrated to a copula and its continuous marginal. In a bivariate setup, copulas are used to link two marginal distributions with a joint distribution such that, for every bivariate distribution function $F(x, y)$ with continuous marginal $F(x), F(y)$, there exists a unique copula function C given by

$$\begin{aligned} F(x, y) &= C\{F(x), F(y)\}, \\ (x, y) &\in (-\infty, \infty) X (-\infty, \infty). \end{aligned} \tag{32}$$

The associated density function of bivariate distribution will be

$$f(x, y) = f(x) f(y) c(F(x), F(y)), \tag{33}$$

where $c(F(x), F(y))$ is the density function of a copula; for further details, see the work of Nelsen [19, 22]. A plethora of choices are available to construct BMEE distributions via copula using EE marginals as given in (1). Here, the Gaussian copula is utilized to construct BMEE distribution.

The Gaussian copula has the following form:

$$C(u, v) = \varphi_{\Sigma}\left(\varphi^{-1}(u), \varphi^{-1}(v)\right), \tag{34}$$

where φ_{Σ} denotes the distribution function of a bivariate standard normal random variable and φ^{-1} represents its inverse. The joint pdf of X_d and Y_d based on Gaussian copula becomes

$$f(x_d, y_d) = f(x_d) f(y_d) \left\{ \frac{1}{\sqrt{1-\rho^2}} \left(\exp\left[\frac{-\rho}{2\left(1-\rho^2\right)} \left\{ \rho\left(z_1^2 + z_2^2\right) - 2 z_1 z_2 \right\} \right] \right) \right\}, \tag{35}$$

where $\rho \in [-1, 1]$ is a dependence parameter and $f(x_d)$ and $f(y_d)$ are the density function of EE distributions given in (1). Suppose that the marginals are EE distribution; then, the bivariate exponentiated exponential (BEE) distribution pdf is

$$\begin{aligned} f(x_d, y_d, \theta, \lambda, \beta, \delta) = {} & \theta \lambda e^{-\lambda x_d} \left(1 - \exp\{-\lambda x_d\}\right)^{\theta - 1} \delta \beta e^{-\beta y_d} \left(1 - \exp\{-\beta y_d\}\right)^{\delta - 1} \\ & \cdot \left\{ \frac{1}{\sqrt{1-\rho^2}} \left(\exp\left[\frac{-\rho}{2\left(1-\rho^2\right)} \left\{ \rho\left(z_1^2 + z_2^2\right) - 2 z_1 z_2 \right\} \right] \right) \right\}, \quad x_d, y_d, \theta, \lambda, \beta, \delta > 0. \end{aligned} \tag{36}$$

Table 1: MLE and Bayesian estimation with different sample sizes in the independent case.

			$\theta = 1.8$, $\lambda = 3.7$, $\beta = 3.8$, $\delta = 2$, $\rho = 0.5$				
n			θ	λ	β	δ	ρ
50	MLE	Bias	−0.7241	−0.4413	−0.3785	−0.5216	0.4645
		MSE	0.8039	0.4955	0.2015	0.6152	0.2158
		L.CI	2.0737	2.1507	0.9461	2.2974	0.0305
	Bayes	Bias	−0.4340	−0.2691	−0.2255	−0.3190	0.4434
		MSE	0.3449	0.2371	0.0988	0.3074	0.1967
		L.CI	1.5517	1.5913	0.8587	1.7783	0.0469
100	MLE	Bias	−0.7098	−0.4190	−0.3405	−0.5176	0.4565
		MSE	0.7900	0.4215	0.1977	0.5552	0.2159
		L.CI	1.5387	1.6693	0.7195	1.8540	0.0231
	Bayes	Bias	−0.3439	−0.2276	−0.2151	−0.3021	0.4251
		MSE	0.2693	0.1321	0.0887	0.2093	0.1920
		L.CI	1.0875	1.1111	0.6302	1.2774	0.0265
200	MLE	Bias	−0.6857	−0.4056	−0.3143	−0.5364	0.4264
		MSE	0.7283	0.4137	0.1820	0.5035	0.2106
		L.CI	1.1900	1.2744	0.5133	1.3715	0.0169
	Bayes	Bias	−0.3431	−0.2142	−0.2055	−0.2931	0.4153
		MSE	0.2366	0.0949	0.0789	0.1564	0.1912
		L.CI	0.8818	0.8686	0.4612	0.9703	0.0181

Table 2: MLE and Bayesian estimation with different sample sizes in the independent case.

			$\theta = 2.8$, $\lambda = 3.7$, $\beta = 3.7$, $\delta = 1.5$, $\rho = 0.5$				
n			θ	λ	β	δ	ρ
50	MLE	Bias	−0.4552	−0.4981	−0.5148	−0.4116	0.4647
		MSE	0.3038	0.5974	0.3828	0.5288	0.2160
		L.CI	1.2187	2.3179	1.3460	2.3510	0.0295
	Bayes	Bias	−0.0552	−0.0646	−0.0605	−0.0486	0.3255
		MSE	0.0262	0.0310	0.0275	0.0311	0.1142
		L.CI	0.5972	0.6423	0.6050	0.6651	0.3554
100	MLE	Bias	−0.3504	−0.4569	−0.4568	−0.3750	0.4064
		MSE	0.2830	0.4972	0.3485	0.3641	0.1922
		L.CI	0.8395	1.6351	0.9777	1.7062	0.0224
	Bayes	Bias	−0.0547	−0.0617	−0.0607	−0.0356	0.3188
		MSE	0.0218	0.0262	0.0231	0.0233	0.1514
		L.CI	0.5172	0.5711	0.5190	0.5823	0.1170
200	MLE	Bias	−0.3053	−0.4262	−0.4160	−0.3533	0.3846
		MSE	0.2306	0.3503	0.3094	0.3927	0.1522
		L.CI	0.6446	1.3235	0.7225	1.2913	0.0162
	Bayes	Bias	−0.0490	−0.0611	−0.0600	−0.0266	0.3041
		MSE	0.0204	0.0218	0.0241	0.0172	0.1723
		L.CI	0.4365	0.5190	0.4659	0.5040	0.0492

Table 3: MLE and Bayesian estimation with different sample sizes in dependent case (Case I).

			$\theta = 2.8, \lambda_1 = 2.5, \beta_1 = 2.3, \lambda_2 = 1.2, \beta_2 = 1.4$, $\delta = 1.5, \rho_{11} = 0.2, \rho_{12} = 0.15$, $\rho_{21} = 0.05$, and $\rho_{22} = 0.1$									
n			θ	λ_1	β_1	λ_2	β_2	δ	ρ_{11}	ρ_{12}	ρ_{21}	ρ_{22}
50	MLE	Bias	−0.5957	−0.0916	−0.2343	−0.0876	−0.2470	−0.0543	0.7093	0.7731	0.5538	0.7953
		MSE	0.5730	0.1729	0.1336	0.0552	0.2841	0.0836	0.5085	0.6064	0.5713	0.6626
		L.CI	1.8320	1.5909	1.1001	0.8547	1.8526	1.1137	0.2867	0.3659	2.0173	0.6805
	Bayes	Bias	0.0322	0.0108	0.0748	−0.0658	−0.2786	0.0142	0.1108	0.1094	0.0433	0.0975
		MSE	0.0298	0.0350	0.0334	0.0286	0.2324	0.0270	0.0498	0.0479	0.0334	0.0468
		L.CI	0.6656	0.7326	0.6537	0.6111	1.3276	0.6425	0.7601	0.7439	0.6966	0.7571

Table 3: Continued.

n			θ	λ_1	β_1	λ_2	β_2	δ	ρ_{11}	ρ_{12}	ρ_{21}	ρ_{22}
			$\theta = 2.8, \lambda_1 = 2.5, \beta_1 = 2.3, \lambda_2 = 1.2, \beta_2 = 1.4, \delta = 1.5, \rho_{11} = 0.2, \rho_{12} = 0.15, \rho_{21} = 0.05$, and $\rho_{22} = 0.1$									
100	MLE	Bias	−0.5762	−0.0823	−0.2263	−0.0867	−0.2029	−0.0499	0.7120	0.7732	0.5603	0.7975
		MSE	0.5342	0.1315	0.1006	0.0397	0.2027	0.0599	0.5081	0.6094	0.5719	0.6610
		L.CI	1.7637	1.3854	0.8717	0.7038	1.5765	0.9399	0.1328	0.3204	1.9921	0.6190
	Bayes	Bias	0.0215	0.0207	0.0349	−0.0399	−0.7671	0.0263	0.0449	0.0466	0.0222	0.0397
		MSE	0.0279	0.0290	0.0235	0.0168	0.5680	0.0208	0.0326	0.0325	0.0320	0.0333
		L.CI	0.6495	0.6631	0.5848	0.4842	1.1856	0.5560	0.6857	0.6834	0.6963	0.6984
200	MLE	Bias	−0.5672	−0.0468	−0.2164	−0.0944	−0.1768	−0.0482	0.7108	0.7750	0.5295	0.7978
		MSE	0.4720	0.1020	0.0814	0.0374	0.1643	0.0395	0.5061	0.6048	0.5534	0.6601
		L.CI	1.5204	1.2392	0.7294	0.6626	1.4303	0.7566	0.1079	0.2523	2.0494	0.6131
	Bayes	Bias	0.0160	0.0190	−0.0367	−0.0215	−0.6883	0.0228	0.2514	0.2327	0.0973	0.1968
		MSE	0.0261	0.0255	0.0186	0.0088	0.5463	0.0167	0.0423	0.0462	0.0342	0.0427
		L.CI	0.6540	0.6222	0.5145	0.3587	1.0562	0.4984	0.7673	0.8275	0.7702	0.8576

