Research Article

Bivariate Mixture of Inverse Weibull Distribution: Properties and Estimation

A. S. AL-Moisheer 1, Refah Mohammed Alotaibi 2, Ghadah A. Alomani 2, and H. Rezk 2,3

1 Mathematics Department, College of Science, Jouf University, P.O. Box 848, Sakaka 72351, Saudi Arabia
2 Mathematical Sciences Department, College of Science, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
3 Statistics Department, Al-Azhar University, Cairo, Egypt

Correspondence should be addressed to A. S. AL-Moisheer; asalmoisheer@ju.edu.sa

Received 5 January 2020; Accepted 20 February 2020; Published 28 March 2020

Academic Editor: Leandro F. F. Miguel

Copyright © 2020 A. S. Al-Moisheer et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this study, we construct a mixture of bivariate inverse Weibull distribution. We assumed that the parameters of two marginals have Bernoulli distributions. Several properties of the proposed model are obtained, such as probability marginal density function, probability marginal cumulative function, the product moment, the moment of the two variables \(x\) and \(y\), the joint moment-generating function, and the correlation between \(x\) and \(y\). The real dataset has been analyzed. We observed that the mixture bivariate inverse Weibull distribution provides a better fit than the other model.

1. Introduction

In the history of statistics, the use of finite mixture models is very old. It was particularly used to model population heterogeneity and generalize distributional assumptions, clustering and classification, and so on. Eight years later, Pearson [1] considered a mixture of two univariate Gaussian distributions to estimate the parameters of the model using the method of moments (MOM) to analyse a dataset containing ratios of forehead to body lengths for 1,000 crabs. Since then several authors have studied finite mixture models under different scenarios. Mendenhall and Hader [2] considered exponentially distributed failure time distributions based on censored lifetime data to estimate the model parameters using the maximum likelihood method. In their study, they divided the failure population into two subpopulations, each representing a different cause or type of failure. Radhakrishna et al. [3] in their paper considered both moment and maximum likelihood estimators of the unknown parameters of two-component mixture of the generalized gamma distribution. Ahmed et al. [4] obtained approximate Bayes estimators for parameters of mixture of two Weibull distributions under type-II censoring.

Inverse Weibull (IW) distribution has been used quite successfully to analyze lifetime data which has a non-monotone hazard function. The IW distribution can be readily applied to a wide range of situations including applications in medicine, reliability, and ecology. Keller et al. [5] obtained the IW model by investigating failures of mechanical components subject to degradation phenomena such as the dynamic components of diesel engines; see, for example, Murthy et al. [6]. The physical failure process given by Erto and Rapone [7] also leads to the IW model. Erto and Rapone [7] showed that the IW model provides a good fit to survival data such as the times to breakdown of an insulating fluid subject to the action of constant tension, see [8]; Calabria and Pulcini [9] provided an interpretation of the IW distribution in the context of a load strength relationship for a component.

There have been several attempts made in the last few years to introduce bivariate distribution. For example, Sarhan and Balakrishnan [10] suggested a bivariate distribution that is more flexible than the bivariate exponential
distribution. Later, this distribution was modified by Kundu and Gupta [11]. Kundu and Gupta [11, 12] introduced the bivariate generalized exponential and bivariate proportional reversed hazard distributions, respectively, and argued its different properties. Sarhan et al. [13] proposed bivariate generalized linear failure rate distributions and discussed its several properties. Sarabia et al. [14] proposed three new classes of bivariate beta-generated distributions, and these classes were created using three alternative definitions of bivariate distributions with classical beta marginals and different covariance structures. Mohammed [15] proposed the bivariate inverse Weibull distribution, and Al-Mutairi et al. [16] introduced bivariate and multivariate weighted Weibull distribution with weighted Weibull marginals and established their several properties.

The mixture of inverse Weibull distribution has also been found useful in some applications. For example, Al Moisheer [17] used an inverse Weibull mixture distribution in measuring the carbon monoxide level in different locations of the Jeddah city. Theoretical work on mixture inverse Weibull was introduced by Sultan et al. [18]; they discussed some properties of the mixture of the inverse Weibull model with some graphs of the density and hazard function. Sultan and Moisheer [19] found the maximum likelihood estimates of the parameters of the mixture of two inverse Weibull distributions by using classified and unclassified observations. The notion of mixing may be regarded as a special case of compounding. In the bivariate case, let X and Y be two random variables with parameters \( \theta_1 \) and \( \theta_2 \), respectively. For given fixed values of \( \theta_1 \) and \( \theta_2 \), X and Y may or may not be independent. The idea of compounding is to say that \( \theta_1 \) and \( \theta_2 \) are themselves random variables which are not constant, and the observed (marginal) distribution of X and Y results from integrating over the joint distribution of \( \theta_1 \) and \( \theta_2 \), i.e.,

\[
h(x, y) = \int h(x, y, \theta_1, \theta_2) \, d\theta_1 \, d\theta_2.
\]

The primary objective of the paper is to construct a bivariate of inverse Weibull mixture distribution by assuming two independent inverse Weibull distributions with the scale parameters having a generalized bivariate Bernoulli distribution.

