J. of Ramanujan Society of Math. and Math. Sc. **ISSN : 2319-1023**
Vol.6, No.2 (2017), pp. 01-22

MARSHALL-OLKIN-RATHIE-SWAMEE DISTRIBUTION AND APPLICATIONS

Ronald T. Nojosa and Pushpa N. Rathie

Department of Statistics and Applied Mathematics,
Federal University of Ceara, 60440-900, Fortaleza, Brazil
E-mail: rtargino@ufc.br, pushpanrathie@yahoo.com

Abstract: The new tilted generalized logistic distribution, also called Marshall-Olkin-Rathie-Swamee (MORS) distribution, is studied in some detail. This distribution is important because it is multimodal and generalizes the logistic distribution among others. Moments and order statistics are given. The reliability $P(X < Y)$, for X and Y independent generalized logistic and beta-generated MORS distributions, is obtained along with its particular cases. Also it is proved that the beta-generated MORS distribution is an infinite linear combination of a new distribution which is obtained from powers of MORS distribution. Maximum likelihood method is used to estimate the parameters of the distribution. Four applications, with real data, are presented to illustrate the applicability of the proposed distribution. For corresponding non-negative random variable, which resulted in a generalization of the Harris extended exponential (HEE) distribution, moments are obtained extending a recent result given for Marshall-Olkin exponential Weibull (MOEW) distribution.

Keywords and Phrases: MORS, MOEW and HEE distributions, Generalized hypergeometric functions, Reliability analysis, Beta-generated distributions.

2010 Mathematics Subject Classification: 60E05, 62E99, 33C60.

1. Introduction

In this article, we introduce a new tilted (skew) multimodal distribution with four parameters illustrating its usefulness in modeling four real data sets. Since it unifies a few previously available distributions, we hope that this new flexible distribution will attract more future research work.

Estimation of parameters is done by likelihood maximization procedure. Our model is very flexible to fit different unimodal and bimodal data. We could not provide a practical application of our distribution for trimodal data because no such data was available for analysis. Bimodality generally result from a binary causative factor which has not been measured in a given data set. So the data may not be separated into two groups due to the unobservability of the causative factor.

In some cases, more than one model provides an adequate fit with only slight changes in BIC or MSE values. In such cases, we may take into account the interpretability of parameters or the standard errors of the estimates and get rid of a parameter with high error. The heavy tailed generalized student-t distribution, as an alternative to Lévy distribution has been used by Rathie et al. [15] to analyse some stock exchange data.

The rest of this paper is divided as follows: we provide some definitions and known results in Section 2. Section 3 deals with the tilted generalized logistic distribution obtained by applying Marshall-Olkin procedure to the Rathie-Swamee generalized logistic distribution. Hazard rate function and order statistics are also included here. In Section 4, n-th moments and two particular cases are derived. Reliability probability $P(X<Y)$ is calculated in Section 5, utilizing beta-generated distribution. Section 6 deals with Marshall-Olkin exponential Weibull (MOEW) distribution studied by Pogány et al. [9] and we derived n-th moments extending their result. Beta-generated distribution, with MOEW as a base distribution, is also derived which resulted in a generalization of the Harris extended exponential (HEE) distribution. In Section 7, we use MLE procedure to estimate parameters of the MORS distribution and apply to real data sets involving (a) Old Faithfull Geyser data: waiting time and eruption duration; (b) Shrimp weight data; and (c) Environmental Performance Index (EPI) data. In the last Section, we give a brief summary concluding the work.

2. Generalized hypergeometric functions and known results

In this section, we give some definitions and results which will be used in this article.

The G-function is defined as

$$G^{m,n}_{p,q}\left[x\middle|{}^{a_1,\ldots,a_p}_{b_1,\ldots,b_q}\right]=\frac{1}{2\pi i}\int_L \frac{\prod_{j=1}^{m}\Gamma(b_j-s)\prod_{j=1}^{n}\Gamma(1-a_j+s)}{\prod_{j=m+1}^{q}\Gamma(1-b_j+s)\prod_{j=n+1}^{p}\Gamma(a_j-s)}x^s ds, \quad (1)$$

where $x\neq 0$, an empty product is interpreted as unity, $0\le m\le q$ and $0\le n\le p$ (not both m and n zeros simultaneously). The parameters b_j, $j=1,2,\ldots,m$ and a_j, $j=1,2,\ldots,n$, are such that no pole of $\prod_{j=1}^{m}\Gamma(b_j-s)$ coincides with any pole of $\prod_{j=1}^{n}\Gamma(1-a_j+s)$. See Luke [4, pp. 143-144] for details about the contour L and conditions of convergence of the integral.

The H-function, which is a generalization of the G-function, is defined as

$$H^{m,n}_{p,q}\left[x\Big|^{(a_1,A_1),\dots,(a_n,A_n),(a_{n+1},A_{n+1}),\dots,(a_p,A_p)}_{(b_1,B_1),\dots,(b_m,B_m),(b_{m+1},B_{m+1}),\dots,(b_q,B_q)}\right] = \frac{1}{2\pi i}\int_L \frac{\prod_{j=1}^{m}\Gamma(b_j-B_js)\prod_{j=1}^{n}\Gamma(1-a_j+A_js)}{\prod_{j=m+1}^{q}\Gamma(1-b_j+B_js)\prod_{j=n+1}^{p}\Gamma(a_j-A_js)}x^s ds. \quad (2)$$

For more details about H-function, see Mathai et al. [6].

The I-function, a generalization of the H-function, is defined as

$$I^{m,n}_{p,q}\left[x\Big|^{(a_1,A_1,\alpha_1),\dots,(a_n,A_n,\alpha_n),(a_{n+1},A_{n+1},\alpha_{n+1}),\dots,(a_p,A_p,\alpha_p)}_{(b_1,B_1,\beta_1),\dots,(b_m,B_m,\beta_m),(b_{m+1},B_{m+1},\beta_{m+1}),\dots,(b_q,B_q,\beta_q)}\right] = \frac{1}{2\pi i}\int_L \frac{\prod_{j=1}^{m}\Gamma^{\beta_j}(b_j-B_js)\prod_{j=1}^{n}\Gamma^{\alpha_j}(1-a_j+A_js)}{\prod_{j=m+1}^{q}\Gamma^{\beta_j}(1-b_j+B_js)\prod_{j=n+1}^{p}\Gamma^{\alpha_j}(a_j-A_js)}x^s ds, \quad (3)$$

where α_j, $j=1,\dots,p$, and β_j, $j=1,\dots,q$, are positive quantities. For more details about I-function, see Rathie [11].

For completeness, we give below a few results which will be used later on:

$$G^{1,1}_{1,1}\left[x\Big|^{1-a}_{0}\right] = \Gamma(a)(1+x)^{-a}. \quad (4)$$

$$\exp(x) = \sum_{r=0}^{\infty}\frac{x^r}{r!}. \quad (5)$$

$$(1+x)^{-2} = \sum_{r=0}^{\infty}(-1)^r(1+r)x^r.\ |x|<1. \quad (6)$$

$$\int_0^{\infty} x^{\alpha-1}\exp(-\beta x^{\gamma})dx = \frac{\Gamma\left(\frac{\alpha}{\gamma}\right)}{\gamma\beta^{\frac{\alpha}{\gamma}}}, \quad \alpha,\beta,\gamma>0. \quad (7)$$

$${}_2F_1(a,b;c;x) = \sum_{r=0}^{\infty}\frac{(a)_r(b)_r}{r!(c)_r}x^r, \text{ for } |x|<1. \quad (8)$$

$${}_2F_1(a,b;c;1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}, \quad (9)$$

for c not a negative integer or zero and $\text{Re}(c-a-b)>0$.

