Brazilian Journal of Probability and Statistics
2018, Vol. 32, No. 1, 20–43
https://doi.org/10.1214/16-BJPS330

Weighted Weibull distribution: Bivariate and multivariate cases

D. K. Al-Mutairi[a], M. E. Ghitany[a] and Debasis Kundu[b]

[a]*Kuwait University*
[b]*Indian Institute of Technology Kanpur*

Abstract. Gupta and Kundu (*Statistics* **43** (2009) 621–643) introduced a new class of weighted exponential distribution and established its several properties. The probability density function of the proposed weighted exponential distribution is unimodal and it has an increasing hazard function. Following the same line Shahbaz, Shahbaz and Butt (*Pak. J. Stat. Oper. Res.* **VI** (2010) 53–59) introduced weighted Weibull distribution, and we derive several new properties of this weighted Weibull distribution. The main aim of this paper is to introduce bivariate and multivariate distributions with weighted Weibull marginals and establish their several properties. It is shown that the hazard function of the weighted Weibull distribution can have increasing, decreasing and inverted bathtub shapes. The proposed multivariate model has been obtained as a hidden truncation model similarly as the univariate weighted Weibull model. It is observed that to compute the maximum likelihood estimators of the unknown parameters for the proposed p-variate distribution, one needs to solve $(p+2)$ non-linear equations. We propose to use the EM algorithm to compute the maximum likelihood estimators of the unknown parameters. We obtain the observed Fisher information matrix, which can be used for constructing asymptotic confidence intervals. One data analysis has been performed for illustrative purposes, and it is observed that the proposed EM algorithm is very easy to implement, and the performance is quite satisfactory.

1 Introduction

Using the similar idea as of Azzalini (1985), a new class of weighted exponential distribution has been introduced recently by Gupta and Kundu (2009). The random variable X is said to have a weighted exponential (WEX) distribution with the shape parameter $\beta > 0$ and scale parameter $\lambda > 0$, if the probability density function (PDF) of X for $x > 0$, is

$$f_X(x;\beta,\lambda) = \frac{\beta+1}{\beta}\lambda e^{-\lambda x}\left(1-e^{-\beta\lambda x}\right), \tag{1}$$

and 0, otherwise. From now on, a WEX distribution with the PDF (1) will be denoted by WEX(β,λ). Gupta and Kundu (2009) developed several interesting

Key words and phrases. Hidden truncation model, maximum likelihood estimator, failure rate, EM algorithm, bootstrap confidence intervals.

Received January 2016; accepted July 2016.

properties of the WEX distribution. The PDF of the WEX is always unimodal and the hazard function (HF) is an increasing function. It can be obtained as a hidden truncation model, and it is observed that the WEX distribution can be used very efficiently to analyze skewed data. Al-Mutairi et al. (2011) provided a bivariate extension of the WEX distribution which has an absolute continuous bivariate joint PDF. Jamalizadeh and Kundu (2013) proposed a bivariate WEX distribution which has a singular component.

Following the same line as in Gupta and Kundu (2009), Shahbaz et al. (2010) introduced three-parameter weighted Weibull (WWE) distribution, which is a natural generalization of the WEX model. The random variable X is said to have a WWE distribution with parameters, $\alpha > 0$, $\beta > 0$ and $\lambda > 0$, if X has the PDF

$$f_X(x; \alpha, \beta, \lambda) = \frac{\beta + 1}{\beta} \alpha \lambda x^{\alpha - 1} e^{-\lambda x^{\alpha}} \left(1 - e^{-\beta \lambda x^{\alpha}}\right), \tag{2}$$

and 0 otherwise. We will denote this model as WWE(α, β, λ). Here α and β are shape parameters and λ is the scale parameter. The authors obtained the moment generating function of WWE distribution and showed graphically that the PDF is unimodal for different values of α and β. They have provided plots of the hazard function also for different values of α and β, and claimed that the hazard function is either increasing or decreasing depending on the shape parameters. In this paper, we have proved that the PDF is a decreasing function if $\alpha \leq \frac{1}{2}$, and unimodal if $\alpha > \frac{1}{2}$, for all $\beta > 0$. We have further showed that the hazard function of WWE distribution can be increasing, decreasing and bathtub shape depending on the values of α. Therefore, the shape of the PDF and hazard function of a WWE distribution depends only on the values of α, not on the parameter β.

The main aim of this paper is to introduce bivariate and multivariate WWE distributions which have WWE marginals. First, we introduce a four-parameter bivariate weighted Weibull (BWWE) distribution. The proposed BWWE distribution has closed form expressions for its joint PDF and the joint CDF. The joint PDF can take variety of shapes. It is shown that the joint PDF of a BWWE distribution can be either a decreasing or an unimodal function. Hence, it will be very useful for analyzing different bivariate data sets in practice. Although, the joint CDF may not be inverted easily, but using a structural representation, it is observed that the generation from a BWWE distribution is quite simple. Due to this reason, simulation experiments related to the BWWE distribution can be performed quite conveniently. It is observed that the BWWE distribution has the total positivity of order two (TP_2) property in the sense of Karlin and Rinott (1980). The distribution function of the minimum or the maximum can be obtained in explicit forms. The stress-parameter also can be obtained in a compact form. We finally introduce multivariate weighted Weibull (MWWE) distribution, and established its several properties.

The proposed p-variate MWWE distribution has $p + 2$ unknown parameters. The maximum likelihood estimators (MLEs) of the unknown parameters do not exist in explicit forms, as expected. It involves solving a $p + 2$ dimensional optimization problem, which may not be a trivial issue particularly if p is large. In finding the MLEs in this case, we treat this problem as a missing value problem, and propose to use the EM algorithm to compute the MLEs. It is observed that in the proposed EM algorithm, in each "E" step, the corresponding "M" step can be performed by a simple one dimensional optimization process. It can be solved by using the standard Newton–Rapson type algorithm. In this case, the EM algorithm can be implemented quite conveniently. Since it involves only a one dimensional optimization problem, the convergence can be assessed quite easily. One trivariate data set has been analyzed for illustrative purposes and also to show the usefulness of the proposed model. Therefore, here we have a multivariate distribution whose joint PDF can take variety of shapes, whose marginals can have both monotone and non-monotone hazard functions, on the other hand the implementation of the proposed model is quite simple even in a higher dimensional case. Hence, the proposed model will provide another choice to the practitioner to use it in practice for multivariate data analysis purposes.

Rest of the paper is organized as follows. In Section 2, we provide some new properties of a WWE distribution. The BWWE is introduced and its several properties are discussed in Section 3. In Section 4, we introduce the MWWE distribution. In Section 5, we present the EM algorithm. The analysis of a real data set is presented in Section 6, and finally the conclusions appear in Section 7.

2 WWE: Some new results

In this section, we establish different properties of a WWE distribution of Shahbaz et al. (2010). Shahbaz et al. (2010) provided PDF plots of a WWE distribution for different values of α and β, when $\lambda = 1$. They mentioned that the PDF of WWE distribution is unimodal for different values of α and β. We have the following result which indicates that the PDF of a WWE distribution will be either a decreasing or an unimodal function depending on the values of α, and it does not depend on the values of β or λ.

Theorem 2.1. *$f_X(x; \alpha, \beta, \lambda)$ is a decreasing function of x, if $\alpha \le \frac{1}{2}$, and unimodal if $\alpha > \frac{1}{2}$, for all $\beta > 0$ and $\lambda > 0$.*

Proof. See in the Appendix. □

In Figure 1, we provide the plots of PDFs for different values of α and β, when $\lambda = 1$. It is clear that it can take different shapes, and the shape depends only on α.

We introduce the following notation. The PDF of a Weibull distribution with the shape parameter $\alpha > 0$ and the scale parameter $\lambda > 0$ is

$$f_{\mathrm{WE}}(u; \alpha, \lambda) = \alpha\lambda u^{\alpha-1} e^{-\lambda u^{\alpha}}, \tag{3}$$

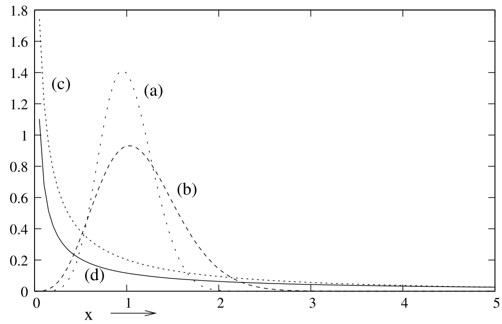

Figure 1 *PDF of WWE distribution for different α and β, when $\lambda = 1$*: (a) $\alpha = 2$, $\beta = 2$, (b) $\alpha = 3$, $\beta = 3$, (c) $\alpha = 0.5$, $\beta = 2$, (d) $\alpha = 0.25$, $\beta = 1$.

for $u > 0$ and 0 otherwise. We will denote this as $\text{WE}(\alpha, \lambda)$. We will show that the weighted Weibull model as defined in (2) can be obtained as a hidden truncation model. Suppose $U \sim \text{WE}(\alpha, \lambda)$, $V \sim \text{WE}(\alpha, \theta)$ and they are independently distributed. Let us define the random variable $X = U$, if $U > V$. Then it easily follows that X has the $\text{WWE}(\alpha, \frac{\theta}{\lambda}, \lambda)$ distribution.