Table 4: MLE and Bayesian estimation with different sample sizes in the dependent case (Case II).

n			θ	λ_1	β_1	λ_2	β_2	δ	ρ_{11}	ρ_{12}	ρ_{21}	ρ_{22}
			$\theta = 2.8, \lambda_1 = 2.5, \beta_1 = 2.3, \lambda_2 = 1.2, \beta_2 = 1.4, \delta = 1.5, \rho_{11} = 0.5, \rho_{12} = 0.4, \rho_{21} = 0.2$, and $\rho_{22} = 0.3$									
50	MLE	Bias	−0.4624	−0.0246	−0.2569	−0.0415	−0.2156	−0.0937	0.4020	0.5533	0.4352	0.6387
		MSE	0.4118	0.1316	0.1287	0.0407	0.2198	0.0720	0.1654	0.3066	0.3893	0.4132
		L.CI	1.7452	1.4193	0.9823	0.7742	1.6326	0.9860	0.2407	0.0783	1.7538	0.2861
	Bayes	Bias	0.0165	−0.0031	0.0702	−0.0744	−0.7825	0.0261	0.1596	0.1794	0.0700	0.1406
		MSE	0.0326	0.0372	0.0327	0.0275	0.7174	0.0257	0.0484	0.0661	0.0430	0.0566
		L.CI	0.7053	0.7567	0.6533	0.5818	1.2712	0.6205	0.5941	0.7227	0.7655	0.7529
100	MLE	Bias	−0.4742	−0.0217	−0.2463	−0.0580	−0.2125	−0.0840	0.3971	0.5514	0.4487	0.6365
		MSE	0.3738	0.0999	0.1126	0.0314	0.1860	0.0576	0.1605	0.3043	0.3836	0.4105
		L.CI	1.5137	1.2367	0.8941	0.6571	1.4721	0.8821	0.2097	0.0612	1.6744	0.2871
	Bayes	Bias	0.0005	0.0147	0.0232	−0.0402	−0.2729	0.0313	0.0857	0.0935	0.0298	0.0569
		MSE	0.0295	0.0294	0.0219	0.0152	0.1619	0.0226	0.0325	0.0420	0.0330	0.0349
		L.CI	0.6736	0.6700	0.5729	0.4572	1.1604	0.5763	0.6218	0.7152	0.7026	0.6974
200	MLE	Bias	−0.4674	−0.0309	−0.2227	−0.0681	−0.1844	−0.0643	0.3963	0.5512	0.4006	0.6404
		MSE	0.3412	0.0788	0.0860	0.0266	0.1446	0.0430	0.1590	0.3040	0.3639	0.4081
		L.CI	1.3740	1.0943	0.7488	0.5815	1.3044	0.7731	0.1723	0.0479	1.7688	0.1133
	Bayes	Bias	−0.0175	0.0344	−0.0357	−0.0153	−0.5980	0.0324	0.2680	0.3121	0.1397	0.2743
		MSE	0.0236	0.0229	0.0170	0.0088	0.1422	0.0182	0.0482	0.0612	0.0416	0.0511
		L.CI	0.5982	0.5776	0.4924	0.3635	0.9984	0.5130	0.3896	0.5891	0.7677	0.7211

The pdf of BMEE distribution is defined as

$$f\left(x_d x, \theta,\right) = \sum_{i=1}^{2} p_i f_i\left(\left(x_d\right)_i, \theta_i\right), \qquad (37)$$

where p_i, are the mixing proportions, and it must satisfy $\sum_{i=1}^{2} p_i = 1$ and $p_i \geq 0$, and all of them are unknown. The pdf of first component of EE is given by (1):

$$f\left(x_d, \theta, \lambda_i\right) = \theta \lambda_i e^{-\lambda_i x_d}\left(1 - e^{-\lambda_i x_d}\right)^{\theta - 1}, \quad x_d > 0, \qquad (38)$$

with fixed shape parameter $\theta > 0$ and random scale parameter $\lambda > 0$, which take two distinct values λ_1 and λ_2, respectively. Similarly, for fixed shape parameter θ_2, let Y_d have an EE mixture density, and the pdf of second component EE is given by

$$g\left(y_d, \phi, \beta\right) = \delta \beta_i e^{-\beta_i y_d}\left(1 - e^{-\beta_i y_d}\right)^{\delta - 1}, \quad y_d > 0, \qquad (39)$$

with β being a random scale parameter taking values β_1 and β_2.

For a given values (λ, β), we assume that X_d and Y_d are dependent, and λ and β are correlated through their generalized bivariate distribution with the probability matrix given by

$$p = \begin{array}{c c} & \begin{array}{cc} \beta_1 & \beta_2 \end{array} \\ \begin{array}{c} \lambda_1 \\ \lambda_2 \end{array} & \begin{bmatrix} p_{\lambda_1 \beta_1} & p_{\lambda_1 \beta_2} \\ p_{\lambda_2 \beta_1} & p_{\lambda_2 \beta_2} \end{bmatrix} \end{array}. \qquad (40)$$

Let $f\left(x_d, y_d\right)$ be the joint pdf of $\left(X_d, Y_d\right)$; then,

$$f(x_d, y_d) = f(x_d|\theta,\lambda_1)g(y_d|\delta,\beta_1)c(F(x_d|\theta,\lambda_1), G(y_d|\delta,\beta_1))p_{\lambda_1\beta_1} + f(x_d|\theta,\lambda_1)g(y_d|\delta,\beta_2)c(F(x_d|\theta,\lambda_1), G(y_d|\delta,\beta_2))p_{\lambda_1\beta_2}$$
$$+ f(x_d|\theta,\lambda_2)g(y_d|\delta,\beta_1)c(F(x_d|\theta,\lambda_2), G(y_d|\delta,\beta_1))p_{\lambda_2\beta_1} + f(x_d|\theta,\lambda_2)g(y_d|\delta,\beta_2)c(F(x_d|\theta,\lambda_2), G(y_d|\delta,\beta_2))p_{\lambda_2\beta_2}. \tag{41}$$

Like the joint pdf in (8), the joint pdf in (41) can assume several different shapes as well. Let $x_d = x$ and $y_d = y$. Consequently, the associated BMEE distribution pdf will be

$$\begin{aligned}
f(x, y, \theta, \lambda, \beta, \delta) = {} & \theta\lambda_1 e^{-\lambda_1 x}\left(1 - \exp\{-\lambda_1 x\}\right)^{\theta-1}\delta\beta_1 e^{-\beta_1 y}\left(1 - \exp\{-\beta_1 y\}\right)^{\delta-1} \\
& \cdot\left\{\frac{1}{\sqrt{1-\rho_{11}^2}}\left(\exp\left[\frac{-\rho_{11}}{2\left(1-\rho_{11}^2\right)}\left\{\rho_{11}\left(z_1^2 + z_2^2\right) - 2z_1z_2\right\}\right]\right)\right\}p\lambda_1\beta_1 \\
& + \theta\lambda_1 e^{-\lambda_1 x}\left(1 - \exp\{-\lambda_1 x\}\right)^{\theta-1}\delta\beta_2 e^{-\beta_2 y}\left(1 - \exp\{-\beta_2 y\}\right)^{\delta-1} \\
& \cdot\left\{\frac{1}{\sqrt{1-\rho_{12}^2}}\left(\exp\left[\frac{-\rho_{12}}{2\left(1-\rho_{12}^2\right)}\left\{\rho_{12}\left(z_1^2 + z_2^2\right) - 2z_1z_2\right\}\right]\right)\right\}p\lambda_1\beta_2 \\
& + \lambda_2 e^{-\lambda_2 x}\left(1 - \exp\{-\lambda_2 x\}\right)^{\theta-1}\delta\beta_1 e^{-\beta_1 y}\left(1 - \exp\{-\beta_1 y\}\right)^{\delta-1} \\
& \cdot\left\{\frac{1}{\sqrt{1-\rho_{21}^2}}\left(\exp\left[\frac{-\rho_{21}}{2\left(1-\rho_{21}^2\right)}\left\{\rho_{21}\left(z_1^2 + z_2^2\right) - 2z_1z_2\right\}\right]\right)\right\}p\lambda_2\beta_1 \\
& + \lambda_2 e^{-\lambda_2 x}\left(1 - \exp\{-\lambda_2 x\}\right)^{\theta-1}\delta\beta_2 e^{-\beta_2 y}\left(1 - \exp\{-\beta_2 y\}\right)^{\delta-1} \\
& \cdot\left\{\frac{1}{\sqrt{1-\rho_{22}^2}}\left(\exp\left[\frac{-\rho_{22}}{2\left(1-\rho_{22}^2\right)}\left\{\rho_{22}\left(z_1^2 + z_2^2\right) - 2z_1z_2\right\}\right]\right)\right\}p\lambda_2\beta_2, \quad x, y, \theta, \lambda, \beta, \delta > 0,
\end{aligned} \tag{42}$$

where $\rho_{ij} \in [-1, 1]$ is the dependence parameter.