Using Mathai et al. [6],

$$I(\alpha,\beta,\gamma,\delta) = \int_0^{\infty} x^{\alpha}\exp[-x(\beta+\gamma x^{\delta})]dx = \beta^{-\alpha-1}H^{1,1}_{1,1}\left[\frac{\beta^{\delta+1}}{\gamma}\Big|^{(1,1)}_{(\alpha+1,\gamma+1)}\right], \quad (10)$$

and

$$\int_0^\infty x^{s-1} H_{p_1,q_1}^{m_1,n_1}\left[\eta x\left|\begin{matrix} {}_1(d_j,D_j)_{p_1} \\ {}_1(e_j,E_j)_{q_1}\end{matrix}\right.\right] H_{p,q}^{m,n}\left[z x^{\sigma}\left|\begin{matrix} {}_1(a_j,A_j)_{p} \\ {}_1(b_j,B_j)_{q}\end{matrix}\right.\right] dx$$
$$= \eta^{-s} H_{p+q,\,q+p_1}^{m+n,\,n+m_1}\left[z\eta^{-\sigma}\left|\begin{matrix} {}_1(a_j,A_j)_n,\ {}_1(1-e_j-sE_j,\sigma E_j)_{q_1},\ {}_{(n+1)}(a_j,A_j)_p \\ {}_1(b_j,B_j)_m,\ {}_1(1-d_j-sD_j,\sigma D_j)_{p_1},\ {}_{(m+1)}(b_j,B_j)_q\end{matrix}\right.\right]. \quad (11)$$

For conditions of existence etc., see Mathai et al. [6].

3. Marshall-Olkin-Rathie-Swamee distribution

In 2006, Rathie and Swamee [14] defined a multimodal distribution, for a random variable $X \sim \mathrm{RS}(a,b,p)$, as

$$F_0(x) = \frac{1}{1+\exp[-x(a+b|x|^p)]}, \quad (12)$$

with the corresponding density function given by

$$f_0(x) = \frac{[a+b(p+1)|x|^p]\exp[-x(a+b|x|^p)]}{\{1+\exp[-x(a+b|x|^p)]\}^2}, \quad (13)$$

for $x \in (-\infty,\infty)$, $a,b \geq 0$ (both a and b are not zeros simultaneously) and $p \geq -1$. Clearly, the above distribution is a generalization of the well-known logistic distribution. Also, for certain values of the parameters a, b and p, it approximates well the normal distribution [12].

Applying the Marshall and Olkin [5] expression

$$F(x) = \frac{F_0(x)}{F_0(x)+\alpha\overline{F}_0(x)}, \alpha > 0, x \in \mathbb{R}, \quad (14)$$

for $F_0(x)$ given in (12), the following tilted generalized logistic distribution is generated:

$$F(x) = \frac{1}{1+\alpha\exp[-x(a+b|x|^p)]}, \quad (15)$$

where the parameter α may be regarded as a tilt parameter.

For various values of a, b, p and α, the density function

$$f(x) = \frac{\alpha[a+b(p+1)|x|^p]\exp[-x(a+b|x|^p)]}{\{1+\alpha\exp[-x(a+b|x|^p)]\}^2}, \quad (16)$$

is plotted in Figure 1 showing multimodality of the distribution. We write $X \sim \mathrm{MORS}(\alpha,a,b,p)$ to indicate that X follows (15) and (16). We may also call X

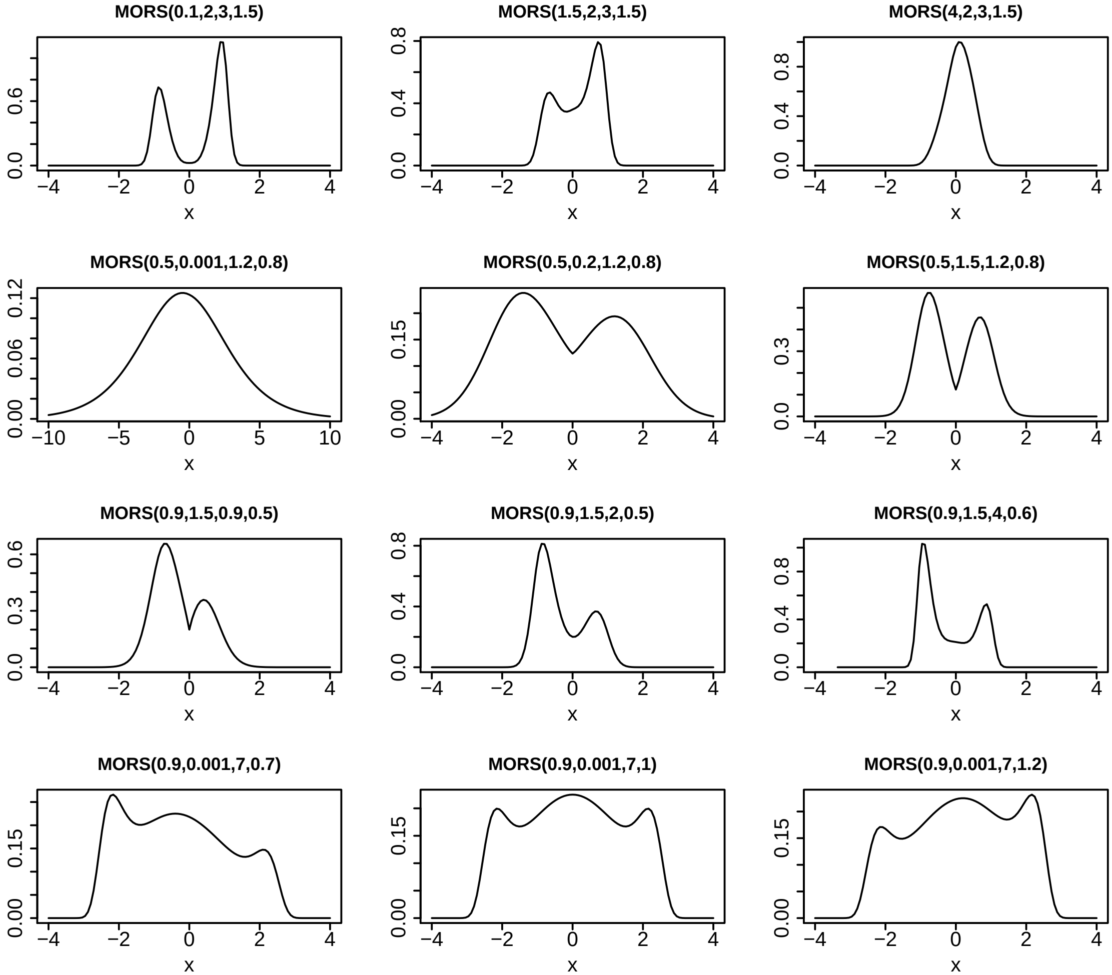

Figure 1: Some shapes for the MORS density.

having Marshall-Olkin-Rathie-Swamee (MORS) distribution. For $\alpha = 1$, (15) and (16) yield (12) and (13) respectively.