The organization of the paper is as follows. The description of the proposed mixture of bivariate inverse Weibull distribution along with basic properties is reported in Section 2. In Sections 2 and 3, we give the moment and correlation definition for the bivariate mixture inverse Weibull model. In Section 5, we use the EM algorithm method. The application study is carried out in Section 6. Finally, concluding remarks are presented in Section 7.

2. Mixture of Bivariate Inverse Weibull Distribution

A random variable with an inverse Weibull distribution (BIWM) has the cumulative distribution function (cdf) and the probability density function (pdf) in the following from:

\[
F(x, \theta, \lambda) = e^{-\lambda x^{-\theta}}, \quad x > 0,
\]

\[
f(x, \theta, \lambda) = \theta \lambda x^{-\theta-1} e^{-\lambda x^{-\theta}}, \quad x > 0,
\]

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

In the bivariate case, let X and Y be two random variables with parameters \( \theta_1 \) and \( \theta_2 \), respectively. For given fixed values of \( \theta_1 \) and \( \theta_2 \), X and Y are independent.

The probability density function of bivariate inverse Weibull mixture distribution is given as

\[
F(x, \theta_1, \theta_2) = \sum_{i=1}^{2} p_i f_i(x, \theta_i), \quad 0 \leq p_1 \leq p_2, \quad p_1 + p_2 = 1,
\]

where pdf of the first component of inverse Weibull is given by equation (3) with fixed shape parameter \( \theta > 0 \) and random scale parameter \( \lambda > 0 \) taking two distinct values \( \lambda_1 \) and \( \lambda_2 \).

Similarly for fixed shape parameter \( \theta_2 \), let Y have an inverse Weibull distribution and the pdf of the second component (inverse Weibull) is given by

\[
g(y, \varphi, \beta) = \beta y^{-\varphi-1} e^{-\beta y^{-\varphi}}, \quad y > 0,
\]

with \( \beta \) being a random scale parameter taking values \( \beta_1 \) and \( \beta_2 \).

For given values \( (\lambda, \beta) \), we assume that X and Y are independent, but \( \lambda \) and \( \beta \) are correlated through their generalized bivariate distribution with the following probability matrix:

\[
p = \begin{pmatrix}
\lambda_1 & \beta_1 \\
\lambda_2 & \beta_2
\end{pmatrix}
\]

where \( \lambda_1, \lambda_2, \beta_1, \beta_2 \) are elements of the probability matrix and \( \lambda, \beta \) are parameters.

Let \( h(x, y) \) be the joint probability density function of \( (X, Y) \), then

\[
h(x, y) = f(x|\theta, \lambda_1) g(y|\varphi, \beta_1) p_{\lambda_1, \beta_1} + f(x|\theta, \lambda_2) g(y|\varphi, \beta_2) p_{\lambda_2, \beta_2} + f(x|\theta, \lambda_2) g(y|\varphi, \beta_1) p_{\lambda_2, \beta_1} + f(x|\theta, \lambda_1) g(y|\varphi, \beta_2) p_{\lambda_1, \beta_2} + \beta_1 p_{\lambda_1, \beta_1} + \beta_2 p_{\lambda_2, \beta_2} + \beta_1 p_{\lambda_2, \beta_1} + \beta_2 p_{\lambda_1, \beta_2}.
\]

Suppose that

\[
a = p_{\lambda_1, \beta_1}, \\
b = p_{\lambda_2, \beta_1}, \\
c = p_{\lambda_1, \beta_2}, \\
d = p_{\lambda_2, \beta_2} = 1 - a - b - c.
\]

Substituting \( a, b, c, \) and \( d \) in equation (7), we obtain

\[
h(x, y) = \theta \varphi x^{-\theta-1} y^{-\varphi-1} \left[ \lambda_1 \beta_1 ae^{-\lambda_1 x^{-\theta} \beta_1 y^{-\varphi}} + \lambda_1 \beta_2 be^{-\lambda_1 x^{-\theta} \beta_2 y^{-\varphi}} + \lambda_2 \beta_1 ce^{-\lambda_2 x^{-\theta} \beta_1 y^{-\varphi}} + \lambda_2 \beta_2 de^{-\lambda_2 x^{-\theta} \beta_2 y^{-\varphi}} \right].
\]
The model can be seen to be capable of producing a wide variety of bivariate shapes. The joint density can be seen to be a mixture of four bivariate inverse Weibull distributions, requiring a total of nine parameters for its specification. For modeling purposes, it may be found that not all of the four components are required, so we may want to consider restrictions such as $b = c = 0$, $a = d = 0$, or $a = b = c = 0$, corresponding to correlations between the scale parameters of $+1$, $-1$, and $0$, respectively. In Figure 1, we show some examples of contour plots for various values of $a$, $b$, $c$, and $d$ with the shape and scale parameters fixed at $\theta = 8$, $\varphi = 3$, $\lambda_1 = 0.25$, $\lambda_2 = 0.4$, $\beta_1 = 1.1$, and $\beta_2 = 6.2$.

The marginal densities of $X$ and $Y$, respectively, are

$$h_X(x) = \pi_1 f_1(x) + (1 - \pi_1) f_2(x),$$

where $\pi_1 = a + b$, and

$$h_Y(y) = \pi_2 g_1(y) + (1 - \pi_2) g_2(y),$$

where $\pi_2 = a + c$.