5.2. EM Algorithm under Gaussian Copula. The EM algorithm is introduced as a method of estimation. To apply the EM algorithm, as before, we augment the data (x_k, y_k), $k = 1, \ldots, n$, with the group membership variables (a_k, b_k, c_k), $k = 1, \ldots, n$, where a_k with the group membership variables (a_k, b_k, c_k), $k = 1, \ldots, n$, where a_k is one if the k^{th} observation is in $f_{ij}(x, y, \theta, \lambda_1, \beta_1, \delta, \rho_{11})$, and zero otherwise. Similarly, for b_k and c_k, we have four groups G_{ij}, $i, j = 1, 2$, for which the densities are

$$\begin{aligned}
f_{ij}\left(x, y, \theta, \lambda_i, \beta_j, \delta, \rho_{ij}\right) = {} & \theta\lambda_i e^{-\lambda_i x}\left(1 - \exp\{-\lambda_i x\}\right)^{\theta-1}\delta\beta_j e^{-\beta_j y}\left(1 - \exp\left\{-\beta_j y\right\}\right)^{\delta-1} \\
& \left\{\frac{1}{\sqrt{1-\rho_{ij}^2}}\left(\exp\left[\frac{-\rho_{ij}}{2\left(1-\rho_{ij}^2\right)}\left\{\rho_{ij}\left(z_1^2 + z_2^2\right) - 2z_1z_2\right\}\right]\right)\right\}, \quad x, y, \theta, \lambda_i, \beta_j, \delta > 0,
\end{aligned} \tag{43}$$

where $\rho_{ij} \in [-1, 1]$ is a dependence parameter.

The mixing proportions are given as follows: $P(G_{11}) = a$, $P(G_{12}) = b$, $P(G_{21}) = c$, and $P(G_{22}) = 1 - a - b - c$. We define $\ell_{ij}(x, y) = \log f_{ij}(x, y, \theta, \lambda_i, \beta_j, \delta, \rho_{ij})$; then, the EM algorithm as the method of estimation is given by finding the complete log likelihood, ℓ, as follows:

$$\ell=\sum_{k=1}^{n}a_k\ell_{11}(x_k,\ y_k)+\sum_{k=1}^{n}b_k\ell_{12}(x_k,\ y_k)+\sum_{k=1}^{n}c_k\ell_{12}(x_k,\ y_k)+\sum_{k=1}^{n}(1-\ a_k-b_k-c_k)\ell_{22}(x_k,\ y_k). \tag{44}$$

This is linear in the group membership variables $(a_k,\ b_k,\ c_k)$; consequently, in the E step, we enter into (26); their expected values, given the current estimates $(\widehat{\theta},\widehat{\lambda}_1,\widehat{\lambda}_2,\ \widehat{\beta}_1,\ \widehat{\beta}_2,\widehat{\delta},\widehat{\rho_{11}},\widehat{\rho_{12}},\ \widehat{\rho_{21}},\widehat{\rho_{22}},\ a,\ b,\ c)$ of the parameter, are calculated as

$$\widehat{a_k}=\frac{\widehat{a}\,f_{11}(x_k,\ y_k)}{\widehat{a}\,f_{11}(x_k,\ y_k)+\widehat{b}\,f_{12}(x_k,\ y_k)+\widehat{c}\,f_{21}(x_k,\ y_k)+(1-\widehat{a}\ -\widehat{b}-\widehat{c})f_{22}(x_k,\ y_k)}. \tag{45}$$

Similarly, for b_k and c_k, we follow the same strategies. Note that the algebraic simplification of the above might be necessary to avoid numerical problems. For the M-step, we need to maximize (27) over $(\theta,\lambda_1,\beta_1,\delta,\rho_{11},\rho_{12},\ \rho_{21},\rho_{22})$, for fixed values of (a_k,b_k,c_k). This is achieved by the conditional dependence of X and Y, given the group membership. We can essentially deal with the univariates and the Gaussian copula parameter separately. Differentiating (25) gives

$$\begin{aligned}\frac{\partial l}{\partial\lambda_1}=\frac{n}{\lambda_1}&\left\{\sum_{k=1}^{n}\left\{a_k\left\{\frac{1}{\sqrt{1-\rho_{11}^2}}\left(\exp\left[\frac{-\rho_{11}}{2\left(1-\rho_{11}^2\right)}\left\{\rho_{11}\left(z_1^2+z_2^2\right)-2z_1z_2\right\}\right]\right)\right\}\right.\right.\\&\left.\left.+b_k\left\{\frac{1}{\sqrt{1-\rho_{12}^2}}\left(\exp\left[\frac{-\rho_{12}}{2\left(1-\rho_{12}^2\right)}\left\{\rho_{12}\left(z_1^2+\ z_2^2\right)-2z_1z_2\right\}\right]\right)\right\}\right\}\right\}\\&+\sum_{k=1}^{n}\frac{x(\theta\ \exp\{-\lambda_1x\}-1)}{(1-\exp\{-\lambda_1x\})}\left\{\sum_{k=1}^{n}\left\{a_k\left\{\frac{1}{\sqrt{1-\rho_{11}^2}}\left(\exp\left[\frac{-\rho_{11}}{2\left(1-\rho_{11}^2\right)}\left\{\rho_{11}\left(z_1^2+z_2^2\right)-2z_1z_2\right\}\right]\right)\right\}\right.\right.\\&\left.\left.+b_k\left\{\frac{1}{\sqrt{1-\rho_{12}^2}}\left(\exp\left[\frac{-\rho_{12}}{2\left(1-\rho_{12}^2\right)}\left\{\rho_{12}\left(z_1^2+\ z_2^2\right)-2z_1z_2\right\}\right]\right)\right\}\right\}\right\}=0,\end{aligned} \tag{46}$$

$$\begin{aligned}\frac{n}{\lambda_2}&\left\{\sum_{k=1}^{n}\left\{c_k\left\{\frac{1}{\sqrt{1-\rho_{21}^2}}\left(\exp\left[\frac{-\rho_{21}}{2\left(1-\rho_{21}^2\right)}\left\{\rho_{21}\left(z_1^2+\ z_2^2\right)-2z_1z_2\right\}\right]\right)\right\}+(1-a_k-b_k-c_k)\right.\right.\\&\left.\left.\cdot\left\{\frac{1}{\sqrt{1-\rho_{22}^2}}\left(\exp\left[\frac{-\rho_{22}}{2\left(1-\rho_{22}^2\right)}\left\{\rho_{22}\left(z_1^2+\ z_2^2\right)-2z_1z_2\right\}\right]\right)\right\}\right\}\right\}\\&+\sum_{k=1}^{n}\frac{x(\theta\ \exp\{-\lambda_2x_k\}-1)}{(1-\exp\{-\lambda_2x_k\})}\left\{\sum_{k=1}^{n}\left\{c_k\left\{\frac{1}{\sqrt{1-\rho_{21}^2}}\left(\exp\left[\frac{-\rho_{21}}{2\left(1-\rho_{21}^2\right)}\left\{\rho_{21}\left(z_1^2+\ z_2^2\right)-2z_1z_2\right\}\right]\right)\right\}\right.\right.\\&\left.\left.+(1-a_k-b_k-c_k)\left\{\frac{1}{\sqrt{1-\rho_{22}^2}}\left(\exp\left[\frac{-\rho_{22}}{2\left(1-\rho_{22}^2\right)}\left\{\rho_{22}\left(z_1^2+\ z_2^2\right)-2z_1z_2\right\}\right]\right)\right\}\right\}\right\}=0,\end{aligned} \tag{47}$$

$$
\begin{aligned}
\frac{\partial l}{\partial \beta_1} = \frac{n}{\beta_1} \Bigg\{ \sum_{k=1}^{n} \Bigg\{ & a_k \Bigg\{ \frac{1}{\sqrt{1-\rho_{11}^2}} \left(\exp\left[\frac{-\rho_{21}}{2\left(1-\rho_{11}^2\right)} \left\{ \rho_{11}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \Bigg\} \\
& + c_k \Bigg\{ \frac{1}{\sqrt{1-\rho_{21}^2}} \left(\exp\left[\frac{-\rho_{21}}{2\left(1-\rho_{21}^2\right)} \left\{ \rho_{21}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \Bigg\} \Bigg\} \Bigg\} \\
& + \sum_{k=1}^{n} \frac{x\left(\theta \exp\{-\beta_1 y_k\} - 1\right)}{\left(1 - \exp\{-\beta_1 y_k\}\right)} \Bigg\{ \sum_{k=1}^{n} \Bigg\{ a_k \Bigg\{ \frac{1}{\sqrt{1-\rho_{11}^2}} \left(\exp\left[\frac{-\rho_{21}}{2\left(1-\rho_{11}^2\right)} \left\{ \rho_{11}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \Bigg\} \\
& + c_k \Bigg\{ \frac{1}{\sqrt{1-\rho_{21}^2}} \left(\exp\left[\frac{-\rho_{21}}{2\left(1-\rho_{21}^2\right)} \left\{ \rho_{21}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \Bigg\} \Bigg\} \Bigg\} = 0,
\end{aligned}
\tag{48}
$$

