CrossMark

Marshall-Olkin generalized exponential distribution

Miroslav M. Ristić · Debasis Kundu

Abstract Marshall and Olkin (Biometrika 641–652, 1997) introduced a new way of incorporating a parameter to expand a family of distributions. In this paper we adopt the Marshall-Olkin approach to introduce an extra shape parameter to the two-parameter generalized exponential distribution. It is observed that the new three-parameter distribution is very flexible. The probability density functions can be either a decreasing or an unimodal function. The hazard function of the proposed model, can have all the four major shapes, namely increasing, decreasing, bathtub or inverted bathtub types. Different properties of the proposed distribution have been established. The new family of distributions is analytically quite tractable, and it can be used quite effectively, to analyze censored data also. Maximum likelihood method is used to compute the estimators of the unknown parameters. Two data sets have been analyzed, and the results are quite satisfactory.

Keywords Generalized exponential distribution · Hazard function · Probability density function · Maximum likelihood estimator · Fisher information matrix

1 Introduction

Exponential distribution has been used quite effectively to analyze lifetime data, mainly due to its analytical tractability. Although, one parameter exponential distribution has several interesting such as 'lack of memory property, one of the major disadvantages of the exponential distribution is that it has a constant hazard function. Moreover, the probability density func-

M. M. Ristić
Department of Mathematics and Informatics, Faculty of Sciences and Mathematics,
University of Nis, Nis, Serbia
e-mail: miristic@ptt.rs

D. Kundu (✉)
Department of Mathematics and Statistics, Indian Institute of Technology Kanpur,
Kanpur 208016, India
e-mail: kundu@iitk.ac.in

tion (PDF) of the exponential distribution is always a decreasing function. Due to this reason several generalizations of the exponential distribution have been suggested in the literature. For example, Weibull, gamma, generalized exponential (GE) distribution as considered by Gupta and Kundu [11] are different extensions of the exponential distribution, which contain exponential distribution as a special case. All the three distributions can have increasing or unimodal PDFs, and monotone hazard functions. Unfortunately, none of them can have non-monotone hazard functions. In many practical situations, one might observe non-monotone hazard functions, and clearly in those cases, none of these distribution functions can be used.

In the last decade Marshall and Olkin [16] introduced a general method to introduce a shape parameter mainly to expand a family of distributions. They have used their method to the one-parameter exponential distribution and created a two-parameter exponential distribution. They have also indicated to apply their method to the two-parameter Weibull distribution, but did not pursue further.

The main aim of this paper, is to apply the Marshall-Olkin method to the two-parameter generalized exponential distribution. In this paper, we introduce a new distribution function for $\alpha > 0, \lambda > 0, \theta > 0$,

$$G(x; \lambda, \alpha, \theta) = \frac{(1 - e^{-\lambda x})^\alpha}{\theta + (1 - \theta)(1 - e^{-\lambda x})^\alpha}, \quad \text{for} \quad x > 0, \tag{1}$$

and 0 otherwise. Clearly, (1) is a proper distribution function, and it generalizes the generalized exponential distribution. From now on a random variable X with the distribution function (1) will be denoted by MOGE(α, λ, θ).

It may be observed that several special cases can be obtained from (1). For example, if we set $\theta = 1$ in (1), then we obtain the generalized exponential distribution as introduced by Gupta and Kundu [11]. It will be denoted by GE(α, λ). For $\alpha = 1$ we obtain the Marshall-Olkin exponential distribution introduced by Marshall and Olkin [16]. For $\alpha = 1$ and $\theta = 1$ we obtain the exponential distribution with parameter λ. Now we provide some physical justification of the proposed model, see also Marshall and Olkin [16] in this respect.

First, let us consider a series system with N independent components. Suppose that a random variable N has the probability mass function $P(N = n) = \theta(1-\theta)^{n-1}, n = 1, 2, \ldots$ and $0 < \theta < 1$. Let $X_1, X_2, \ldots$ represent the lifetimes of each component and suppose they are independent and identically distributed (i.i.d.) GE random variables with parameters λ and α. Then a random variable $Y = \min(X_1, \ldots, X_N)$ represents the time to the first failure with distribution function

$$G(x) = 1 - \sum_{n=1}^{\infty} P(\min(X_1, \ldots, X_n) > x)\theta(1-\theta)^{n-1}$$

$$= 1 - \theta(1 - (1 - e^{-\lambda x})^\alpha) \sum_{n=1}^{\infty} \left((1-\theta)(1 - (1 - e^{-\lambda x})^\alpha)\right)^{n-1}$$

$$= \frac{(1 - e^{-\lambda x})^\alpha}{\theta + (1 - \theta)(1 - e^{-\lambda x})^\alpha}, \quad \text{for} \quad x > 0.$$

Thus we obtain the distribution function given by (1).

Second, let us consider now a parallel system with N independent components and suppose that a random variable N has the probability mass function $P(N = n) = \theta^{-1}(1 - \theta^{-1})^{n-1}$, $n = 1, 2, \ldots$ and $\theta > 1$. Let $X_1, X_2, \ldots$ represent the lifetimes of each component and suppose they are generalized exponential distributed with parameters λ and α. Then a random

variable $Z = \max(X_1, \ldots, X_N)$ represents the lifetime of the system. The distribution function of the random variable Z is given as (1).

Third, let $\theta > 1/2$. Using the series expansion

$$(\theta + (1 - \theta)(1 - e^{-\lambda x})^\alpha)^{-1} = \theta^{-1} \sum_{k=0}^{\infty} (1 - \theta^{-1})^k (1 - e^{-\lambda x})^{\alpha k},$$

we obtain that the distribution function (1) can be rewritten as

$$G(x; \lambda, \alpha, \theta) = \theta^{-1} \sum_{k=0}^{\infty} (1 - \theta^{-1})^k F_{GE(\alpha(k+1), \lambda)}(x).$$

On the other hand, if $0 < \theta < 2$, by using the series expansion

$$(\theta + (1 - \theta)(1 - e^{-\lambda x})^\alpha)^{-1} = \theta^{-1} \sum_{k=0}^{\infty} (1 - \theta)^k \left(1 - (1 - e^{-\lambda x})^\alpha\right)^k,$$

we obtain that (1) can be rewritten as

$$G(x; \lambda, \alpha, \theta) = \sum_{k=0}^{\infty} (k + 1)(1 - \theta)^k \sum_{j=0}^{k} \binom{k}{j} (-1)^j (1 - e^{-\lambda x})^{\alpha j}.$$

Thus it follows that the distribution function given by (1) can be represented as a generalized mixture of generalized exponential distribution function. It may be mentioned that generalized mixture distribution has received some attention recently, see for example [9]. Since it allows negative weights also, it has more flexibility than the mixture models.

We call this new three-parameter extension of the GE distribution as the Marshall-Olkin generalized exponential (MOGE) distribution. As expected this new three parameter distribution has two shape parameters and one scale parameter. It is observed that the proposed MOGE distribution can have decreasing or unimodal PDFs. It is interesting to observe that the hazard function can take four different major shapes. It can have increasing, decreasing, bathtub or inverted tub shaped. Therefore it can be used quite extensively to analyze life time data. Since it has only three unknown parameters, the estimation of the unknown parameter is also not very difficult. It may be mentioned that not too many three parameter distributions can have all the three possible hazard functions, therefore, the introduction of the proposed three-parameter MOGE distribution will be quite useful. Moreover, since MOGE distribution has a compact distribution function, it can be used very effectively to analyze censored data, and the generation from a MOGE distribution is also very straight forward.