The cumulative distribution function is given by

$$F(x, y) = \int_0^y \int_0^x h(t, s) \, dt \, ds$$

where

$$R(x, y) = 1 - F(x, y).$$

The survival function of BIWM distribution is given by

$$R(x, y) = 1 - a \left[ e^{-(\lambda_1 x^{\phi^{\lambda_1}})} - 1 \right] \left[ e^{-(\beta_1 y^{\phi^{\beta_1}})} - 1 \right] - b \left[ e^{-(\lambda_1 x^{\phi^{\lambda_1}})} - 1 \right] \left[ e^{-(\beta_2 y^{\phi^{\beta_2}})} - 1 \right]$$

$$- c \left[ e^{-(\lambda_2 x^{\phi^{\lambda_2}})} - 1 \right] \left[ e^{-(\beta_1 y^{\phi^{\beta_1}})} - 1 \right] - d \left[ e^{-(\lambda_2 x^{\phi^{\lambda_2}})} - 1 \right] \left[ e^{-(\beta_2 y^{\phi^{\beta_2}})} - 1 \right].$$

The hazard rate function (hrf) is given as follows:

$$hrf(x, y) = \frac{\theta \varphi x^{-\theta - 1} y^{-\varphi - 1} \left[ \lambda_1 \beta_1 ae^{-(\lambda_1 x^{\phi^{\lambda_1}} + \beta_1 y^{\phi^{\beta_1}})} + \lambda_1 \beta_2 be^{-(\lambda_1 x^{\phi^{\lambda_1}} + \beta_2 y^{\phi^{\beta_2}})} + \lambda_2 \beta_1 ce^{-(\lambda_2 x^{\phi^{\lambda_2}} + \beta_1 y^{\phi^{\beta_1}})} + \lambda_2 \beta_2 be^{-(\lambda_2 x^{\phi^{\lambda_2}} + \beta_2 y^{\phi^{\beta_2}})} \right]}{1 - a \left[ e^{-(\lambda_1 x^{\phi^{\lambda_1}})} - 1 \right] \left[ e^{-(\beta_1 y^{\phi^{\beta_1}})} - 1 \right] - b \left[ e^{-(\lambda_1 x^{\phi^{\lambda_1}})} - 1 \right] \left[ e^{-(\beta_2 y^{\phi^{\beta_2}})} - 1 \right] - c \left[ e^{-(\lambda_2 x^{\phi^{\lambda_2}})} - 1 \right] \left[ e^{-(\beta_1 y^{\phi^{\beta_1}})} - 1 \right] - d \left[ e^{-(\lambda_2 x^{\phi^{\lambda_2}})} - 1 \right] \left[ e^{-(\beta_2 y^{\phi^{\beta_2}})} - 1 \right]}.$$

The conditional probability function of $x$ for given $y$ is given by

$$h(x|y) = \frac{h(x, y)}{h(y)} = x > 0$$

$$= \frac{\theta \varphi x^{-\theta - 1} y^{-\varphi - 1} \left[ \lambda_1 \beta_1 ae^{-(\lambda_1 x^{\phi^{\lambda_1}} + \beta_1 y^{\phi^{\beta_1}})} + \lambda_1 \beta_2 be^{-(\lambda_1 x^{\phi^{\lambda_1}} + \beta_2 y^{\phi^{\beta_2}})} + \lambda_2 \beta_1 ce^{-(\lambda_2 x^{\phi^{\lambda_2}} + \beta_1 y^{\phi^{\beta_1}})} + \lambda_2 \beta_2 be^{-(\lambda_2 x^{\phi^{\lambda_2}} + \beta_2 y^{\phi^{\beta_2}})} \right]}{\pi_2 g_1(y) + (1 - \pi_2) g_2(y)}.$$
and the conditional probability function of $y$ for given $x$ is given by

\[
h(y|x) = \frac{h(x, y)}{h(x)}, \quad x > 0
\]

\[
= \left\{ \theta \varphi x^{-\theta - 1} y^{-\varphi - 1} \left[ \lambda_1 \beta_1 \alpha e^{-\left( \lambda_1 x^{-a} + \beta_1 y^{-b} \right)} + \lambda_2 \beta_2 be^{-\left( \lambda_2 x^{-a} + \beta_2 y^{-b} \right)} + \lambda_3 \beta_1 ce^{-\left( \lambda_3 x^{-a} + \beta_3 y^{-b} \right)} + \lambda_4 \beta_2 de^{-\left( \lambda_4 x^{-a} + \beta_4 y^{-b} \right)} \right] \right\}
\]

\[
\left\{ \pi_1 f_1(x) + (1 - \pi_1) f_2(x) \right\}
\]

\[\text{Figure 1: Contours of mixture of the bivariate inverse Weibull model with } \theta = 8, \varphi = 3, \lambda_1 = 0.25, \lambda_2 = 0.4, \beta_1 = 1.1, \beta_2 = 6.2, \text{ and selected choices of the mixing parameters.}\]
3. Moments