$$
\begin{aligned}
\frac{\partial l}{\partial \beta_2} = \frac{n}{\beta_2} \Bigg\{ \sum_{k=1}^{n} \Bigg\{ & b_k \Bigg\{ \frac{1}{\sqrt{1-\rho_{12}^2}} \left(\exp\left[\frac{-\rho_{12}}{2\left(1-\rho_{12}^2\right)} \left\{ \rho_{12}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \Bigg\} \\
& + \left(1 - a_k - b_k - c_k\right) \Bigg\{ \frac{1}{\sqrt{1-\rho_{22}^2}} \left(\exp\left[\frac{-\rho_{22}}{2\left(1-\rho_{22}^2\right)} \left\{ \rho_{22}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \Bigg\} \Bigg\} \Bigg\} \\
& + \sum_{k=1}^{n} \frac{y_k\left(\theta \exp\{-\beta_2 y_k\} - 1\right)}{\left(1 - \exp\{-\beta_2 y_k\}\right)} \Bigg\{ \sum_{k=1}^{n} \Bigg\{ b_k \Bigg\{ \frac{1}{\sqrt{1-\rho_{12}^2}} \left(\exp\left[\frac{-\rho_{12}}{2\left(1-\rho_{12}^2\right)} \left\{ \rho_{12}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \Bigg\} \\
& + \left(1 - a_k - b_k - c_k\right) \Bigg\{ \frac{1}{\sqrt{1-\rho_{22}^2}} \left(\exp\left[\frac{-\rho_{22}}{2\left(1-\rho_{22}^2\right)} \left\{ \rho_{22}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \Bigg\} \Bigg\} \Bigg\} = 0,
\end{aligned}
\tag{49}
$$

$$
\begin{aligned}
\frac{\partial l}{\partial \theta} = & \left\{ \frac{n}{\theta} + \sum_{k=1}^{n} \log\left(1 - \exp\{-\lambda_1 x_k\}\right) \right\} \\
& \left\{ \begin{array}{c} \left\{ \sum_{k=1}^{n} \left\{ a_k \left\{ \frac{1}{\sqrt{1-\rho_{11}^2}} \left(\exp\left[\frac{-\rho_{11}}{2\left(1-\rho_{11}^2\right)} \left\{ \rho_{11}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \right\} + b_k \left\{ \frac{1}{\sqrt{1-\rho_{12}^2}} \left(\exp\left[\frac{-\rho_{12}}{2\left(1-\rho_{12}^2\right)} \left\{ \rho_{12}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \right\} \right\} \right\} + \left\{ \frac{n}{\theta} + \sum_{k=1}^{n} \log\left(1 - \exp\{-\lambda_2 x_k\}\right) \right\} \\
\left\{ \sum_{k=1}^{n} \left\{ c_k \left\{ \frac{1}{\sqrt{1-\rho_{11}^2}} \left(\exp\left[\frac{-\rho_{21}}{2\left(1-\rho_{21}^2\right)} \left\{ \rho_{11}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \right\} + \left(1 - a_k - b_k - c_k\right) \left\{ \frac{1}{\sqrt{1-\rho_{22}^2}} \left(\exp\left[\frac{-\rho_{22}}{2\left(1-\rho_{22}^2\right)} \left\{ \rho_{22}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \right\} \right\} \right\} \end{array} \right\} = 0,
\end{aligned}
\tag{50}
$$

$$
\frac{\partial l}{\partial \delta} = \frac{\partial l}{\partial \delta} \left\{ \begin{array}{c} \left\{ \frac{n}{\delta} + \sum_{k=1}^{n} \log\left(1 - \exp\{-\beta_1 y_k\}\right) \right\} \left\{ \sum_{k=1}^{n} a_k \left\{ \frac{1}{\sqrt{1-\rho_{11}^2}} \left(\exp\left[\frac{-\rho_{11}}{2\left(1-\rho_{11}^2\right)} \left\{ \rho_{11}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \right\} + c_k \left\{ \frac{1}{\sqrt{1-\rho_{21}^2}} \left(\exp\left[\frac{-\rho_{21}}{2\left(1-\rho_{21}^2\right)} \left\{ \rho_{21}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \right\} \right\} \\
+ \left\{ \frac{n}{\delta} + \sum_{k=1}^{n} \log\left(1 - \exp\{-\beta_2 y_k\}\right) \right\} \left\{ \sum_{k=1}^{n} b_k \left\{ \frac{1}{\sqrt{1-\rho_{12}^2}} \left(\exp\left[\frac{-\rho_{12}}{2\left(1-\rho_{12}^2\right)} \left\{ \rho_{12}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \right\} + 1 - a_k - b_k - c_k \left\{ \frac{1}{\sqrt{1-\rho_{22}^2}} \left(\exp\left[\frac{-\rho_{22}}{2\left(1-\rho_{22}^2\right)} \left\{ \rho_{22}\left(z_1^2 + z_2^2\right) - 2z_1 z_2 \right\} \right] \right) \right\} \right\} \end{array} \right\} = 0.
\tag{51}
$$

The approach involves a two-step procedure in estimating the marginal of X and Y and the copula function independently that gives the maximum likelihood estimation of $\theta, \lambda_1, \beta_1, \lambda_2, \beta_2$, δ, ρ_{11}, ρ_{12}, ρ_{21}, and ρ_{22}. The solution of the nonlinear (46)–(51) gives the MLE of $\theta, \lambda_1, \beta_1, \lambda_2, \beta_2$, and δ and $\theta, \lambda_1, \beta_1, \lambda_2, \beta_2$, δ, ρ_{11}, ρ_{12}, ρ_{21}, and ρ_{22}.

Then, copula density is estimated as

$$
\log L_{ij}(\rho) = \sum_{k=1}^{n} \log c\left(\widehat{F}_i(x_k), \widehat{F}_j(y_k)\right),
\tag{52}
$$

where $\widehat{F}_i(x_k)$ and ,$\widehat{F}_j(y_k)$ denote the maximum likelihood estimates of the pdf from the first step. The solution of the nonlinear equation (34) gives the MLE of ρ_{11}, ρ_{12}, ρ_{21}, and ρ_{22}.

The M-step is completed by setting

$$
\widehat{a} = \frac{1}{n} \sum_{k=1}^{n} a_k .
\tag{53}
$$

We use the copula R package to solve these equations numerically. After the maximum likelihood estimators for

$\theta, \lambda_1, \beta_1, \lambda_2, \beta_2, \delta, \rho_{11}, \rho_{12}, \rho_{21}$, and ρ_{22} are obtained, next, we substitute these estimates in (a_k, b_k, c_k). We complete the M-step by setting $\hat{a} = 1/n\sum_{k=1}^{n} a_k$, etc.

Initial values of the parameters for the mixing proportions are obtained by the method of matching moments that are obtained from the marginal univariate EEM and the Gaussian copula parameter separately, Then, we take the resulting estimates of the BEE parameters as starting values for the EM algorithm. Next, we merge the moment estimators of the marginal mixing parameters to obtain initial values for the bivariate mixing parameters, assuming the dependence between two variables X and Y. We apply this method in application as mentioned in Section 9, specifically, in Tables 1–4. For more details, see the work of Kosmidis and Karlis [31].

Next, we provide the estimation procedure of the unknown parameters for the density in (44). In the copula-based estimation, we adopt two approaches that are termed as parametric and semiparametric.

5.3. Maximum Likelihood Estimation (MLE). Here, we discuss the estimation of the unknown parameters of BEE distributions by the approach of the maximum likelihood, by using the two-step estimation. It involves a two-step procedure by which we estimate the marginal and the copula function separately.

The log-likelihood function is expressed as

$$\begin{aligned}\log L = \sum_{i=1}^{n} [\log f_1(x_i) + \log f_2(y_i) \\ + \log c(F_1(x_i), F_2(y_i))].\end{aligned} \tag{54}$$

The log-likelihood function in (35) can be re-expressed as

$$\begin{aligned}logL = \sum_{i=1}^{n} \log f_1(x_i) + \sum_{i=1}^{n} \log f_2(x_{2i}) \\ + \sum_{i=1}^{n} \log c(F_1(x_{1i}), F_2(x_{2i})).\end{aligned} \tag{55}$$

The first step involves estimating the parameters of marginals distribution F_1 and F_2 by MLE, separately given as follows:

$$\begin{aligned}\log L_1 &= \sum_{i=1}^{n} \log \ \log f_1(x_i), \\ \log L_2 &= \sum_{i=1}^{n} \log f_2(y_i).\end{aligned} \tag{56}$$

Then, estimating copula parameters by maximizing the copula density, we will obtain

$$\log L = \sum_{i=1}^{n} \log c(F_1(x_i), F_2(y_i)). \tag{57}$$

By considering the first step with EE distributions, the parameters of each marginal distribution will be estimated by the MLE. If $x_1, \ldots, x_n$ is a random sample from $\mathrm{EE}(\theta, \lambda)$ and $y_1, \ldots, y_n$ is a random sample from $\mathrm{EE}(\delta, \beta)$, then the log-likelihood functions are, respectively, given by

$$\log L_1(x, \theta, \lambda) = n \log(\theta) + n \log(\lambda) - \sum_{i=1}^{n} x_i + (\theta - 1) \sum_{i=1}^{n} \log(1 - \exp\{-\lambda x_i\}), \tag{58}$$

$$\log L_2(y, \delta, \beta) = n \log(\delta) + n \log(\beta) - \sum_{i=1}^{n} y_i + (\delta - 1) \sum_{i=1}^{n} \log(1 - \exp\{-\beta y_i\}). \tag{59}$$