Now to establish different properties of a $\text{WWE}(\alpha, \beta, \lambda)$ without loss of generality it is assumed that $\lambda = 1$. We denote this as $\text{WWE}(\alpha, \beta)$. If $X \sim \text{WWE}(\alpha, \beta)$, then it has the cumulative distribution function as

$$F_X(x; \alpha, \beta) = \frac{\beta + 1}{\beta}\left[1 - e^{-x^\alpha} - \frac{1}{\beta + 1}(1 - e^{-(1+\beta)x^\alpha})\right]; \tag{4}$$

for $x > 0$, and 0 otherwise. The hazard function for $x > 0$, is

$$\begin{aligned} h_X(x; \alpha, \beta) &= \alpha(\beta + 1)x^{\alpha-1} \times \frac{(1 - e^{-\beta x^\alpha})}{(\beta + 1 - e^{-\beta x^\alpha})} \\ &= \alpha(\beta + 1)x^{\alpha-1} \times \frac{1}{(\beta(1 - e^{-\beta x^\alpha})^{-1} + 1)}. \end{aligned} \tag{5}$$

The hazard functions for different values of α and β are plotted in Figure 2. The hazard function of WWE distribution can be increasing, decreasing and upside down functions. The following result provides the shape of the hazard function for different values of α. Theorem 2.2 establishes that the shape of the hazard function of a WWE distribution depends only on α, and it does not depend on β.

Theorem 2.2. *The hazard function is decreasing* (*upside-down bathtub shape*) (*increasing*) *if* $\alpha \le \frac{1}{2}$ $(\frac{1}{2} < \alpha < 1)$ $(\alpha \ge 1)$, *for all* $\beta > 0$.

Proof. See in the Appendix. □

Now we provide the following representation of X.

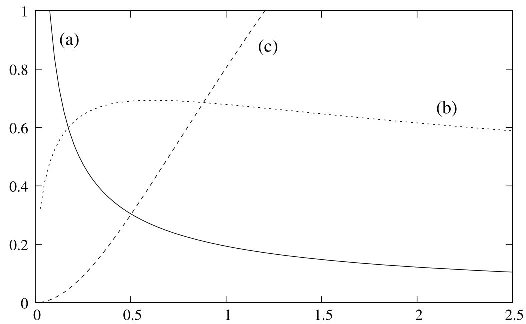

Figure 2 *HF of WWE distribution for different values of α and β when $\lambda = 1$*: (a) $\alpha = 0.25$, $\beta = 1.0$, (b) $\alpha = 0.75$, $\beta = 5.0$, (c) $\alpha = 1.5$, $\beta = 0.1$.

Theorem 2.3. *$X \sim \text{WWE}(\alpha, \beta)$ if and only if $X^\alpha \stackrel{d}{=} U + V$. Here U and V are independent exponential random variables with means 1 and $\frac{1}{(1+\beta)}$ respectively, and "$\stackrel{d}{=}$" means equal in distribution.*

Proof. The moment generating function of X^α for $|t| < 1$ is

$$\begin{aligned} M_{X^\alpha}(t) = Ee^{tX^\alpha} &= \frac{\beta+1}{\beta}\alpha \int_0^\infty e^{tx^\alpha} x^{\alpha-1} e^{-x^\alpha} \left(1 - e^{-\beta x^\alpha}\right) dx \\ &= \frac{\beta+1}{\beta} \int_0^\infty e^{tu} e^{-u} \left(1 - e^{-\beta u}\right) du \\ &= \frac{\beta+1}{\beta} \left[(1-t)^{-1} - \frac{1}{1+\beta} \left(1 - \frac{t}{1+\beta}\right)^{-1} \right] \\ &= \left(1 - \frac{t}{1+\beta}\right)^{-1} (1-t)^{-1}. \end{aligned}$$

Therefore, the result immediately follows. □

Theorem 2.3 becomes very useful in generating WWE distribution, using independent exponential distributions.

3 BWWE distribution

In this section, we introduce the bivariate weighted Weibull (BWWE) distribution and establish its different properties. Suppose $X_1 \sim \text{WE}(\alpha, \lambda_1)$, $X_2 \sim \text{WE}(\alpha, \lambda_2)$ and $X_3 \sim \text{WE}(\alpha, \lambda_3)$ and they are independently distributed. Consider the following bivariate random variables $X = X_1$ and $Y = X_2$, if $X_1 > X_3$ and $X_2 > X_3$. The new random variables (X, Y) is called the bivariate weighted Weibull distribution with parameters $\alpha, \lambda_1, \lambda_2, \lambda_3$, and it will be denoted by BWWE $(\alpha, \lambda_1, \lambda_2, \lambda_3)$. The joint CDF and the joint PDF can be obtained as follows.

Theorem 3.1. *If* $(X, Y) \sim \text{BWWE}(\alpha, \lambda_1, \lambda_2, \lambda_3)$, *then the joint CDF of* (X, Y) *is*

$$\begin{aligned} F_{X,Y}(x, y) = \left(1 - e^{-\lambda z^{\alpha}}\right) &- \frac{\lambda}{\lambda_1 + \lambda_3} e^{-\lambda_2 y^{\alpha}} \left(1 - e^{-(\lambda_1+\lambda_3) z^{\alpha}}\right) \\ &- \frac{\lambda}{\lambda_2 + \lambda_3} e^{-\lambda_1 x^{\alpha}} \left(1 - e^{-(\lambda_2+\lambda_3) z^{\alpha}}\right) \\ &+ \frac{\lambda}{\lambda_3} e^{-\lambda_1 x^{\alpha}} e^{-\lambda_2 y^{\alpha}} \left(1 - e^{-\lambda_3 z^{\alpha}}\right), \end{aligned} \tag{6}$$

where $z = \min\{x, y\}$, *and* $\lambda = \lambda_1 + \lambda_2 + \lambda_3$.

Proof.

$$\begin{aligned} F_{X,Y}(x, y) &= P(X \le x, Y \le y) = P(X_1 \le x, X_2 \le y | X_1 > X_3, X_2 > X_3) \\ &= \frac{P(X_1 \le x, X_2 \le y, X_3 \le \min\{X_1, X_2\})}{P(X_3 \le \min\{X_1, X_2\})}. \end{aligned} \tag{7}$$

The numerator of (7) is

$$\int_0^z \alpha \lambda_3 u^{\alpha-1} e^{-\lambda_3 u^{\alpha}} \left(e^{-\lambda_1 u^{\alpha}} - e^{-\lambda_1 x^{\alpha}}\right) \times \left(e^{-\lambda_2 u^{\alpha}} - e^{-\lambda_2 y^{\alpha}}\right) du. \tag{8}$$

Making the transformation $v = u^{\alpha}$ to perform the integration, and using the fact $P(X_3 \le \min\{X_1, X_2\} = \frac{\lambda_3}{\lambda})$, the result follows. □

The following results provide the joint PDF, marginals and the conditional PDF of a BWWE distribution.

Theorem 3.2. *Suppose* $(X, Y) \sim \text{BWWE}(\alpha, \lambda_1, \lambda_2, \lambda_3)$.

(a) *The joint PDF of* (X, Y) *is*

$$f_{X,Y}(x, y) = \frac{\alpha^2 \lambda (\lambda_1 \lambda_2)}{\lambda_3} x^{\alpha-1} y^{\alpha-1} e^{-\lambda_1 x^{\alpha} - \lambda_2 y^{\alpha}} \left(1 - e^{-\lambda_3 z^{\alpha}}\right). \tag{9}$$

(b) *The PDFs of X and Y are*

$$f_X(x) = \frac{\alpha \lambda \lambda_1}{(\lambda_2 + \lambda_3)} x^{\alpha-1} e^{-\lambda_1 x^{\alpha}} \left(1 - e^{-(\lambda_2+\lambda_3) x^{\alpha}}\right) \tag{10}$$

and

$$f_Y(y) = \frac{\alpha \lambda \lambda_2}{(\lambda_1 + \lambda_3)} y^{\alpha-1} e^{-\lambda_2 y^{\alpha}} \left(1 - e^{-(\lambda_1+\lambda_3) y^{\alpha}}\right), \tag{11}$$

respectively.

(c) *The conditional PDF of X given $Y = y$ is*

$$f_{X|Y=y}(x) = \begin{cases} \dfrac{\alpha\lambda_1(\lambda_1+\lambda_3)}{\lambda_3} x^{\alpha-1} e^{-\lambda_1 x^\alpha} \times \dfrac{(1-e^{-\lambda_3 y^\alpha})}{(1-e^{-(\lambda_1+\lambda_3)y^\alpha})} & \text{if } y < x, \\ \dfrac{\alpha\lambda_1(\lambda_1+\lambda_3)}{\lambda_3} x^{\alpha-1} e^{-\lambda_1 x^\alpha} \times \dfrac{(1-e^{-\lambda_3 x^\alpha})}{(1-e^{-(\lambda_1+\lambda_3)y^\alpha})} & \text{if } y > x. \end{cases} \tag{12}$$

Proof. The proofs can be obtained by routine calculations, and they are avoided. □

In Figure 3, we provide the surface plots of the joint PDF of (X, Y) for different parameters values. The joint PDF surface of (X, Y) is either a decreasing or an unimodal shape.

The following theorem provides the shape of the joint PDF of a BWWE distribution.

Theorem 3.3. *If* $(X, Y) \sim \text{BWW}(\alpha, \lambda_1, \lambda_2, \lambda_3)$, *then* $f_{X,Y}(x, y)$ *is decreasing* (*unimodal*) *if* $\alpha \le \frac{2}{3}$ $(\alpha > \frac{2}{3})$, *for all* $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0$.