In this section, we derive the statistical properties of the BIWM distribution. The product moments about zero is given by

\begin{equation}
\mu_{ij} = \int_0^\infty \int_0^\infty xy h(x, y) dx \, dy
= a \int_0^\infty \theta x e^{-\theta (1, x)} dx \cdot \int_0^\infty \beta_1 y e^{-\beta_1 y} dy
+ b \int_0^\infty \theta x e^{-\theta (1, x)} dx \cdot \int_0^\infty \beta_2 y e^{-\beta_2 y} dy
+ c \int_0^\infty \theta x e^{-\theta (1, x)} dx \cdot \int_0^\infty \beta_1 y e^{-\beta_1 y} dy
+ d \int_0^\infty \theta x e^{-\theta (1, x)} dx \cdot \int_0^\infty \beta_2 y e^{-\beta_2 y} dy.
\end{equation}

Let \( z = \lambda_1 x^\theta \), \( w = \beta_1 y^\varphi \), and \( dz = -\lambda_1 x^{\theta-1} dx \), at \( x = 0 \), \( z = 0 \), \( x = \infty \), \( z = \infty \), \( y = 0 \), \( w = \infty \), and \( w = \infty \); hence, we obtain

\begin{equation}
\mu_{ij} = a \int_0^\infty z^{(-1/\varphi)} \lambda_1 e^{-z} dz \cdot \int_0^\infty w^{(-1/\varphi)} \beta_1 e^{-w} dw
+ b \int_0^\infty z^{(-1/\varphi)} \lambda_1 e^{-z} dz \cdot \int_0^\infty w^{(-1/\varphi)} \beta_2 e^{-w} dw
+ c \int_0^\infty z^{(-1/\varphi)} \lambda_2 e^{-z} dz \cdot \int_0^\infty w^{(-1/\varphi)} \beta_1 e^{-w} dw
+ d \int_0^\infty z^{(-1/\varphi)} \lambda_2 e^{-z} dz \cdot \int_0^\infty w^{(-1/\varphi)} \beta_2 e^{-w} dw,
\end{equation}

and then

\begin{equation}
\mu_{ij} = \Gamma \left( \frac{\theta - 1}{\theta} \right) \Gamma \left( \frac{\varphi - 1}{\varphi} \right) \left[ a \beta_1^{1/\varphi} \lambda_1^{1/\theta} + b \beta_2^{1/\varphi} \lambda_1^{1/\theta} \right]
+ c \beta_1^{1/\varphi} \lambda_2^{1/\theta} + d \beta_2^{1/\varphi} \lambda_2^{1/\theta}.
\end{equation}

The expected values of \( X \) and \( Y \) are

\begin{align*}
E(X) &= -\lambda_1^{2(1/\theta)} \Gamma \left( -\frac{1 + \theta}{\theta} \right), \quad -1 \leq \theta \leq 0, \lambda_1 > 0, \\
E(Y) &= -\beta_1^{2(1/\varphi)} \Gamma \left( -\frac{1 + \varphi}{\varphi} \right), \quad -1 \leq \varphi \leq 0, \beta_1 > 0.
\end{align*}

The joint moment-generating function of the bivariate mixture is given by

\begin{align*}
\mu_{xy}(s, t) &= \int_0^\infty \int_0^\infty e^{s x + t y} h(x, y) dx \, dy \\
&= \int_0^\infty \int_0^\infty e^{s x + t y} (-\theta - 1) (-\varphi - 1) \left[ \lambda_1 \beta_1 e^{-\lambda_1 x^{\theta} \beta_1 y^{\varphi}} + \lambda_1 \beta_2 e^{-\lambda_1 x^{\theta} \beta_2 y^{\varphi}} + \lambda_2 \beta_1 e^{-\lambda_2 x^{\theta} \beta_1 y^{\varphi}} + \lambda_2 \beta_2 e^{-\lambda_2 x^{\theta} \beta_2 y^{\varphi}} \right] dx \, dy, \\
\mu_{xy}(s, t) &= a \int_0^\infty \lambda_1 \beta_1 e^{-\lambda_1 x^{\theta} \beta_1 y^{\varphi}} dx \cdot \int_0^\infty \beta_1 \varphi e^{s y} y^{(-\varphi - 1)} e^{-\beta_1 y} dy
+ b \int_0^\infty \lambda_1 \beta_2 e^{-\lambda_1 x^{\theta} \beta_2 y^{\varphi}} dx \cdot \int_0^\infty \beta_2 \varphi e^{s y} y^{(-\varphi - 1)} e^{-\beta_2 y} dy
+ c \int_0^\infty \lambda_2 \beta_1 e^{-\lambda_2 x^{\theta} \beta_1 y^{\varphi}} dx \cdot \int_0^\infty \beta_1 \varphi e^{s y} y^{(-\varphi - 1)} e^{-\beta_1 y} dy
+ d \int_0^\infty \lambda_2 \beta_2 e^{-\lambda_2 x^{\theta} \beta_2 y^{\varphi}} dx \cdot \int_0^\infty \beta_2 \varphi e^{s y} y^{(-\varphi - 1)} e^{-\beta_2 y} dy.
\end{align*}
By using exponential expansion,
\[
e^{tx} = \sum_{j=0}^{\infty} \frac{(tx)^j}{j!},
\]
\[
e^{(\lambda x)^j}\sum_{j=0}^{\infty} \frac{(-\lambda x)^j}{j!},
\]
\[
e^{sy} = \sum_{k=0}^{\infty} \frac{(sy)^k}{k!},
\]
\[
e^{-(\beta y)^m} = \sum_{m=0}^{\infty} \frac{(-\beta y)^m}{m!}.
\]

