An Extended Lindley Distribution with Application to Lifetime Data with Bayesian Estimation

Morad Alizadeh [1,*], Vahid Ranjbar [2], Abbas Eftekharian [3], Omid Kharazmi [4]

[1]*Department of Statistics, Persian Gulf University, Bushehr, 75169, Iran*
[2]*Department of Statistics, Faculty Science, Golestan University, Gorgan, Iran*
[3]*Department of Statistics, School of Sciences, University of Hormozgan, Bandar Abbas, Iran*
[4]*Department of Statistics, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran*

Abstract   A four-parameter extended of Lindley distribution with application to lifetime data is introduced. It is called extended Marshal-Olkin generalized Lindley distribution. Some mathematical properties such as moments, skewness, kurtosis and extreme value are derived. These properties with plots of density and hazard functions are shown the high flexibility of the mentioned distribution. The maximum likelihood estimations of proposed distribution parameters with asymptotic properties of these estimations are examined. A simulation study to investigate the performance of maximum likelihood estimations is presented. Moreover, the performance and flexibility of the new distribution are investigated by comparing with several generalizations of Lindley distribution through two real data sets. Finally, Bayesian analysis and efficiency of Gibbs sampling are provided based on the two real data sets.

1. Introduction

The statistical analysis and modeling of complex real data sets such as lifetime data are incapable to do with well-known standard distributions. This matter is created some grave concerns among researchers who attempt for modeling using new distributions and a lot of attention has been recently paid in researches. Nevertheless, many types researches in their pursuit have attempted for modeling complex real data by introducing new family distributions with three or four parameters. To this end, researchers have recently suggested diverse methods for generating new continuous family distributions in concerning to lifetime data analysis such that these distributions are able to fit various lifetime real data sets that have very complex with a high degree of kurtosis and skewness. Such distributions are usually obtained based on the T-X class which is defined by Alzaatreh et al. [8]. The distribution theory has widely extended using the T-X class. Given the vast amount of papers published after 2013, we can only mention a few of the most recent contributions: Kumaraswamy generated, beta generated, odd log-logistic family.

The most well-known one-parameter statistical distributions that use for modeling lifetime data are as: exponential, gamma, Weibull, Lindley and lognormal. In the recent past, since the cumulative distribution function

---

*Correspondence to: Morad Alizadeh(Email: m.alizadeh@pgu.ac.ir). Department of Statistics, Persian Gulf University, Bushehr, 75169, Iran.

(cdf) and survival function of the gamma and the lognormal do not have any closed form, so most researchers chose exponential, Weibull and Lindley distributions for their researches. Among these distributions, although working with the exponential distribution is easier than that of Weibull and Lindley distributions, but since the hazard rate function (hrf) of exponential distribution is fixed and we usually interest to apply a distribution in which has different shapes of hrf and on the other hand, the Lindley distribution has hrf with increasing, decreasing, unimodal and bathtub shapes, therefore, these reasons can motivate use of Lindley distribution to analyze the real lifetime data. The Lindley distribution as one parameter distribution with only one scale parameter was first presented by Lindley [22] in the Bayesian statistics literature. Accordingly Lindley distribution has hrf with various shapes, so it will be capable of modeling the data with a non-monotonic failure rate as well as with monotonic increasing failure rate. It is used to analyze failure time data sets, particularly for modeling stress-strength reliability. As mentioned widely in the literature, the shape parameters play an important role in flexibility behavior of the distribution and it has caused that researchers to present many extensions of the Lindley distribution. Among these extensions, some works are more excellent than others: Zakerzadeh and Dolati [29] proposed a generalization of the Lindley distribution with three parameters that its hrf can be plotted as increasing, bathtub-shaped and decreasing. A two-parameter generalized Lindley distribution due to Nadarajah et al. [27]. An extension of the Lindley distribution was obtained by Bakouch et al. [12] by considering a particular exponentiation. A new generalized Lindley distribution due to Abouammoh et al. [1] on the basis of the weighted mixture of two gamma distributions. Asgharzadeh et al. [9] investigated the properties of a weighted Lindley distribution that it can be expressed as a mixture of weighted exponential and gamma distributions. Odd Burr Lindley distribution was suggested by Altun et al. [7]. Asgharzadeh et al. [10] proposed Weibull Lindley distribution based on three parameters. Recently, using odd log-logistic family and by considering Lindley distribution as underlying distribution several contributions are presented by Alizadeh et al. [4], [5] and [6]. Also, we can mention [14] and [16] as other contributions in related to the Lindley distribution.

Authors have presented new various families of distributions based on a cdf as $G(\cdot)$ during the last decade, e.g. Kumaraswamy generated, Exponentiated-G, gamma generated, proportional odds and generalized beta generated. In the following, Alizadeh et al. [3] presented another class of distributions which is called extended exponentiated-G (EE-G) family. The cdf of EE-G family is

$$F(x;\boldsymbol{\xi}) = \frac{G(x;\boldsymbol{\theta})^\alpha}{G(x;\boldsymbol{\theta})^\alpha + [1 - G(x;\boldsymbol{\theta})^\gamma]}, \tag{1}$$

where $\boldsymbol{\xi} = (\boldsymbol{\theta}, \alpha, \gamma)$, $G(x;\boldsymbol{\theta})$ is the baseline cdf with a parameter vector $\boldsymbol{\theta}$ and $\alpha, \gamma > 0$ are additional shape parameters. Also, the pdf of the family is

$$f(x;\boldsymbol{\xi}) = \frac{g(x;\boldsymbol{\theta})\, G(x;\boldsymbol{\theta})^{\alpha-1}\, [\alpha + (\gamma - \alpha)G(x;\boldsymbol{\theta})^\gamma]}{\{G(x;\boldsymbol{\theta})^\alpha + [1 - G(x;\boldsymbol{\theta})^\gamma]\}^2}, \tag{2}$$

where $g(x;\boldsymbol{\theta})$ is the baseline pdf. It is clear that the EE-G is a two-parameter family without considering the parameters of the baseline distribution. Since the EE-G family is more flexible than some other families, so this family can be proposed as a serious competitor for other extended families of distributions.

In order to obtain more flexibility in modeling observed data in comparison with other generalizations that already introduced, we introduce another class of distributions using the ideas of EE-G and Marshal-Olkin families. In fact, we can add another parameter, which is called the Marshal-Olkin parameter, to EE-G family. This family is called extended Marshal-Olkin generalized-G (EMOG-G). Based on the T-X class and given a continuous baseline cdf $G(x;\boldsymbol{\theta})$ with a parameter vector $\boldsymbol{\theta}$, the cdf of the EMOG-G family is given by

$$F(x;\boldsymbol{\xi}) = \int_0^{\frac{G(x;\boldsymbol{\theta})^\alpha}{1 - G(x;\boldsymbol{\theta})^\gamma}} \frac{\beta\, dt}{(1 + \beta\, t)^2} = \frac{G(x;\boldsymbol{\theta})^\alpha}{G(x;\boldsymbol{\theta})^\alpha + \beta(1 - G(x;\boldsymbol{\theta})^\gamma)}, \tag{3}$$

where $\beta > 0$ is the Marshal-Olkin parameter and $\boldsymbol{\xi} = (\boldsymbol{\theta}, \alpha, \gamma, \beta)$. It is obviously that in special case, the EMOG-G family reduces to EE-G family when $\beta = 1$. For $\alpha = \gamma = 1$, it reduces to Marshal-olkin family. If $\beta = 1$ and $\alpha = \gamma$, then it reduces to Exp-G family. By considering $\alpha = \beta = \gamma = 1$, we obtain the baseline distribution G.

Since the EMOG-G is a three-parameter family without considering the parameters of the baseline distribution, so use of the baseline distribution with only one parameter such as exponential and Lindley is more efficient than that of with two or three parameters. On the other hand, because of the Lindley distribution has hrf with increasing, decreasing, unimodal and bathtub shapes and is usually applied for the lifetime real data, so we consider the Lindley distribution as the baseline distribution in this paper. Suppose that the cdf of Lindley distribution is as follows

$$G(x;\lambda) = 1 - (1 + \frac{\lambda x}{1 + \lambda})e^{-\lambda x}, \quad x > 0, \tag{4}$$

where $\lambda > 0$ is the scale parameter and pdf corresponding to (4) is given by

$$g(x;\lambda) = \frac{\lambda^2}{1 + \lambda}(1 + x)\,e^{-\lambda x}. \tag{5}$$

By considering $G(x;\boldsymbol{\theta})$ as (4) in Eq. (3) where $\boldsymbol{\theta} = \lambda$, we obtain a new distribution that is called extended Marshal-Olkin generalized Lindley (EMOGL) distribution. In the next section, we present the cdf, pdf and some properties of EMOGL distribution.