We have derived several properties of the MOGE distribution. The PDF of the proposed MOGE is either a decreasing or an unimodal function. Interestingly, because of the introduction of a new shape parameter, the MOGE can have an increasing, decreasing, unimodal or bathtub shaped hazard functions. The median and mode can be obtained in explicit forms. The moments cannot be obtained explicitly, we have obtained the moments in terms of infinite series. A small table is provided indicating the first four moments of MOGE distribution for different values of the shape parameters. We have obtained the density function of the i-th order statistics, and it is observed that it can be represented as an infinite mixture of the beta generalized exponential density function. We have also provided the Renyi's entropy, which measures the uncertainty of variation. Since MOGE distribution has been obtained as a geometric maxima or minima of i.i.d. GE distributions, several ordering properties can be easily established.

The maximum likelihood estimators (MLEs) cannot be obtained in explicit form. Three dimensional optimization procedure is needed to compute the MLEs. We propose to use the EM algorithm, see [8], to compute the MLEs of the unknown parameters. Two data analysis are performed for illustrative purposes.

The paper is organized as follows. In Sect. 2 we derive the probability density function and discuss its shapes. The hazard function is considered in Sect. 3. In Sect. 4 we give some expressions for the moments. The order statistics and the limiting distribution of sample extremes are considered in Sect. 5. In Sect. 6 we derive two entropies, the Rényi's and the Shannon's entropy. We have derived several ordering relations of MOGE in Sect. 7. The maximum likelihood estimation and an EM algorithm are provided in Sect. 8. Analysis of two data sets are provided in Sect. 9, and finally we conclude the paper in Sect. 10.

2 The probability density function

If the random variable X has a distribution function (1), the corresponding probability density function (PDF) for $\alpha > 0$, $\lambda > 0$ and $\theta > 0$, is

$$g(x; \alpha, \lambda, \theta) = \frac{\alpha\lambda\theta e^{-\lambda x}(1 - e^{-\lambda x})^{\alpha-1}}{(\theta + (1 - \theta)(1 - e^{-\lambda x})^\alpha)^2}, \quad x > 0, \tag{2}$$

see also [3]. Suppose that $0 < \theta < 2$. Then the denominator in (2) can be expressed as

$$(\theta + (1 - \theta)(1 - e^{-\lambda x})^\alpha)^{-2} = \sum_{k=0}^{\infty}(k + 1)(1 - \theta)^k(1 - (1 - e^{-\lambda x})^\alpha)^k$$

$$= \sum_{k=0}^{\infty}(k + 1)(1 - \theta)^k \sum_{j=0}^{k}\binom{k}{j}(-1)^j(1 - e^{-\lambda x})^{\alpha j}.$$

Using this result, we obtain that the pdf given by (2) can be expressed in the generalized mixture form as

$$g(x; \alpha, \lambda, \theta) = \theta \sum_{k=0}^{\infty}\sum_{j=0}^{k}(-1)^j(1 - \theta)^k\binom{k+1}{j+1}f_{GE(\alpha(j+1),\lambda)}(x), \tag{3}$$

where $f_{GE(\alpha(j+1),\lambda)}(x)$ denotes the pdf of a random variable with generalized exponential distribution with parameters $\alpha(j + 1)$ and λ, see also [3]. Note that the density $g(x; \alpha, \lambda, \theta)$ can be represented in the generalized mixture form of beta generalized exponential probability density functions, Barreto-Souza et al. [4], as $g(x; \alpha, \lambda, \theta) = \theta \sum_{k=0}^{\infty}(1 - \theta)^k f_{BGE(1,k+1,\lambda,\alpha)}(x)$.

Similarly, if $\theta > 1/2$ and using the expansion

$$(\theta + (1 - \theta)(1 - e^{-\lambda x})^\alpha)^{-2} = \theta^{-2}\sum_{k=0}^{\infty}(k + 1)\left(1 - \frac{1}{\theta}\right)^k(1 - e^{-\lambda x})^{\alpha k},$$

we obtain the expression

$$g(x; \alpha, \lambda, \theta) = \theta^{-1}\sum_{k=0}^{\infty}\left(1 - \frac{1}{\theta}\right)^k f_{GE(\alpha(k+1),\lambda)}(x). \tag{4}$$
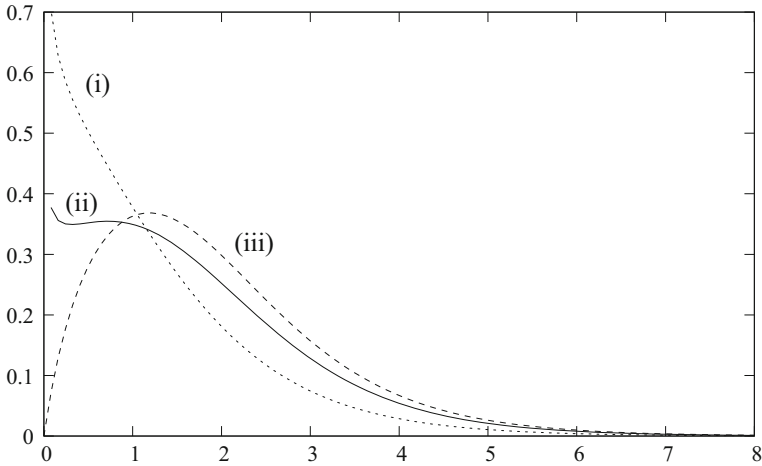
Fig. 1 The PDF of the MOGE distribution for different values of α and θ when $\lambda = 1$. *(i)* $\alpha = 0.8, \theta = 2.0$, *(ii)* $\alpha = 0.4, \theta = 4.0$, *(iii)* $\alpha = 2.0, \theta = 2.0$

Let us consider the shape of the PDF of MOGE distribution. Since λ is the scale parameter, the shape of the PDF of MOGE distribution does not depend on λ. It can be easily shown that for (i) $0 < \alpha \le 1$ and $0 < \theta \le 1$, the PDF of MOGE decreases with $g(0) = \infty$ and $g(\infty) = 0$, (ii) $0 < \alpha \le 1$ and $\theta > 1$, for some $x_1 < x_2$, the probability density function $g(x; \alpha, \theta)$ decreases on $(0, x_1) \cup (x_2, \infty)$ and increases on $[x_1, x_2]$. Furthermore, $g(0) = \infty$ and $g(\infty) = 0$. (iii) For $\alpha > 1$, it follows that the PDF $g(x; \alpha, \theta)$ has a single mode and $g(0) = g(\infty) = 0$.

We can conclude that the shape of the PDF of the MOGE is different than the shape of the PDF of the GE distribution. The PDF of the GE distribution is a decreasing function for $0 < \alpha < 1$, while for $\alpha > 1$ is an increasing function. Some possible shapes of the probability density function $g(x; \alpha, \theta)$ are presented in Fig. 1.