So, the maximum likelihood equations are

$$\frac{\partial l}{\partial \lambda} = \frac{n}{\lambda} + \sum_{i=1}^{n} \frac{x_i(\theta \ \exp\{-\lambda x_i\} - 1)}{(1 - \exp\{-\lambda x_i\})} = 0, \tag{60}$$

$$\frac{\partial l}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^{n} \frac{y_i(\delta \ \exp\{-\beta y_i\} - 1)}{(1 - \exp\{-\beta y_i\})} = 0, \tag{61}$$

$$\frac{\partial l}{\partial \theta} = \left\{\frac{n}{\theta} + \sum_{i=1}^{n} \log(1 - \exp\{-\lambda x_i\})\right\} = 0, \tag{62}$$

$$\frac{\partial l}{\partial \delta} = \left\{\frac{n}{\delta} + \sum_{i=1}^{n} \log(1 - \exp\{-\beta y_i\})\right\} = 0, \tag{63}$$

The solution of the system of nonlinear equations (60)–(63) gives the MLE of θ, λ, β, and δ. Then, copula density is estimated as follows:

$$\log L(\gamma) = \sum_{i=1}^{n} \log c\left(\hat{F}_1(x_i), \hat{F}_2(y_i)\right), \tag{64}$$

where $\hat{F}_1(x)$ and $\hat{F}_2(y)$ denote the ML estimates of the parameters from the first step.

The solution of the nonlinear equation (64) gives the MLE of γ.

5.4. Semiparametric Methods of Estimation. Two semiparametric methods are used to estimate the copula parameter in the copula models and are compared with the two methods of moments approaches which are the inversion Kendall's τ and inversion of Spearman's ρ, respectively.

5.5. Methods of Moments. From the moment's method of inversion of Kendall's τ and the inversion of Spearman's ρ mentioned in Kojadinovic and Yan [32], we provide a brief details which are given as follows. Let c be a bivariate random sample from a cdf $C_\gamma\,[F_1(x),\,F_2(y)]$, where F_1 and F_2 are continuous cdf's and C_γ is an absolutely continuous copula such that $\gamma \in \mathcal{O}$, where $\mathcal{O}$ is an open subset of R^2. Furthermore, let $R_1, \ldots, R_n$ are the vectors of ranks associated with $x_1, \ldots, x_n$ unless otherwise stated. In what follows, all vectors are row vectors. Moment's approaches are based on the inversion of a consistent estimator of a moment of the copula C_γ. The two best-known moments, Spearman's rho and Kendall's tau, are, respectively, given by

$$\rho(\gamma) = 12\int\int_{[0,1]^2} u\,v\;dC_\gamma(u,v) - 3, \tag{65}$$

$$\tau(\gamma) = 4\int\int_{[0,1]^2} C_\gamma(u,v)dC_\gamma(u,v) - 1. \tag{66}$$

Consistent estimators of these two moments can be expressed as

$$\rho_n = \frac{12}{n(n+1)(n-1)}\sum_{i=1}^{n} R_{i,1}R_{i,2} - 3\frac{n+1}{n-1}, \tag{67}$$

$$\tau_n = \frac{4}{n(n-1)}\sum_{i=1}^{n} 1\left[x_{i,1} \le x_{j,1}\right]1\left[x_{i,2} \le x_{j,2}\right] - 1. \tag{68}$$

If ρ and τ are one-to-one, consistent estimators of γ will be $\gamma_{n,\rho} = \rho^{-1}(\rho_n)$ and $\gamma_{n,\tau} = \tau^{-1}(\tau_n)$, respectively.

It can be called inversion of Kendall's τ and inversion of Spearman's ρ, respectively. For more information, see the work of Kojadinovic and Yan [32] and the references cited therein.

As explained above, the moment's method of τ and ρ estimation for copula may be considered under the umbrella of semiparametric approach estimation.

5.6. Goodness of Fit Tests for Copula. We want to compare the empirical copula with the parametric estimator derived under the null hypothesis; for details, see the work of Fermanian [33]. Theory suggests a test if C is well-represented by a specific copula C_γ:

$$H_0\colon C = C_\gamma \text{ Vs. } H_1\colon C \ne C_\gamma. \tag{69}$$

Several well-known approaches are available in the literature; for example, see the work of Genest and Rémillard [34], or the fast multiplier approach, Genest et al. [35], and Kojadinovic et al. [36]. The goodness of fit tests based on the empirical process is given as

$$\mathbb{C}_n(u,v) = \sqrt{n}\left\{C_n(u,v) - C_{\gamma_n}(u,v)\right\}, \tag{70}$$

where $C_n(u,v)$ is the empirical copula of the data of X and Y

$C_n(u,v) = 1/n\sum_{i=1}^{n} 1\,(U_{i,n} \le u, V_{i,n} \le v), \quad u, v \in [0,1]$,

and $U_{i,n}$ and $V_{i,n}$ are pseudoobservations from C calculated from data as follows.

$U_{i,n} = R_{1i}/n+1$, $V_{i,n} = R_{2i}/n+1$, and R_{1i} and R_{2i} are, respectively, the ranks of X_i and Y_i.

Here, $C_n(u,v)$ is a consistent estimator, and θ_n is an estimator of γ obtained using the pseudoobservations. According to Genest et al. [35], the appropriate test statistics is the Cramer-von Miss and is defined as

$$S_n = \sum_{i=1}^{n}\left\{C_n\left(U_{i,n}, V_{i,n}\right) - C_{\gamma_n}\left(U_{i,n}, V_{i,n}\right)\right\}^2. \tag{71}$$

6. Bayesian Estimation

In this section, the Bayes estimates of the model parameters of the joint pdf (51) are obtained under the assumption that the random variables, $\Phi = (\theta, \lambda_1, \beta_1, \lambda_2, \beta_2,\, \delta)$, have an independent gamma prior distributions with hyperparameter w_k and m_k, $k = 1, 2, 3, 4, 5, 6$, given by

$$f(\Phi; w, m) = \frac{m_k w_k}{\Gamma w_k}\Phi^{w_k - 1}e^{-m_k\Phi}, \quad \Phi > 0, \tag{72}$$

and $\rho_{11}, \rho_{12}, \rho_{21}$, and ρ_{22} have a noninformative prior.

By multiplying (23) or (44) with (72), the joint posterior density for the vector Φ, given the data, becomes

$$\pi(\Phi|\,\underline{x}) \propto L(x|\Phi)f(\Phi; w_k, s_k). \tag{73}$$

Marginal distributions of Φ can be obtained by integrating out the (nuisance) hyperparameters. Thus, the Bayesian estimators of the parameters Φ under square error loss function can be calculated as follows:

$$\widehat{\Phi} \propto \int_{\widehat{\Phi}}^{\infty} \Phi\pi(\Phi|\,\underline{x})d\Phi. \tag{74}$$

The integrals in (74) cannot be obtained in a closed form, so the Markov chain Monte Carlo (MCMC) technique is used. In MCMC methods, the posterior distribution and the intractable integrals using simulated samples from the posterior distribution are obtained. Also, Gibbs sampling and the Metropolis–Hastings (MH) algorithm as a MCMC technique are used. For more details, see Metropolis et al. [37], Hastings [38], and Mohsin et al. [39]. The M-H algorithm considers that, to each iteration of the algorithm, an applicant value can be generated from a proposed distribution. Thus, the applicant value is allowed according to a sufficient approval probability. This technique assurances the convergence of the Markov chain for the target density. Finally, we can investigate that the advantage of the MCMC method over the MLE method is that we can always obtain a reasonable interval estimate of the parameters by constructing the probability intervals based on empirical posterior distribution. This is often unavailable in MLE.

6.1. Credible Intervals. In this section, a symmetric $100(1-\varepsilon)\%$ two-sided Bayes probability interval estimate of Φ, denoted by $[\mathrm{L}_\Phi, \mathrm{U}_\Phi]$, can be obtained as satisfying the following expression:

$$p\left[L(\underline{t}) < \Phi < U(\underline{t})\right] = \int_{L(\underline{t})}^{U(\underline{t})} \pi(\theta, \beta, \lambda|\,\underline{t})d\Phi = 1 - \varepsilon, \tag{75}$$

Since it is difficult to find the interval L_Φ and U_Φ analytically, therefore, we apply suitable numerical techniques to solve this nonlinear equation.