Using a location parameter $\mu \in (-\infty,\ \infty)$, (16) may be changed to $g(x)$ as

$$g(x) = \frac{\alpha[a + b(p+1)|x-\mu|^p]\exp[-(x-\mu)(a+b|x-\mu|^p)]}{\{1+\alpha\exp[-(x-\mu)(a+b|x-\mu|^p)]\}^2}. \tag{17}$$

Notationally, (17) may be denoted by $X \sim \text{MORS}(\alpha, a, b, p, \mu)$. There is no need to introduce a scale parameter σ, otherwise the density function will become non-identifiable. In that case, $\frac{a}{\sigma}$ and $\frac{b}{\sigma^{p+1}}$ may be changed to parameters A and B respectively to avoid non-identifiability.

We may observe that a and α control the peak at $x = \mu$, the location parameter. The shape parameters are a, b and p. Also, a and b have scale parameter included non-explicitely. Roughly, b may affect the number of modes (for example, $b = 0$ implies unimodality) in conjunction with p; p separates the modes; and α introduces skewness (tilt). The maximum number of three modes has been observed but there is no mathematical proof so far.

The hazard rate function is defined by

$$h(x) = \frac{f(x)}{1-F(x)},$$

which is an important concept for applications in life phenomena. In our case, it is given by

$$h(x) = \frac{[a + b(p+1)|x|^p]}{\{1+\alpha\exp[-x(a+b|x|^p)]\}},\ x \neq 0.$$

It is interesting to note that

$$\frac{\partial h(x)}{\partial \alpha} = -\frac{f(x)}{\alpha}.$$

Shapes of the hazard rate function are shown in Figure 2 for certain values of the parameters.

3.1. Order statistics

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from cumulative distribution $F(.)$. The cumulative distribution and density functions of the n-th order statistics are given by

$$F_n(x) = [F(x)]^n = \frac{1}{\{1+\alpha\exp[-x(a+b|x|^p)]\}^n} \tag{18}$$

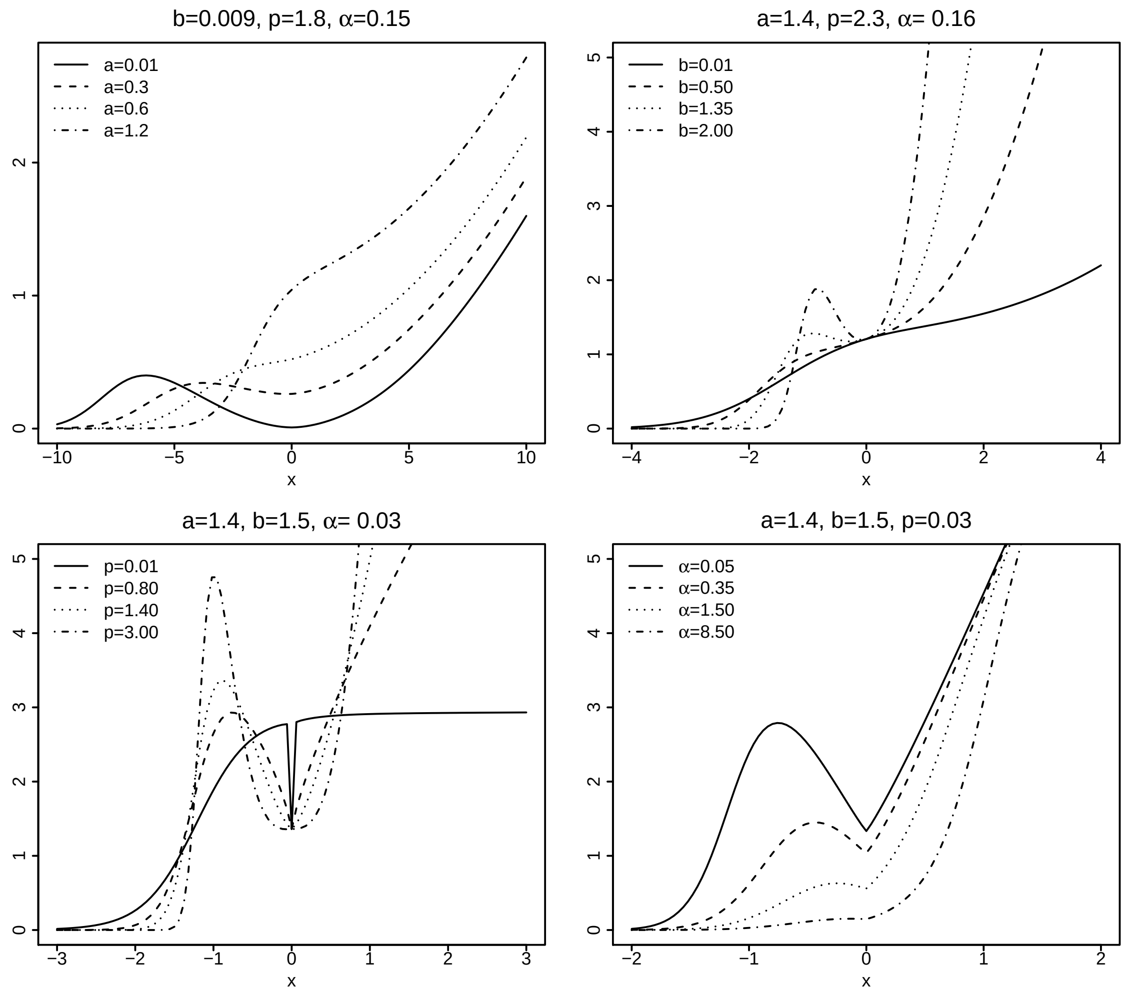

Figure 2: Some shapes for the hazard rate function.

and

$$f_n(x) = n[F(x)]^{n-1} f(x) = \frac{n\alpha[a + b(p+1)|x|^p] \exp[-x(a + b|x|^p)]}{\{1 + \alpha \exp[-x(a + b|x|^p)]\}^{n+1}} \tag{19}$$

respectively. For the 1st order statistics the cumulative distribution and density functions are given by

$$\begin{aligned} F_1(x) &= 1 - [1 - F(x)]^n = 1 - \left[\frac{\alpha \exp[-x(a + b|x|^p)]}{1 + \alpha \exp[-x(a + b|x|^p)]}\right]^n \\ &= 1 - \frac{\alpha^n \exp[-nx(a + b|x|^p)]}{\{1 + \alpha \exp[-x(a + b|x|^p)]\}^n} \end{aligned} \tag{20}$$

and

$$f_1(x) = n[1 - F(x)]^{n-1} f(x) = \frac{n\alpha^n \exp[-nx(a + b|x|^p)][a + b(p+1)|x|^p]}{\{1 + \alpha \exp[-x(a + b|x|^p)]\}^{n+1}} \tag{21}$$

respectively.