Proof. See in the Appendix. □

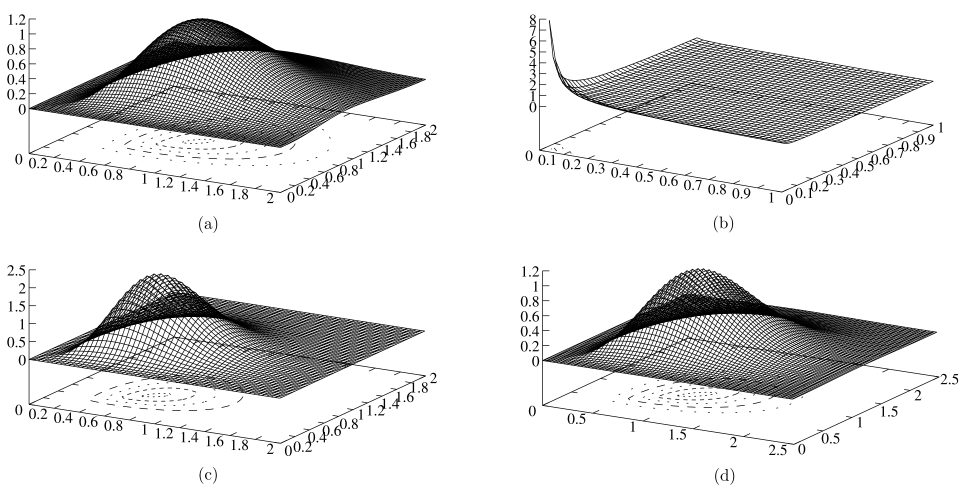

Figure 3 *Surface plots of* $f_{X,Y}(x, y)$ *for different values of* α, λ_1, λ_2 *and* λ_3. (a) $\alpha_1 = 2.0$, $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = 2.00$, (b) $\alpha_1 = 0.25$, $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = 2.00$, (c) $\alpha_1 = \lambda_1 = \lambda_2 = \lambda_3 = 2.00$, (d) $\alpha_1 = 2.0$, $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = 0.25$.

Using the joint PDF (9), we can calculate the product raw moment as follows.

$$E(XY) = \int\int_{x<y} xy f_{X,Y}(x,y)\,dx\,dy + \int\int_{x>y} xy f_{X,Y}(x,y)\,dx\,dy$$

$$= \frac{\lambda\lambda_1\lambda_2}{\lambda_3}\alpha\bigg\{\int_0^\infty y^\alpha e^{-\lambda_2 y^\alpha}\bigg[\frac{\gamma(1+\frac{1}{\alpha},\lambda_1 y^\alpha)}{\lambda_1^{1+\frac{1}{\alpha}}} - \frac{\gamma(1+\frac{1}{\alpha},(\lambda_1+\lambda_3)y^\alpha)}{(\lambda_1+\lambda_3)^{1+\frac{1}{\alpha}}}\bigg]dy$$

$$+ \int_0^\infty x^\alpha e^{-\lambda_1 x^\alpha}\bigg[\frac{\gamma(1+\frac{1}{\alpha},\lambda_2 x^\alpha)}{\lambda_2^{1+\frac{1}{\alpha}}} - \frac{\gamma(1+\frac{1}{\alpha},(\lambda_2+\lambda_3)x^\alpha)}{(\lambda_2+\lambda_3)^{1+\frac{1}{\alpha}}}\bigg]dx\bigg\}$$

$$= \frac{\lambda\lambda_1\lambda_2}{\lambda_3}\{I(\alpha,\lambda_2,\lambda_1) - I(\alpha,\lambda_2,\lambda_1+\lambda_3)$$

$$+ I(\alpha,\lambda_1,\lambda_2) - I(\alpha,\lambda_1,\lambda_2+\lambda_3)\},$$

where $\gamma(a,z) = \int_0^z t^{a-1}e^{-t}\,dt$, $a, z > 0$, is the lower incomplete gamma function and, for $\alpha, b, c > 0$,

$$I(\alpha,b,c) = \frac{\alpha}{c^{1+\frac{1}{\alpha}}}\int_0^\infty x^\alpha e^{-bx^\alpha}\gamma\Big(1+\frac{1}{\alpha}, cx^\alpha\Big)dx$$

$$= \frac{1}{c^{1+\frac{1}{\alpha}}}\int_0^\infty t^{1/\alpha}e^{-bt}\gamma\Big(1+\frac{1}{\alpha}, ct\Big)dt$$

$$= \frac{\Gamma(2+\frac{2}{\alpha})}{(1+\frac{1}{\alpha})(b+c)^{2+\frac{2}{\alpha}}}\,{}_2F_1\Big(2+\frac{2}{\alpha}, 2+\frac{2}{\alpha}; 2+\frac{1}{\alpha}; \frac{c}{b+c}\Big),$$

where

$${}_2F_1(a,b;c;z) = \sum_{n=0}^\infty \frac{(a)_n(b)_n}{(c)_n}\frac{z^n}{n!}, \qquad |z|<1,$$

$$(a)_0 = 1, \qquad (a)_n = a(a+1)\cdots(a+n-1),$$

is the hypergeometric function, see Erdélyi (1954, p. 308).

In Figure 4, we provide the correlation coefficient of X and Y for different values of α and λ_3, when $\lambda_1 = 1$ and $\lambda_2 = 1$. For each λ_3, as α increases correlation coefficient increases, and for fixed α, as λ_3 increases the correlation coefficient decreases.

Now we discuss how a random sample can be generated from the BWWE distribution. It is possible to generate random sample from a BWWE distribution using the acceptance rejection principle. To avoid that the following representation of the BWWE may be used.

Theorem 3.4. $(X, Y) \sim \text{BWWE}(\alpha, \lambda_1, \lambda_2, \lambda_3)$ *if and only if*

$$X^\alpha \stackrel{d}{=} U_1 + V \quad \textit{and} \quad Y^\alpha \stackrel{d}{=} U_2 + V, \tag{13}$$

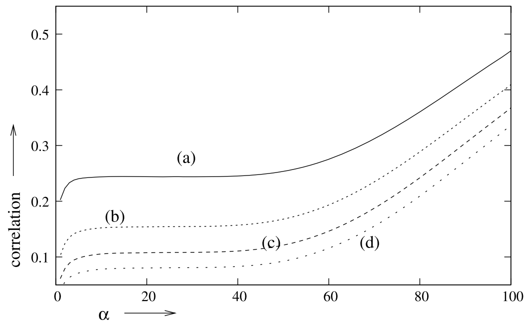

Figure 4 *Correlation coefficient between X and Y for different values of α and λ_3 when $\lambda_1 = 1$, $\lambda_2 = 1$*: (a) $\lambda_3 = 2.0$, (b) $\lambda_3 = 3.0$, (c) $\lambda_3 = 4.0$, (d) $\lambda_3 = 5.0$.

where U_1, U_2 and V are independent exponential random variables with means $\frac{1}{\lambda_1}$, $\frac{1}{\lambda_2}$ and $\frac{1}{\lambda}$, respectively.

Proof. Using (9), the joint PDF of X and Y, the joint moment generating function of X^α and Y^α can be easily obtained for $|t_1| < \lambda_1$ and $|t_2| < \lambda_2$, as

$$M_{X^\alpha, Y^\alpha}(t_1, t_2) = \left(1 - \frac{t_1}{\lambda_1}\right)^{-1} \left(1 - \frac{t_2}{\lambda_2}\right)^{-1} \left(1 - \frac{t_1 + t_2}{\lambda}\right)^{-1}.$$

Hence, the result follows immediately. □

It is clear that using Theorem 3.4, the BWWE random deviates can be generated from independent exponential random deviates directly. Moreover, Theorem 3.4 is a characterization of the BWWE distribution. Several structural properties of a BWWE can be obtained using Theorem 3.4. Using Theorem 3.4, it immediately follows that if (X, Y) is same as defined in Theorem 3.4, the stress strength parameter $R = P(X < Y)$ can be easily obtained as

$$R = P(X < Y) = P\big(X^\alpha < Y^\alpha\big) = P(U_1 < U_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

Now we study some of the stochastic monotonicity and dependence properties.

Theorem 3.5. *Suppose* $(X, Y) \sim \text{BWWE}(\alpha, \lambda_1, \lambda_2, \lambda_3)$, *then for all values of* $\alpha > 0, \lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0$, X (Y) *is neither stochastically increasing or decreasing in* Y (X).

Proof. X is stochastically increasing (decreasing) in Y, if $P(X > x|Y = y)$ is a non-decreasing (non-increasing) function of y for all x, see Shaked (1977). From

Theorem 3.2(c), we obtain

$$P(X > x|Y = y) = \begin{cases} g_1(x)\dfrac{(1 - e^{-\lambda_3 y^\alpha})}{(1 - e^{-(\lambda_1+\lambda_3)y^\alpha})} & \text{if } y < x, \\ g_2(x)\dfrac{1}{(1 - e^{-(\lambda_1+\lambda_3)y^\alpha})} & \text{if } y > x. \end{cases} \quad (14)$$

Here $g_1(\cdot)$ and $g_2(\cdot)$ are functions of x only. For any fixed x, since $\frac{(1-e^{-\lambda_3 y^\alpha})}{(1-e^{-(\lambda_1+\lambda_3)y^\alpha})}$ is an increasing and $\frac{1}{(1-e^{-(\lambda_1+\lambda_3)y^\alpha})}$ is a decreasing function of y, for all $\alpha > 0$, $\lambda_1 > 0$ and $\lambda_3 > 0$, the result follows. □

Theorem 3.6. *Suppose* $(X, Y) \sim \text{BWWE}(\alpha, \lambda_1, \lambda_2, \lambda_3)$, *then* (X, Y) *has total positivity of order two* (TP_2) *property.*

Proof. Note that (X, Y) has TP_2 property if and only if for any $t_{11}, t_{12}, t_{21}, t_{22}$, whenever $0 < t_{11} < t_{12}$ and $0 < t_{21} < t_{22}$, we have

$$f_{X,Y}(t_{11}, t_{21}) f_{X,Y}(t_{12}, t_{22}) - f_{X,Y}(t_{12}, t_{21}) f_{X,Y}(t_{11}, t_{22}) \geq 0. \quad (15)$$

Now we will consider all the cases separately. For example, if $t_{11} < t_{21} < t_{12} < t_{22}$, then left side of (15) becomes

$$e^{-\lambda_3 t_{21}^\alpha} - e^{-\lambda_3 t_{12}^\alpha}. \quad (16)$$

Since $t_{21} < t_{12}$, (16) is true. Similarly, it can be proved other cases also. □

The distribution of minimum and maximum of two random variables, say X and Y, play an important role in various statistical applications. For example, in the competing risks problems when the item can fail by two failures only, one observes only $T_1 = \min\{X, Y\}$, not both X and Y. In the reliability studies, when the components are arranged in a series system, only $T_1 = \min\{X, Y\}$ is observed. Similarly, in the complementary risks analysis, or when the components are arranged in parallel, one observes only $T_2 = \max\{X, Y\}$. If X and Y are independent and identically distributed random variables, T_1 and T_2 represent the two order statistics in a random sample of size 2. In practice, the independent assumptions may not very reasonable. Now we study different properties of T_1 and T_2, when (X, Y) follows BWWE.