3 The hazard rate function

Now we study the shapes of the hazard function of MOGE distribution for different values of α and θ. Since λ is the scale parameter, the shape of the hazard function does not depend on λ. So without loss of generality we assume that $\lambda = 1$. Therefore, the hazard function of the MOGE is of the form

$$h(x; \alpha, \theta) = \frac{\alpha e^{-x}(1 - e^{-x})^{\alpha-1}}{(\theta + (1 - \theta)(1 - e^{-x})^\alpha)(1 - (1 - e^{-x})^\alpha)}, \quad \text{for} \quad x > 0.$$

Since the shape of $h(x; \alpha, \theta)$ is same as the shape of $\ln h(x; \alpha, \theta)$, we study the shape of $\ln h(x; \alpha, \theta)$ only. The first derivative of $\ln h(x; \alpha, \lambda)$ is

$$\frac{d \log h(x; \alpha, \lambda)}{dx} = \frac{s(x)}{(1 - e^{-x})(\theta + (1 - \theta)(1 - e^{-x})^\alpha)(1 - (1 - e^{-x})^\alpha)},$$

where

$$s(x) = -\theta + (2\theta - 1)(1 - e^{-x})^\alpha + \alpha\theta e^{-x} + (1 - \theta)(1 - e^{-x})^{2\alpha}(1 + \alpha e^{-x}).$$

Four shapes of the hazard rate function are possible:

- If $0 < \alpha < 1$ and $0 < \theta < (1 + \alpha)/(2\alpha)$, then the function $s(x)$ is negative for $x > 0$ and it follows that the hazard function is a decreasing function with $h(0) = \infty$ and $h(\infty) = 1$.
- If $0 < \alpha < 1$ and $\theta > (1 + \alpha)/(2\alpha)$, then the function s has one root x_0 with $s(0) = \theta(\alpha - 1) < 0$ and $s(\infty) = 0$. Thus we obtain that the hazard rate function $h(x)$ decreases on $(0, x_0)$ and increases on (x_0, ∞) with $h(0) = \infty$ and $h(\infty) = 1$.
- If $\alpha > 1$ and $\theta > (1 + \alpha)/(2\alpha)$, then the function $s(x)$ is positive for $x > 0$ and it follows that the hazard function is an increasing function with $h(0) = 0$ and $h(\infty) = 1$.
- If $\alpha > 1$ and $0 < \theta < (1 + \alpha)/(2\alpha)$, then the function $s(x)$ has one root x_0 with $s(0) = \theta(\alpha - 1) > 0$ and $s(\infty) = 0$. Thus we obtain that the hazard function $h(x)$ increases on $(0, x_0)$ and decreases on (x_0, ∞) with $h(0) = 0$ and $h(\infty) = 1$.

In comparison with the hazard rate function of the Weibull, gamma or GE distributions, the hazard rate function of the proposed MOGE distribution has two more possible shapes. Therefore it becomes more flexible for analyzing lifetime data. Some possible shapes of the hazard function $h(x; \alpha, \theta)$ for different values of α and θ, are presented in Fig. 2.

Let us derive now the reverse hazard rate function. As was noted in [19], the reverse hazard rate function is useful in constructing the information matrix and in estimating the survival function for censored data. The reverse hazard function of MOGE distribution is given as

$$r(x) = \frac{g(x)}{G(x)} = \frac{\alpha\theta e^{-x}}{(1 - e^{-x})(\theta + (1 - \theta)(1 - e^{-x})^\alpha)}, \quad x, \alpha, \lambda, \theta > 0.$$

The reverse hazard rate function decreases on $(0, \infty)$ with $r(0) = \infty$ and $r(\infty) = 0$. We can see that the reverse hazard function for $\theta \neq 1$ is not a linear function of α as the reverse hazard function of GE distribution.
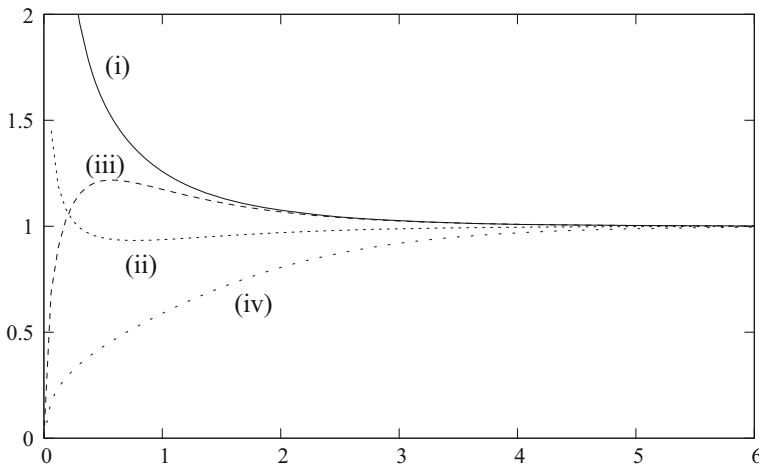

Fig. 2 The hazard function of the MOGE distribution for different values of α and θ when $\lambda = 1$. *(i)* $\alpha = 0.5$, $\theta = 0.5$, *(ii)* $\alpha = 0.5$, $\theta = 2.0$, *(iii)* $\alpha = 1.5$, $\theta = 0.5$, *(iv)* $\alpha = 1.5$, $\theta = 2.0$

4 Moments

In this section we derive the n-th moments of a random variable $X \sim \text{MOGE}(\alpha, \lambda, \theta)$. Let $Y_{\alpha, \lambda} \sim \text{GE}(\alpha, \lambda)$. We will first consider the case $0 < \theta < 2$. By using (3), we obtain that the n-th moment of a random variable X as

$$E(X^n) = \theta \sum_{k=0}^{\infty} \sum_{j=0}^{k} (-1)^j (1-\theta)^k \binom{k+1}{j+1} E(Y_{\alpha(j+1), \lambda}^n). \tag{5}$$

Nadarajah and Kotz [17] derived the n-th moment of a random variable $Y_{\alpha, \lambda}$ as

$$E\left(Y_{\alpha, \lambda}^n\right) = \frac{\alpha(-1)^n}{\lambda^n} \frac{\partial^n}{\partial p^n} B(\alpha, p)\Big|_{p=1}. \tag{6}$$

Here for $u > 0$ and $v > 0$, $B(u, v)$ is the beta function defined as follows: $B(u, v) = \int_0^1 x^{u-1}(1-x)^{v-1} dx$. Now by combining (5) and (6), the n-th moment of a random variable X can be calculated as

$$E(X^n) = \frac{\alpha\theta(-1)^n}{\lambda^n} \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^j (1-\theta)^k (k+1)!}{j!(k-j)!} \frac{\partial^n}{\partial p^n} B(\alpha(j+1), p)\Big|_{p=1}.$$

In particular, the expectation is

$$E(X) = \frac{\theta}{\lambda} \sum_{k=0}^{\infty} \sum_{j=0}^{k} (-1)^k (1-\theta)^k \binom{k+1}{j+1} \left[\Psi(\alpha(j+1)+1) - \Psi(1)\right],$$

and the second moment is

$$E(X^2) = \frac{\theta}{\lambda^2} \sum_{k=0}^{\infty} \sum_{j=0}^{k} (-1)^k (1-\theta)^k \binom{k+1}{j+1} \left[\Psi^2(1) + \Psi'(1)\right.$$
$$\left. -2\Psi(1)\Psi(\alpha(j+1)+1) - \Psi'(\alpha(j+1)+1) + \Psi^2(\alpha(j+1)+1)\right],$$

where $\Psi(x) = \frac{d \log \Gamma(x)}{dx}$ is the Euler's psi function.