4. Moments

The n-th moment about the origin are giving by

$$\begin{aligned} E(X^n) &= \alpha \int_{-\infty}^{\infty} \frac{x^n[a + b(p+1)|x|^p] \exp[-x(a + b|x|^p)]}{\{1 + \alpha \exp[-x(a + b|x|^p)]\}^2} dx \\ &= \alpha \int_0^{\infty} \frac{x^n[a + b(p+1)x^p] \exp[-x(a + bx^p)]}{\{1 + \alpha \exp[-x(a + bx^p)]\}^2} dx \\ &\quad + (-1)^n \alpha \int_0^{\infty} \frac{x^n[a + b(p+1)x^p] \exp[x(a + bx^p)]}{\{1 + \alpha \exp[x(a + bx^p)]\}^2} dx \\ &= \alpha \int_0^{\infty} x^n g_\alpha(x) dx + (-1)^n \frac{1}{\alpha} \int_0^{\infty} x^n g_{\frac{1}{\alpha}}(x) dx, \end{aligned}$$

where

$$g_\alpha(x) = \frac{x^n[a + b(p+1)x^p] \exp[-x(a + bx^p)]}{\{1 + \alpha \exp[-x(a + bx^p)]\}^2}. \tag{22}$$

Let

$$J_\alpha = \int_0^{\infty} x^n g_\alpha(x) dx. \tag{23}$$

Then

$$
\begin{aligned}
J_\alpha &= \int_0^\infty x^n[a+b(p+1)x^p]\exp[-x(a+bx^p)]G_{1,1}^{1,1}\left[\alpha\exp[-x(a+bx^p)]\Big|_0^{-1}\right]dx \\
&= \int_0^\infty x^n[a+b(p+1)x^p]\exp[-x(a+bx^p)] \\
&\quad \times\left\{\frac{1}{2\pi i}\int_L \Gamma(-s)\Gamma(2+s)\{\alpha\exp[-x(a+bx^p)]\}^s ds\right\}dx, (\text{Real}(s)>0) \\
&\quad (\text{using } (1)) \\
&= \frac{1}{2\pi i}\int_L \Gamma(-s)\Gamma(2+s)\alpha^s \\
&\quad \times\left\{\int_0^\infty x^n[a+b(p+1)x^p]\exp[-(1+s)x(a+bx^p)]dx\right\}ds
\end{aligned}
\tag{24}
$$

$$
\begin{aligned}
&= \sum_{r=0}^\infty \frac{(-b)^r}{r!}\frac{1}{2\pi i}\int_L \Gamma(-s)\Gamma(2+s)\alpha^s(1+s)^r \\
&\quad \times\left(\int_0^\infty x^n[a+b(p+1)x^p]x^{(p+1)r}\exp[-(1+s)ax]dx\right)ds \\
&\quad (\text{using } (5)) \\
&= \sum_{r=0}^\infty \frac{(-b)^r}{r!}\frac{1}{2\pi i}\int_L \Gamma(-s)\Gamma(2+s)\alpha^s(1+s)^r \\
&\quad \left\{a\frac{\Gamma[n+(p+1)r+1]}{[a(1+s)]^{n+(p+1)r+1}}+b(p+1)\frac{\Gamma[n+(p+1)r+p+1]}{[a(1+s)]^{n+(p+1)r+p+1}})\right\}ds \\
&\quad (\text{using } (7)) \\
&= \sum_{r=0}^\infty \frac{(-b)^r}{r!}\left\{\frac{\Gamma[n+(p+1)r+1]}{a^{n+(p+1)r}}I_1+b(p+1)\frac{\Gamma[n+(p+1)(r+1)]}{a^{n+(p+1)(r+1)}}I_{p+1}\right\},
\end{aligned}
\tag{25}
$$

where

$$
\begin{aligned}
I_m &= \frac{1}{2\pi i}\int_L \frac{\Gamma(-s)\Gamma(2+s)\alpha^s}{(1+s)^{n+pr+m}}ds \\
&= \frac{1}{2\pi i}\int_L \frac{\Gamma(-s)[\Gamma(1+s)]^{n+pr+m}\alpha^s}{[\Gamma(2+s)]^{n+pr+m-1}}ds \\
&= G_{n+pr+m,\,n+pr+m}^{1,\,n+pr+m}\left[\alpha\Big|_{0,-1,\ldots,-1}^{0,\ldots,0}\right].(\text{using } (1))
\end{aligned}
$$

Hence

$$E(X^n) = \alpha J_\alpha + \frac{(-1)^n}{\alpha} J_{\frac{1}{\alpha}}. \tag{26}$$

For $a = 0$ or $b = 0$, the n-th moments are given by (26) where J_α is calculated below for the two particular cases:

(i) For $a = 0$, in (24), we have

$$\begin{aligned} J_\alpha &= b(p+1)\frac{1}{2\pi i}\int_L \Gamma(-s)\Gamma(2+s)\alpha^s \left\{\int_0^\infty x^{n+p}\exp[-(1+s)bx^{p+1}]dx\right\} ds \\ &= b(p+1)\frac{1}{2\pi i}\int_L \Gamma(-s)\Gamma(2+s)\alpha^s \frac{\Gamma(\frac{n+p+1}{p+1})}{(p+1)[b(1+s)]^{\frac{n+p+1}{p+1}}} ds \\ &\quad \text{(using (7))} \\ &= \frac{\Gamma(\frac{n}{p+1}+1)}{b^{\frac{n}{p+1}}}\frac{1}{2\pi i}\int_L \Gamma(-s)\frac{\Gamma^{\frac{n}{p+1}+1}(1+s)}{\Gamma^{\frac{n}{p+1}}(2+s)}\alpha^s ds \\ &= \frac{\Gamma(\frac{n}{p+1}+1)}{b^{\frac{n}{p+1}}} I^{1,1}_{1,2}\left[\alpha\middle|^{(0,1,\frac{n}{p+1}+1)}_{(0,1,1),(-1,1,\frac{n}{p+1})}\right]. \\ &\quad \text{(using(3))} \end{aligned} \tag{27}$$

For $\frac{n}{p+1}$ an integer, the I-function reduces to a H-function which can also be written as a G-function.

(ii) For $b = 0$, in (24), we get

$$\begin{aligned} J_\alpha &= \frac{1}{2\pi i}\int_L \Gamma(-s)\Gamma(2+s)\alpha^s \left\{\int_0^\infty ax^n \exp[-(1+s)ax]dx\right\} ds \\ &= \frac{a}{2\pi i}\int_L \Gamma(-s)\Gamma(2+s)\alpha^s \frac{\Gamma(n+1)}{[(1+s)a]^{n+1}} ds \\ &\quad \text{(using (7))} \\ &= \frac{n!}{a^n}\frac{1}{2\pi i}\int_L \Gamma(-s)\frac{\alpha^s\Gamma^{n+1}(1+s)}{\Gamma^n(2+s)} ds \\ &= \frac{n!}{a^n} G^{1,n+1}_{n+1,n+1}\left[\alpha\middle|^{0,\ldots,0}_{0,-1,\ldots,-1}\right]. \\ &\quad \text{(using(1))} \end{aligned} \tag{28}$$

5. Reliability

From Andrade and Rathie [1], for any base distribution $F(x)$, we have

$$\begin{aligned} F_{\text{asy}}(x) &= \frac{1}{B(\alpha_1,\beta_1)}\int_0^{F(x)} u^{\alpha_1-1}(1-u)^{\beta_1-1}du \\ &= \frac{F^{\alpha_1}(x)}{\alpha_1 B(\alpha_1,\beta_1)}\ {}_2F_1(\alpha_1, 1-\beta_1; 1+\alpha_1; F(x)) \qquad (29) \\ &= \frac{1}{\alpha_1 B(\alpha_1,\beta_1)}\sum_{r=0}^{\infty}\frac{(\alpha_1)_r(1-\beta_1)_r}{r!(1+\alpha_1)_r}F^{\alpha_1+r}(x). \qquad (30) \\ &\text{(using (8))} \end{aligned}$$