Table 5: MLE and Bayesian estimation with different sample sizes in dependent case (Case III).

n			θ	λ_1	β_1	λ_2	β_2	δ	ρ_{11}	ρ_{12}	ρ_{21}	ρ_{22}
			$\theta = 1.8, \lambda_1 = 1.5, \beta_1 = 1.3, \lambda_2 = 2.2, \beta_2 = 2.5,\ \delta = 2, \rho_{11} = 0.2, \rho_{12} = 0.15,\ \rho_{21} = 0.05$, and $\rho_{22} = 0.1$									
50	MLE	Bias	−0.2789	−0.2044	−0.4097	−0.1416	−0.0470	−0.2871	0.6815	0.7838	0.6638	0.8068
		MSE	0.1968	0.1373	0.3126	0.2071	0.0898	0.2926	0.5099	0.6165	0.6176	0.6765
		L.CI	1.3529	1.2123	1.4922	1.6964	1.1610	1.7979	0.8371	0.1837	1.6502	0.6260
	Bayes	Bias	0.0048	0.0282	0.2110	−0.2575	−0.4632	0.0392	0.1802	0.2637	0.1278	0.2249
		MSE	0.0304	0.0275	0.0829	0.1301	0.2832	0.0323	0.0831	0.1151	0.0576	0.0965
		L.CI	0.6831	0.6410	0.7681	0.9907	1.0279	0.6882	0.8825	0.8369	0.7972	0.8401
100	MLE	Bias	−0.2655	−0.1979	−0.4074	−0.1566	−0.0580	−0.2896	0.6933	0.7816	0.7166	0.8126
		MSE	0.1589	0.1101	0.2651	0.1787	0.0615	0.2322	0.5080	0.6114	0.6043	0.6766
		L.CI	1.1662	1.0443	1.2345	1.5397	0.9454	1.5106	0.6493	0.0901	1.4128	0.5019
	Bayes	Bias	0.0084	0.0387	0.1821	−0.2361	−0.3387	0.0504	0.0882	0.1205	0.0521	0.1102
		MSE	0.0251	0.0263	0.0680	0.1058	0.1540	0.0319	0.0482	0.0528	0.0403	0.0508
		L.CI	0.6204	0.6176	0.7321	0.8775	0.7772	0.6721	0.7882	0.7668	0.7598	0.7716
200	MLE	Bias	−0.2401	−0.1699	−0.4124	−0.1019	−0.0545	−0.2706	0.6920	0.7819	0.7175	0.8152
		MSE	0.1206	0.0708	0.2375	0.1395	0.0405	0.1849	0.4945	0.6012	0.6014	0.6727
		L.CI	0.9844	0.8035	1.0189	1.4095	0.7602	1.3105	0.4901	0.0842	1.3941	0.3565
	Bayes	Bias	−0.0205	0.0173	0.1272	−0.2147	−0.2106	0.0545	0.1802	0.2637	0.1278	0.2249
		MSE	0.0203	0.0166	0.0449	0.0844	0.0604	0.0274	0.0374	0.0567	0.0226	0.0359
		L.CI	0.5536	0.5010	0.6641	0.7673	0.4963	0.6127	0.0366	0.0371	0.0312	0.0327

Table 6: MLE and Bayesian estimation with different sample sizes in dependent case (Case IV).

n			θ	λ_1	β_1	λ_2	β_2	δ	ρ_{11}	ρ_{12}	ρ_{21}	ρ_{22}
			$\theta = 1.8, \lambda_1 = 1.5, \beta_1 = 1.3, \lambda_2 = 2.2, \beta_2 = 2.5,\ \delta = 2, \rho_{11} = 0.5, \rho_{12} = 0.4,\ \rho_{21} = 0.2$, and $\rho_{22} = 0.3$									
50	MLE	Bias	−0.2722	−0.2106	−0.3576	−0.1175	−0.0275	−0.2251	0.4088	0.5547	0.5258	0.6412
		MSE	0.1722	0.1153	0.2311	0.1548	0.0602	0.1793	0.1705	0.3081	0.4137	0.4138
		L.CI	1.2287	1.0445	1.2600	1.4728	0.9565	1.4068	0.2289	0.0805	1.4525	0.2035
	Bayes	Bias	0.0091	0.0273	0.2060	−0.2488	−0.2623	0.0337	0.2394	0.3242	0.1440	0.2928
		MSE	0.0266	0.0316	0.0800	0.1152	0.0593	0.0310	0.0714	0.1244	0.0633	0.1180
		L.CI	0.6392	0.6885	0.7608	0.9056	0.9579	0.6777	0.4659	0.5448	0.8091	0.7048
100	MLE	Bias	−0.2827	−0.2017	−0.3537	−0.1082	−0.0122	−0.2140	0.4068	0.5526	0.5548	0.6393
		MSE	0.1422	0.0880	0.2032	0.1193	0.0421	0.1431	0.1678	0.3056	0.4197	0.4115
		L.CI	0.9789	0.8537	1.0958	1.2863	0.8035	1.2235	0.1903	0.0584	1.3120	0.2063
	Bayes	Bias	−0.0005	0.0252	0.1543	−0.2224	−0.3411	0.0487	0.1335	0.1967	0.0710	0.1565
		MSE	0.0221	0.0243	0.0567	0.0915	0.0582	0.0298	0.0443	0.0742	0.0452	0.0664
		L.CI	0.5827	0.6037	0.7111	0.8041	0.7456	0.6499	0.6379	0.7396	0.7856	0.8025
200	MLE	Bias	−0.2756	−0.1894	−0.3556	−0.0970	−0.0193	−0.1949	0.4045	0.5524	0.5770	0.6418
		MSE	0.1281	0.0742	0.1840	0.0928	0.0317	0.1230	0.1659	0.3053	0.4360	0.4124
		L.CI	0.8956	0.7681	0.9407	1.1326	0.6946	1.1433	0.1877	0.0488	1.2592	0.0907
	Bayes	Bias	−0.0140	0.0009	0.0934	−0.1745	−0.1933	0.0461	0.0572	0.0956	0.0262	0.0594
		MSE	0.0183	0.0169	0.0350	0.0590	0.0535	0.0252	0.0284	0.0447	0.0324	0.0390
		L.CI	0.5283	0.5104	0.6359	0.6628	0.4978	0.5952	0.6217	0.7399	0.6980	0.7383

7. Simulation Study

Here, a simulation study is conducted to see the efficacy of the proposed model in two cases, independent case and the dependent case. Monte Carlo simulation is done for comparison between maximum likelihood and Bayesian estimation methods, for estimating parameters of BEEM distribution using *R* language such as (bbmle). The MLE estimation methods are done based on Newton–Raphson algorithm by using "maxLik" package. The Bayesian estimation by using Markov Chain Monte Carlo (MCMC) approach and Metropolis–Hastings (MH) algorithm is carried out by using R-program. Monte Carlo simulations are considered based on a data that is generated from a Gaussian copula by using copula package in R. We generate 10000 random samples of sizes $n = 50$, 100, and 200, and different cases of actual values of the parameters are listed below:

Case I: $\theta = 2.8, \lambda_1 = 2.5, \beta_1 = 2.3, \lambda_2 = 1.2, \beta_2 = 1.4$, $\delta = 1.5,\ \rho_{11} = 0.2, \rho_{12} = 0.15,\ \rho_{21} = 0.05$, and $\rho_{22} = 0.1$

Case II: $\theta = 2.8, \lambda\ 1 = 2.5, \beta_1 = 2.3, \lambda_2 = 1.2, \beta_2 = 1.4$, $\delta = 1.5,\ \rho_{11} = 0.5, \rho_{12} = 0.4,\ \rho_{21} = 0.2$, and $\rho_{22} = 0.3$

Case III: $\theta = 1.8, \lambda_1 = 1.5, \beta_1 = 1.3, \lambda_2 = 2.2,\ \beta_2 = 2.5$, $\delta = 2,\ \rho_{11} = 0.2, \rho_{12} = 0.15,\ \rho_{21} = 0.05$, and $\rho_{22} = 0.1$

Table 7: MLE and Bayesian estimation for the BMEE distribution using the dataset.

		θ	λ	β	δ	ρ
MLE	Coefficient	6.5195	0.0128	8.4580	0.0129	0.9791
	S.E	0.0755	0.0042	0.1170	0.0110	0.1975
Bayesian	Coefficient	6.5882	0.0130	8.5762	0.0133	0.9576
	S.E	0.0650	0.0016	0.0703	0.0015	0.0205

Table 8: KS distance and its p value and MLE.

	$\widehat{\theta}$	$\widehat{\lambda}_1$	KS-d	p value	L
X	11.4721	0.0151	0.2664	0.1553	102.1654
	5.5896	0.0029			
Y	9.8507	0.0138	0.1001	0.9853	103.8703
	4.7174	0.0028			

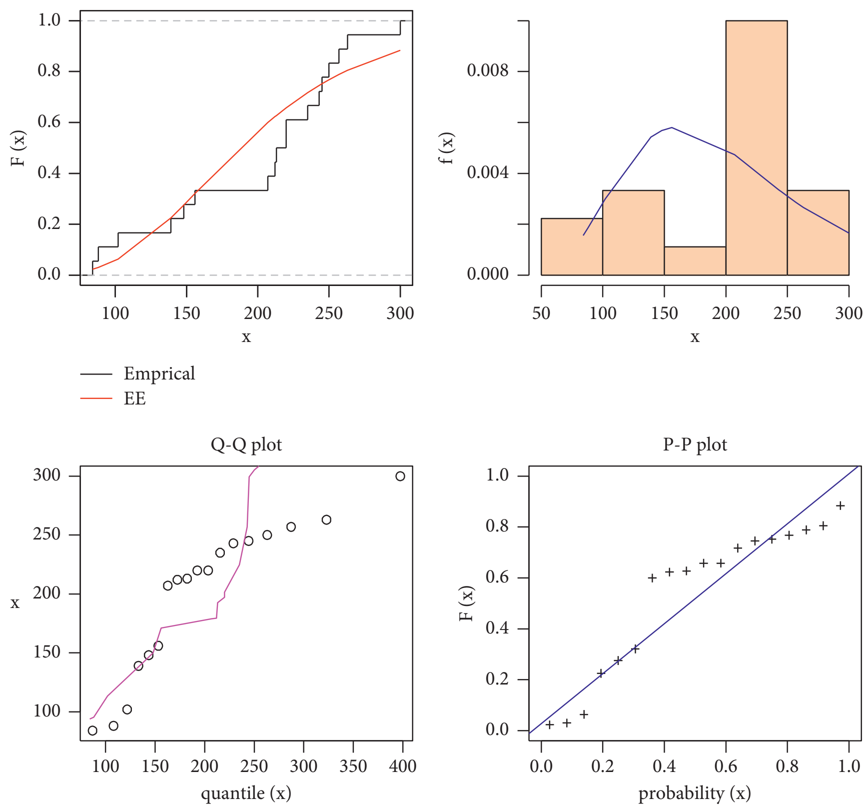

Figure 3: The fitted pdf, cdf, PP plot, and QQ plot for EE distribution for the first variable.