Theorem 3.7. *Suppose* $(X, Y) \sim \text{BWWE}(\alpha, \lambda_1, \lambda_2, \lambda_3)$.

(a) *If* $T_1 = \min\{X, Y\}$, *then* T_1 *has* $\text{WWE}(\alpha, \frac{\lambda_3}{\lambda_1+\lambda_2}, (\lambda_1 + \lambda_2))$ *distribution with the following PDF*;

$$f_{T_1}(x) = \frac{\alpha\lambda(\lambda_1 + \lambda_2)}{\lambda_3} x^{\alpha-1} e^{-(\lambda_1+\lambda_2)x^\alpha} (1 - e^{-\lambda_3 x^\alpha}); \qquad x > 0. \quad (17)$$

(b) *If* $T_2 = \max\{X, Y\}$, *then the PDF of* T_2 *is*

$$f_{T_2}(x) = c_1 f_{\mathrm{WE}}(x; \alpha, \lambda) + c_2 f_{\mathrm{WE}}(x; \alpha, \lambda_1) + c_3 f_{\mathrm{WE}}(x; \alpha, \lambda_2) - c_4 f_{\mathrm{WE}}\big(x; \alpha, (\lambda_1 + \lambda_2)\big) \tag{18}$$

here

$$c_2 = \frac{\lambda}{\lambda_2 + \lambda_3}, \qquad c_3 = \frac{\lambda}{\lambda_1 + \lambda_3}, \qquad c_4 = \frac{\lambda}{\lambda_3}, \qquad c_1 = 1 - c_2 - c_3 + c_4$$

and $f_{\mathrm{WE}}(\cdot)$ *is the Weibull PDF as defined in* (3).

Proof. (a) Since

$$T_1 = \min\{X, Y\} \quad \Rightarrow \quad T_1^{\alpha} = \min\{X^{\alpha}, Y^{\alpha}\} = \min\{U_1, U_2\} + V.$$

Here U_1, U_2 and V are same as defined in Theorem 3.3. Since U_1 and U_2 are independent, $\min\{U_1, U_2\}$ has exponential distribution with mean $1/(\lambda_1 + \lambda_2)$. Since V is also exponential it follows that T_1 has WWE distribution. The PDF of T_1 can be obtained using the moment generating function.

(b) Note that using Theorem 3.1 and properly arranging the terms, $P(T_2 \le x)$ can be written as follows;

$$\begin{aligned} F_{T_2}(x) &= P(T_2 \le x) = P(X \le x, Y \le x) \\ &= 1 - c_1 e^{-\lambda x^{\alpha}} - c_2 e^{-\lambda_1 x^{\alpha}} - c_3 e^{-\lambda_2 x^{\alpha}} + c_4 e^{-(\lambda_1 + \lambda_2) x^{\alpha}}. \end{aligned}$$

Hence, the result can be easily obtained by taking $\frac{d}{dx} F_{X,Y}(x, x)$. □

4 Multivariate weighted Weibull distribution

In this section, we introduce the the multivariate weighted Weibull (MWWE) distribution using the same idea as the BWWE model. Suppose $U_1 \sim \mathrm{WE}(\alpha, \lambda_1), \ldots, U_p \sim \mathrm{WE}(\alpha, \lambda_p)$, and $V \sim \mathrm{WE}(\alpha, \lambda_{p+1})$, and they are independently distributed. Then we define the MWWE as follows:

Definition. Define $X_1 = U_1, \ldots, X_p = U_p$, if $\min\{U_1, \ldots, U_p\} \ge V$, where $U_1, \ldots, U_p, V$ are same as defined above. The random vector $(X_1, \ldots, X_p)$ is said to have MWWE distribution with parameters $\alpha, \lambda_1, \ldots, \lambda_{p+1}$, and it will be denoted by $\mathrm{MWWE}(\alpha, \lambda_1, \ldots, \lambda_{p+1})$.

To provide the distribution function of a MWWE, we need the following notations.

$$I = \{1, \ldots, p\}, \qquad A_{i_1,\ldots,i_j} = \{i_1, \ldots, i_j\} \subset I,$$
$$A^c_{i_1,\ldots,i_j} = I \setminus A_{i_1,\ldots,i_j}, \qquad \lambda = \sum_{k=1}^{p+1} \lambda_k. \tag{19}$$

The joint CDF and the joint PDF are provided in the following theorem.

Theorem 4.1. *Suppose* $\boldsymbol{X} = (X_1, \ldots, X_p) \sim \mathrm{MWWE}(\alpha, \lambda_1, \ldots, \lambda_{p+1})$.

(a) *The joint CDF of* $\boldsymbol{X} = (X_1, \ldots, X_p)$ *for* $z = \min\{x_1, \ldots, x_p\}$, *is*

$$\begin{aligned} &P(X_1 \leq x_1, \ldots, X_p \leq x_p) \\ &\quad = F_{\boldsymbol{X}}(x_1, \ldots, x_p) \\ &\quad = \lambda \sum_{j=0}^{p} \sum_{1 \leq i_1 < \cdots < i_j \leq p} \frac{(-1)^{p-j} e^{-\sum_{k \in A^c_{i_1,\ldots i_j}} (\lambda_k x_k^\alpha)}}{\lambda_{p+1} + \sum_{k \in A_{i_1,\ldots,i_j}} \lambda_k} \left(1 - e^{-z^\alpha \sum_{k \in A_{i_1,\ldots,i_j}} \lambda_k}\right), \end{aligned}$$

for $x_1 > 0, \ldots, x_p > 0$, *when* $\lambda = \lambda_1 + \cdots + \lambda_{p+1}$, *and* $z = \min\{x_1, \ldots, x_p\}$.

(b) *The joint PDF of* $\boldsymbol{X} = (X_1, \ldots, X_p)$ *is*

$$f_{\boldsymbol{X}}(x_1, \ldots, x_p) = \frac{\alpha^p \lambda}{\lambda_{p+1}} \prod_{j=1}^{p} \left\{\lambda_j x_j^{\alpha-1} e^{-\lambda_j x_j^\alpha}\right\} \left(1 - e^{-\lambda_{p+1} z^\alpha}\right), \tag{20}$$

for $x_1 > 0, \ldots, x_p > 0$, *when* $\lambda = \lambda_1 + \cdots + \lambda_{p+1}$, *and* $z = \min\{x_1, \ldots, x_p\}$.

Proof. (a)

$$\begin{aligned} &P(X_1 \leq x_1, \ldots, X_p \leq x_p) \\ &\quad = P\big(U_1 \leq x_1, \ldots, U_p \leq x_p | V < \min\{U_1, \ldots, U_p\}\big) \\ &\quad = \frac{P(U_1 \leq u_1, \ldots, U_p \leq u_p, V \leq \min\{U_1, \ldots, U_p\})}{P(V \leq \min\{U_1, \ldots, U_p\})}. \end{aligned} \tag{21}$$

The denominator of (21) is $\frac{\lambda_{p+1}}{\lambda}$, and the numerator can be written as

$$\int_0^z \alpha \lambda_{p+1} u^{\alpha-1} e^{-\lambda_{p+1} u^\alpha} \left(e^{-\lambda_1 u^\alpha} - e^{-\lambda_1 x_1^\alpha}\right) \cdots \left(e^{-\lambda_p u^\alpha} - e^{-\lambda_1 x_p^\alpha}\right) du. \tag{22}$$

Making the transformation $v = u^\alpha$, in (22), we can write (21) as

$$\lambda \int_0^{z^\alpha} e^{-\lambda_{p+1} v} \prod_{k=1}^{p} \left(e^{-\lambda_k v} - e^{-\lambda_k x_p^\alpha}\right) dv. \tag{23}$$

Now the result follows by expanding

$$\begin{aligned} &\prod_{k=1}^{p} \left(e^{-\lambda_k v} - e^{-\lambda_k x_p^\alpha}\right) \\ &\quad = \sum_{j=0}^{p} \sum_{1 \leq i_1 < \cdots < i_j \leq p} (-1)^{p-j} e^{-\sum_{k \in A^c_{i_1,\ldots i_j}} (\lambda_k x_k^\alpha)} e^{-v \sum_{k \in A_{i_1,\ldots,i_j}} \lambda_k}, \end{aligned} \tag{24}$$

and performing the integration.