Similarly for the case $\theta > 1/2$, it can be shown in this case that the n-th moment of a random variable X can be calculated as

$$E(X^n) = \frac{\alpha(-1)^n}{\theta \lambda^n} \sum_{k=0}^{\infty} \left(1 - \frac{1}{\theta}\right)^k (k+1) \frac{\partial^n}{\partial p^n} B(\alpha(k+1), p)\Big|_{p=1},$$

see [3]. The first two moments can be written as

$$E(X) = \frac{1}{\theta\lambda} \sum_{k=0}^{\infty} (1 - \theta^{-1})^k \left[\Psi(\alpha(k+1)+1) - \Psi(1)\right],$$

and the second moment is

$$E(X^2) = \frac{1}{\theta\lambda^2} \sum_{k=0}^{\infty} (1 - \theta^{-1})^k \left[\Psi^2(1) + \Psi'(1) - 2\Psi(1)\Psi(\alpha(k+1)+1)\right.$$
$$\left. -\Psi'(\alpha(k+1)+1) + \Psi^2(\alpha(k+1)+1)\right].$$

5 Order statistics

In this section we consider the order statistics $X_{1:n}$, $X_{2:n}$, …, $X_{n:n}$, from a random sample X_1, X_2, …, X_n from the MOGE distribution. Let us derive the density function of the i-th order statistics $X_{i:n}$, $1 \leq i \leq n$. We have that

$$
\begin{aligned}
g_{i:n}(x) &= \frac{n!}{(i-1)!(n-i)!} \cdot g(x)(G(x))^{i-1}(1-G(x))^{n-i} \\
&= \frac{n!}{(i-1)!(n-i)!} \cdot \frac{\alpha\lambda\theta^{n-i+1}e^{-\lambda x}(1-e^{-\lambda x})^{\alpha i-1}(1-(1-e^{-\lambda x})^{\alpha})^{n-i}}{(\theta+(1-\theta)(1-e^{-\lambda x})^{\alpha})^{n+1}}.
\end{aligned}
$$

This density function can be represented as an infinite weighted sum of beta generalized exponential density function. Consider the case when $\theta > 1/2$. Using the series expansion $(1-z)^{-k} = \sum_{j=0}^{\infty} \frac{\Gamma(k+j)}{j!\Gamma(k)} z^j$, $k > 0$, we obtain the representation

$$
g_{i:n}(x) = \frac{1}{\theta^i} \sum_{j=0}^{\infty} \binom{i+j-1}{j} \left(1-\frac{1}{\theta}\right)^j f_{BGE}(x; i+j, n-i+1, \lambda, \alpha).
$$

Similarly, in the case $0 < \theta < 2$, we obtain the representation

$$
g_{i:n}(x) = \theta^{n-i+1} \sum_{j=0}^{\infty} \binom{n+j-i}{j} (1-\theta)^j f_{BGE}(x; i, n+j-i+1, \lambda, \alpha).
$$

Barreto-Souza et al. [4] derived the moments of the i-th order statistics from beta generalized exponential distribution. Let $\mu_{i:n}^r(a, b)$ represents the r-th moment of the i-th order statistics from the BGE(a, b, λ, α) distribution. Then the r-th moment of the i-th order statistics from the MOGE distribution can be derived as

$$
E(X_{i:n}^r) = \begin{cases} \frac{1}{\theta^i} \sum_{j=0}^{\infty} \binom{i+j-1}{j} \left(1-\frac{1}{\theta}\right)^j \mu_{i:n}^r(i+j, n-i+1), & \theta > 1/2, \\ \theta^{n-i+1} \sum_{j=0}^{\infty} \binom{n+j-i}{j} (1-\theta)^j \mu_{i:n}^r(i, n+j-i+1), & 0 < \theta < 2, \end{cases}
$$

see Barreto-Souza et al. [3].

Now we discuss the asymptotic distributions of the order statistics. First we consider the sample maxima $X_{n:n}$. Since $G^{-1}(1) = \infty$, $\lim_{x\to\infty} h(x) = \lambda$ and $\lim_{x\to\infty} \frac{g'(x)}{g(x)} = -\lambda$, the von Mises' condition (iii) from Arnold et al. [2, Theorem 8.3.3] is satisfied. This implies that

$$
\frac{X_{n:n} - a_n}{b_n} \xrightarrow{d} e^{-e^{-x}}, \quad x \in \mathbf{R},
$$

where the normalizing constants a_n and b_n can be derived by Arnold et al. [2, Theorem 8.3.4 (iii)].

Second we consider the sample minimum $X_{1:n}$. Since $G^{-1}(0) = 0$ and $\lim_{\varepsilon\to 0+} \frac{G(\varepsilon x)}{G(\varepsilon)} = x^{\alpha}$, we obtain from Arnold et al. [2, Theorem 8.3.6 (ii)] that

$$
\frac{X_{1:n} - a_n^*}{b_n^*} \xrightarrow{d} 1 - e^{-(-x)^{\alpha}}, \quad x < 0, \alpha > 0,
$$

where $a_n^* = 0$ and $b_n^* = G^{-1}(1/n)$.

Finally, the asymptotic distribution of the order statistics $X_{n-i+1:n}$ follows from the asymptotic distribution of the sample maxima. Thus

$$\frac{X_{n-i+1:n} - a_n}{b_n} \xrightarrow{d} e^{-e^{-x}} \sum_{j=0}^{i-1} \frac{e^{-jx}}{j!}, \quad x \in \mathbf{R},$$

where a_n and b_n are the normalizing constants derived by Arnold et al. [2, Theorem 8.3.4 (iii)].

6 Rényi entropy

The entropy is a measure of diversity, uncertainty or randomness of a system. A popular entropy is the Rényi entropy, see [20], which generalizes the well known Shannon entropy.

The Rényi entropy is given by $I_R(\xi) = \frac{1}{1-\xi} \log \int_0^\infty g^\xi(x) dx$, where $\xi > 0$ and $\xi \neq 1$. For $\theta > 1/2$ the function $g^\xi(x)$ can be expanded as

$$g^\xi(x) = \frac{\lambda^\xi \alpha^\xi}{\theta^\xi} \sum_{j=0}^\infty \frac{\Gamma(2\xi+j)}{\Gamma(2\xi)j!} \left(1 - \frac{1}{\theta}\right)^j e^{-\lambda\xi x}(1 - e^{-\lambda x})^{\xi(\alpha-1)+\alpha j}.$$

Now using the fact that $\int_0^\infty e^{-\lambda\xi x}(1 - e^{-\lambda x})^{\xi(\alpha-1)+\alpha j} dx = \lambda^{-1} B(\xi, \xi(\alpha-1) + \alpha j + 1)$, we obtain that for $\theta > 1/2$ the Rényi entropy is

$$I_R(\xi) = \frac{1}{1-\xi} \log \left\{ \frac{\lambda^{\xi-1}\alpha^\xi}{\theta^\xi} \sum_{j=0}^\infty \frac{\Gamma(2\xi+j)}{\Gamma(2\xi)j!} \left(1 - \frac{1}{\theta}\right)^j B(\xi, \xi(\alpha-1) + \alpha j + 1) \right\}.$$

Let us consider the case when $0 < \theta < 2$. The the function $g^\xi(x)$ can be expanded as

$$g^\xi(x) = \lambda^\xi \alpha^\xi \theta^\xi \sum_{j=0}^\infty \sum_{k=0}^j \frac{\Gamma(2\xi+j)}{\Gamma(2\xi)j!} \binom{j}{k} (-1)^k (1-\theta)^j e^{-\lambda\xi x}(1 - e^{-\lambda x})^{\xi(\alpha-1)+\alpha k},$$

which implies that the Rényi entropy for $0 < \theta < 1$ is

$$I_R(\xi) = \frac{1}{1-\xi} \log \left\{ \lambda^{\xi-1}\alpha^\xi \theta^\xi \sum_{j=0}^\infty \sum_{k=0}^j \frac{\Gamma(2\xi+j)}{\Gamma(2\xi)j!} \binom{j}{k} (-1)^k (1-\theta)^j B(\xi, \xi(\alpha-1) + \alpha k + 1) \right\}.$$

A special case of the Rényi entropy is the Shannon entropy defined as $E(-\log g(X))$, where X is a random variable. The Shannon entropy represents the limit of $I_R(\xi)$ when $\xi \uparrow 1$. If we suppose that a random variable X has the MOGE, then the Shannon entropy is $E(-\log g(X)) = -\log(\alpha\lambda\theta) + \lambda E(X) - (\alpha-1)E(\log(1 - e^{-\lambda X})) + 2E(\log(\theta + (1-\theta)(1 - e^{-\lambda X})^\alpha))$. Replacing $E(\log(1 - e^{-\lambda X})) = \log\theta/(2(1-\theta))$ and $E(\log(\theta + (1-\theta)(1 - e^{-\lambda X})^\alpha)) = 1 + \log\theta/(1-\theta)$ in the last equation, we obtain that the Shannon entropy is

$$E(-\log g(X)) = 2 - \log(\alpha\lambda) + \lambda E(X) + \frac{(\alpha\theta + 1)\log\theta}{\alpha(1-\theta)}.$$