Then, we have
\[
\mu_{xy}(s,t) = a \int_0^{\infty} \int_0^{\infty} \frac{\lambda_1 \sum_{i=0}^{\infty} \frac{(tx)^i}{i!} \sum_{j=0}^{\infty} \frac{(-\lambda x)^j}{j!} dx}{m!}.
\]
\[
\cdot \int_0^{\infty} \frac{\beta_1 y}{m!} \sum_{k=0}^{\infty} \frac{(sy)^k}{k!} \sum_{m=0}^{\infty} \frac{(-\beta y)^m}{m!} dy
\]
\[
+ b \int_0^{\infty} \frac{\lambda_1 \sum_{i=0}^{\infty} \frac{(tx)^i}{i!} \sum_{j=0}^{\infty} \frac{(-\lambda x)^j}{j!} dx}{m!}.
\]
\[
\cdot \int_0^{\infty} \frac{\beta_2 \sum_{k=0}^{\infty} \frac{(sy)^k}{k!} \sum_{m=0}^{\infty} \frac{(-\beta y)^m}{m!} dy}{m!}
\]
\[
+ c \int_0^{\infty} \frac{\lambda_1 \sum_{i=0}^{\infty} \frac{(tx)^i}{i!} \sum_{j=0}^{\infty} \frac{(-\lambda x)^j}{j!} dx}{m!}.
\]
\[
\cdot \int_0^{\infty} \frac{\lambda_1 \sum_{i=0}^{\infty} \frac{(tx)^i}{i!} \sum_{j=0}^{\infty} \frac{(-\lambda x)^j}{j!} dx}{m!}.
\]
\[
+ d \int_0^{\infty} \frac{\lambda_1 \sum_{i=0}^{\infty} \frac{(tx)^i}{i!} \sum_{j=0}^{\infty} \frac{(-\lambda x)^j}{j!} dx}{m!}.
\]
\[
\cdot \int_0^{\infty} \frac{\lambda_1 \sum_{i=0}^{\infty} \frac{(tx)^i}{i!} \sum_{j=0}^{\infty} \frac{(-\lambda x)^j}{j!} dx}{m!}.
\]
\[
+ e \int_0^{\infty} \frac{\lambda_1 \sum_{i=0}^{\infty} \frac{(tx)^i}{i!} \sum_{j=0}^{\infty} \frac{(-\lambda x)^j}{j!} dx}{m!}.
\]
\[
\cdot \int_0^{\infty} \frac{\lambda_1 \sum_{i=0}^{\infty} \frac{(tx)^i}{i!} \sum_{j=0}^{\infty} \frac{(-\lambda x)^j}{j!} dx}{m!}.
\]

and then
\[
= a \lambda_1 \sum_{i=0}^{\infty} \frac{(t)^i}{i!} \lambda_1 \sum_{j=0}^{\infty} \frac{(-\lambda)^j}{j!} \int_0^{\infty} x^{j-\theta-1-\theta} dx \cdot \beta_1 \varphi
\]
\[
\cdot \sum_{k=0}^{\infty} \frac{(s)^k}{k!} \sum_{m=0}^{\infty} \frac{(-\beta y)^m}{m!} \int_0^{\infty} y^{j-\varphi-1-\varphi} dy
\]
\[
+ b \lambda_1 \sum_{i=0}^{\infty} \frac{(t)^i}{i!} \lambda_1 \sum_{j=0}^{\infty} \frac{(-\lambda)^j}{j!} \int_0^{\infty} x^{j-\theta-1-\theta} dx \cdot \beta_2 \varphi
\]
\[
\cdot \sum_{k=0}^{\infty} \frac{(s)^k}{k!} \sum_{m=0}^{\infty} \frac{(-\beta y)^m}{m!} \int_0^{\infty} y^{j-\varphi-1-\varphi} dy
\]
\[
+ c \lambda_1 \sum_{i=0}^{\infty} \frac{(t)^i}{i!} \lambda_1 \sum_{j=0}^{\infty} \frac{(-\lambda)^j}{j!} \int_0^{\infty} x^{j-\theta-1-\theta} dx \cdot \beta_1 \varphi \sum_{k=0}^{\infty} \frac{(s)^k}{k!}
\]
\[
\cdot \sum_{m=0}^{\infty} \frac{(-\beta y)^m}{m!} \int_0^{\infty} y^{j-\varphi-1-\varphi} dy
\]
\[
+ d \lambda_1 \sum_{i=0}^{\infty} \frac{(t)^i}{i!} \lambda_1 \sum_{j=0}^{\infty} \frac{(-\lambda)^j}{j!} \int_0^{\infty} x^{j-\theta-1-\theta} dx \cdot \beta_2 \varphi \sum_{k=0}^{\infty} \frac{(s)^k}{k!}
\]
\[
\cdot \sum_{m=0}^{\infty} \frac{(-\beta y)^m}{m!} \int_0^{\infty} y^{j-\varphi-1-\varphi} dy.
\]