From (30), we conclude that $F_{\text{asy}}(x)$ is an infinite linear combination (or mixture) of the generalized MORS distributions with parameters a, b, p, α and α_1+r whose distribution function is

$$[F(x)]^{\alpha_1+r} = \frac{1}{\{1+\alpha\exp[-x(a+b|x|^p)]\}^{\alpha_1+r}}. \qquad (31)$$

The corresponding density function is given by

$$\begin{aligned} f_{\text{asy}}(x) &= \frac{f(x)[F(x)]^{\alpha_1-1}[1-F(x)]^{\beta_1-1}}{B(\alpha_1,\beta_1)} \qquad (32) \\ &= \frac{\alpha_1^{\beta}[a+b(p+1)|x|^p]\exp[-\beta_1 x(a+b|x|^p)]}{B(\alpha_1,\beta_1)\{1+\alpha\exp[-x(a+b|x|^p)]\}^{\alpha_1+\beta_1}}, \qquad (33) \end{aligned}$$

when (15) and (16) are used.

Let the density functions $f_{\text{asy}}(x)$ of X and $f(y)$ of Y, then the reliability $P(X<Y)$, when X and Y are independent, is given by

$$\begin{aligned} P(X<Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{y} f(y)f_{\text{asy}}(x)dxdy \\ &= \int_{-\infty}^{\infty} f(y)F_{\text{asy}}(y)dy. \end{aligned}$$

For $f(y)=f_0(y)$ given by (13) and $F(y)$ given by (15), where $Y\sim$ MORS $(\alpha_1, a_1, b_1, p_1)$, we have

$$P(X<Y) = \frac{1}{\alpha_1 B(\alpha_1,\beta_1)} \sum_{r=0}^{\infty} \frac{(\alpha_1)_r(1-\beta_1)_r}{r!(1+\alpha_1)_r} \int_{-\infty}^{\infty} f_0(y) F^{\alpha_1+r}(y)dy.$$

Case 1: When $a=a_1$, $b=b_1$, $p=p_1$, we have

$$P(X<Y) = \frac{1}{\alpha_1 B(\alpha_1,\beta_1)} \sum_{r=0}^{\infty} \frac{(\alpha_1)_r(1-\beta_1)_r}{r!(1+\alpha_1)_r} I_1,$$

where

$$\begin{aligned} I_1 &= \int_{-\infty}^{\infty} f_0(y) F^{\alpha_1+r}(y)dy \\ &= \int_{-\infty}^{\infty} \frac{[a+b(p+1)|y|^p]\exp[-y(a+b|y|^p)]}{\{1+\exp[-y(a+b|y|^p)]\}^2\{1+\alpha\exp[-y(a+b|y|^p)]\}^{\alpha_1+r}}dy \quad (34) \\ &= \int_0^{\infty} \frac{dz}{(1+z)^2(1+\alpha z)^{\alpha_1+r}} \quad (35) \\ &\quad \text{(on substituting } \exp[-y(a+b|y|^p)]=z) \\ &= \frac{1}{\Gamma(\alpha_1+r)} \int_0^{\infty} G_{1,1}^{1,1}\left[z\middle|{-1 \atop 0}\right] G_{1,1}^{1,1}\left[\alpha z\middle|{1-\alpha_1-r \atop 0}\right] dz \\ &\quad \text{(using(4))} \\ &= \frac{1}{\Gamma(\alpha_1+r)} G_{2,2}^{2,2}\left[\alpha\middle|{0,1-\alpha_1-r \atop 1,0}\right]. \\ &\quad \text{(using(11))} \end{aligned}$$

For various conditions of validity, see Luke [4, pp. 162-164].

In this case,

$$\begin{aligned} P(X<Y) &= \frac{1}{\alpha_1 B(\alpha_1,\beta_1)} \sum_{r=0}^{\infty} \frac{(\alpha_1)_r(1-\beta_1)_r}{r!(1+\alpha_1)_r\Gamma(\alpha_1+r)} G_{2,2}^{2,2}\left[\alpha\middle|{0,1-\alpha_1-r \atop 0,1}\right] \\ &= \frac{1}{B(\alpha_1,\beta_1)} \sum_{r=0}^{\infty} \frac{(1-\beta_1)_r}{r!\Gamma(1+\alpha_1+r)} G_{2,2}^{2,2}\left[\alpha\middle|{0,1-\alpha_1-r \atop 0,1}\right]. \end{aligned}$$

It may observed that the reliability expression does not contain a, b and p.

For $\alpha=1$, (35) gives

$$I_1 = \int_0^{\infty} \frac{dz}{(1+z)^{2+\alpha_1+r}} = \frac{1}{1+\alpha_1+r}.$$

Thus

$$\begin{aligned} P(X<Y) &= \frac{1}{\alpha_1 B(\alpha_1,\beta_1)}\sum_{r=0}^{\infty}\frac{(\alpha_1)_r(1-\beta_1)_r}{r!(1+\alpha_1)_r(1+\alpha_1+r)} \\ &= \frac{\Gamma(\alpha_1)}{\Gamma(\alpha_1+2)B(\alpha_1,\beta_1)}\sum_{r=0}^{\infty}\frac{(\alpha_1)_r(1-\beta_1)_r}{r!(\alpha_1+2)_r} \\ &= \frac{\Gamma(\alpha_1+\beta_1)}{\Gamma(\alpha_1+2)\Gamma(\beta_1)}\ {}_2F_1(\alpha_1,1-\beta_1;2+\alpha_1;1) \\ &= \frac{\beta_1}{\alpha_1+\beta_1}. \qquad (36) \\ &\quad \text{(using (9))} \end{aligned}$$

Case 2: In the general case, we have

$$P(X<Y) = \frac{1}{\alpha_1 B(\alpha_1,\beta_1)}\sum_{r=0}^{\infty}\frac{(\alpha_1)_r(1-\beta_1)_r}{r!(1+\alpha_1)_r} I, \qquad (37)$$

where

$$\begin{aligned} I &= \int_{-\infty}^{\infty} f_0(y)F^{\alpha_1+r}(y)dy \\ &= \int_{-\infty}^{\infty}\frac{[a+b(p+1)|y|^p]\exp[-y(a+b|y|^p)]}{\{1+\exp[-y(a+b|y|^p)]\}^2\{1+\alpha\exp[-y(a_1+b_1|y|^{p_1})]\}^{\alpha_1+r}}dy. \qquad (38) \end{aligned}$$

As the above integral is not easy to evaluate, we consider two interesting particular cases:

Subcase 2.1: When $a=0=a_1$, $p=p_1$ and denoting I by I_1, we have

$$\begin{aligned} I_1 &= \int_{-\infty}^{\infty}\frac{b(p+1)|y|^p\exp(-by|y|^p)}{[1+\exp(-by|y|^p)]^2[1+\alpha\exp(-b_1y|y|^p)]^{\alpha_1+r}}dy \\ &= \int_0^{\infty}\frac{dz}{(1+z)^2(1+\alpha z^{\frac{b_1}{b}})^{\alpha_1+r}}, \end{aligned}$$

where

$$\exp(-by|y|^p) = z.$$

Thus

$$P(X<Y)=\frac{1}{\alpha_1 B(\alpha_1,\beta_1)}\sum_{r=0}^{\infty}\frac{(\alpha_1)_r(1-\beta_1)_r}{r!(1+\alpha_1)_r}\frac{1}{\Gamma(\alpha_1+r)}H_{2,2}^{2,2}\left[\alpha\middle|\begin{matrix}(1-\alpha_1-r,1),(0,\frac{b_1}{b})\\(0,1),(1,\frac{b_1}{b})\end{matrix}\right]. \tag{39}$$