(b) The joint PDF of $\boldsymbol{X} = (X_1, \ldots, X_p)$ can be obtained as

$$f_{\boldsymbol{X}}(x_1, \ldots, x_p) = \frac{\partial^p}{\partial x_1 \cdots \partial x_p} F_{X_1,\ldots,X_p}(x_1, \ldots, x_p). \tag{25}$$

From (23), it is clear that if $z = x_1$, then

$$\begin{aligned}
&\frac{\partial^p}{\partial x_1 \cdots \partial x_p} F_{\boldsymbol{X}}(x_1, \ldots, x_p) \\
&\quad = (-1)^p \frac{\partial^p}{\partial x_1 \cdots \partial x_p} \lambda e^{-\sum_{k=1}^p \lambda_k x_k^\alpha} \int_0^{x_1^\alpha} e^{-\lambda_{p+1} v}\, dv \\
&\qquad + (-1)^{p-1} \frac{\partial^p}{\partial x_1 \cdots \partial x_p} \lambda e^{-\sum_{k=2}^p \lambda_k x_k^\alpha} \int_0^{x_1^\alpha} e^{-(\lambda_1+\lambda_{p+1}) v}\, dv \\
&\quad = (-1)^p \frac{\lambda}{\lambda_{p+1}} \frac{\partial^p}{\partial x_1 \cdots \partial x_p} e^{-\sum_{k=1}^p \lambda_k x_k^\alpha} \left(1 - e^{-\lambda_{p+1} x_1^\alpha}\right) \\
&\qquad + (-1)^{p-1} \frac{\lambda}{\lambda_1 + \lambda_{p+1}} \frac{\partial^p}{\partial x_1 \cdots \partial x_p} e^{-\sum_{k=2}^p \lambda_k x_k^\alpha} \left(1 - e^{-(\lambda_1+\lambda_{p+1}) x_1^\alpha}\right).
\end{aligned} \tag{26}$$

The right-hand side of (26) can be written as

$$\begin{aligned}
&\frac{\lambda}{\lambda_{p+1}} \alpha^p \prod_{k=1}^p \{\lambda_k x_k^{\alpha-1} e^{-\lambda_k x_k^\alpha}\} \\
&\qquad - \frac{\lambda}{\lambda_{p+1}} \alpha^p (\lambda_1 + \lambda_{p+1}) x_1^{\alpha 1} e^{-(\lambda_1+\lambda_{p+1}) x_1^\alpha} \prod_{k=2}^p \{\lambda_k x_k^{\alpha-1} e^{-\lambda_k x_k^\alpha}\} \\
&\qquad + \lambda \alpha^p x_1^{\alpha 1} e^{-(\lambda_1+\lambda_{p+1}) x_1^\alpha} \prod_{k=2}^p \{\lambda_k x_k^{\alpha-1} e^{-\lambda_k x_k^\alpha}\}.
\end{aligned} \tag{27}$$

Note that (27) can be written as

$$\begin{aligned}
&\frac{\lambda}{\lambda_{p+1}} \alpha^p \prod_{k=1}^p \{\lambda_k x_k^{\alpha-1} e^{-\lambda_k x_k^\alpha}\} (1 - e^{-\lambda_{p+1} x_1^\alpha}) \\
&\qquad = \frac{\lambda}{\lambda_{p+1}} \alpha^p \prod_{k=1}^p \{\lambda_k x_k^{\alpha-1} e^{-\lambda_k x_k^\alpha}\} (1 - e^{-\lambda_{p+1} z^\alpha}).
\end{aligned} \tag{28}$$

Similarly, it can be shown if $z = x_2, \ldots, x_p$, and the result follows. □

It can be shown that the moment generating function of $X_1^\alpha, \ldots, X_p^\alpha$, for $|t_1| < \lambda_1, \ldots, |t_p| < \lambda_p$, is

$$M_{X_1^\alpha,\ldots,X_p^\alpha}(t_1, \ldots, t_p) = \left(1 - \frac{t_1}{\lambda_1}\right)^{-1} \cdots \left(1 - \frac{t_p}{\lambda_p}\right)^{-1} \left(1 - \frac{\sum_{k=1}^p t_k}{\lambda}\right)^{-1}. \tag{29}$$

Clearly, (29) can be used quite effectively for generating MWWE random deviates, because from (29) it easily follows that

$$X_1^\alpha \stackrel{d}{=} U_1 + V, \qquad \ldots, \qquad X_p^\alpha \stackrel{d}{=} U_p + V. \tag{30}$$

$U_1, \ldots, U_P$ and V are independent exponential random variables with means $\frac{1}{\lambda_1}$, $\ldots$, $\frac{1}{\lambda_p}$, $\frac{1}{\lambda}$ respectively. It is immediate from the moment generating function (29) that if $(X_1, \ldots, X_p) \sim \text{MMWE}(\alpha, \lambda_1, \ldots, \lambda_{p+1})$, then for $q < p$, $(X_1, \ldots, X_q) \sim \text{MMWE}(\alpha, \lambda_1, \ldots, \lambda_q, \widetilde{\lambda}_{q+1})$, where $\widetilde{\lambda}_{q+1} = (\lambda_{q+1} + \cdots + \lambda_{p+1})$.

Now we provide the distributions of the minimum and the maximum of $X_1, \ldots, X_p$, for some particular cases.

Theorem 4.2. *Suppose* $(X_1, \ldots, X_p) \sim \text{MWWE}(\alpha, \lambda_1, \ldots, \lambda_{p+1})$.

(a) *If* $T_1 = \min\{X_1, \ldots, X_p\}$, *then* T_1 *has the* $\text{WWE}(\alpha, \frac{\lambda_{p+1}}{\sum_{j=1}^p \lambda_j}, \sum_{j=1}^p \lambda_j)$ *distribution with the following PDF for* $x > 0$;

$$f_{T_1}(x) = \frac{\alpha\lambda(\sum_{j=1}^p \lambda_j)}{\lambda_{p+1}} x^{\alpha-1} e^{-x^\alpha \sum_{j=1}^p \lambda_j} \left(1 - e^{-\lambda_{p+1} x^\alpha}\right). \tag{31}$$

(b) *If* $T_p = \max\{X_1, \ldots, X_p\}$, *and* $\lambda_1 = \cdots = \lambda_p = \theta$, $\lambda_{p+1} = \beta$, *then the PDF of* T_p *for* $x > 0$, *is as follows*;

$$\begin{aligned} f_{T_p}(x) &= \alpha x^{\alpha-1} p\theta(p\theta + \beta) \\ &\quad \times \sum_{j=0}^{p-1} \binom{p-1}{j} (-1)^{p-1-j} \frac{(1 - e^{-(\theta-2+p-j)x^\alpha})}{(\theta - 2 + p - j)}. \end{aligned} \tag{32}$$

Proof. (a) It mainly follows from the representation (30).

(b) Using (30), we obtain

$$T_p^\alpha = \max\{X_1^\alpha, \ldots, X_p^\alpha\} = \max\{U_1, \ldots, U_p\} + V = W + V. \tag{33}$$

Note that W has a generalized exponential distribution (see Gupta and Kundu (1999)) with the PDF

$$f_W(w) = p\theta e^{-\theta w}(1 - e^{-\theta w})^{p-1}; \qquad w > 0. \tag{34}$$

Therefore, the PDF of $W + V = T_p^\alpha$ can be obtained as

$$f_{W+V}(y) = p\theta(p\theta + \beta) e^{-(p\theta+\beta)y} \int_0^y e^{-(\theta-1)x} (1 - e^{-\theta x})^{p-1}\, dx. \tag{35}$$

Since p is an integer, making the binomial expansion of $(1 - e^{-\theta x})^{p-1}$, and using the transformation $(W + V)^{1/\alpha}$, the result follows. □

Theorem 4.3. *Suppose* $\boldsymbol{X} = (X_1, \ldots, X_p) \sim \text{MWWE}(\alpha, \lambda_1, \ldots, \lambda_{p+1})$, *the joint PDF of* $\boldsymbol{X} = (X_1, \ldots, X_p)$ *has the multivariate total positivity of order two* (MTP_2) *property.*

Proof. Recall that a random vector $(X_1, \ldots, X_p)$ has the MTP_2 property if the joint PDF of $(X_1, \ldots, X_p)$ satisfies the following;

$$\begin{aligned} &f_{\boldsymbol{X}}(x_1, \ldots, x_p) f_{\boldsymbol{X}}(y_1, \ldots, y_p) \\ &\quad \le f_{\boldsymbol{X}}(x_1 \vee y_1, \ldots, x_p \vee y_p) f_{\boldsymbol{X}}(x_1 \wedge y_1, \ldots, x_p \wedge y_p), \end{aligned} \tag{36}$$

for $x_i, y_i \ge 0$, $x_i \vee y_i = \max(x_i, y_i)$ and $x_i \wedge y_i = \min(x_i, y_i)$ for $i = 1, \ldots, p$. Let us use the following notations;

$$\begin{aligned} &u = \min\{x_1, \ldots, x_p\}, \qquad v = \min\{y_1, \ldots, y_p\}, \\ &a = \max\{x_1 \vee y_1, \ldots, x_p \vee y_p\}, \qquad b = \min\{x_1 \wedge y_1, \ldots, x_p \wedge y_p\}. \end{aligned}$$

Clearly $b = \min\{u, v\} \le \max\{u, v\} \le a$.

First, consider the case $u \le v$. Therefore, $b = u \le v \le a$. Now proving (36) is equivalent in proving

$$\left(e^{-\lambda_{p+1} v} - e^{-\lambda_{p+1} a}\right)\left(1 - e^{-\lambda_{p+1} u}\right) \ge 0. \tag{37}$$

Since $v \le a$, (36) is true. Similarly, it can be proved for $u > v$ also. □

5 Maximum likelihood estimators

In this section, we consider the maximum likelihood estimators of the unknown parameters of the multivariate weighted Weibull distribution. We can state the problem as follows. Suppose we have the following p-variate sample of size n, $\{(x_{11}, \ldots, x_{p1}), \ldots, (x_{1n}, \ldots, x_{pn})\}$ from $\text{MWWE}(\alpha, \lambda_1, \ldots, \lambda_{p+1})$. Based on the sample, we want to estimate the unknown parameters, $\alpha, \lambda_1, \ldots, \lambda_{p+1}$. Based on the above data, from (20), the log-likelihood function can be written as

$$\begin{aligned} l(\alpha, \lambda_1, \ldots, \lambda_{p+1} | data) = n&\left(p \ln \alpha + \ln \lambda + \sum_{j=1}^{p} \ln \lambda_j - \ln \lambda_{p+1}\right) \\ &+ (\alpha - 1) \sum_{i=1}^{n} \sum_{j=1}^{p} \ln x_{ji} \\ &- \sum_{i=1}^{n} \sum_{j=1}^{p} \lambda_j x_{ji}^{\alpha} - \sum_{i=1}^{n} \ln\left(1 - e^{-\lambda_{p+1} z_i^{\alpha}}\right). \end{aligned} \tag{38}$$

Here $z_i = \min\{x_{1i}, \ldots, x_{pi}\}$ and $\lambda = \sum_{j=1}^{p+1} \lambda_j$. The MLEs can be obtained by maximizing (38) with respect to the unknown parameters. One needs to solve

$(p+2)$ non-linear equations to compute the MLEs of the unknown parameters. To avoid that, we propose to use the EM algorithm, which can be obtained by solving one non-linear equation at "E" step.