Finally, we get
\[
\mu_{xy}(s,t) = \beta_1 e^{-(\beta y)^m} \left( a \lambda_1 e^{-(\lambda x)^m} + b \lambda_2 e^{-(\lambda x)^m} \right)
\]
\[
- \beta_2 e^{-(\beta y)^m} \left( c \lambda_1 e^{-(\lambda x)^m} + d \lambda_2 e^{-(\lambda x)^m} \right).
\]

4. Correlation

The covariance between X and Y for the BIWM model is given by
\[
\text{Cov}(X,Y) = E(XY) - E(X)E(Y).
\]

So,
\[
\text{Cov}(X,Y) = \Gamma \left( \frac{\theta - 1}{\theta} \right) \left( \frac{\varphi - 1}{\varphi} \right) \left[ a \beta_1 \lambda_1^1 \beta_1 \lambda_1^1 \theta + b \beta_1 \lambda_1^1 \beta_1 \lambda_1^1 \theta 
\]
\[
+ c \beta_1 \lambda_1^1 \beta_1 \lambda_1^1 \theta + d \beta_2 \lambda_1^1 \beta_1 \lambda_1^1 \theta 
\]
\[
- \Gamma \left( \frac{\theta - 1}{\theta} \right) \left( \frac{\varphi - 1}{\varphi} \right) \beta_1 \lambda_1^1 \beta_1 \lambda_1^1 \theta + 2 \right].
\]
It is noted that the covariance between two variables $X$ and $Y$ depends only on the parameters, $\theta, \phi, \lambda_1, \beta_1, \lambda_2, \beta_2, a, b, c,$ and $d$. The correlation coefficient of $X$ and $Y$ is given as follows:

$$
\rho(XY) = \frac{\text{Cov}(XY)}{\sigma_X \sigma_Y},
$$

where $\sigma_X$ and $\sigma_Y$ are the standard deviation for $X$ and $Y$, respectively. Then, the variance of $X$ and $Y$ are shown as follows:

$$
\sigma_X^2 = E(X^2) - [E(X)]^2,
$$

$$
\sigma_Y^2 = E(Y^2) - [E(Y)]^2,
$$

$$
= \lambda_1^{(2/\theta)+2} \left(\frac{-\theta - 2}{\theta}\right) - \left[ \Gamma\left(\frac{-\theta - 1}{\theta}\right) \lambda_1^{(1/\theta)+2} \right],
$$

$$
\sigma_Y^2 = \beta_1^{(1/\phi)+2} \left(\frac{-\phi - 2}{\phi}\right) - \left[ \Gamma\left(\frac{-\phi - 1}{\phi}\right) \beta_1^{(1/\phi)+2} \right],
$$

$$
\theta > 2, \lambda_1 > 0,
$$

$$
\phi > 2, \beta_1 > 0,
$$

where

$$
E(X^2) = \lambda_1^{(2/\theta)+2} \left(\frac{-\theta - 2}{\theta}\right),
$$

$$
E(Y^2) = \beta_1^{(2/\phi)+2} \left(\frac{-\phi - 2}{\phi}\right).
$$

Then, we can obtain the standard deviation for $X$ and $Y$ by taking the positive squared root of $\sigma_X^2$ and $\sigma_Y^2$, respectively, as follows:

$$
\sigma_X = \sqrt{\sigma_X^2},
$$

$$
\sigma_Y = \sqrt{\sigma_Y^2}.
$$

We compute numerically the correlation coefficient $\rho(XY)$. The numerical results are listed in Table 1 for different values of the parameters.

5. EM Algorithm

McLachlan and Krishnan [20] introduced the EM algorithm as a method of estimation; to apply the EM algorithm, 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 $a_k$ is one if the $k$th observation is in $(X, \lambda_1, \beta_1)$ and zero otherwise. Similarly, for $b_k, c_k$, we have four groups $G_{ij}, i, j = 1, 2$, for which the densities are

$$
f_{ij}(X, Y) = f_i(X)f_j(Y) = \theta \lambda x^{-\theta - 1} e^{-\lambda x^\theta} \phi \beta y^{-\phi - 1} e^{-\beta y^\phi}.
$$

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

We define $\ell_i(x, y) = \log f_i(x, y)$; then, the EM algorithm as a method of estimation is given by finding the complete log-likelihood $\ell$ as follows:

$$
\ell = \sum_{k=1}^{n} a_k \ell_{11}(x_k, y_k) + \sum_{k=1}^{n} b_k \ell_{12}(x_k, y_k)
$$