The rest of the paper is organized as follows. In Section 2, the EMOGL distribution is introduced. Some properties of the EMOGL distribution such as the nth moment, the nth incomplete moment, moment generating function and Bonferroni and Lorenz curves are presented in Section 3. In Section 4, some asymptotic properties and extreme values of the EMOGL distribution are investigated. The maximum likelihood estimators for the unknown parameters of the EMOGL distribution are obtained in Section 5. A simulation study is presented in Section 6. In Section 7, the applications using two real data sets are reported. Bayesian inference and a Gibbs sampling procedure for the considered data sets are investigated in Section 8. Finally, some conclusions are stated in Section 9.

2. The extended Marshal-Olkin generalized lindley distribution

By substituting (4) in (3), the distribution function of the EMOGL is derived as

$$F(x;\boldsymbol{\xi}) = \frac{\left[1 - (1 + \frac{\lambda x}{1+\lambda})e^{-\lambda x}\right]^\alpha}{\left[1 - (1 + \frac{\lambda x}{1+\lambda})e^{-\lambda x}\right]^\alpha + \beta\left[1 - (1 - (1 + \frac{\lambda x}{1+\lambda})e^{-\lambda x})^\gamma\right]}, \tag{6}$$

and pdf

$$f(x;\boldsymbol{\xi}) = \frac{\beta\,\lambda^2(1+x)e^{-\lambda x}\left[1 - (1 + \frac{\lambda x}{1+\lambda})e^{-\lambda x}\right]^{\alpha-1}\left\{\alpha + (\gamma - \alpha)\left[1 - (1 + \frac{\lambda x}{1+\lambda})e^{-\lambda x}\right]^\gamma\right\}}{\left\{\left[1 - (1 + \frac{\lambda x}{1+\lambda})e^{-\lambda x}\right]^\alpha + \beta\left[1 - (1 - (1 + \frac{\lambda x}{1+\lambda})e^{-\lambda x})^\gamma\right]\right\}^2}. \tag{7}$$

Throughout this paper, the EMOGL distribution with parameters $\boldsymbol{\xi} = (\lambda, \alpha, \beta, \gamma)$ will be denoted by $EMOGL(\lambda, \alpha, \beta, \gamma)$.

Special cases: Let $X \sim EMOGL(\lambda, \alpha, \beta, \gamma)$.

- If $\beta = 1$, then X reduces to the extended generalized Lindley (EG-L).
- For $\alpha = \gamma$, X coincides with Marshall- Olkin generalized Lindley.
- If $\alpha = \gamma = 1$, then X reduces to Marshall- Olkin Lindley.
- By taking $\alpha = \gamma$ and $\beta = 1$, X coincides with generalized Lindley.
- For $\alpha = \beta = \gamma = 1$, X is ordinary Lindley.

Some of the possible shapes of the density function (27) for the selected parameter values are illustrated in Figure 1(a). As seen in Figure 1(a), the density function can take various forms depending on the parameter values. It is evident that the EMOGL distribution is much more flexible than the Lindley distribution, i.e. the additional shape parameters cause high flexibility of the EMOGL distribution. Both unimodal and monotonically decreasing and increasing shapes appear to be possible.

By definition of the survival function, we set $S(x; \boldsymbol{\xi}) = 1 - F(x; \boldsymbol{\xi})$ and then we get

$$S(x; \boldsymbol{\xi}) = \frac{\beta \left[1 - (1 - (1 + \frac{\lambda x}{1+\lambda}) e^{-\lambda x})^{\gamma} \right]}{\left[1 - (1 + \frac{\lambda x}{1+\lambda}) e^{-\lambda x} \right]^{\alpha} + \beta \left[1 - (1 - (1 + \frac{\lambda x}{1+\lambda}) e^{-\lambda x})^{\gamma} \right]}. \tag{8}$$

Now, using (27) and (8) the hrf of EMOGL is given by

$$h(x; \boldsymbol{\xi}) = \frac{\beta \lambda^2 (1+x) e^{-\lambda x} \left[1 - (1 + \frac{\lambda x}{1+\lambda}) e^{-\lambda x} \right]^{\alpha-1} \left\{ \alpha + (\gamma - \alpha) \left[1 - (1 + \frac{\lambda x}{1+\lambda}) e^{-\lambda x} \right]^{\gamma} \right\}}{\left[(1 - (1 + \frac{\lambda x}{1+\lambda}) e^{-\lambda x})^{\alpha} + \beta [1 - (1 - (1 + \frac{\lambda x}{1+\lambda}) e^{-\lambda x})^{\gamma})] \right] \beta \left[1 - (1 - (1 + \frac{\lambda x}{1+\lambda}) e^{-\lambda x})^{\gamma} \right]} \tag{9}$$

By choosing different values of parameters, it observes that the EMOGL distribution has various shapes of the hrf. Some of these shapes are drawn in Figure 1(b). Figure 1(b) shows that the hrf of the EMOGL distribution can have very flexible shapes, such as increasing, decreasing, bathtub followed by upside-down bathtub, and bathtub shapes for the selected values of the model parameters. This attractive flexibility makes the hrf of the EMOGL distribution useful and suitable for non-monotone empirical hazard behaviors which are more likely to be encountered or observed in real life situations.
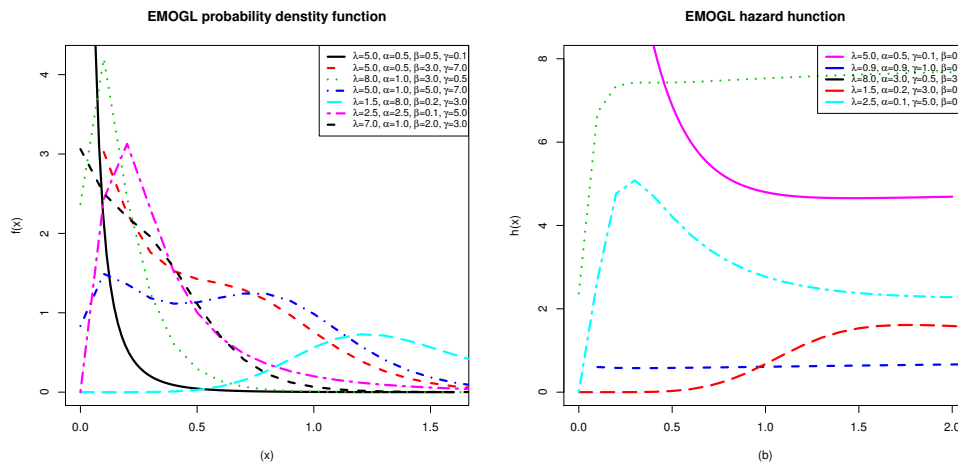

Figure 1. Plots of pdf and hrf of the EMOGL distribution for selected λ, α, γ and β.

2.1. Mixture representations for the pdf and cdf

The cdf and pdf of the EMOGL can be written as mixture representations and such forms of cdf and pdf can be used to derive some mathematical properties, e.g. moments, moments of residual life and incomplete moments. To this purpose, first, let us remind inverse of a power series using the following Remark.

Remark 1
[17, p.18]
Inverse of a power series $\sum_{k=0}^{\infty} b_k x^k$ is

$$\frac{1}{\sum_{k=0}^{\infty} b_k x^k} = \sum_{k=0}^{\infty} c_k x^k,$$

where $c_0 = \frac{1}{b_0}$ and $c_k = -\frac{1}{b_0^2} \sum_{r=1}^{k} c_{k-r} b_r$ for $k \geq 1$.