We will treat this problem as a missing value problem. It is clear from the definition of the MWWE that $X_1 = U_1, \ldots, X_p = U_p$ are observable and V is missing. But it is known that when we observe $X_1 = U_1, \ldots, X_p = U_p$, $V \leq \min\{U_1, \ldots, U_p\}$. Suppose we have the observed data and also the missing data $\{(x_{11}, \ldots, x_{p1}, v_1), \ldots, (x_{1n}, \ldots, x_{pn}, v_n)\}$, then the log-likelihood function based on the complete observations (CO) can be written as

$$\begin{aligned} l_C(\alpha, \lambda_1, \ldots, \lambda_{p+1}|CO) = n\left((p+1)\ln\alpha + \sum_{j=1}^{p+1}\ln\lambda_j\right) \\ + (\alpha-1)\left\{\sum_{i=1}^{n}\sum_{j=1}^{p}\ln x_{ji} + \sum_{i=1}^{n}\ln v_i\right\} \\ - \sum_{i=1}^{n}\sum_{j=1}^{p}\lambda_j x_{ji}^{\alpha} - \lambda_{p+1}\sum_{i=1}^{n} v_i^{\alpha}. \end{aligned} \tag{39}$$

From (39) based on the complete observations, for fixed α, the MLEs of $\lambda_1, \ldots \lambda_{p+1}$, say $\widehat{\lambda}_1(\alpha), \ldots, \widehat{\lambda}_{p+1}(\alpha)$ can be obtained as

$$\begin{aligned} &\widehat{\lambda}_1(\alpha) = \frac{n}{\sum_{i=1}^{n} x_{1i}^{\alpha}}, \quad \ldots, \quad \widehat{\lambda}_p(\alpha) = \frac{n}{\sum_{i=1}^{n} x_{pi}^{\alpha}}, \quad \text{and} \\ &\widehat{\lambda}_{p+1}(\alpha) = \frac{n}{\sum_{i=1}^{n} v_i^{\alpha}}, \end{aligned} \tag{40}$$

respectively. By maximizing $l_C(\alpha, \widehat{\lambda}_1(\alpha), \ldots, \widehat{\lambda}_{p+1}(\alpha)|CO)$ with respect to α, the MLE of α can be obtained by, say $\widehat{\alpha}$. Finally, the MLEs of $\lambda_1, \ldots, \lambda_{p+1}$ become $\widehat{\lambda}_1(\widehat{\alpha}), \ldots, \widehat{\lambda}_{p+1}(\widehat{\alpha})$, respectively.

We use the following result, whose proof is trivial, for further development.

Result 5.1. If $U \sim \text{WE}(\alpha, \theta)$, then for any fixed $c > 0$,

$$E(U|U \leq c) = \frac{\alpha\theta\int_0^c x^{\alpha} e^{-\theta x^{\alpha}}\,dx}{1 - e^{-\theta c^{\alpha}}}. \tag{41}$$

The following approximation of (41) will be useful.

$$E(U|U \leq c) \approx \left(E\left(U^{\alpha}|U^{\alpha} \leq c^{\alpha}\right)\right)^{\frac{1}{\alpha}} = \left(\frac{1}{\theta} - \frac{c^{\alpha} e^{-\theta c^{\alpha}}}{1 - e^{-\theta c^{\alpha}}}\right)^{\frac{1}{\alpha}} = A(\alpha, \theta, c). \tag{42}$$

Now we are in a position to provide the EM algorithm. Suppose at the kth stage of the EM algorithm the values of $\alpha, \lambda_1, \ldots, \lambda_{p+1}$ are $\alpha^{(k)}, \lambda_1^{(k)}, \ldots, \lambda_{p+1}^{(k)}$ respectively, we will show how to obtain $\alpha^{(k+1)}, \lambda_1^{(k+1)}, \ldots, \lambda_{p+1}^{(k+1)}$.

The "E"-step or the "pseudo" log-likelihood function at the kth stage can be formed by writing the log-likelihood function of the complete observations where the missing values are replaced by their expectation. Since v_i's are missing, we replace v_i by its expectation, namely $A(\alpha^{(k)}, \sum_{j=1}^{k} \lambda_j^{(k)}, z_i) = A_i^{(k)}$. At the "M"-step the "pseudo" log-likelihood function needs to be maximized with respect to the unknown parameters to compute $\alpha^{(k+1)}$, $\lambda_1^{(k+1)}, \ldots, \lambda_{p+1}^{(k+1)}$. By maximizing $l_C(\alpha, \lambda_1^{(k+1)}(\alpha), \ldots, \lambda_{p+1}^{(k+1)}(\alpha)|CO)$ with respect to α, $\alpha^{(k+1)}$ can be obtained, where

$$\begin{aligned}
\lambda_1^{(k+1)}(\alpha) &= \widehat{\lambda}_1(\alpha) = \frac{n}{\sum_{i=1}^{n} x_{1i}^{\alpha}}, \qquad \ldots, \\
\lambda_p^{(k+1)}(\alpha) &= \widehat{\lambda}_p(\alpha) = \frac{n}{\sum_{i=1}^{n} x_{pi}^{\alpha}}, \\
\lambda_{p+1}^{(k+1)}(\alpha) &= \frac{n}{\sum_{i=1}^{n} [A_i^{(k)}]^{\alpha}}.
\end{aligned} \tag{43}$$

The maximization of $l_C(\alpha, \lambda_1^{(k+1)}(\alpha), \ldots, \lambda_{p+1}^{(k+1)}(\alpha)|CO)$ with respect to α needs to be performed numerically. The following result provides the uniqueness of the solution.

Theorem 5.1. *The function* $g(\alpha) = l_C(\alpha, \lambda_1^{(k+1)}(\alpha), \ldots, \lambda_{p+1}^{(k+1)}(\alpha)|CO)$ *is unimodal.*

Proof. Along the same line as the proof of Theorem 2 of Kundu (2008), it can be shown that $g(\alpha)$ is log-concave. Now the result follows by observing the fact $g(\alpha)$ goes to $-\infty$ as $\alpha \to 0$ or $\alpha \to \infty$. □

6 Data analysis

In this section, we perform the analysis of a data set to see how the proposed model works in practice. The data set represents the marks of Physics (P), Chemistry (C) and Mathematics (M) of 30 students who had qualified Joint Entrance Examination (JEE) 2009 and are studying in a particular branch of Indian Institute of Technology (IIT) Kanpur. The JEE is a nation wide examination in India, which is being conducted at the class 12th level for entries in different IITs in India. Before 2007, these marks were not available, but after that, due to Right to Information (RTI) act, these marks are available. One of the authors have collected these marks of 30 students, from a particular class. The marks are presented in Table 1.

It is known that for qualifying the JEE, it is not only the total marks, but each subject has a individual cutoff, i.e. for each qualified student it is known that $X_1 > u_1$, $X_2 > u_2$, and $X_3 > u_3$. Here u_1, u_2, u_3 are the cutoff marks for Mathematics, Physics and Chemistry, and they are not fixed. They vary in each year,

Table 1 *The marks of Mathematics (M), Physics (P) and Chemistry (C) of* 30 *students*

No.	M X_1	P X_2	C X_3	No.	M X_1	P X_2	C X_3	No.	M X_1	P X_2	C X_3
1	153	149	122	2	149	148	120	3	143	144	126
4	136	156	117	5	141	151	115	6	153	154	97
7	151	127	118	8	117	149	120	9	156	124	105
10	132	146	106	11	126	140	118	12	137	131	114
13	132	118	132	14	131	142	109	15	145	131	105
16	123	136	122	17	139	132	109	18	145	132	101
19	138	121	118	20	145	130	101	21	134	120	121
22	143	131	100	23	124	131	118	24	115	136	121
25	127	141	104	26	113	141	118	27	151	139	82
28	119	141	111	29	135	124	112	30	128	126	117

Table 2 *The basic statistics of the subject wise scores of* 30 *students*

Var.	Mean	SD	Median	Minimum	Maximum	Q_1	Q_3	IQR
X_1	136.03	12.06	136.5	113	156	127.25	145	17.75
X_2	136.37	10.45	136	118	156	130.25	143.5	13.25
X_3	112.63	10.25	116	82	132	105.25	119.5	14.25

therefore, we can take them as random variables. We are making the assumption that $u_1 = u_2 = u_3$, therefore, our model can be used to analyze this data set. We may have another interpretation of our model for fitting this data set as follows. It may be assumed that the marks (or a transformed version of the marks) of a student for a specific subject depend on his/ her overall knowledge plus the subject specific knowledge, that is,

$$X^{\alpha} = V + U_1, \qquad X_2^{\alpha} = V + U_2, \qquad X_3^{\alpha} = V + U_3.$$

Here V is the contribution due to overall knowledge, X_i is the contribution due to subject specific knowledge. Therefore, based on the above assumptions, our proposed model is applicable for this data set.