Case IV: $\theta = 1.8, \lambda_1 = 1.5, \beta_1 = 1.3, \lambda_2 = 2.2, \beta_2 = 2.5, \delta = 2, \rho_{11} = 0.5, \rho_{12} = 0.4, \rho_{21} = 0.2$, and $\rho_{22} = 0.3$

We claim the best performance as the method which minimizes the mean squared error (MSE), bias of estimation, and length of confidence interval (L.CI) of the estimator. The two-sided confidence limit with confidence level $\gamma = 0.95$ of the parameters is constructed as well.

From the reported bias, MSE, and L.CI, in Tables 3–6, it appears that the efficiency of the estimation under the Bayesian paradigm is quite evident. In particular, the MSE values in all parametric combinations tried in this article support in favor of this statement. We have the following observation on the simulation study as follows:

(i) As sample size (n) increases and for the same case (fixed actual parameters), the bias, MSE, and L.CI

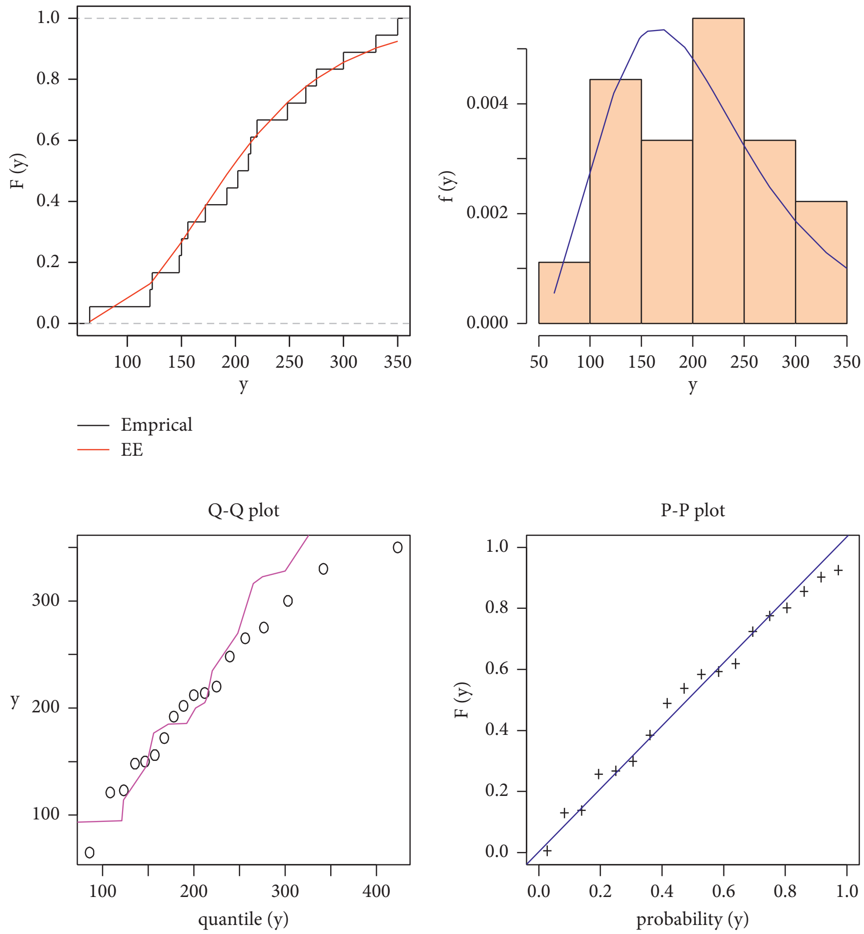

Figure 4: The fitted pdf, cdf, PP plot, and QQ plot for EE distribution for the second variable.

associated with the parameter estimates decrease for both methods of estimation.

(ii) In majority of the situations, for fixed n and for the correlation matrix parameters ρ_{11}, ρ_{12}, ρ_{21} , and ρ_{22}, the bias, MSE, and L.CI associated with them decrease under both the classical approach (MLE method) as well as under the Bayesian paradigm. As expected, when the sample size increases, a greater efficiency is observed in the overall estimation procedure (i.e., low bias and smaller values of the MSE).

(iii) From the reported bias, MSE, and L.CI, it appears that the efficiency of the estimation under the Bayesian paradigm is quite evident for both independent and dependent cases. In particular, the MSE values in all parametric combinations tried in this article support in favor of this statement. For reference, see Tables 1–6 in the revised manuscript.

(iv) Credible interval(s) constructed appear to be the best indicator on the estimation of the model parameters as expected.

(v) Bias measures whether, over many replications, the estimator yields result that is correct on an average. A statistic is positively biased if it tends to overestimate the parameter; a statistic is negatively biased if it tends to underestimate the parameter. Negative bias means that the estimator is too small on an average compared to the true value.

8. Real Dataset (Motor Data)

The data represent the failure times of a parallel system constituted by two identical motors in days. These data are reported in Relia Soft (2003), where $X = (102, 84, 88, 156, 148, 139, 245, 235, 220, 207, 250, 212, 213, 220, 243, 300, 257, 263)$. $Y = (65, 148, 202, 121, 123, 150, 156, 172, 192, 214, 212, 220, 265, 275, 300, 248, 330, 350)$. We fit at first the marginals of X and Y separately on the motor data. The MLE and Bayesian estimation for the BMEE distribution using dataset are shown in Table 7. The MLE of the parameters Kolmogorov–Smirnov distance (KS-d) and its p value for the marginals are listed in Table 8.

The empirical cdf, the histogram of the pdf, PP plots, and QQ plots are displayed for the first variable in Figure 3. The empirical cdf, the histogram of the pdf, PP plots, and QQ plots are displayed for the second variable in Figure 4. The

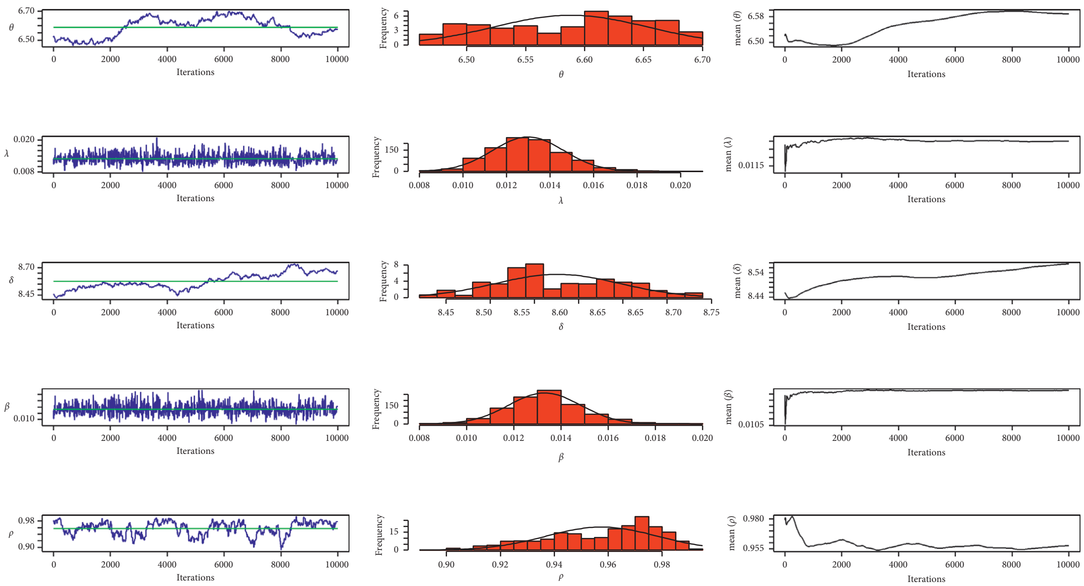

Figure 5: The Markov Chain Monte Carlo (MCMC) plots for motor data using the BMEE model.

EE distribution is fitted to real data using Kolmogorov–Smirnov goodness-of-fit test. The estimated cdf with empirical cdf, the histogram of the pdf, PP plots, and QQ plots are displayed for the first variable in Figure 3. The estimated cdf with empirical cdf, the histogram of the pdf, PP plots, and QQ plots are displayed for the second variable in Figure 4. These figures show that the tow variables are fitted for the marginal EE distribution. Also, Table 8 confirmed this conclusion by using Kolmogorov–Smirnov goodness-of-fit test, where p values are more than 0.05.