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

These group membership variables $(a_k, b_k, c_k)$ so in the E step have a linear relation, so $(\theta, \phi, \lambda_1, \lambda_2, \beta_1, \beta_2, a, b, c)$ are calculated as

$$
\frac{\partial \ell}{\partial \beta_1} = \frac{n}{\beta_1} \sum_{k=1}^{n} (a_k + c_k) - \sum_{k=1}^{n} (a_k + c_k) y_k^{-\phi},
$$

$$
\frac{\partial \ell}{\partial \beta_2} = \frac{n}{\beta_2} \sum_{k=1}^{n} (b_k + d_k) - \sum_{k=1}^{n} (b_k + d_k) y_k^{-\phi},
$$

$$
\frac{\partial \ell}{\partial \theta} = 4n - 4 \sum_{k=1}^{n} \log x_k - 2 \sum_{k=1}^{n} (\lambda x_k^{\theta}) \log x_k - 2 \sum_{k=1}^{n} (\lambda x_k^{\theta}) \log x_k,
$$

$$
\frac{\partial \ell}{\partial \phi} = 4n - 4 \sum_{k=1}^{n} \log y_k - 2 \sum_{k=1}^{n} (\beta y_k^{\phi}) \log y_k - 2 \sum_{k=1}^{n} (\beta y_k^{\phi}) \log y_k.
$$
The M step is completed by setting

$$\hat{a} = \frac{1}{n} \sum_{k=1}^{n} a_k.$$ (42)

We equate the systems of equations (36)–(41) to zero. We use the mixtools R package to solve these equations numerically. After the maximum likelihood estimators for \(\lambda_1, \lambda_2, \beta_1, \beta_2, \theta, \) and \(\varphi\) are obtained, substitute these estimates in \((a_k, b_k, c_k)\). We complete the M step by setting \(\hat{a} = (1/n) \sum_{k=1}^{n} a_k\) and so on.

Initial values for the mixing proportions are obtained by using the moment’s method of the marginal univariate inverse Weibull mixtures, see [3, 21]. Then, we take the resulting estimates of the inverse Weibull 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 the independence between two variables \(X\) and \(Y\). We apply this method in application as mentioned in Section 7, specifically, in Table 2.

### 6. Application

In this section, we use the data of cows, i.e., the data concerning the milk yield of dairy cows to illustrate the flexibility of the new model. Cow data are introduced by Tocher [22]. In Figure 2, we plot the marginal and joint contours, histogram cow data fitted for bivariate mixture of inverse Weibull distribution. We compare the fits of the new bivariate inverse Weibull mixture with the other competitive models, such as BIWM, bivariate gamma mixture (MBG), mixture inverse Weibull (MIW), and bivariate inverse Weibull (BIW) distributions. The comparison is done based on some measures of goodness of fit, the maximized log-likelihood under the \((-\ell)\), Akaike information criterion (AIC), Bayesian information criterion (BIC), consistent Akaike information criterion (CAIC), Hannan–Quinn information criterion (HQIC), and Kolmogorov–Smirnov (KS) statistic and its \(P\) value (PV). Cow data represent age and lactation period of cows. The first variable \(X\) is age (years), and the second variable \(Y\) is lactation period (weeks). We obtain the maximum likelihood estimators and the corresponding log values for the inverse Weibull mixture distribution of the cow data using the EM algorithm. The examination of the marginal density mixture \(f(x)\) of the model is given in Table 3. The examination of the marginal density \(f(y)\) of the model is given Table 4. Figures 3–6 give the estimated CDF (E-CDF), the estimated density (E-PDF), P-P plot (P-P), and Q-Q plot for \(f(x)\) or \(f(y)\). Now, we fit the data to MBIW, BGM, MIW, and BIW distributions. It is noted that the proposed model fits the cow data better than the BGM, MIW, and BIW distributions. Jones et al. [23] presented that the BGM is well fit of the cow data. The parameter estimates from the model are illustrated in Table 2. The comparison among BIWM and other three models are given in Table 5.

### 7. Conclusions

In this paper, we suggest BIWM distribution as a new finite mixture bivariate model. Some properties of the model have been derived such as the marginal probability density of \(X\) or \(Y\), cumulative distribution function, product moments about zero, the expected value of \(X\) or \(Y\), the joint moment-generating function, the variance of \(X\) or \(Y\), and the covariance between \(X\) and \(Y\). The correlation coefficient has been studied for the two variables \(X\) and \(Y\). It is noted that the correlation coefficient depends on the parameters. The EM algorithm
Figure 2: Cow data: marginal and joint contours, histogram fitted for bivariate mixture of inverse Weibull distribution. (a) Age: marginal contour $h(x)$. (b) Location: marginal contour $h(y)$. (c) Perspective joint contour. (d) Cow data. (e) Image joint contour.
Table 3: One-sample Kolmogorov–Smirnov test, AIC, CAIC, BIC, and HQIC of \( f(x) \).

| K-S | \( P \) value | AIC  | CAIC | BIC  | HQIC |
|-----|----------------|------|------|------|------|
| \( D \) = 0.16982 | 0.6842 | 104.5092 | 105.4323 | 106.0544 | 104.5883 |

Table 4: One-sample Kolmogorov–Smirnov test, AIC, CAIC, BIC, and HQIC of \( f(y) \).

| K-S | \( P \) value | AIC  | CAIC | BIC  | HQIC |
|-----|----------------|------|------|------|------|
| \( D \) = 0.19619 | 0.5075 | 141.6888 | 142.6119 | 143.234 | 141.07679 |

Figure 3: Empirical CDF, fitted IW CDF, density of histogram, and fitted IW density for the data \( x \).