We present the mean, standard deviation (SD), median, lowest score, highest score, first quartile (Q_1), third quartile (Q_3) and inter quartile range (IQR) of the subject wise scores of the 30 students in Table 2. We have also provided the scaled total time on test (TTT) transform as suggested by Aarset (1987) in Figure 5, where

$$T\left(\frac{r}{n}\right) = \frac{\sum_{i=1}^{r} t_{(i)} + (n-r)t_{(r)}}{\sum_{i=1}^{n} t_i}, \qquad r = 1, 2, \ldots, n,$$

and $t_{(1)} \leq t_{(2)} \leq \cdots \leq t_{(n)}$ are the order statistics of a data set $t_1, t_2, \ldots, t_n$.

Since each of them is concave in nature, we conclude that the hazard rate functions are increasing, see Aarset (1987).

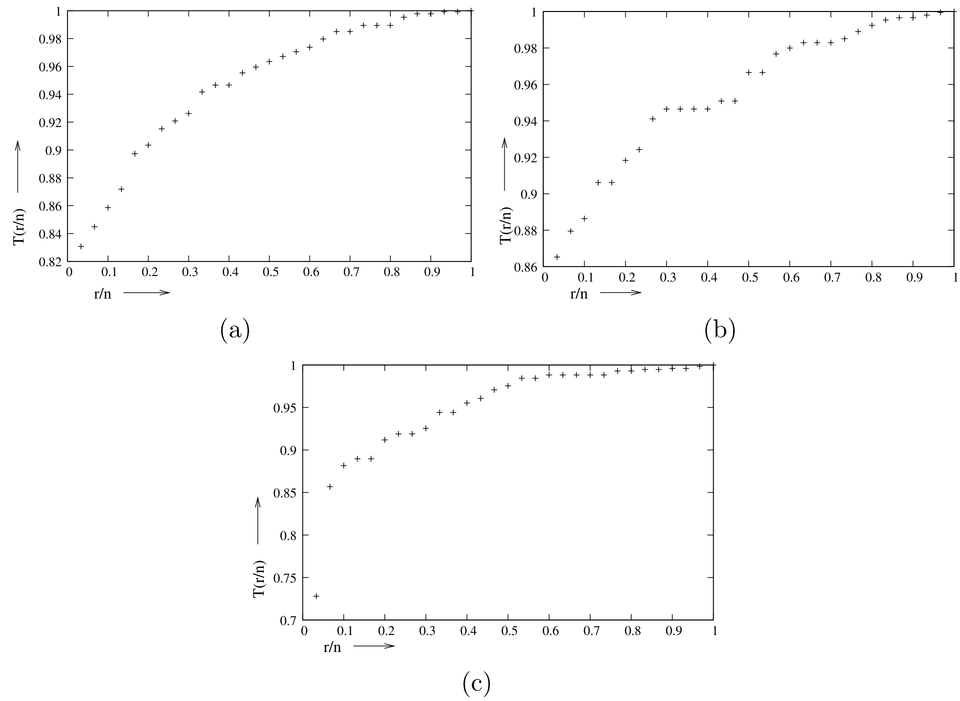

Figure 5 *Scaled TTT transform of* (a) X_1, (b) X_2, (c) X_3.

First, we fit the univariate weighted Weibull models to X_1, X_2 and X_3. We use the EM algorithm, to compute the MLEs of α, $\lambda_1 = \lambda$, $\lambda_2 = \lambda_1/\beta$ and we have computed the associated 95% confidence intervals also based on bootstrapping. In each case, we have also computed the Kolmogorov–Smirnov distances between the empirical distribution function and the fitted distribution functions, and the associated p values. The results are presented in Table 3. From the p values, it is clear that we cannot reject the null hypotheses that X_1, X_2 and X_3 are coming from weighted Weibull distributions.

Now we fit the trivariate weighted Weibull distribution to the data set. We obtain the estimates of α, λ_1, λ_2, λ_3 and λ_4 as 13.8399, 0.0081, 0.0086, 0.1141 and 0.0077, respectively. The corresponding log-likelihood value becomes -232.131. The associated 95% confidence intervals are (12.6964, 16.1787), (0.0038, 0.0144), (0.0038, 0.0146), (0.0677, 0.1532), (0.0022, 0.0366), respectively.

Some of the simple interpretations can be provided from the fitted model. For example, the estimate of the probability that in this group a student gets more marks in Mathematics than Chemistry, that is, $P(X_1 > X_3)$, is 0.94, similarly, a student gets more marks in Mathematics than Physics, $P(X_1 > X_2)$, is 0.52. Moreover, for the transformed data say for Mathematics, X_1^{α}, the mean contribution due to subject knowledge is $(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)/\lambda_1 \approx 17.1$ more than the mean contribution due to overall knowledge.

Table 3 *The parameter estimates of the weighted Weibull distribution fitted to the subject wise scores of* 30 *students*

Var.	Estimates (confidence intervals)			Goodness of fit	
	α	λ_1	λ_2	K-S	p
X_1	13.320 (10.700, 17.461)	0.0099 (0.0015, 0.0292)	0.0005 (0.0000, 0.0018)	0.1096	0.864
X_2	14.467 (12.250, 18.549)	0.0068 (0.0009, 0.0165)	0.0004 (0.0000, 0.0012)	0.1918	0.220
X_3	13.756 (10.350, 20.427)	0.1159 (0.0330, 0.2257)	0.0047 (0.0015, 0.0127)	0.1189	0.790

Table 4 *AIC and BIC values for the different models*

Model	No. of parameters	Log-likelihood value	AIC value	BIC value
TWE	4	−345.17	698.34	703.94
BBTW	5	−271.56	553.12	560.13
MWWE	5	−232.13	474.26	481.27

For comparison purposes, we have fitted four-parameter trivariate weighted exponential (TWE) distribution as provided in Al-Mutairi et al. (2011) and five-parameter Block and Basu trivariate Weibull distribution (BBTW) as proposed by Pradhan and Kundu (2016). We use the same set of notations as they have been used in those respective papers, and we present the MLEs and the corresponding log-likelihood (ll) values:

TWE: $\widehat{\lambda}_1 = 0.7351$, $\widehat{\lambda}_2 = 0.7333$, $\widehat{\lambda}_3 = 0.8878$, $\widehat{\lambda}_4 = 9.4251$, $\text{ll} = -345.17$.

BBTW: $\widehat{\lambda}_0 = 0.1345$, $\widehat{\lambda}_1 = 0.4561$, $\widehat{\lambda}_2 = 0.6751$, $\widehat{\lambda}_3 = 0.4519$, $\widehat{\alpha} = 5.1756$, $\text{ll} = -271.56$.

Now we would like to choose the best fitted model based on Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC). The AIC and BIC values for the different models are provided in Table 4. Therefore, based on AIC and BIC values, it is observed that for this data set MWWE provides a better fit than the other two models.

7 Conclusions

In this paper, we have studied different properties of the weighted Weibull distribution proposed by Shahbaz et al. (2010). It is observed that the proposed model can be obtained as a hidden truncation model, similarly as the skewed normal distribution proposed by Azzalini (1985). This three parameter weighted Weibull model is very flexible in terms of the different shapes of its PDF and HRF, and therefore it can be used very effectively to analyze failure time data. We have also proposed to use the EM algorithm to compute the maximum likelihood estimators, and the implementation of the EM algorithm is also quite simple since it involves just solving one non-linear equation at each "M" step.

We further extend the model to bivariate and multivariate cases. We have observed that the bivariate and multivariate models enjoy several interesting properties. The generation from the bivariate or multivariate weighted Weibull distribution is quite straight forward. The MLEs of the unknown parameters for the multivariate weighted Weibull distribution can be obtained using the EM algorithm, and it also involves just solving one non-linear equation at each "M"-step. We have analyzed one trivariate data set, and it is observed that the proposed model and the EM algorithm work very well in this case.

Appendix

Proof of Theorem 2.1. Without loss of generality we assume $\lambda = 1$, and we denote in this proof only $f_X(x; \alpha, \beta, 1) = f(x)$. Now note that

$$f(0) = \begin{cases} \infty, & \text{if } \alpha < \frac{1}{2}, \\ 1, & \text{if } \alpha = \frac{1}{2}, \\ 0, & \text{if } \alpha > \frac{1}{2}, \end{cases} \qquad f(\infty) = 0.$$

$f'(x) = 0$ implies that

$$w(y) = \left(\alpha - 1 - \frac{\alpha}{\beta} y\right) e^y + \alpha\left(1 + \frac{1}{\beta}\right) y + 1 - \alpha = 0, \qquad y = \beta x^\alpha > 0.$$

Note that $w(0) = 0$, and $w(\infty) = -\infty$,

$$w'(y) = \left[\alpha - 1 - \frac{\alpha}{\beta}(1 + y)\right] e^y + \alpha\left(1 + \frac{1}{\beta}\right), \qquad y > 0.$$

Note that $w'(0) = 2\alpha - 1$ and $w'(\infty) = -\infty$.

Now, since

$$w''(y) = \left[\alpha - 1 - \frac{\alpha}{\beta}(2 + y)\right] e^y, \qquad y > 0$$

it follows that $w'(y)$ is monotonically decreasing. So, $w'(y)$ is $-$ve (if $\alpha \le \frac{1}{2}$) and has a unique maximum changing sign from $+$ve to $-$ve (if $\alpha > \frac{1}{2}$). Since $w(0) = 0$ and $w(\infty) = -\infty$, it follows that $w(y)$ is $-$ve (if $\alpha \le \frac{1}{2}$) and has a unique maximum changing sign from $+$ve to $-$ve (if $\alpha > \frac{1}{2}$). $\square$

To prove Theorem 2.2, we need the following lemma.

Lemma 1. *The sign of the function*

$$g(y) = 2(\alpha - 1)(\beta + 1)e^y + \alpha\beta y + \beta - 2(\alpha - 1), \qquad y > 0, \alpha, \beta > 0,$$

is negative (*positive then negative*) (*positive*) *if* $\alpha \leq \frac{1}{2}$ $(\frac{1}{2} < \alpha < 1)$ $(\alpha \geq 1)$.