To obtain the mixture representation of the cdf of EMOGL, note that for any $0 < u < 1$,

$$
\begin{aligned}
u^\alpha &= \sum_{i=1}^\infty (-1)^i \binom{\alpha}{i} (1-u)^i = \sum_{i=1}^\infty \sum_{k=0}^i (-1)^{i+k} \binom{\alpha}{i} \binom{i}{k} u^k \\
&= \sum_{k=0}^\infty \sum_{i=k}^\infty (-1)^{i+k} \binom{\alpha}{i} \binom{i}{k} u^k = \sum_{k=0}^\infty a_k u^k,
\end{aligned}
$$

where $a_k = a_k(\alpha) = \sum_{i=k}^\infty (-1)^{i+k} \binom{\alpha}{i} \binom{i}{k}$. By similar argument, we have

$$
[1 - (1 + \tfrac{\lambda}{1+\lambda}x)e^{-\lambda x}]^\alpha + \beta \left\{ 1 - [1 - (1 + \tfrac{\lambda}{1+\lambda}x)e^{-\lambda x}]^\gamma \right\} = \sum_{k=0}^\infty b_k [1 - (1 + \tfrac{\lambda}{1+\lambda}x)e^{-\lambda x}]^k,
$$

where $b_0 = a_0(\alpha) + \beta - a_0(\gamma)$ and $b_k = a_k(\alpha) - \beta a_k(\gamma)$, for $k \geq 1$. Now, using Remark 1, we get

$$
\begin{aligned}
F(x) &= \frac{[1 - (1 + \tfrac{\lambda}{1+\lambda}x)e^{-\lambda x}]^\alpha}{\sum_{k=0}^\infty b_k [1 - (1 + \tfrac{\lambda}{1+\lambda}x)e^{-\lambda x}]^k} \\
&= \sum_{k=0}^\infty c_k [1 - (1 + \frac{\lambda}{1+\lambda}x)e^{-\lambda x}]^{k+\alpha} \\
&= \sum_{k=0}^\infty c_k G(x; \lambda)^{k+\alpha},
\end{aligned}
\tag{10}
$$

where $c_0 = \frac{1}{b_0}$ and $c_k = -\frac{1}{b_0^2} \sum_{r=1}^k b_r c_{k-r}$, for $k \geq 1$.

The Eq. (10) can be interpreted as a linear combination of generalized Lindley distribution. Using this equation, the mixture representation of pdf is given by

$$
f(x) = \sum_{k=0}^\infty (k + \alpha) c_k \, g(x; \lambda) G(x; \lambda)^{k+\alpha-1}.
\tag{11}
$$

3. Mathematical properties

The formulae obtained in this paper can be simply used in some mathematical and statistical software such as Mathematica, Maple and R. These software be able to deal with complex analytic expressions. To determine mathematical properties of EMOGL distribution, use of some algebraic expansions can be more efficient than computing these properties directly. In what follows, we derive the nth moment, kth central moment and moment generating function of EMOGL distribution. In addition, we provide the nth incomplete moment, mean deviations, Bonferroni and Lorenz curves and present numerical values of skewness and kurtosis using the first four ordinary moments. First of all, assume that $X \sim EMOGL(\lambda, \alpha, \beta, \gamma)$. Using (11), we define

$$
A(a_1, a_2, a_3, a_4; \lambda) = \int_0^\infty x^{a_1} (1+x)^{a_2} e^{-a_3 x} \left[1 - (1 + \frac{\lambda}{1+\lambda} x)e^{-\lambda x} \right]^{a_4} dx.
$$

By using generalized binomial expansion it can be shown that

$$
A(a_1, a_2, a_3, a_4; \lambda) = \sum_{l,r=0}^\infty \sum_{k=0}^l (-1)^l \binom{a_4}{l} \binom{l}{k} \binom{a_2}{r} (\frac{\lambda}{1+\lambda})^l \times \frac{\Gamma(a_1 + 1 + k + r)}{(\lambda l + a_3)^{a_1+1+k+r}}.
\tag{12}
$$

So, the nth moment of EMOGL distribution is given by

$$
E[X^n] = \frac{\lambda^2}{1+\lambda} \sum_{k=0}^\infty (k + \alpha) c_k \, A(n, 1, \lambda, k + \alpha - 1; \lambda).
\tag{13}
$$

The central moments $\mu_k = E(X - \mu)^k$ of EMOGL distribution can be derived from (13) as

$$\mu_k = E(X - \mu)^k = \sum_{r=0}^{k} \binom{k}{r} \mu_r'(-\mu)^{k-r}. \tag{14}$$

where $\mu_k' = E(X^k)$, $\mu = \mu_1' = E(X)$ and k is an integer value.

The mean and variance of X can be particularly obtained using Eq.s (13) and (14). Furthermore, these equations are used to derive the skewness as

$$S = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3}{(\mu_2' - \mu_1'^2)^{3/2}},$$

and the kurtosis as

$$K = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2' - 3\mu_1'^4}{\mu_2' - \mu_1'^2}.$$

It is to highlight that the Eq. (13) can be easily computed numerically using the mathematical or statistical software. For this purpose, one can compute this equation for a large natural number, say N, instead of infinity in the sums. Therefore, several quantities of X such as moments, skewness and kurtosis can be computed numerically using (13). Table 1 shows numerical values of the first four ordinary moments, skewness and kurtosis of the EMOGL distribution for different values of parameters $\lambda, \alpha, \beta, \gamma$. Also, the skewness and kurtosis plots of the EMOGL distribution for selected values of α, β and γ when $\lambda = 2$ are drawn in Figure 2.

Moreover, it is easy to verify that the moment generating function for EMOGL distribution is given by

$$M_X(t) = E\left[e^{tX}\right] = \frac{\lambda^2}{1 + \lambda} \sum_{k=0}^{\infty} (k + \alpha)\, c_k\, A(0, 1, \lambda - t, k + \alpha - 1; \lambda).$$

In order to obtain the nth incomplete moment of the EMOGL distribution let us define

$$B(a_1, a_2, a_3, a_4; y, \lambda) = \int_0^y x^{a_1} (1 + x)^{a_2} e^{-a_3 x} \left[1 - (1 + \frac{\lambda}{1+\lambda} x)e^{-\lambda x}\right]^{a_4} dx$$

$$= \sum_{l,r=0}^{\infty} \sum_{k=0}^{l} (-1)^l \binom{a_4}{l}\binom{l}{k}\binom{a_2}{r} (\frac{\lambda}{1+\lambda})^l \times \frac{\gamma(a_1 + 1 + k + r, \frac{y}{\lambda l + a_3})}{(\lambda l + a_3)^{a_1 + 1 + k + r}}, \tag{15}$$

where $\gamma(\lambda, z) = \int_0^z t^{\lambda - 1} e^{-t}\, dt$ stands for the incomplete gamma function. Note that the second equality of (15) is obtained by generalized binomial expansion. Hence, using (15) the nth incomplete moment of the EMOGL distribution is derived by

$$m_n(y) = E\left[X^n\, |X < y\right] = \frac{\lambda^2}{1 + \lambda} \sum_{k=0}^{\infty} (k + \alpha)\, c_k\, B(n, 1, \lambda, k + \alpha - 1, y; \lambda). \tag{16}$$

In what follows, we provide two measures of deviation, i.e. mean deviation about the mean (δ_1) and the mean deviation about the median (δ_2). By definition of these measures, it is easy to show that

$$\delta_1(X) = 2\mu\, F(\mu) - 2 \int_0^\mu x\, f(x)\, dx,$$

and

$$\delta_2(X) = \mu - 2 \int_0^M x\, f(x)\, dx,$$

where M denotes the median of X. Therefore, it can be verified that measures $\delta_1(X)$ and $\delta_2(X)$ are given by

$$\delta_1(X) = 2\mu\, F(\mu) - \frac{\lambda^2}{1 + \lambda} \sum_{k=0}^{\infty} (k + \alpha) c_k\, A(\beta, 1, \lambda, k + \alpha - 1; \lambda, \beta),$$

and

$$\delta_2(X) = \mu - \frac{2\beta\lambda^2}{1 + \lambda} \sum_{k=0}^{\infty} (k + \alpha) c_{k+1}\, B(\beta, 1, \lambda, k + \alpha - 1; M, \lambda, \beta).$$

Table 1. Moments, skewness, and kurtosis of EMOGL distribution for some parameter values.