Figure 4: PP-Plot fitted IW CDF, the QQ-Plot, and quantile of IW for the data \( x \).
estimation has been used to estimate the parameters of the BIWM. Cow data have been used to illustrate the importance and flexibility of the BIWM model. It is found that the BIWM model performs the best among the competitive distributions such as BGM, MIW, and BIW distributions.

Data Availability

The data used in the example application are given in the paper link: https://doi.org/10.1080/03610920008832636.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This project was supported by Jouf University, Deanship of Scientific Research, under the research project (40/327).

References

[1] K. Pearson, “Contributions to the mathematical theory of evolution,” *Philosophical Transactions, A.* vol. 185, pp. 71–110, 1894.
[2] W. Mendenhall and R. J. Hader, “Estimation of parameters of mixed exponentially distributed failure time distributions from censored life test data,” *Biometrika*, vol. 45, no. 3/4, pp. 504–520, 1958.
[3] C. Radhakrishna, A. V. Dattatreya Rao, and G. V. S. R. Anjaneyulu, “Estimation of parameters in a two-component mixture generalized gamma distribution,” *Communications in Statistics—Theory and Methods*, vol. 21, no. 6, pp. 1799–1805, 1992.
[4] K. E. Ahmed, H. M. Moustafa, and A. M. Abd-Elrahman, “Approximate Bayes estimation for mixture of two Weibull distributions under type-2 censoring,” *Journal of Statistical Computation and Simulation*, vol. 58, pp. 269–285, 1997.
[5] A. Z. Keller, A. R. R. Kamath, and U. D. Perera, “Reliability analysis of CNC machine tools,” *Reliability Engineering*, vol. 3, no. 6, pp. 449–473, 1982.
[6] D. N. P. Murthy, M. Xie, and R. Jiang, *Weibull Models*, Wiley, New York, NY, USA, 2003.
[7] P. Erto and M. Rapone, “Non-informative and practical Bayesian confidence bounds for reliable life in the Weibull model,” *Reliability Engineering*, vol. 7, no. 3, pp. 181–191, 1984.
[8] W. B. Nelson, *Applied Life Data Analysis*, Vol. 577, John Wiley & Sons, Hoboken, NJ, USA, 2005.
[9] R. Calabria and G. Pulcini, “An engineering approach to Bayes estimation for the Weibull distribution,” *Microelectronics Reliability*, vol. 34, no. 5, pp. 789–802, 1994.
[10] 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.
[11] D. Kundu and R. D. Gupta, “Modified Sarhan-Balakrishnan singular bivariate distribution,” *Journal of Statistical Planning and Inference*, vol. 140, no. 2, pp. 526–538, 2010.
[12] D. Kundu and R. D. Gupta, “Bivariate generalized exponential distribution,” *Journal of Multivariate Analysis*, vol. 100, no. 4, pp. 581–593, 2009.
[13] A. M. Sarhan, D. C. Hamilton, B. Smith, and D. Kundu, “The bivariate generalized linear failure rate distribution and its multivariate extension,” *Computational Statistics & Data Analysis*, vol. 55, no. 1, pp. 644–654, 2011.
[14] J. M. Sarabia, F. Prieto, and V. Jordà, “Bivariate beta-generated distributions with applications to well-being data,” *Journal of Statistical Distributions and Applications*, vol. 1, no. 15, pp. 1–26, 2014.
[15] H. Z. Muhammed, “Bivariate inverse Weibull distribution,” *Journal of Statistical Computation and Simulation*, vol. 86, no. 12, pp. 2335–2345, 2016.
[16] D. K. Al-Mutairi, M. E. Ghitany, and D. Kundu, “Weighted Weibull distribution: bivariate and multivariate cases,” *Brazilian Journal of Probability and Statistics*, vol. 32, no. 1, pp. 20–43, 2018.
[17] A. S. Al-Moisheer, “Updating a nonlinear discriminant function estimated from a mixture of two Burr Type III distributions,” *Journal of Applied Statistics*, vol. 44, no. 15, pp. 2685–2696, 2017.
[18] K. S. Sultan, M. A. Ismail, and A. S. Al-Moisheer, “Mixture of two inverse Weibull distributions: properties and estimation,” *Computational Statistics & Data Analysis*, vol. 51, no. 11, pp. 5377–5387, 2007.
[19] K. S. Sultan and A. S. Al-Moisheer, “Estimation of a discriminant function from a mixture of two inverse Weibull distributions,” *Journal of Statistical Computation and Simulation*, vol. 83, no. 3, pp. 405–416, 2013.
[20] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, John Wiley & Sons, New York, NY, USA, 1997.
[21] S. John, “On identifying the population of origin of each Observation in a mixture of observations from two gamma populations,” *Technometrics*, vol. 12, no. 3, pp. 565–568, 1970.
[22] J. F. Tocher, “An investigation of the milk yield of dairy cows: being a statistical analysis of the data of the Scottish milk records association for the years 1908, 1909, 1911, 1912, 1920 and 1923,” *Biometrika*, vol. 20B, no. 2, pp. 105–240, 1928.
[23] 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.