7 Random minima, maxima and different ordering relations

In reliability and survival analysis the occurrence of a series or parallel system with random number of components is very common, see for example [12]. In many agricultural and biological experiments it is impossible to have a fixed sample size as some of the observations often get lost due to different reasons. In many situations the sample size may depend on

the occurrence of some specific event, which makes the sample size random. For example, a common dose of radiation is given to a group of animals, often the interest is in the times that the first and last expire, see [7]. In actuarial science, the claims received by an insurer in a certain time interval often make up of a sample of random size, and the largest claim amount is of chief interest there, see [15]. It has already been observed that the proposed MOGE distribution can be obtained as the random minima or random maxima of the GE distributions depending on $0 < \theta < 1$ or $1 < \theta < \infty$. Therefore, the proposed model may be used quite effectively in these cases. In this section we establish different results based on this property.

Let us recall the following definitions. Suppose U and V be two continuous random variables with PDFs f_U and f_V, respectively. The corresponding cumulative distribution functions (CDF) will be denoted by F_U and F_V, respectively. The random variable U is said to be smaller than the random variable V in the likelihood ratio ordering (denoted by $U \leq_{lr} V$) if $f_U(x) f_V(y) \geq f_U(y) f_V(x)$, for all $0 < x \leq y < \infty$. The random variable U is said to be smaller than the random variable V in stochastic order (denoted by $U \leq_{st} V$) if $P(U \geq x) \leq P(V \geq x)$, for all $x > 0$. The random variable U is said to be smaller than the random variable V in dispersive order (denoted by $U \leq_{disp} V$) if $F_U^{-1}(\beta) - F_U^{-1}(\alpha) \leq F_V^{-1}(\beta) - F_V^{-1}(\alpha)$, for all $0 < \alpha \leq \beta < 1$. The random variable U is said to be smaller than the random variable V in hazard rate order (denoted by $U \leq_{hr} V$), if $P(V > x)/P(U > x)$ is an increasing function of x. The random variable U is said to be smaller than the random variable V in the convex transform order (denoted by $U \leq_c V$), if $F_V^{-1} F_U$ is a convex function in $(0, \infty)$. Further, the random variable U is said to be smaller than the random variable V in star order (denoted by $U \leq_* V$), if $F_V^{-1} F_U(x)/x$ is an increasing function of $x \in (0, \infty)$. We have the following results.

Result 1 Let $X \sim \text{GE}(\alpha, \lambda)$, $Y \sim \text{MOGE}(\alpha, \lambda, \theta)$, $Z \sim \text{MOGE}(\alpha, \lambda, 1/\theta)$, where $\alpha > 0$, $\lambda > 0$ and $0 < \theta < 1$, then

$$Y \leq_{lr} X \leq_{lr} Z.$$

Proof The result mainly follows from Corollary 2.5 of [21]. □

Result 2 (a) If $Y_1 \sim \text{MOGE}(\alpha_1, \lambda, \theta)$ and $Y_2 \sim \text{MOGE}(\alpha_2, \lambda, \theta)$, where $0 < \alpha_1 < \alpha_2$, $\theta > 0$ and $\lambda > 0$, then $Y_1 \leq_{st} Y_2$.
(b) If $Y_1 \sim \text{MOGE}(\alpha, \lambda_1, \theta)$ and $Y_2 \sim \text{MOGE}(\alpha, \lambda_2, \theta)$, where $0 < \lambda_1 < \lambda_2$, $\theta > 0$ and $\alpha > 0$, then $Y_2 \leq_{st} Y_1$.

Proof Note that if $U \sim \text{GE}(\alpha_1, \lambda)$ and $V \sim \text{GE}(\alpha_2, \lambda)$, then for $\alpha_1 \leq \alpha_2$, $U \leq_{st} V$. Similarly, if $U \sim \text{GE}(\alpha, \lambda_1)$ and $V \sim \text{GE}(\alpha, \lambda_2)$, then for $\lambda_1 \leq \lambda_2$, $V \leq_{st} U$. Hence both (a) and (b) follow using Theorem 3.1 of [21]. □

Result 3 (a) If $Y_1 \sim \text{MOGE}(\alpha_1, \lambda, \theta)$ and $Y_2 \sim \text{MOGE}(\alpha_2, \lambda, \theta)$, where $0 < \alpha_1 < \alpha_2$, $\theta > 0$ and $\lambda > 0$, then $Y_1 \leq_{disp} Y_2$.
(b) If $Y_1 \sim \text{MOGE}(\alpha_1, \lambda, \theta)$ and $Y_2 \sim \text{MOGE}(\alpha_2, \lambda, \theta)$, where $0 < \alpha_1 < \alpha_2$, $\theta > 0$ and $\lambda > 0$, then $Y_1 \leq_{hr} Y_2$.

Proof Suppose $U \sim \text{GE}(\alpha_1, \lambda)$ and $V \sim \text{GE}(\alpha_2, \lambda)$, then for $\alpha_1 \leq \alpha_2$, $U \leq_{disp} V$ and $U \leq_{hr} V$, see [11]. Hence (a) and (b) follow using Theorem 3.2 and Theorem 3.3, respectively of [21]. □

Result 4 (a) If $Y_1 \sim \text{MOGE}(\alpha_1, \lambda, \theta)$ and $Y_2 \sim \text{MOGE}(\alpha_2, \lambda, \theta)$, where $0 < \alpha_1 < \alpha_2$, $\theta > 0$ and $\lambda > 0$, then $Y_1 \leq_c Y_2$.

(b) If $Y_1 \sim \mathrm{MOGE}(\alpha_1, \lambda, \theta)$ and $Y_2 \sim \mathrm{MOGE}(\alpha_2, \lambda, \theta)$, where $0 < \alpha_1 < \alpha_2, \theta > 0$ and $\lambda > 0$, then $Y_1 \leq_* Y_2$.

Proof Suppose $U \sim \mathrm{GE}(\alpha_1, \lambda)$ and $V \sim \mathrm{GE}(\alpha_2, \lambda)$, then for $\alpha_1 \leq \alpha_2$, $U \leq_c V$, see [11]. Hence (a) follows using Theorem 2 (a) of [5]. Since convex ordering implies start ordering, (b) follows from (a). □

8 Estimation

In this section we consider the maximum likelihood estimation of the unknown parameters based on a complete sample. Let us assume that we have a sample of size n, namely $\{x_1, \ldots, x_n\}$ from $\mathrm{MOGE}(\alpha, \lambda, \theta)$ distribution. The log-likelihood function is given by

$$l(\alpha, \theta, \lambda | Data) = \log L(\alpha, \theta, \lambda)$$

$$= n \log(\alpha \lambda \theta) - \lambda \sum_{i=1}^{n} x_i + (\alpha - 1) \sum_{i=1}^{n} \log(1 - e^{-\lambda x_i})$$

$$- 2 \sum_{i=1}^{n} \log(\theta + (1 - \theta)(1 - e^{-\lambda x_i})^\alpha).$$

Normal equations can be obtained by taking the first derivatives of the log-likelihood function with respect to λ, α and θ are equate them to zeros as follows;

$$\frac{\partial \log L(\alpha, \theta, \lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i + (\alpha - 1) \sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{1 - e^{-\lambda x_i}}$$

$$- 2\alpha(1 - \theta) \sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}(1 - e^{-\lambda x_i})^{\alpha-1}}{\theta + (1 - \theta)(1 - e^{-\lambda x_i})^\alpha}$$

$$\frac{\partial \log L(\alpha, \theta, \lambda)}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} \log(1 - e^{-\lambda x_i}) + 2\theta \sum_{i=1}^{n} \frac{\log(1 - e^{-\lambda x_i})}{\theta + (1 - \theta)(1 - e^{-\lambda x_i})^\alpha}$$

$$\frac{\partial \log L(\alpha, \theta, \lambda)}{\partial \theta} = \frac{n}{\theta} - 2 \sum_{i=1}^{n} \frac{1 - (1 - e^{-\lambda x_i})^\alpha}{\theta + (1 - \theta)(1 - e^{-\lambda x_i})^\alpha} .$$