Proof. First, note that $g(0) = (2\alpha - 1)\beta$ and

$$g'(y) = 2(\alpha - 1)(\beta + 1)e^y + \alpha\beta, \qquad y > 0, \alpha, \beta > 0.$$

(i) $0 < \alpha \leq \frac{1}{2}$: Here, $g'(y) \leq 2(\alpha - 1)\beta + \alpha\beta = (3\alpha - 2)\beta < 0$ and $g(0) \leq 0$, implying that $g(y) \leq 0$.
(ii) $\frac{1}{2} < \alpha < 1$: Here, $g'(y_0) = 0$ implies that $y_0 = \ln[\frac{\alpha\beta}{2(1-\alpha)(\beta+1)}]$. Also, $g''(y) = 2(\alpha - 1)(\beta + 1)e^y > 0$, implying that $g(y)$ has a unique maximum at y_0. Now, in this case, $g(0) > 0$ and $g(\infty) = -\infty$, implying that $g(y)$ changes sign from positive to negative.
(iii) $\alpha \geq 1$: Here, $g'(y) > 0$ and $g(0) > 0$, implying that $g(y) > 0$. □

Proof of Theorem 2.2. Without loss of generality, we assume $\lambda = 1$. In this proof, only we denote $h_X(x; \alpha, \beta) = h(x)$. Now note that

$$h(0) = f(0) = \begin{cases} \infty, & \text{if } \alpha < \dfrac{1}{2}, \\ 1, & \text{if } \alpha = \dfrac{1}{2}, \\ 0, & \text{if } \alpha > \dfrac{1}{2}, \end{cases} \qquad h(\infty) = \begin{cases} 0, & \text{if } \alpha < 1, \\ 1, & \text{if } \alpha = 1, \\ \infty, & \text{if } \alpha > 1. \end{cases}$$

$h'(x) = 0$ implies that

$$u(y) = (\alpha - 1)(\beta + 1)e^{2y} - [(\alpha - 1)(\beta + 2) - \alpha\beta y]e^y + \alpha - 1, \qquad y = \beta x^\alpha > 0.$$

Note that the sign of $h'(x)$ is the sign of $u(\beta x^\alpha)$ and $u(0) = 0$.

Since

$$u'(y) = e^y[2(\alpha - 1)(\beta + 1)e^y + \alpha\beta y + \beta - 2(\alpha - 1)] = e^y g(y),$$

it follows that the sign of $u'(y)$ is the sign of $g(y)$. Now by Lemma 1, we have

(i) $0 < \alpha \leq \frac{1}{2}$: Here, $u(y) \leq 0$, for all $y > 0$, i.e. $h'(x) \leq 0$ implying that $h(x)$ is a decreasing function.
(ii) $\frac{1}{2} < \alpha < 1$: Here, $u(y)$ changes sign from positive to negative, i.e. $h'(x)$ changes sign from positive to negative implying that $h(x)$ is an increasing-decreasing (upside-down bathtub shape) function.
(iii) $\alpha \geq 1$: Here $u(y) > 0$, for all $y > 0$, i.e. $h'(x) > 0$ implying that $h(x)$ is an increasing function. □

Proof of Theorem 3.3. *Case 1*: $R_1 = \{(x, y) : 0 < x < y\}$: Here, we have

$$f_{X,Y}(x, y) = a_1 f_1(x) g_1(y), \qquad x < y,$$

where a_1 is a normalizing constant, $f_1(x)$ and $g_1(y)$ are the PDFs of WWE($\alpha, \lambda_1, \frac{\lambda_3}{\lambda_1}$) and WE($\alpha, \lambda_2$), respectively.

We know, using Theorem 2.1, that WWE PDF $f_1(x)$ has a unique critical point x_1 if $\alpha > \frac{1}{2}$. Also, the Weibull PDF $g_1(y)$ has a unique critical point $y_1 = (\frac{\alpha-1}{\alpha\lambda_2})^{1/\alpha}$ if $\alpha > 1$. It follows that (x_1, y_1) is the unique critical point of $f_{X,Y}(x, y)$ provided that $\alpha > 1$ and $x_1 < y_1$. Finally, $(x_1, y_1) \in R_1$ maximizes $f_{X,Y}(x, y)$, since, for $\alpha > 1$, $f_{X,Y}(0, 0) = 0$ and $f_{X,Y}(x, y) > 0$.

Case 2: $R_2 = \{(x, y) : x > y > 0\}$: Here, we have

$$f_{X,Y}(x, y) = a_2 f_2(x) g_2(y), \qquad x > y,$$

where a_2 is a normalizing constant, $f_2(x)$ and $g_2(y)$ are the PDFs of WE(α, λ_1) and WWE($\alpha, \lambda_2, \frac{\lambda_3}{\lambda_2}$), respectively.

By similar arguments as in Case 1, $f_2(x)$ has a unique critical point $x_2 = (\frac{\alpha-1}{\alpha\lambda_1})^{1/\alpha}$ if $\alpha > 1$ and $g_2(y)$ has a unique critical point y_2 if $\alpha > \frac{1}{2}$. The unique critical point $(x_2, y_2) \in R_2$ maximizes $f_{X,Y}(x, y)$ provided that $\alpha > 1$ and $x_2 > y_2$.

Before we proceed further, we show that, under the conditions of Cases 1 and 2, the critical points $(x_1, y_1) \in R_1$ and $(x_2, y_2) \in R_2$ cannot occur simultaneously.

First, note that $f_1(x) = b_1 f_2(x) F_3(x)$, where b_1 is a normalizing constant and $F_3(x)$ is the CDF of $W(\alpha, \lambda_3)$. Since

$$\begin{aligned} f_1'(x)\big|_{x=x_2} &= b_1\big\{f_2(x) F_3'(x) + f_2'(x) F_3(x)\big\}\big|_{x=x_2} \\ &= b_1 f_2(x_2) F_3'(x_2) > 0, \end{aligned}$$

it follows that $x_1 > x_2$. Similarly, by noting that $g_2(y) = b_2 g_1(y) F_3(y)$, where b_2 is a normalizing constant, we have $y_1 < y_2$. Figure 4 shows that the critical points $(x_1, y_1) \in R_1$ and $(x_2, y_2) \in R_2$ cannot occur simultaneously.

Case 3: $R_3 = \{(x, y) : x = y > 0\}$: Here, we have

$$f_{X,Y}(x, x) = \frac{\lambda\lambda_1\lambda_2}{\lambda_3}\alpha^2 x^{2(\alpha-1)}\, e^{-(\lambda_1+\lambda_2)x^\alpha}\big(1 - e^{-\lambda_3 x^\alpha}\big), \qquad x > 0,$$

which is a unimodal function if $\alpha > \frac{2}{3}$, the proof is similar to that of Theorem 2.1.

Clearly, if $\alpha > 1$ and there does not exist a critical point in R_1 (R_2), that is, if $x_1 \geq y_1$ ($x_2 \leq y_2$), then $f_{X,Y}(x, y)$ has a global maximum on the boundary R_3, since, in this case, $f_{X,Y}(0, 0) = f_{X,Y}(\infty, \infty) = 0$ and $f_{X,Y}(x, y) > 0$. Also, when $\frac{2}{3} < \alpha \leq 1$, $f_{X,Y}(x, y)$ is maximized on the boundary R_3. Finally, when $\alpha \leq \frac{2}{3}$ no critical point exist for $f_{X,Y}(x, y)$ in the space $(0, \infty) \times (0, \infty)$. In fact, $f_{X,Y}(x, y)$ is a decreasing function in (x, y), since, in this case, $f(0, 0) = \infty$ and $f(\infty, \infty) = 0$. □

Acknowledgments

The authors would like to thank two unknown referees for many constructive suggestions which had helped us to improve the manuscript significantly.

References

Aarset, M. V. (1987). How to identify a bathtub shaped hazard rate? *IEEE Trans. Reliab.* **36**, 106–108.

Al-Mutairi, D. K., Ghitany, M. E. and Kundu, D. (2011). A new bivariate distribution with weighted exponential marginals and its multivariate generalization. *Statist. Papers* **52**, 921–936. MR2846694

Azzalini, A. (1985). A class of distribution which includes the normal ones. *Scand. J. Stat.* **12**, 171–178. MR0808153

Erdélyi, A. (1954). *Tables of Integral Transforms, Vol. II*. New York: McGraw-Hill.

Gupta, R. D. and Kundu, D. (1999). Generalized exponential distribution. *Aust. N. Z. J. Stat.* **41**, 171–188. MR1705342

Gupta, R. D. and Kundu, D. (2009). A new class of weighted exponential distribution. *Statistics* **43**, 621–643. MR2588273

Jamalizadeh, A. and Kundu, D. (2013). Weighted Marshall–Olkin bivariate exponential distribution. *Statistics* **47**, 917–928. MR3175724

Karlin, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive dependence. *J. Multivariate Anal.* **10**, 467–498. MR0599685

Kundu, D. (2008). Bayesian inference and life testing plans for the Weibull distribution in presence of progressive censoring. *Technometrics* **50**, 144–154. MR2439875

Pradhan, B. and Kundu, D. (2016). Bayes estimation for the Block and Basu bivariate and multivariate Weibull distributions. *J. Stat. Comput. Simul.* **86**, 170–182. MR3403631

Shahbaz, S., Shahbaz, M. Q. and Butt, N. S. (2010). A class of weighted Weibull distribution. *Pak. J. Stat. Oper. Res.* **VI**, 53–59.

Shaked, M. (1977). A family of concepts of dependence for bivariate distribution. *J. Amer. Statist. Assoc.* **72**, 642–650. MR0652948

D. K. Al-Mutairi
M. E. Ghitany
Department of Statistics
and Operation Research
Faculty of Science
Kuwait University
P.O. Box 5969
Safat
Kuwait 13060
E-mail: mutairi@yahoo.com
mghitany@yahoo.com

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