| α	β	γ	λ	μ'_1	μ'_2	μ'_3	μ'_4	$Skewness$	$Kurtosis$
0.5	0.5	0.5	0.5	1.396277	5.928780	40.532130	373.99553	2.663484	51.660924
0.5	0.5	0.5	2.0	0.263184	0.253219	0.4109636	0.9153027	3.136901	3.1176889
0.5	0.5	2.0	0.5	2.452451	15.01126	125.69743	1298.3496	1.658489	55.405829
0.5	0.5	2.0	2.0	0.495022	0.686221	1.3388687	3.2811819	1.919220	3.3068528
0.5	2.0	0.5	0.5	3.067353	17.58323	138.73628	1377.3817	1.482694	49.202303
0.5	2.0	2.0	2.0	1.063595	1.902021	4.2792546	11.456124	0.911271	3.0115553
3.0	1.5	0.5	0.5	4.224679	22.60495	153.88846	1329.7518	1.753745	40.844270
1.5	0.5	2.0	1.5	0.959045	1.622247	3.9336620	12.353666	1.750124	5.2358174
2.0	2.5	1.5	3.0	0.789251	0.858329	1.1827318	1.9781510	1.170519	1.2242069
2.0	0.5	0.5	1.0	1.144500	2.066153	5.6910130	22.502296	2.425476	9.9698970
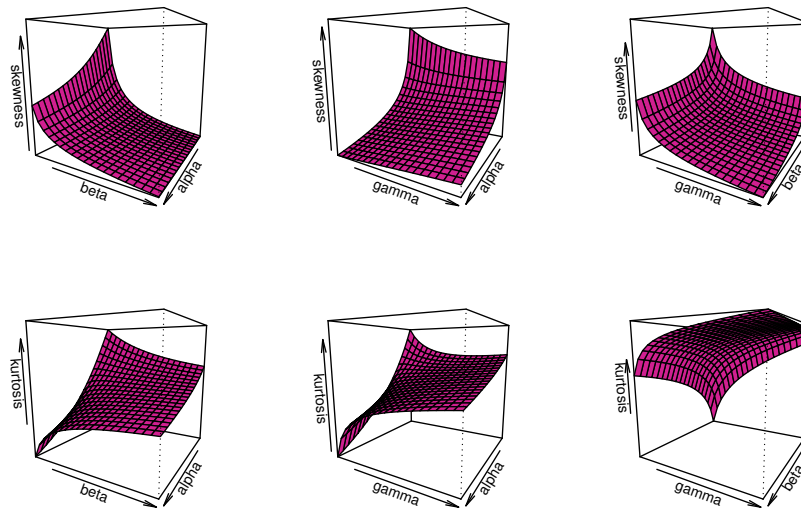
Figure 2. The skewness and kurtosis plots of EMOGL distribution for selected values of α, β and γ when $\lambda = 2$.

3.1. Bonferroni and Lorenz curves

Bonferroni and Lorenz curves were first presented by Bonferroni [13] and Lorenz [23] to measure the inequality of the distribution for a random variable, respectively. These curves are widely used in the literature related to some fields such as reliability, economics, insurance, and etc. The Bonferroni and Lorenz indexes are defined as

$$B(p) = \frac{1}{p\mu} \int_0^q x f(x) \mathrm{d}x,$$

and

$$L(p) = \frac{1}{\mu} \int_0^q x f(x) \mathrm{d}x,$$

respectively, where $q = F^{-1}(p)$ is the quantile function. If $X \sim EMOGL(\lambda, \alpha, \beta, \gamma)$, then it can be shown that the Bonferroni curve of the EMOGL distribution is as follows

$$B(p) = \frac{1}{\mu\, p} \times \frac{\beta\, \lambda^2}{1+\lambda} \sum_{k=0}^{\infty} (k+\alpha) c_k\, B(\beta, 1, \lambda, k+\alpha-1; q, \lambda, \beta).$$

Moreover, the Lorenz curve of the EMOGL distribution can be easily demonstrated as

$$L(p) = \frac{1}{\mu} \times \frac{\beta \lambda^2}{1 + \lambda} \sum_{k=0}^{\infty} (k + \alpha) c_k \, B(\beta, 1, \lambda, k + \alpha - 1; q, \lambda, \beta).$$

In order to use the Lorenz curve as a measure of inequality of income, one should investigate the area between the Lorenz curve and the line $L(p) = p$ that is called the area of concentration and such area is important in economics, reliability, insurance, and medicine.

4. Asymptotic properties and extreme value

In this section, we present some asymptotic properties and extreme values of EMOGL distribution. For simplicity, assume that $F(x)$, $f(x)$ and $h(x)$ are cdf, pdf, and hrf of EMOGL distribution, respectively. If $X_1, ..., X_n$ is a random sample from EMOGL distribution, then asymptotic properties and extreme values based on the EMOGL distribution are obtained as follows.

4.1. Asymptotic properties

The asymptotic of cdf, pdf and hrf of the EMOGL distribution as $x \to 0$ are, respectively, given by

$$F(x) \sim \frac{(\lambda x)^{\alpha}}{\beta} \quad as \quad x \to 0,$$

$$f(x) \sim \frac{\alpha \lambda^{\alpha} x^{\alpha - 1}}{\beta} \quad as \quad x \to 0,$$

$$h(x) \sim \frac{\alpha \lambda^{\alpha} x^{\alpha - 1}}{\beta} \quad as \quad x \to 0.$$

The asymptotic of cdf, pdf and hrf of the EMOGL distribution as $x \to \infty$ are, respectively, as follows

$$1 - F(x) \sim \frac{\beta \gamma \lambda}{1 + \lambda} x e^{-\lambda x} \quad as \quad x \to \infty,$$

$$f(x) \sim \frac{\beta \gamma \lambda^2}{1 + \lambda} x e^{-\lambda x} \quad as \quad x \to \infty,$$

$$h(x) \sim \lambda \quad as \quad x \to \infty.$$

These equations show the effect of parameters on the tails of the EMOGL distribution.

4.2. Extreme value

If $\bar{X} = (X_1 + ... + X_n)/n$ denotes the sample mean, then by the usual central limit theorem, $\sqrt{n}(\bar{X} - E(X))/\sqrt{\mathrm{Var}(X)}$ approaches the standard normal distribution as $n \to \infty$. One may be interested in the asymptotic of the extreme values $M_n = \max(X_1, ..., X_n)$ and $m_n = \min(X_1, ..., X_n)$. Let $\tau(x) = \frac{1}{\lambda}$, we obtain following equations for the cdf in (6) as

$$\lim_{t \to 0} \frac{F(t\,x)}{F(t)} = \lim_{t \to 0} \frac{G(t\,x)^{\alpha}}{G(t)^{\alpha}} = \lim_{t \to 0} \frac{\left[1 - \left(1 + \frac{\lambda t x}{1+\lambda}\right) e^{-\lambda t x}\right]^{\alpha}}{\left[1 - \left(1 + \frac{\lambda t}{1+\lambda}\right) e^{-\lambda t}\right]^{\alpha}} = \lim_{t \to 0} \frac{\left[1 - e^{-\lambda t x}\right]^{\alpha}}{\left[1 - e^{-\lambda t}\right]^{\alpha}}$$

$$= \lim_{t \to 0} \frac{(\lambda\,t\,x)^{\alpha}}{(\lambda\,t)^{\alpha}} = x^{\alpha} \tag{17}$$

and

$$\lim_{t \to \infty} \frac{1 - F(t + x\,\tau(t))}{1 - F(t)} = \lim_{t \to \infty} \frac{1 - G(t + x\,\tau(t))^{\gamma}}{1 - G(t)^{\gamma}} = \mathrm{e}^{-x}. \tag{18}$$

Thus, from [20], there must be norming constants $a_n > 0$, $\quad b_n$, $\quad c_n > 0$ and d_n such that

$$Pr\left[a_n(M_n - b_n) \le x\right] \to \mathrm{e}^{-\mathrm{e}^{-\mathrm{x}}},$$

and

$$Pr\left[c_n(m_n - d_n) \le x\right] \to 1 - \mathrm{e}^{-\mathrm{x}^{\alpha}},$$

as $n \to \infty$. The form of the norming constants can also be determined. For instance, using Corollary 1.6.3 in [20], one can see that $b_n = F^{-1}(1 - \frac{1}{n})$ and $a_n = \lambda$, where $F^{-1}(\cdot)$ denotes the inverse function of $F(\cdot)$.