(using (4) and (11))

Subcase 2.2: When $b=0=b_1$ and denoting I by I_2, we get

$$\begin{aligned}I_2 &= \int_{-\infty}^{\infty}\frac{a\exp(-ay)}{[1+\exp(-ay)]^2[1+\alpha\exp(-a_1y)]^{\alpha_1+r}}dy\\ &= \int_0^{\infty}\frac{dz}{(1+z)^2(1+\alpha z^{\frac{a_1}{a}})^{\alpha_1+r}}\\ &\quad \text{(substituting } \exp(-ay)=z)\\ &= \frac{1}{\Gamma(\alpha_1+r)}\int_0^{\infty}G_{1,1}^{1,1}\left[z\middle|\begin{matrix}-1\\0\end{matrix}\right]G_{1,1}^{1,1}\left[\alpha z^{\frac{a_1}{a}}\middle|\begin{matrix}1-\alpha_1-r\\0\end{matrix}\right]dz.\\ &\quad \text{(using (4))}\end{aligned}$$

Thus

$$P(X<Y)=\frac{1}{\alpha_1 B(\alpha_1,\beta_1)}\sum_{r=0}^{\infty}\frac{(\alpha_1)_r(1-\beta_1)_r}{r!(1+\alpha_1)_r}\frac{1}{\Gamma(\alpha_1+r)}H_{2,2}^{2,2}\left[\alpha\middle|\begin{matrix}(1-\alpha_1-r,1),(0,\frac{a_1}{a})\\(0,1),(1,\frac{a_1}{a})\end{matrix}\right]. \tag{40}$$

(using (11))

6. MOEW distribution

The corresponding distribution and density functions for $X>0$ are given by

$$F(x)=\frac{1-\exp[-x(a+bx^p)]}{1+\alpha\exp[-x(a+bx^p)]} \tag{41}$$

and

$$f(x)=\frac{(1+\alpha)[a+b(p+1)x^p]\exp[-x(a+bx^p)]}{\{1+\alpha\exp[-x(a+bx^p)]\}^2}, \tag{42}$$

respectively, where $x>0$, $a\geq 0, b\geq 0$ (both not zero simultaneously) and $p\geq -1$.

This distribution, under the name Marshall-Olkin exponential Weibull (MOEW) distribution, was studied recently by Pogány et al. [9]. They obtained the n-th moments for $|\alpha| < 1$. We derive the n-th moments for $|\alpha| < 1$ in terms of an infinite series involving H-function. Also, we give the n-th moments for $\alpha > 1$, extending their result.

6.1 Moments

Case 1: For $|\alpha| < 1$, we have

$$\begin{aligned} E(X^n) &= (1+\alpha)\int_0^\infty \frac{x^n[a+b(p+1)x^p]\exp[-x(a+bx^p)]}{\{1+\alpha\exp[-x(a+bx^p)]\}^2}dx \\ &= (1+\alpha)\sum_{r=0}^{\infty}(-1)^r(1+r)\alpha^r\int_0^\infty x^n[a+b(p+1)x^p] \\ &\quad \times \exp[-(1+r)x(a+bx^p)]dx \quad (\text{using}(6)) \\ &= (1+\alpha)\sum_{r=0}^{\infty}(-1)^r(1+r)\alpha^r\big[aI(n,(1+r)a,(1+r)b,p) \\ &\quad +\ b(p+1)I(n+p,(1+r)a,(1+r)b,p)\big]. \\ &\quad (\text{using}(10)) \end{aligned} \tag{43}$$

Case 2: For $\alpha > 1$, we have

$$\begin{aligned} E(X^n) &= (1+\alpha)\int_0^\infty \frac{x^n[a+b(p+1)x^p]\exp[-x(a+bx^p)]}{\{1+\alpha\exp[-x(a+bx^p)]\}^2}dx \\ &= (1+\alpha)\int_0^\infty x^n[a+b(p+1)x^p]\exp[-x(a+bx^p)] \\ &\quad \times G_{1,1}^{1,1}\left[\alpha\exp[-x(a+bx^p)]\,\middle|\,\begin{matrix}-1\\0\end{matrix}\right]dx \quad (\text{using}(1)) \\ &= (1+\alpha)J_\alpha, \end{aligned} \tag{44}$$

where J_α is given in (25).

6.2 Beta-generated MOEW distribution (BMOEW distribution)

Using (41) and (42) in (29) and (32) respectively, we get the BMOEW distribution as

$$G(x) = \frac{F^{\alpha_1}(x)}{\alpha_1 B(\alpha_1,\beta_1)}\ {}_2F_1(\alpha_1, 1-\beta_1; 1+\alpha_1; F(x)),\ x>0, \tag{45}$$

and

$$g(x) = \frac{(1+\alpha)^{\beta_1}[a+b(p+1)x^p]\exp[-\beta_1 x(a+bx^p)]\{1-\exp[-x(a+bx^p)]\}^{\alpha_1-1}}{B(\alpha_1,\beta_1)\{1+\alpha\exp[-x(a+bx^p)]\}^{\alpha_1+\beta_1}}, \tag{46}$$

$x > 0$.
For $\alpha_1 = 1$, $b = 0$, we have

$$g(x) = \frac{\beta_1 a(1+\alpha)^{\beta_1} \exp(-\beta_1 a x)}{[1+\alpha\exp(-ax)]^{1+\beta_1}},\ x > 0. \tag{47}$$

This is the HEE distribution [3] studied by Pinho et al. [8].

7. Parameter estimation and applications

The parameter vector $\boldsymbol{\theta} = (a,\, b,\, p,\, \mu,\, \alpha)^\top$ of (17) can be estimated by the method of maximum likelihood. The likelihood is given by $L(\boldsymbol{\theta}) = \prod_{i=1}^n g(x_i)$, where $g(x_i)$ is given in (17). Thus,

$$\begin{aligned} L(\boldsymbol{\theta}) &= \prod_{i=1}^n \frac{\alpha[a+b(p+1)|x_i-\mu|^p]\exp[-(x_i-\mu)(a+b|x_i-\mu|^p)]}{\{1+\alpha\exp[-(x_i-\mu)(a+b|x_i-\mu|^p)]\}^2} \\ &= \frac{\alpha^n \prod_{i=1}^n [a+b(p+1)|x_i-\mu|^p] \exp\{-\sum_{i=1}^n [(x_i-\mu)(a+b|x_i-\mu|^p)]\}}{\prod_{i=1}^n \{1+\alpha\exp[-(x_i-\mu)(a+b|x_i-\mu|^p)]\}^2} \end{aligned} \tag{48}$$

and the log-likelihood function is given by

$$\begin{aligned} l(\boldsymbol{\theta}) &= n\ln(\alpha) + \sum_{i=1}^n \ln[a+b(p+1)|x_i-\mu|^p] - \sum_{i=1}^n [(x_i-\mu)(a+b|x_i-\mu|^p)] \\ &\quad -2\sum_{i=1}^n \ln\{1+\alpha\exp[-(x_i-\mu)(a+b|x_i-\mu|^p)]\}. \end{aligned} \tag{49}$$

The estimates for $\hat{\alpha}$, $\hat{a}$, $\hat{b}$, $\hat{p}$ and $\hat{\mu}$ are obtained by solving simultaneously the five equations obtained by equating to zero the partial derivatives of right hand side of the (49) with respect to α, a, b, p and μ respectively. These equations cannot be solved analytically and R program [10] has been used to solve them numerically. We use the `constrOptim` routine to execute the maximization process. This routine maximize the log-likelihood function subject to linear inequality constraints using an adaptive barrier algorithm. In our experience, the convergence of the maximization procedure involving multimodal distributions (generalized flexible distributions) almost alway depends on the initial guesses. We used multiple starts and the `GenSA` routine that searches for global minimum (or maximum) of a function using simulated annealing.