In majority of the cases, one may observe that the estimates under the Bayesian method are preferable than the measures of MLE estimates as the standard error of the estimates are smaller in all estimates of the parameters. History plots, approximate marginal posterior density, and MCMC convergence of θ, λ, δ, β, and ρ are represented in Figure 5.

9. Conclusions

In this paper, we have proposed and studied a new class of BEE whose marginals are EE distributions. The proposed class of distribution is constructed via two different types of mixture: (a) type I: starting with two independent EE distributions and (b) type II: using a bivariate Gaussian copula. Estimation of the model parameters for both types of MBEE distribution are conducted using classical (the method of moments and the method of maximum likelihood) and under the Bayesian paradigm using independent gamma priors. Since the joint distribution function and the joint density function are in closed forms, consequently, this distribution can be used in practice for nonnegative and positively correlated random variables. Since the maximum likelihood estimators of the unknown parameters cannot be obtained in the closed form, we consider the EM algorithm that works quite well, and it can be effectively used to compute the MLEs. Since the choice of hyperparameters for a prior in a Bayesian paradigm is of paramount importance, as a continuation of this work in future, we will be focusing on (including but not limited to) the following:

(i) Exploring various strategies (for example, matching conditional moments, or conditional percentiles information to be provided by our expert with the corresponding theoretical moments and percentiles and subsequently assuming, say, Euclidean distance) to estimate/guess best choice(s) of the hyperparameters for the priors.

(ii) In the current work, our prior choices are mostly conjugate in nature. However, in a real-life scenario, we might not have such an information on the prior always. Also, in the presence of more concrete information, one might consider a more precise prior for the model parameter(s), possibly a partially informative improper prior. It would be interesting to see the effect on the overall efficiency of the MCMC and Gibbs sampling in such a setting. We are currently working on it, and it will be reported somewhere else.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the Deanship of Scientific Research at Princess Nourah Bint Abdulrahman University through the Fast-track Research Funding Program.

References

[1] S. Newcomb, "A generalized theory of the combination of observations so as to obtain the best result," *American Journal of Mathematics*, vol. 8, no. 4, pp. 343–366, 1886.

[2] K. Pearson, "Contributions to the mathematical theory of evolution," *Philosophical Transactions*, vol. 185, pp. 71–110, 1894.

[3] D. M. Titterington, A. F. M. Smith, and U. E. Makov, *Statistical Analysis of Finite Mixture Distributions*, Wiley, New York, NY, USA, 1985.

[4] B. S. Everitt and D. J. Hand, *Finite Mixture Distributions*, John Wiley & Sons, Ltd., Amsterdam, Netherlands, 1981.

[5] G. J. McLachlan and K. E. Basford, *Mixture Models Inferences and Applications to Clustering*, Marcel Dekker, New York, NY, USA, 1988.

[6] E. K. Al-Hussaini and K. S. Sultan, "Reliability and hazard based on finite mixture Models," in *Handbook of Statistics*, N. Balakrishnan and C. R. Rao, Eds., pp. 139–183, Elsevier Science, Amsterdam, The Netherlands, 2001.

[7] R. D. Gupta and D. Kundu, "Closeness of gamma and generalized exponential distribution," *Communications in Statistics - Theory and Methods*, vol. 32, no. 4, pp. 705–721, 2003.

[8] R. D. Gupta and D. Kundu, "Generalized exponential distributions," *Australian and New Zealand Journal of Statistics*, vol. 41, no. 2, pp. 173–188, 1999.

[9] R. D. Gupta and D. Kundu, "Exponentiated exponential family: an alternative to gamma and Weibull distributions," *Biometrical Journal*, vol. 43, no. 1, pp. 117–130, 2001.

[10] R. D. Gupta and D. Kundu, "Generalized exponential distribution: different method of estimations," *Journal of Statistical Computation and Simulation*, vol. 69, no. 4, pp. 315–337, 2001.

[11] R. D. Gupta and D. Kundu, "Discriminating between the Weibull and the GE distributions," *Computational Statistics & Data Analysis*, vol. 43, pp. 179–196, 2002.

[12] R. D. Gupta and D. Kundu, "On the comparison of Fisher information of the Weibull and GE distributions," *Journal of Statistical Planning and Inference*, vol. 136, no. 9, pp. 3130–3144, 2006.

[13] D. Kundu and R. D. Gupta, "Bivariate generalized exponential distribution," *Journal of Multivariate Analysis*, vol. 100, no. 4, pp. 581–593, 2009.

[14] G. Jones, C. D. Lai, and J. C. W. Rayner, "A bivariate gamma mixture distribution," *Communications in Statistics-Theory and Methods*, vol. 29, no. 12, pp. 2775–2790, 2000.

[15] A. M. Sarhan and N. Balakrishnan, "A new class of bivariate distributions and its mixture," *Journal of Multivariate Analysis*, vol. 98, no. 7, pp. 1508–1527, 2007.

[16] J. Chen and X. Tan, "Inference for multivariate normal mixtures," *Journal of Multivariate Analysis*, vol. 100, no. 7, pp. 1367–1383, 2009.

[17] M. Khosravi, D. Kundu, and A. Jamalizadeh, "On bivariate and a mixture of bivariate Birnbaum-Saunders distributions," *Statistical Methodology*, vol. 23, pp. 1–17, 2015.

[18] A. S. AL-Moisheer, R. M. Alotaibi, G. A. Alomani, and H. Rezk, "Bivariate mixture of inverse Weibull distribution: properties and estimation," *Mathematical Problems in Engineering*, vol. 2020, pp. 1–12, 2020.

[19] R. B. Nelsen, *An Introduction to Copulas*, Springer-Verlag, New York, NY, USA, 1999.

[20] J. G. McLachlan and E. B. Krishnan, *EM Algorithm and Extensions*, John Wiley & Sons, New York, NY, USA, 1997.

[21] D. Karlis and E. Xekalaki, "Choosing initial values for the EM algorithm for finite mixtures," *Computational Statistics & Data Analysis*, vol. 41, no. 3-4, pp. 577–590, 2003.

[22] R. B. Nelson, *An Introduction to Copulas*, Springer-Verlag, New York, NY, USA, 2006.

[23] P. K. Trivedi and D. H. Zimmer, "Copula modeling: an introduction to practitioners," *Foundations and trends in econometrics*, vol. 1, no. 1, pp. 1–111, 2005.

[24] S. A. Adham and S. G. Walker, "A multivariate Gompertz-type distribution," *Journal of Applied Statistics*, vol. 28, no. 8, pp. 1051–1065, 2001.

[25] D. Kundu, N. Balakrishnan, and A. Jamalizadeh, "Bivariate Birnbaum-Saunders distribution and associated inference," *Journal of Multivariate Analysis*, vol. 101, no. 1, pp. 113–125, 2010.

[26] D. Kundu and R. D. Gupta, "Absolute continuous bivariate generalized exponential distribution," *AStA Advances in Statistical Analysis*, vol. 95, no. 2, pp. 169–185, 2011.

[27] D. Kundu, "Bivariate sinh-normal distribution and a related model," *Brazilian Journal of Probability and Statistics*, vol. 29, pp. 590–607, 2015.

[28] M. El-Morshedy, Z. A. Alhussain, D. Atta, E. M. Almetwally, and M. S. Eliwa, "Bivariate burr X generator of distributions: properties and estimation methods with applications to complete and type-II censored samples," *Mathematics*, vol. 8, pp. 1–31, 2020.

[29] R. M. AlOtaibi, H. R. Rezk, I. Ghosh, and S. Dey, "Bivariate exponentiated half logistic distribution: properties and application," *Communications in Statistics-Theory and Methods*, vol. 2, pp. 1–23, 2020.

[30] A. Sklar, *Fonctions de Répartition à n Dimensions et Leurs Marges*, Publications de l'Institut Statistique de l'Université de Paris, Paris, France, 1959.

[31] I. Kosmidis and D. Karlis, "Model-based clustering using copulas with applications," *Statistics and Computing*, vol. 26, no. 5, pp. 1079–1099, 2016.

[32] I. Kojadinovic and J. Yan, "Comparison of three semiparametric methods for estimating dependence parameters in copula models," *Insurance: Mathematics and Economics*, vol. 47, no. 1, pp. 52–63, 2010.

[33] J. D. Fermanian, "Goodness-of-fit tests for copulas," *Journal of Multivariate Analysis*, 2005.

[34] C. Genest and B. Remillard, "Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models," *Annales de l'Institut Henri Poincare (B): Probability and Statistics*, vol. 44, no. 6, pp. 1096–1127, 2008.

[35] C. Genest, B. Rémillard, and D. Beaudoin, "Goodness-of-fit tests for copulas: a review and a power study," *Insurance: Mathematics and Economics*, vol. 44, no. 2, pp. 199–213, 2009.

[36] I. Kojadinovic, J. Yan, and M. Holmes, "Fast large-sample goodness-of-fit tests for copulas," *Statistica Sinica*, vol. 21, no. 2, pp. 841–871, 2011.

[37] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.

[38] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[39] M. Mohsin, H. Kazianka, J. Pilz, and A. Gebhardt, "A new bivariate exponential distribution for modeling moderately negative dependence," *Statistical Methods and Applications*, vol. 23, no. 1, pp. 123–148, 2014.