5. Estimation

Point estimation is the first step of statistical inference on the unknown parameters of the underlying population. In order to find point estimations there are different methods which the most well-known of them is the maximum likelihood estimation (MLE) method. In the present paper, we obtain the MLEs of the unknown parameters of the EMOGL distribution. To this end, let $X_1, X_2, ..., X_n$ be a random sample from $EMOGL(\lambda, \alpha, \beta, \gamma)$ with observed values as $x_1, x_2, ..., x_n$. According to the sample, it is easy to see that the log-likelihood function of $(\alpha, \gamma, \lambda, \beta)$ is as

$$
\begin{aligned}
\ell_n &= n\log(\frac{\beta\lambda^2}{1+\lambda}) + \sum_{i=1}^{n}\log(1+x_i) - \lambda\sum_{i=1}^{n}x_i + (\alpha-1)\sum_{i=1}^{n}\log k_i \\
&\quad + \sum_{i=1}^{n}\log(\alpha + (\gamma-\alpha)k_i^{\alpha}) - 2\sum_{i=1}^{n}\log(k_i^{\alpha} + \beta(1-k_i^{\gamma})),
\end{aligned}
$$

where

$$k_i = 1 - (1 + \frac{\lambda}{1+\lambda}x_i)e^{-\lambda x_i}.$$

By differentiating from the log-likelihood function with respect to the parameters α, γ, λ and β and after simple algebra, we have

$$\frac{\partial \ell_n}{\partial \alpha} = \sum_{i=1}^{n}\log k_i + \sum_{i=1}^{n}\frac{1-k_i^{\gamma}}{\alpha+(\gamma-\alpha)k_i^{\gamma}} - 2\sum_{i=1}^{n}\frac{k_i^{\alpha}\log(k_i)}{k_i^{\alpha}+\beta(1-k_i^{\gamma})}, \tag{19}$$

$$\frac{\partial \ell_n}{\partial \beta} = \frac{n}{\beta} - 2\sum_{i=1}^{n}\frac{1-k_i^{\gamma}}{k_i^{\alpha}+\beta(1-k_i^{\gamma})}, \tag{20}$$

$$\frac{\partial \ell_n}{\partial \gamma} = \sum_{i=1}^{n}\frac{k_i^{\gamma}}{\alpha+(\gamma-\alpha)k_i^{\gamma}} + \sum_{i=1}^{n}\frac{(\gamma-\alpha)k_i^{\gamma}\log(k_i)}{\alpha+(\gamma-\alpha)k_i^{\gamma}} + 2\beta\sum_{i=1}^{n}\frac{k_i^{\gamma}\log(k_i)}{k_i^{\alpha}+\beta(1-k_i^{\gamma})}, \tag{21}$$

$$
\begin{aligned}
\frac{\partial \ell_n}{\partial \lambda} &= \frac{2n}{\lambda} - \frac{n}{1+\lambda} - \sum_{i=1}^{n}x_i + (\alpha-1)\sum_{i=1}^{n}\frac{k_i^{(\lambda)}}{k_i} + \sum_{i=1}^{n}\frac{\gamma(\gamma-\alpha)k_i^{(\lambda)}k_i^{\gamma-1}}{\alpha+(\gamma+\alpha)k_i^{\gamma}} \\
&\quad - 2\sum_{i=1}^{n}\frac{\alpha k_i^{(\lambda)}k_i^{\alpha-1} - \beta\gamma k_i^{(\lambda)}k_i^{\gamma-1}}{k_i^{\alpha}+\beta(1-k_i^{\gamma})},
\end{aligned}
\tag{22}
$$

where

$$k_i^{(\lambda)} = \frac{\partial k_i}{\partial \lambda} = x_i e^{-\lambda x_i}\left[1 + \frac{\lambda}{\lambda+1}x_i + \frac{1}{(1+\lambda)^2}\right].$$

We denote the MLEs of $(\alpha, \gamma, \lambda, \beta)$, by $(\widehat{\alpha}, \widehat{\gamma}, \widehat{\lambda}, \widehat{\beta})$, where these estimations are derived by solving simultaneously the equations $\frac{\partial \ell_n}{\partial \alpha} = 0; \frac{\partial \ell_n}{\partial \gamma} = 0; \frac{\partial \ell_n}{\partial \lambda} = 0; \frac{\partial \ell_n}{\partial \beta} = 0$. Since Eq.s (19)-(22) are non-linear with respect to parameters, so we can

not obtain the explicit expression of $(\widehat{\alpha}, \widehat{\gamma}, \widehat{\lambda}, \widehat{\beta})$. However, by using some numerical iterative techniques we can solve these equations and compute the global maxima of the log-likelihood. Moreover, using Eq.s (19)-(22), we can present the Fisher information matrix as follows

$$I(\theta) = - \begin{bmatrix} I_{\alpha\alpha} & I_{\alpha\gamma} & I_{\alpha\lambda} & I_{\alpha\beta} \\ I_{\gamma\alpha} & I_{\gamma\gamma} & I_{\gamma\lambda} & I_{\gamma\beta} \\ I_{\lambda\alpha} & I_{\lambda\gamma} & I_{\lambda\lambda} & I_{\lambda\beta} \\ I_{\beta\alpha} & I_{\beta\gamma} & I_{\beta\lambda} & I_{\beta\beta} \end{bmatrix}$$

where $I_{ii} = \frac{\partial^2 \ell_n}{\partial i^2}$; $I_{ij} = \frac{\partial^2 \ell_n}{\partial i \partial j}$ for any $i, j = \alpha, \gamma, \lambda, \beta$. According to importance of the Fisher information matrix for interval estimating, we follow asymptotical properties of the MLEs. Let $\theta = (\alpha, \gamma, \lambda, \beta)^T$, then under standard regularity conditions (see [21, pp. 461-463]), $\sqrt{n}(\theta - \widehat{\theta}) \to^d N_4(0, K(\theta)^{-1})$, where $\widehat{\theta}$ and $K(\theta)$ stand for the MLE of θ and the expected Fisher information matrix, respectively. The asymptotic behavior will be established provided that $K(\theta)$ is substituted by the observed Fisher information matrix which is multiplied by $1/n$, that is $I(\theta)/n$, approximated by $\widehat{\theta}$, i.e. $I(\widehat{\theta})/n$.

6. Simulation study

The performance of the maximum likelihood method is evaluated for estimating the EMOGL parameters using a Monte Carlo simulation study. The mean squared error (MSE) and the bias of the parameter estimates are calculated. We are generated $N = 10,000$ samples of sizes $n = 50, 55, ..., 500$ from the EMOGL distribution with $\alpha = 0.5$, $\beta = 2$, $\gamma = 1.5$, $\lambda = 2.5$. We are applied the following Algorithm 1 for generating random data from EMOGL distribution.

Algorithm 1.

- Step 1. Generate random numbers u_i from $U \sim U(0,1)$ for $i = 1, \cdots, n$.
- Step 2. Select arbitrary values for parameters of EMOGL distribution, i.e. α, β, γ and λ.
- Step 3. Solve numerically the non-linear equation

$$u_i = \frac{\left[1 - (1 + \frac{\lambda x_i^\beta}{1+\lambda}) e^{-\lambda x_i^\beta}\right]^\alpha}{\left[1 - (1 + \frac{\lambda x_i^\beta}{1+\lambda}) e^{-\lambda x_i^\beta}\right]^\alpha + \beta \left[1 - (1 + \frac{\lambda x_i^\beta}{1+\lambda}) e^{-\lambda x_i^\beta}\right]^\gamma}, \tag{23}$$

and compute values of x_i for $i = 1, \cdots, n$.
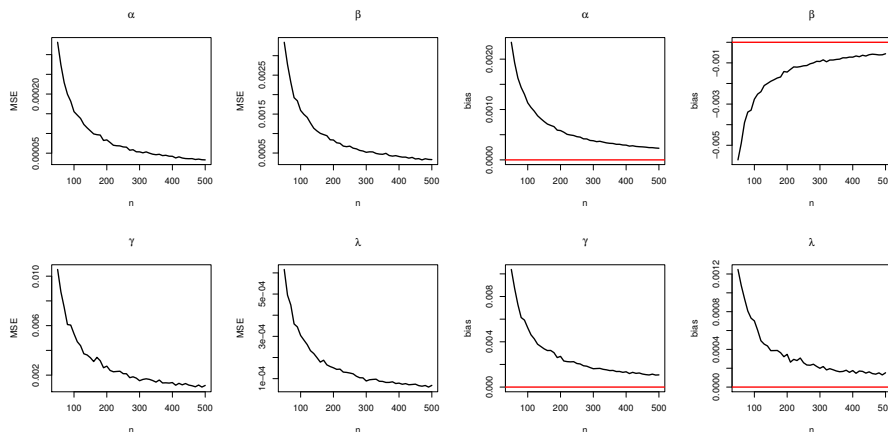- Step 4. Repeat Steps 1 to 3 for N times.