It is clear that the MLEs do not have explicit solutions, and the MLEs can be obtained by solving a three dimensional optimization process. We may use the standard Gauss-Newton or Newton-Raphson methods, but they have their usual problem of convergence. If the initial guesses are not close to the optimal value, the iteration may not converge, see for example [18] for a recent reference on this issue on a related problem. Moreover, choosing a three dimensional initial guesses may not very simple in most of the practical situations. Before progressing further we present the following result related to the MLE.

If the parameters λ and θ are known, the properties of the MLE of the parameter α follow from the following theorem.

Theorem 1 *Let α be the true value of the parameter. If $0 < \theta < 1$, then the equation $\frac{\partial \log L(\alpha, \theta, \lambda)}{\partial \alpha} = 0$ has exactly one root. If $\theta > 1$, then the root of equation $\frac{\partial \log L(\alpha, \theta, \lambda)}{\partial \alpha} = 0$ lies in the interval $[(2\theta - 1)^{-1} \psi_\lambda^{-1}, \psi_\lambda^{-1}]$, where $\psi_\lambda = -n^{-1} \sum_{i=1}^{n} \log(1 - e^{-\lambda x_i})$.*

Proof Let us first consider the case $0 < \theta < 1$. Then the function $\frac{\partial \log L(\alpha,\theta,\lambda)}{\partial\alpha}$ is decreasing with $\lim_{\alpha\to 0} \frac{\partial \log L(\alpha,\theta,\lambda)}{\partial\alpha} = \infty$ and $\lim_{\alpha\to\infty} \frac{\partial \log L(\alpha,\theta,\lambda)}{\partial\alpha} = -n\psi_\lambda < 0$. Thus it follows that exists exactly one root. Now consider the case $\theta > 1$ and let $w(\alpha; \lambda, \theta) = 2\theta \sum_{i=1}^n \frac{\log(1-e^{-\lambda x_i})}{\theta+(1-\theta)(1-e^{-\lambda x_i})^\alpha}$. The function w is increasing. We can see that $\lim_{\alpha\to 0} w = 2\theta \sum_{i=1}^n \log(1-e^{-\lambda x_i})$ and $\lim_{\alpha\to\infty} w = 2 \sum_{i=1}^n \log(1-e^{-\lambda x_i})$. This implies that

$$\frac{n}{\alpha} + (2\theta - 1) \sum_{i=1}^n \log(1 - e^{-\lambda x_i}) < \frac{\partial \log L(\alpha,\theta,\lambda)}{\partial\alpha} < \frac{n}{\alpha} + \sum_{i=1}^n \log(1 - e^{-\lambda x_i}).$$

Then we obtain that $\frac{\partial \log L(\alpha,\theta,\lambda)}{\partial\alpha} > 0$ for $\alpha < (2\theta - 1)^{-1}\psi_\lambda^{-1}$ and $\frac{\partial \log L(\alpha,\theta,\lambda)}{\partial\alpha} < 0$ for $\alpha > \psi_\lambda^{-1}$. This proves the theorem.   □

As it has been mentioned before that it is possible to use the standard three dimensional optimization algorithm to maximize the log-likelihood function (7). We propose a simple iterative technique to compute the MLEs of the unknown parameters, which avoids solving a three dimensional optimization process directly, it needs solving three one dimensional optimization problems. The idea comes from the following observations;

Let us consider the random variables X and Z with the following joint PDF

$$f(x, z; \alpha, \lambda, \theta) = \frac{\alpha\lambda\theta z e^{-\lambda x}(1 - e^{-\lambda x})^{\alpha-1}}{(1 - (1 - e^{-\lambda x})^\alpha)^2} e^{-z(\theta-1+(1-(1-\exp(-\lambda x))^\alpha)^{-1})}, \tag{7}$$

It can be easily observed that the random variable X follows the Marshall-Olkin Generalized exponential distribution. Based on a random sample of size n, say $\{x_i, z_i\}$ from (7), the log-likelihood function can be written as;

$$\ln l(\alpha, \lambda, \theta; Data, z_1, \ldots, z_n) = n \ln \alpha + n \ln \lambda + n \ln \theta + \sum_{i=1}^n \ln z_i - \lambda \sum_{i=1}^n x_i$$

$$+ (\alpha - 1) \sum_{i=1}^n \log(1 - e^{-\lambda x_i}) - 2 \sum_{i=1}^n \log(1 - (1 - e^{-\lambda x_i})^\alpha)$$

$$- \sum_{i=1}^n z_i \left(\theta - 1 + (1 - (1 - e^{-\lambda x_i})^\alpha)^{-1} \right). \tag{8}$$

Note that the maximization of (8) with respect to α, λ and θ can be decoupled. The maximization of (8) with respect to θ can be obtained as $\widehat{\theta} = \frac{n}{\sum_{i=1}^n z_i}$, and the maximization of (8) with respect to α and λ can be obtained by maximizing $g(\alpha, \lambda)$, where

$$g(\alpha, \lambda) = n \ln \alpha + n \ln \lambda - \lambda \sum_{i=1}^n x_i + (\alpha - 1) \sum_{i=1}^n \log(1 - e^{-\lambda x_i})$$

$$- 2 \sum_{i=1}^n \log(1 - (1 - e^{-\lambda x_i})^\alpha) - \sum_{i=1}^n z_i (1 - (1 - e^{-\lambda x_i})^\alpha)^{-1}. \tag{9}$$

The method proposed by Song et al. [22] can be used to maximize (9). The method was used by Kannan et al. [13] in a similar problem, and it can be described as follows. Let us write $g(\alpha, \lambda)$ as;

$$g(\alpha, \lambda) = g_1(\alpha, \lambda) + g_2(\alpha, \lambda), \tag{10}$$

where

$$g_1(\alpha, \lambda) = n \ln \alpha + n \ln \lambda - \lambda \sum_{i=1}^{n} x_i + (\alpha - 1) \sum_{i=1}^{n} \log(1 - e^{-\lambda x_i}), \tag{11}$$

and

$$g_2(\alpha, \lambda) = -2 \sum_{i=1}^{n} \log(1 - (1 - e^{-\lambda x_i})^\alpha) - \sum_{i=1}^{n} z_i (1 - (1 - e^{-\lambda x_i})^\alpha)^{-1}. \tag{12}$$

We need to solve

$$g'(\alpha, \lambda) = g_1'(\alpha, \lambda) + g_2'(\alpha, \lambda) = 0 \quad \Leftrightarrow \quad g_1'(\alpha, \lambda) = -g_2'(\alpha, \lambda), \tag{13}$$