Four applications of the MORS model are presented for the following real data sets:

1. Old Faithful Geyser data: waiting times (in mins) to the next eruption and durations of the eruptions (in mins) available as a built-in data set in R program [10].

2. Shrimp weights data from a study on morphometrical and molecular levels of genetic variability in shrimp from the coast of Rio Grande do Norte, Brazil (see [7]).

3. Environmental performance index (EPI) of 2012 countries from `http://epi.yale.edu/`.

The estimates of the parameters for the MORS(α, a, b, p, μ) through maximum likelihood method are obtained using `constrOptim` function of the R program. The results are presented in Table 1. In Figure 3, we can see the fits of the

Table 1: Maximum likelihood estimates for MORS model.

Data set (sample size)	α	a	b	p	μ
Waiting time (272)	1.7279	0.0265	0.0013	1.5105	66.4339
Eruption duration (272)	1.5191	0.2286	0.4674	2.9833	3.1819
Shrimp weight (120)	0.8105	0.0178	0.0353	0.7091	16.5993
EPI (132)	1.7533	0.1602	0.0000	3.9495	49.8589

MORS(α, a, b, p, μ) density and distribution functions for histogram and empirical cumulative distribution functions (ecfd) for each of the data sets. From these plots, we can say that the MORS model fits the data adequately. We apply Kolmogorov-Smirnov (KS), Anderson-Darling (AD) and Cramér-von-Mises (CvM) tests to assess the goodness of fit of the model. In general, the smaller the values of KS, AD and CvM, the better the fit to the data. Table 2 gives the p-value for the three tests. The tests are described in [16].

Table 2: p-values for KS, AD and CvM test.

	Data set			
Test	Waiting time	Eruption duration	Shrimp weight	EPI
KS	0.864	0.549	0.999	0.963
AD	0.888	0.293	0.980	0.963
CvM	0.895	0.551	0.998	0.998

All the test showed that we do not reject the hypothesis to fit the data using MORS distribution.

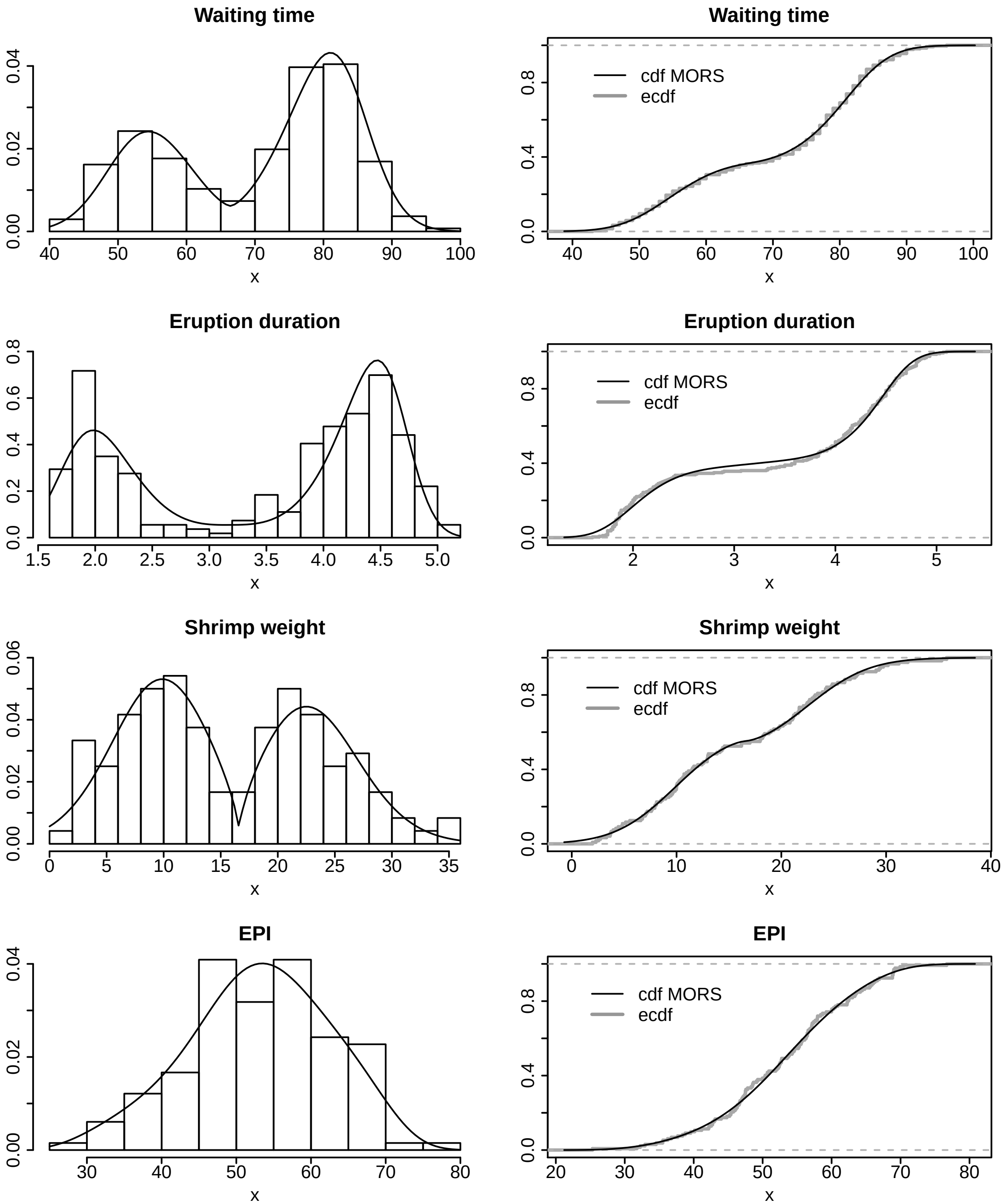

Figure 3: Adjusted MORS distribution for the four data sets: with histograms and with the empirical distributions.

Now, we compare the MORS model with the models presented by Andrade and Rathie [1] for each data set. For EPI data the bimodal asymmetric power normal (BAPN) model and for the other data sets the Rathie-Swamee asymmetric (RSA) model are used. For these two models, only BIC and MSE were given in [1] as indicated in Table 3.