Figure 3. Estimated biases and MSEs for the selected parameter values.

To compute the MSE and bias of the MLEs, let $(\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}, \widehat{\lambda})$ be the MLEs of the new model parameters and $(s_{\widehat{\alpha}}, s_{\widehat{\beta}}, s_{\widehat{\gamma}}, s_{\widehat{\lambda}})$ be the standard errors of the MLEs. The estimated biases and MSEs are given by

$$\widehat{Bias}_\xi(n) = \frac{1}{N} \sum_{i=1}^{N} (\widehat{\xi}_i - \xi)$$

and

$$\widehat{MSE}_\xi(n) = \frac{1}{N} \sum_{i=1}^{N} (\widehat{\xi}_i - \xi)^2,$$

for $\xi = \alpha, \beta, \gamma, \lambda$. Figure 3 displays the numerical results for the above measures. From Figure 3, the following results are deduced:

✓ The estimated biases decrease when the sample size n increases,
✓ The estimated MSEs tend to zero as n increases,

These results reveal the consistency property of the MLEs.

7. Application

In this section, we investigate the performance of fitting the $EMOGL \sim (\lambda, \alpha, \beta, \gamma)$ model for two real data sets. The first data set is the exceedances of flood peaks (in m^3/s) of the Wheaton River near Carcross in Yukon Territory, Canada, and the second data set is the failure times components. Furthermore, we consider several other families that proposed in the literature to compare with the obtained results in this paper. These families are as follows:

- Lindley Distribution, $F_L(x; \lambda) = 1 - (1 + \frac{\lambda x}{1+\lambda}) \mathrm{e}^{-\lambda x}$.
- Power Lindley distribution, $F_{PL}(x; \lambda, \beta) = 1 - (1 + \frac{\lambda x^\beta}{1+\lambda}) \mathrm{e}^{-\lambda x^\beta}$.
- Generalized Lindley, $GL(\alpha, \lambda)$, (see [27]), with distribution function given by

$$F_{GL}(x; \lambda, \alpha) = \left(1 - (1 + \frac{\lambda x}{1+\lambda}) \mathrm{e}^{-\lambda x}\right)^\alpha.$$

- Beta Lindley [25], $BL(\alpha, \beta, \lambda)$, with distribution function

$$F_{BL}(x; \alpha, \beta, \lambda) = \frac{1}{B(\alpha, \beta)} \int_0^{1-(1+\frac{\lambda x}{1+\lambda})\mathrm{e}^{-\lambda x}} t^{\alpha-1}(1-t)^{\beta-1} dt,$$

where $B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\alpha-1} dt$ denotes the beta function.
- Exponentiated power Lindley distribution [11], $EPL(\alpha, \beta, \lambda)$, with distribution function as

$$F_{EPL}(x; \alpha, \beta, \lambda) = \left(1 - (1 + \frac{\lambda x^\beta}{1+\lambda}) \mathrm{e}^{-\lambda x^\beta}\right)^\alpha.$$

- Odd log-logistic power Lindley distribution [3], $OLL - PL(\alpha, \beta, \lambda)$, with distribution function given by

$$F_{OLL-PL}(x; \alpha, \beta, \lambda) = \frac{(1 - (1 + \frac{\lambda x^\beta}{1+\lambda})\mathrm{e}^{-\lambda x^\beta})^\alpha}{(1 - (1 + \frac{\lambda x^\beta}{1+\lambda})\mathrm{e}^{-\lambda x^\beta})^\alpha + [1 - (1 - (1 + \frac{\lambda x^\beta}{1+\lambda})\mathrm{e}^{-\lambda x^\beta})]^\alpha}.$$

- Kumaraswamy Power Lindley [24], $KwPL(\alpha, \beta, \gamma, \lambda)$, has distribution function as follows

$$F_{KwPL}(x; \alpha, \beta, \gamma, \lambda) = 1 - \left[1 - (1 - (1 + \frac{\lambda x^\beta}{1+\lambda})\mathrm{e}^{-\lambda x^\beta})^\alpha\right]^\gamma.$$

- Odd Burr-Power Lindley [7], $OBu - PL(\alpha, \beta, \gamma, \lambda)$, with distribution function

$$F_{OBu-PL}(x; \alpha, \beta, \gamma, \lambda) = 1 - \left(1 - \frac{(1 - (1 + \frac{\lambda x^\beta}{1+\lambda})e^{-\lambda x^\beta})^\alpha}{(1 - (1 + \frac{\lambda x^\beta}{1+\lambda})e^{-\lambda x^\beta})^\alpha + (1 - (1 - (1 + \frac{\lambda x^\beta}{1+\lambda})e^{-\lambda x^\beta}))^\alpha} \right)^\beta.$$

- Extended generalized Lindley [28], $EGL(\alpha, \gamma, \lambda)$, has distribution function as

$$F_{EGL}(x; \alpha, \gamma, \lambda) = \frac{(1 - (1 + \frac{\lambda x}{1+\lambda})e^{-\lambda x})^\alpha}{(1 - (1 + \frac{\lambda x}{1+\lambda})e^{-\lambda x})^\alpha + 1 - (1 - (1 + \frac{\lambda x}{1+\lambda})e^{-\lambda x})^\gamma}.$$
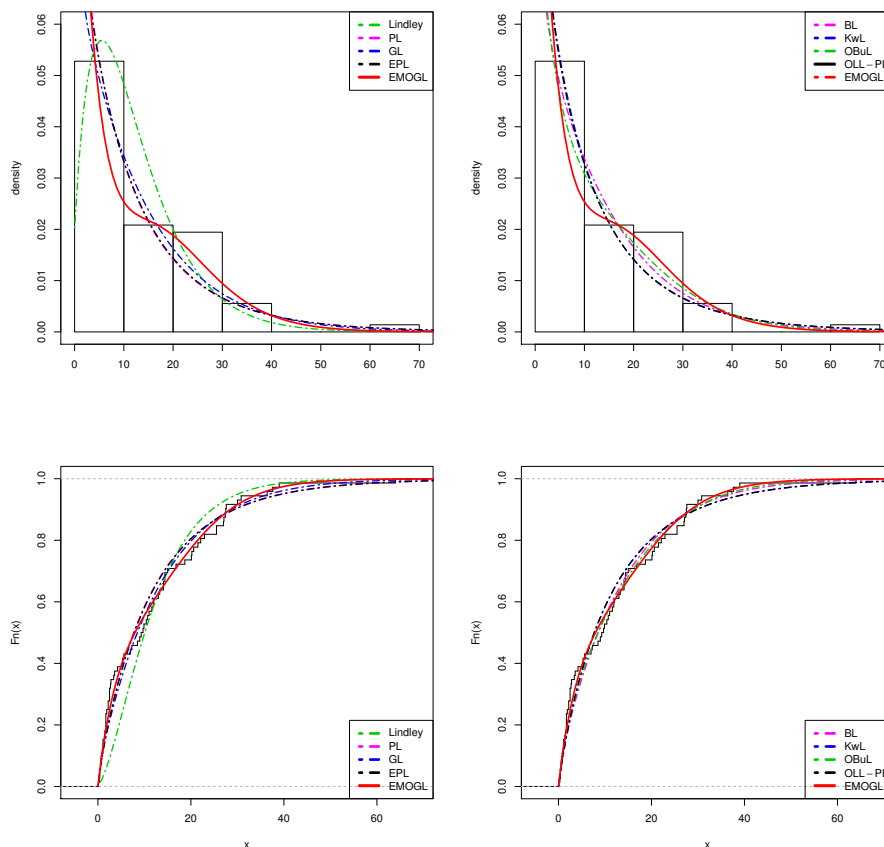

Figure 4. Fitted densities and distribution functions for the first data set.

For both data sets, the MLEs of unknown parameters, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Cramer Von Mises and Anderson-Darling statistics (W^* and A^*) are computed for the EMOGL and all the above distributions. Moreover, the Kolmogorov-Smirnov (K-S) statistic with its corresponding p-value and the maximized log-likelihood function ($l(\hat{\boldsymbol{\xi}}, x)$) are considered. All of the computations were carried out using the software R.

7.1. First application

The first data are the exceedances of flood peaks (in m^3/s) of the Wheaton River near Carcross in Yukon Territory, Canada. The data consist of 72 exceedances for the years 1958-1984 rounded to one decimal place and exist in Table 2. These data were analyzed in [2].