Here $g'(\alpha, \lambda) = \left(\dfrac{\partial g(\alpha, \lambda)}{\partial \alpha}, \dfrac{\partial g(\alpha, \lambda)}{\partial \lambda} \right)$. First solve

$$g_1'(\alpha, \lambda) = 0. \tag{14}$$

using the following fixed point type non-linear equation iteratively

$$\lambda = \left(\frac{1}{n} \sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{(1 - e^{-\lambda x_i})} \left(1 + \frac{n}{\sum_{i=1}^{n} \ln(1 - e^{-\lambda x_i})} \right) + \frac{1}{n} \sum_{i=1}^{n} x_i \right)^{-1}. \tag{15}$$

If $\lambda^{(0)}$ is the solution of (15), then obtain

$$\alpha^{(0)} = -\frac{n}{\sum_{i=1}^{n} \ln(1 - e^{-\lambda^{(0)} x_i})}. \tag{16}$$

Now $\alpha^{(1)}$ and $\lambda^{(1)}$ can be obtained as the solution of the following

$$g_1'(\alpha, \lambda) = -g_2(\alpha^{(0)}, \lambda^{(0)}), \tag{17}$$

similarly, $\alpha^{(2)}$ and $\lambda^{(2)}$ can be obtained as the solution of the following

$$g_1'(\alpha, \lambda) = -g_2(\alpha^{(1)}, \lambda^{(1)}), \tag{18}$$

The iteration continues until converges. Note that the solution $(\widetilde{\alpha}, \widetilde{\lambda})$ of the following equation, for any arbitrary c_1 and c_2

$$g_1'(\alpha, \lambda) = (c_1, c_2) \tag{19}$$

can be obtained as follows. First solve the non-linear equation iteratively

$$\lambda = \left[\frac{c_2}{n} + \frac{1}{n} \sum_{i=1}^{n} x_i + \left(1 - \frac{n}{c_1 - \sum_{i=1}^{n} \ln(1 - e^{-\lambda x_i})} \right) \times \left(\frac{1}{n} \sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{1 - e^{-\lambda x_i}} \right) \right]^{-1} \tag{20}$$

to obtain $\widetilde{\lambda}$, and then obtain

$$\widetilde{\alpha} = \left[\frac{c_1 - \sum_{i=1}^{n} \ln(1 - e^{-\lambda x_i})}{n} \right]^{-1}, \tag{21}$$

see Kannan et al. [13]. Finally for implementation of the EM algorithm we need the following result. The conditional expectation of Z given $X = x$, is

$$E(Z|X = x; \alpha, \lambda, \theta) = \frac{2(1 - (1 - e^{-\lambda x})^\alpha)}{\theta + (1 - \theta)(1 - e^{-\lambda x})^\alpha}. \tag{22}$$

Now we are ready to provide the EM algorithm. Suppose at the k-th stage the value of α, λ and θ are $\alpha^{(k)}$, $\lambda^{(k)}$ and $\theta^{(k)}$ respectively.

E-step': In the E-step obtain the 'pseudo log-likelihood function' (8) replacing z_i by $z_i^{(k)}$, where

$$z_i^{(k)} = E(Z|X = x_i; \alpha^{(k)}, \lambda^{(k)}, \theta^{(k)}); \quad i = 1, \ldots, n. \tag{23}$$

M-step: At the k-th stage, in the M-step, we maximize the 'pseudo-log-likelihood' function with respect to α, λ and θ, to compute the $\alpha^{(k+1)}$, $\lambda^{(k+1)}$ and $\theta^{(k+1)}$. The maximization can be performed, as it has been described before.

9 Data Analysis

For illustrative purposes, in this section we present the analysis of two data sets to show how our proposed model works in practice.

GUINA PIG DATA: This data set has been obtained from Bjerkedal [6]. This data set represents the survival times (in days) of guinea pigs injected with different doses of tuber bacilli. It may be mentioned that guinea pigs have high susceptibility to human tuberculosis, and that is why they are usually used in this kind of study. The data set consists of survival times of 72 animals who were under the regimen 4.342. The regimen number is the common logarithm of the number of bacillary units in 0.5 ml., of challenge solution, *i.e.* regimen 4.342 corresponds to 2.2×10^4 bacillary units per 0.5 ml. ($\log(2.2 \times 10^4) = 4.342$), see Gupta et al. [10].

This data set is available in Gupta et al. [10], and it has been analyzed by them also. The preliminary data analysis by Gupta et al. [10] indicated that the data are right hand skewed, and the empirical hazard function is unimodal. Due to this reason Gupta et al. [10] analyzed the data using the log-normal model. The MLEs of the log-normal parameters are 5.0043 (μ) and 0.6290 (σ) respectively, the associated log-likelihood value is -429.0945. Based on the Kolmogorov-Statistic (KS) distance 0.1298, and the associated p value (0.1765) and also from the quantile plot they claimed that the log-normal model provides a good fit to the data.

Since the proposed MOGE model can have a unimodal hazard function, we analyze the data using the MOGE model also. The MLEs of α, θ and λ are 3.6050, 1.0287 and 0.0113 respectively. The associated 95 % bootstrap confidence intervals are (0.7288, 6.4930), (0.1734, 4.2919), (0.0075, 0.0223) respectively. The corresponding log-likelihood value is -425.8080. The KS distance between the fitted and empirical distribution function is 0.0917, and the associated p values is 0.5803.

For comparison purposes, we have also fitted the Birnbaum-Saunders distribution, which also has unimodal hazard function, see for example Kundu et al. [14]. The MLEs of the unknown Birnbaum-Saunders parameters are 0.7038 (α) and 141.7175 (β), and the associated log-likelihood value is -434.0186. The KS distance between the fitted and the empirical distribution functions is 0.1569 and the associated p value is 0.0576.

If we want to perform the following test:

$$H_0 : \text{log-normal} \quad \text{vs.} \quad H_1 : \text{MOGE}$$

then based on the likelihood ratio test statistic (6.573) the corresponding p value is <0.05, based on the χ_1^2 distribution. Since the p value is quite small we reject the null hypothesis. Similarly, we also reject null hypothesis of the following test:

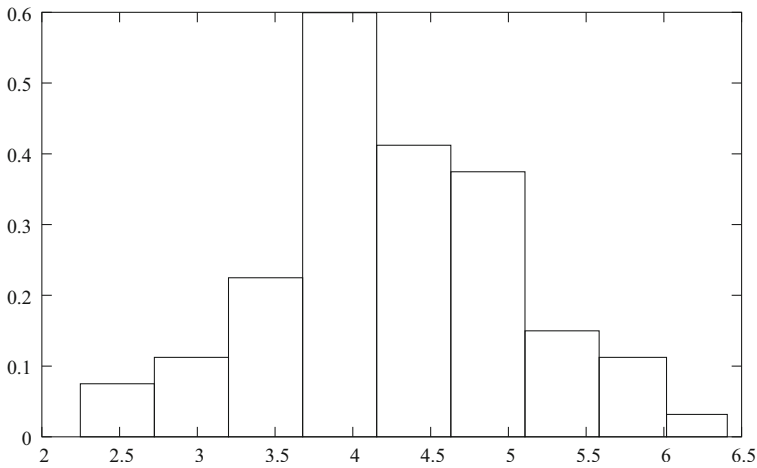$$H_0 : \text{Birnbaum-Saunders} \quad \text{vs.} \quad + H_1 : \text{MOGE}$$

Fig. 3 The histogram of the strength data set

Therefore, based on the KS distance and also based on the likelihood ratio test, we prefer MOGE distribution than log-normal or Birnbaum-Saunders distribution.