1. Bimodal asymmetric power normal (BAPN) model [2]: the recentred-rescaled BAPN model has density given by

$$f(z) = 2\gamma\sigma^{-1}\left(\frac{2^{\gamma-1}}{2^{\gamma}-1}\right)\phi(z)\Phi(|z|)^{\gamma-1}\Phi(\lambda z),$$

   for $z = (x-\mu)/\sigma$, $x \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\sigma > 0$, $\gamma > 0$, $\lambda \in \mathbb{R}$, where Φ and ϕ respectively are cumulative distribution and density of the standard normal.

2. Rathie-Swamee asymmetric (RSA) model [13]: the recentred RSA model has density given by

$$f(y) = \frac{2[a+b(p+1)|y|^p]\exp[-y(a+b|y|^p)]}{\{1+\exp[-y(a+b|y|^p)]\}^2\{1+\exp[-\lambda y(a+b|\lambda y|^p)]\}},$$

   for $y = x-\mu$, $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ and a, b, p all non-negative but not all zeroes simultaneously.

We used the Akaike Information Criterion (AIC), corrected Akaike Information Criterion (AICc) and Bayesian Information Criterion (BIC). These criteria are defined by

$$\begin{aligned}
\text{AIC} &= -2\log(f(\mathbf{x}|\boldsymbol{\theta})) + 2p,\\
\text{BIC} &= -2\log(f(\mathbf{x}|\boldsymbol{\theta})) + p\log(n),\\
\text{AICc} &= -2\log(f(\mathbf{x}|\boldsymbol{\theta})) + 2p + \left[\frac{2p(p+1)}{n-p-1}\right],
\end{aligned}$$

where $\log(f(\mathbf{x}|\boldsymbol{\theta}))$ is the log-likelihood function, p the number of parameters of the model and n the sample size. The best model is one that has least value according to the criterion applied. The Table 3 show all these measures. In [1], the BAPN and RSA models have the least values of BIC.

We also apply the measures of accuracy, Mean Square Error (MSE), Mean Absolute Deviation (MAD) and Maximum Absolute Deviation (MaxAD), which are given by

$$\begin{aligned} \text{MSE} &= \frac{\sum_{i=1}^{n} (F_e(x_i) - \hat{F}(x_i))^2}{n}, \\ \text{MAD} &= \frac{\sum_{i=1}^{n} |F_e(x_i) - \hat{F}(x_i)|}{n}, \\ \text{MaxAD} &= \max(|F_e(x_i) - \hat{F}(x_i)|), \end{aligned}$$

where $F_e(x_i)$, $i = 1, \ldots, n$, is the empirical cumulative distribution and $\hat{F}(x_i)$ the fitted cumulative distribution of the data. The best model has the lowest value (close to zero) according to the criterion used. The results are shown in Table 3. For waiting time data, the BIC values fo MORS and RSA models are close enough

Table 3: Model selection criterion and measures of accuracy.

	Waiting time		Eruption duration		Shrimp weight		EPI	
	MORS	RSA	MORS	RSA	MORS	RSA	MORS	BAPN
AIC	2078.96		568.86		835.56		984.34	
BIC	2096.99	2095	586.89	579	849.49	844	998.76	995
AICc	2079.18		569.09		836.08		984.82	
MSE ($\times 10^3$)	0.42	0.4	0.40	0.9	0.17	0.8	0.30	0.4
MAD ($\times 10^2$)	1.06		1.70		1.06		1.37	
MaxAD ($\times 10^2$)	3.64		4.47		2.74		4.36	

and the MSE values differ at the fifth decimal place. For others data sets, the BIC values for MORS model are slightly higher but the MSE values are lower, especially for shrimp weight data. These results indicate that MORS model is a good alternative for modeling the data.

8. Concluding remarks

The tilted generalized logistic MORS distribution is flexible for statistical analysis of real data. It is multimodal and unifies a few previously proposed distributions, including the logistic distribution. The beta-generated MORS distribution is an infinite linear combination of another distribution. Moments, hazard rate function and order statistics are derived. Maximum likelihood estimate method is used to estimate the parameters. The various result are given in manageable forms which are used in practical applications utilizing the existing numerical capabilities of the computers. The practical data analysis applications demonstrated better fits, in general, compared to existing earlier results.

The density function for the MORS model is very simple as compared to the RSA model which is obtained by applying Azzalini method. We hope that the

MORS distribution will provide wider applicability in general, and in reliability studies, in particular, obtaining the corresponding results for other distributions which are particular cases of our distribution.

Acknowledgement

P. N. Rathie thanks CAPES, Brazil for supporting his Senior National Visiting Professorship.

References

[1] Andrade, B. B. and Rathie, P. N., Fitting asymmetric bimodal data with selected distributions, Journal of Statistical Computation and Simulation, 86(16):3205-3224, 2016. doi: 10.1080/00949655.2016.1157180.

[2] Bolfarine, H., Martínez-Flórez, G., and Salinas, H. S., Bimodal symmetric-asymmetric power-normal families, Communications in Statistics - Theory and Methods, Forthcoming. doi: 10.1080/03610926.2013.765475.

[3] Harris, T. E., Branching processes, The Annals of Mathematical Statistics, 4(19): 474-494, 1948.

[4] Luke, Y. L., The Special Functions and Their Approximations, volume I. New York: Academic Press, 1969.

[5] Marshall, A. W. and Olkin, I., A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families, Biometrika, 84(3):641-652, 1997.

[6] Mathai, A. M., Saxena, R. K., and Haubold, H. J., The H-function: Theory and Applications, Springer, New York, 2010.

[7] Pinheiro, A. Caracterização genética e biometrica das populações de camarão rosa Farfantepenaeus brasiliensis de três localidades da costa do Rio Grande do Norte, PhD thesis, Ecology and Natural Resources, Federal University of São Carlos, SP, Brazil, 2008.

[8] Pinho, L. G. B., Cordeiro, G. M., and Nobre, J. S., The Harris extended exponential distribution, Communications in Statistics - Theory and Methods, 44(16):3486-3502, 2015. doi: 10.1080/03610926.2013.851221.

[9] Pogány, T. K., Saboor, A., and Provost, S., The Marshall-Olkin exponential Weibull distribution, Hacettepe Journal of Mathematics and Statistics, 44(6):1579-1594, 2015.

[10] R Core Team. R, A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2017. URL https://www.R-project.org/.

[11] Rathie, A. K., A new generalization of generalized hypergeometric functions, Le Matematiche, 52(2):297-310, 1997.

[12] Rathie, P. N., Normal distribution, univariate, In Lovric, M., editor, International Encyclopedia of Statistical Science, pages 1012-1013. Springer Berlin Heidelberg, 2011. ISBN 978-3-642-04898-2.

[13] Rathie, P. N. and Coutinho, M., A new skew generalized logistic distribution and approximations to skew normal distribution, Aligarh J Statist, 31:1-12, 2011.

[14] Rathie, P. N. and Swamee, P. K., On a new invertible generalized logistic distribution approximation to normal distribution, Technical report 07, Univ. of Brasilia, Brasilia, Brazil, 2006.

[15] Rathie, P. N., Coutinho, M., Sousa, T. R., Rodrigues, G. S., and Carrijo, T. B., Stable and generalized-t distributions and applications, Communications in Nonlinear Science and Numerical Simulation, 17(12):5088-5096, 2012.

[16] Thas, O., Comparing Distributions, Springer, New York, 2010.