Table 2. The first data set.

1.70	2.20	14.4	1.10	0.40	20.6	5.30	0.70	1.90	13.0	12.0	9.30	1.40
18.7	8.50	25.5	11.6	14.1	22.1	1.10	2.50	14.4	1.70	37.6	0.60	2.20
39.0	0.30	15.0	11.0	7.30	22.9	1.70	0.10	1.10	0.60	9.00	1.70	7.00
20.1	0.40	2.80	14.1	9.90	10.4	10.7	30.0	3.60	5.60	30.8	13.3	4.20
25.5	3.40	11.9	21.5	27.6	36.4	2.70	64.0	1.50	2.50	27.4	1.00	27.1
20.2	16.8	5.30	9.70	27.5	2.50	27.0						

Table 3. Parameter ML estimates and their standard errors (in parentheses) for the first data set.

$Model$	α	β	γ	λ
$Lindley(\lambda)$	–	–	–	0.153 (0.0128)
$GL(\alpha, \lambda)$	0.508 (0.0767)	–	–	0.104 (0.01491)
$PL(\beta, \lambda)$	–	0.700 (0.0570)	–	0.338 (0.0559)
$BL(\alpha, \beta, \lambda)$	0.555 (0.0983)	0.274 (0.2397)	–	0.333 (0.2723)
$EPL(\alpha, \beta, \lambda)$	0.730 (0.2351)	0.915 (0.5956)	–	0.300 (0.2791)
$OLLPL(\alpha, \beta, \lambda)$	0.557 (0.1781)	1.073 (0.2447)	–	0.154 (0.0915)
$KwPL(\alpha, \beta, \gamma, \lambda)$	1.675 (2.4335)	0.453 (0.4323)	7.563 (11.7366)	0.279 (0.5225)
$OBuPL(\alpha, \beta, \lambda)$	0.611 (0.0680)	0.904 (0.4550)	–	0.200 (0.0926)
$EGL(\alpha, \gamma, \lambda)$	0.618 (0.1018)	–	2.770 (1.7047)	0.169 (0.0288)
$EMOGL(\alpha, \gamma, \lambda)$	0.737 (0.1158)	0.517 (0.1706)	3.574 (1.8853)	0.152 (0.0288)

The MLEs of unknown parameters and the goodness-of-fit test statistics for the first data set are presented in Tables 3 and 4, respectively. One can easily observe that the smallest values of AIC, BIC, A^*, W^* and $-l$ statistics and the largest p-value belong to the EMOGL distribution. Therefore, the EMOGL distribution has better performance than the other considered competitors based on these criteria.

Table 4. Goodness-of-fit test statistics for the first data set.

$Model$	AIC	BIC	$p-value$	W^*	A^*	$-l$
$Lindley(\lambda)$	530.423	532.700	0.001	0.139	0.852	264.211
$GL(\alpha, \lambda)$	509.349	513.902	0.276	0.132	0.822	252.674
$PL(\beta, \lambda)$	508.443	512.996	0.405	0.123	0.766	252.103
$BL(\alpha, \beta, \lambda)$	510.206	517.036	0.297	0.150	0.866	252.221
$EPL(\alpha, \beta, \lambda)$	510.425	517.255	0.395	0.147	0.854	252.212
$OLLPL(\alpha, \beta, \lambda)$	507.937	514.767	0.471	0.093	0.592	250.968
$KwPL(\alpha, \beta, \gamma, \lambda)$	512.221	521.328	0.371	0.152	0.866	252.110
$OBuPL(\alpha, \beta, \lambda)$	507.990	514.820	0.401	0.098	0.617	250.995
$EGL(\alpha, \gamma, \lambda)$	508.931	515.761	0.174	0.101	0.662	251.465
$EMOGL(\alpha, \gamma, \lambda)$	506.768	515.875	0.780	0.056	0.365	248.163

In addition, the likelihood ratio (LR) test is applied to compare the EMOGL distribution with its sub-models. For example, a testing hypothesis of $H_0 : \beta = 1$ versus $H_1 : \beta \neq 1$ is equivalent to compare the EMOGL distribution with EGL distribution. To investigate this testing hypothesis, the LR statistic can be computed by the following equation

$$LR = 2 \left[l(\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}, \widehat{\lambda}) - l(\widehat{\alpha^*}, 1, \widehat{\gamma^*}, \widehat{\lambda^*}) \right],$$

Table 5. The LR test results for the first data set.

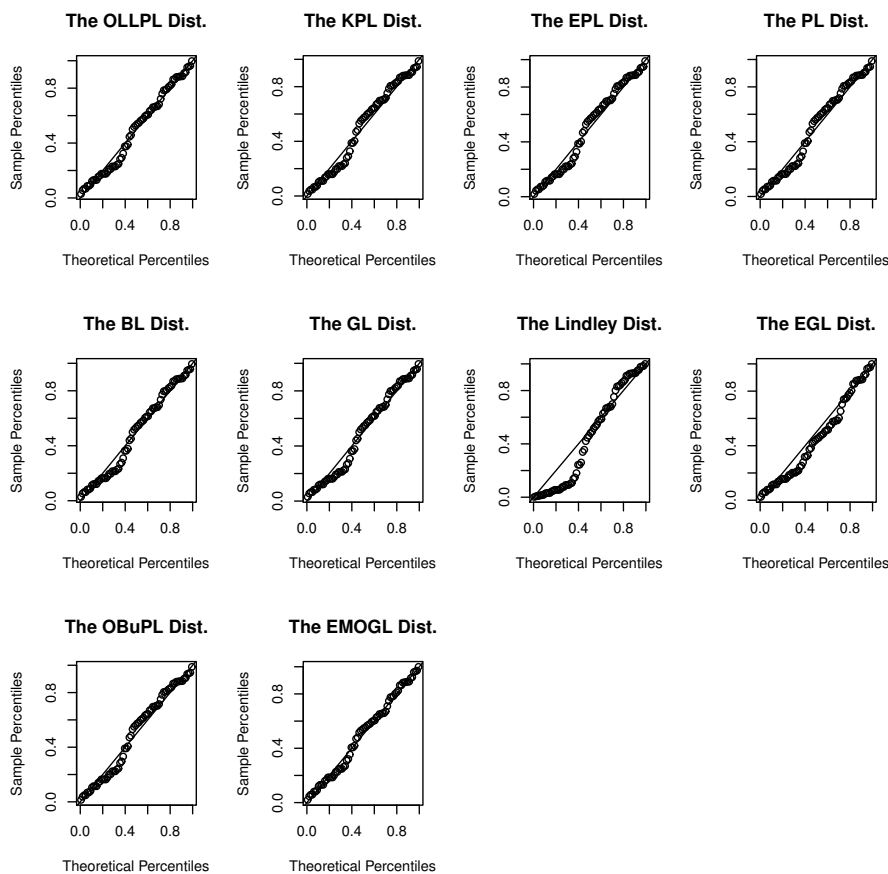|	Hypotheses	LR	p-value
EMOGL versus Lindley	$H_0 : \alpha = \beta = \gamma = 1$	34.0966	< 0.0001
EMOGL versus GL	$H_0 : \alpha = \gamma, \ \beta = 1$	11.0223	0.004
EMOGL versus MOL	$H_0 : \alpha = \gamma = 1$	24.2439	< 0.0001
EMOGL versus MOGL	$H_0 : \alpha = \gamma$	7.6962	0.0055
EMOGL versus EGL	$H_0 : \beta = 1$	4.0458	0.0442

Figure 5. Probability plots for the first data set.

where $\widehat{\alpha^*}, \widehat{\gamma^*}$, and $\widehat{\lambda^*}$ are the MLEs of α, γ, and λ, obtained under H_0, respectively. Under H_0 and by considering the regularity conditions, the LR test statistic converges to χ_r^2 in distribution, where r is equal to difference between the number of parameters estimated under H_0 and the number of parameters estimated in general, that is, under $H_0 : \beta = 1$, we have $r = 1$.