STRENGTH DATA

Now we present the analysis of a data set obtained from Prof. R.G. Surles. It is a strength data measured in GPA, the single carbon fibers, and impregnated 1,000-carbon fiber tows. Single fibers were tested under tension at gauge length 1 mm. The data are provided below:

2.247	2.64	2.908	3.099	3.126	3.245	3.328	3.355	3.383	3.572	3.581	3.681	3.726	3.727
3.728	3.783	3.785	3.786	3.896	3.912	3.964	4.05	4.063	4.082	4.111	4.118	4.141	4.246
4.251	4.262	4.326	4.402	4.457	4.466	4.519	4.542	4.555	4.614	4.632	4.634	4.636	4.678
4.698	4.738	4.832	4.924	5.043	5.099	5.134	5.359	5.473	5.571	5.684	5.721	5.998	6.06

Before progressing further first we provide the histogram of the strength data in Fig. 3. It is immediate that the data are unimodal. We further provide the the scaled TTT transform, see [1], of the data set in Fig. 4.

Since the scaled TTT plot is concave, it indicates that the empirical hazard function is an increasing function. We have subtracted 2.0 from all the data points before analyzing the data set. We have used the proposed MOGE model and estimates of α, θ and λ are 1.5759, 67.6793 and 2.0866 respectively. The associated 95% bootstrap confidence intervals are (0.4821, 2.6123), (55.2345, 82.5123) and (1.6041, 2.6704) respectively. The corresponding log-likelihood value is -67.8507. The KS distance between the fitted and the empirical distribution functions is 0.0474 and the associated p value is 0.9996.

For comparison purposes we have fitted two-parameter Weibull, gamma and GE distributions. The MLEs, the corresponding log-likelihood values, the KS distances between the fitted and the empirical distribution functions and the associated p values are reported in Table 1. From the table values

it is clear that between Weibull, gamma and GE distributions Weibull provides the best fit. Now if we want to test the hypothesis

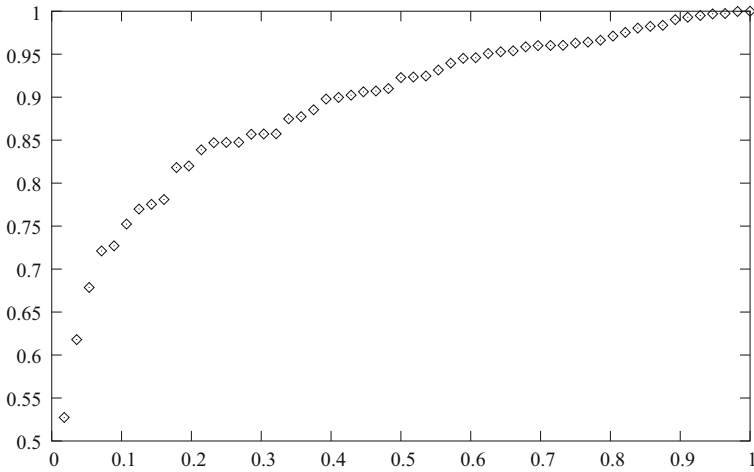$$H_0 : \text{Weibull} \quad \text{vs.} \quad H_1 : \text{MOGE}$$

Fig. 4 The scaled TTT transform of the strength data set

Table 1 Maximum likelihood estimates, maximized log-likelihood values, K-S statistics and the associated p-values for Weibull, gamma and GE distributions while fitting to the strength data

Distribution	Estimates		Log-likelihood	K-S distance	p value
	Shape	Scale			
Gamma	5.9685	2.6401	−71.8824	0.0973	0.6410
Weibull	3.0045	0.3961	−70.3395	0.0648	0.9726
GE	6.9634	1.1193	−74.6607	0.1221	0.3735

then based on the likelihood ratio test, the p value is less than 0.05. Therefore, we reject the null hypothesis. Similarly, if we want to test H_0 : GE or H_0 : Gamma and the alternative in both the cases in H_1 : MOGE, we reject the null hypothesis in both the cases. Therefore, in this case also based on the KS distances, and also based on the likelihood ratio tests, we prefer MOGE distribution, than Weibull, gamma or GE distributions.

10 Conclusions

In this paper we have introduced a new three-parameter distribution by incorporating the Marshall-Olkin method to the generalized exponential distribution. This new three-parameter distribution has an explicit distribution function and the PDF is also in a compact form. It is a very flexible three-parameter distribution, and it can have all possible four different hazard functions depending on the two shape parameters. Finally it should be mentioned that although we have incorporated only the generalized exponential distribution, but many of the properties are valid for a more general class of distributions, namely the proportional reversed hazard class. It will be interesting to see different properties of the general Marshall-Olkin proportional reversed hazard class. More work is needed in that direction.

References

1. Aarset, M.V.: How to identify a bathtub hazard rate? IEEE Trans. Reliab. **36**, 106–108 (1987)
2. Arnold, B.C., Balakrishnan, N., Nagaraja, H.N.: First Course in Order Statistics. Wiley, New York (1992)
3. Barreto-Souza, W., Lemonte, A.J., Cordeiro, G.M.: General results for the Marshall and Olkin's family of distributions. Ann Brazil. Acad. Sci. **85**, 3–21 (2013)
4. Barreto-Souza, W., Santos, A.H.S., Cordeiro, G.M.: The beta generalized exponential distribution. J. Stat. Comput. Simulat. **80**, 159–172 (2010)
5. Bartoszewicz, J.: Stochastic comparisons of random minima and maxima from life distributions. Stat. Probab. Lett. **55**, 107–112 (2001)
6. Bjerkedal, T.: Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. Am. J. Hygenes **72**, 130–148 (1960)
7. Consul, P.C.: On the distrbutions of order statistics for a random sample size. Stat. Neerl. **38**, 249–256 (1984)
8. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Se. B **39**, 1–38 (1977)
9. Franco, M., Balakrishnan, N., Kundu, D., Vivo, J.-M.: Generalized mixture of Weibull distributions. Test **23**, 515–535 (2014)
10. Gupta, R.C., Kannan, N., Raychaudhuri, A.: Analysis of lognormal survival data. Math. Biosci. **139**, 101–115 (1997)
11. Gupta, R.D., Kundu, D.: Generalized exponential distribution. Aust. N. Z. J. Stat. **41**, 173–188 (1999)
12. Hazra, N.K., Nanda, A.K., Shaked, M.: Some aging properties of parallel and series systems with random number of components. Naval Res. Logist. **61**, 238–243 (2014)
13. Kannan, N., Kundu, D., Nair, P., Tripathi, R.C.: The generalized exponential cure rate model with covariates. J. Appl. Stat. **37**, 1625–1636 (2010)
14. Kundu, D., Kannan, N., Balakrishnan, N.: On the hazard function of Birnbaum-Saunders distribution and associated inference. Comput. Stat. Data Anal. **52**, 2692–2702 (2008)
15. Li, X., Zuo, M.J.: Preservation of stochastic orders for random minima and maxima, with applications. Nav. Res. Logist. **51**,332–344
16. Marshall, A.W., Olkin, I.: A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika **84**, 641–652 (1997)
17. Nadarajah, S., Kotz, S.: The beta exponential distribution. Reliab. Eng. Syst. Saf. **91**, 689–697 (2006)
18. Pradhan, B., Kundu, D.: Analysis of interval-censored data with Weibull lifetime distribution. Sankhya. Ser. B **76**, 120–139 (2014)
19. Raqab, M.Z., Kundu, D.: Burr type X distribution: revisited. J. Probab. Stat. Sci. **4**, 179–193 (2006)
20. Rényi A (1961) On a measure of entropy and information. In: Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, vol. 1, Berkeley: University of California Press, pp 547–561
21. Shaked, M., Wong, T.: Stochastic comparisons of random minima and maxima. J. Appl. Probab. **34**, 420–425 (1997)
22. Song, P.X., Fan, Y., Kalbfleish, J.D.: Maximization by parts in likelihood inference (with discussions). J. Am. Stat. Assoc. **100**, 1145–1167 (2005)