   Table 5 gives values of the LR statistics with their corresponding p-values. From Table 5, we can see that the obtained p-values are too small so, we can reject all the null hypotheses. In other words, based on the LR criterion we conclude that fitting of the EMOGL has better performance than the considered sub-models for the first data set. Also, the fitted pdfs, cdfs and P-P plots of the considered models for the sake of visual comparison are provided in Figures 4 and 5, respectively. Figure 4 suggests that the EMOGL fits the skewed data very well. Figures 5 shows that the plotted points for the EMOGL distribution best capture the diagonal line in the probability plots. Therefore, the EMOGL distribution can be considered as an appropriate model for fitting the first data set.
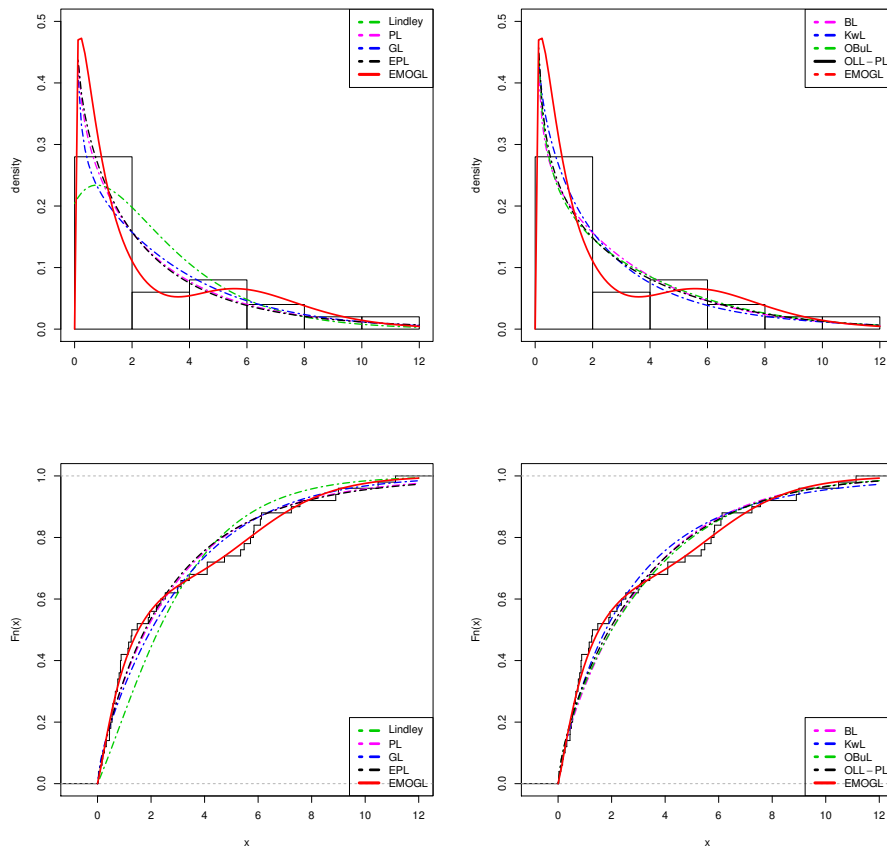
Figure 6. Fitted densities and distribution functions for the second data set.

Table 6. The second data set.

0.032	0.035	0.104	0.169	0.196	0.260	0.326	0.445	0.449	0.496	0.543
0.544	0.577	0.648	0.666	0.742	0.757	0.808	0.857	0.858	0.882	1.138
1.163	1.256	1.283	1.484	1.897	1.944	2.201	2.365	2.531	2.994	3.118
3.424	4.097	4.100	4.744	5.346	5.479	5.716	5.825	5.847	6.084	6.127
7.241	7.560	8.901	9.000	10.482	11.133					

7.2. Second application

The second data set represents the failure times of 50 components, [26, p. 220]. The data set is reported in Table 6. The MLEs of the unknown parameters and the goodness-of-fit test statistics for the second data set are presented in Tables 7 and 8, respectively. It is easy to see that the EMOGL distribution has smaller values of the AIC, BIC, A^*, W^* and $-l$ statistics than that of the other considered distributions. Moreover, the largest p-value belongs to the EMOGL distribution in Table 8. Hence, the EMOGL distribution outperforms the other considered competitors on the basis of these criteria. These conclusions can also be drawn visually in Figures 6 and 7. Furthermore, the values of the LR statistics with their corresponding p-values are reported in Table 9. From Table 9, we observe that the computed p-values are too small in comparison with their counterparts so, we can reject all the null hypotheses. That is, accordingly the LR criterion we deduce that fitting of the EMOGL is more efficient than the considered sub-models for the second data set.

Table 7. Parameter ML estimates and their standard errors (in parentheses) for the second data set.

$Model$	α	β	γ	λ
$Lindley(\lambda)$	–	–	–	0.565 (0.0585)
$GL(\alpha, \lambda)$	0.622 (0.1143)	–	–	0.435 (0.0709)
$PL(\beta, \lambda)$	–	0.759 (0.0792)	–	0.770 (0.1088)
$BL(\alpha, \beta, \lambda)$	0.660 (0.1407)	0.409 (0.5014)	–	0.947 (1.0566)
$EPL(\alpha, \beta, \lambda)$	0.937 (0.6390)	0.682 (0.2687)	–	0.937 (0.6390)
$OLLPL(\alpha, \beta, \lambda)$	0.623 (0.2296)	0.698 (0.2916)	–	0.623 (0.2296)
$KwPL(\alpha, \beta, \gamma, \lambda)$	1.329 (0.8622)	0.533 (0.9283)	0.691 (0.2407)	1.485 (1.7167)
$OBuPL(\alpha, \beta, \lambda)$	0.707 (0.0813)	1.417 (1.0283)	–	0.468 (0.3014)
$EGL(\alpha, \gamma, \lambda)$	0.619 (0.1068)	–	0.414 (0.4174)	0.381 (0.1489)
$EMOGL(\alpha, \gamma, \lambda)$	1.158 (0.2085)	0.388 (0.1345)	14.14 (12.261)	0.701 (0.1488)

Table 8. Goodness-of-fit test statistics for the second data set.

$Model$	AIC	BIC	$p-value$	W^*	A^*	$-l$
$Lindley(\lambda)$	215.880	217.792	0.012	0.135	0.741	106.941
$GL(\alpha, \lambda)$	210.574	214.398	0.333	0.139	0.757	103.287
$PL(\beta, \lambda)$	209.629	213.453	0.510	0.108	0.606	102.814
$BL(\alpha, \beta, \lambda)$	212.145	217.881	0.337	0.131	0.716	103.072
$EPL(\alpha, \beta, \lambda)$	211.548	217.284	0.555	0.099	0.566	102.774
$OLLPL(\alpha, \beta, \lambda)$	211.023	216.759	0.414	0.122	0.656	102.511
$KwPL(\alpha, \beta, \gamma, \lambda)$	213.460	221.108	0.630	0.092	0.536	102.730
$OBuPL(\alpha, \beta, \lambda)$	210.778	216.514	0.523	0.115	0.621	102.389
$EGL(\alpha, \gamma, \lambda)$	212.404	218.140	0.443	0.128	0.704	103.202
$EMOGL(\alpha, \gamma, \lambda)$	205.893	213.541	0.978	0.022	0.167	98.9467

Table 9. The LR test results for the second data set.

	Hypotheses	LR	p-value
EMOGL versus Lindley	$H_0 : \alpha = \beta = \gamma = 1$	15.98658	0.00112
EMOGL versus GL	$H_0 : \alpha = \gamma, \ \beta = 1$	8.680934	0.01303
EMOGL versus MOL	$H_0 : \alpha = \gamma = 1$	9.271497	0.00969
EMOGL versus MOGL	$H_0 : \alpha = \gamma$	6.54374	0.01052
EMOGL versus EGL	$H_0 : \beta = 1$	3.064931	0.07999

8. Bayesian estimation

Bayesian inference procedure has been taken into consideration by many statistical researchers, especially researchers in the field of survival analysis and reliability engineering. In this section, a complete sample data is analyzed through Bayesian point of view. We assume that the parameters α, β, γ and λ of $EMOGL$ distribution have independent prior distributions as

$$\alpha \sim Gamma(a, b), \gamma \sim Gamma(c, d), \lambda \sim Gamma(e, f), \beta \sim Gamma(g, h)$$

where a,b,c,d,e, f, g and h are positive. Hence, the joint prior density function is formulated as follow:

$$\pi(\alpha, \beta, \gamma, \lambda) = \frac{b^a d^c f^e h^g}{\Gamma(a)\Gamma(c)\Gamma(e)\Gamma(g)} \alpha^{a-1} \beta^{h-1} \gamma^{c-1} \lambda^{e-1} e^{-(b\alpha + h\beta + d\gamma + f\lambda)}. \tag{24}$$
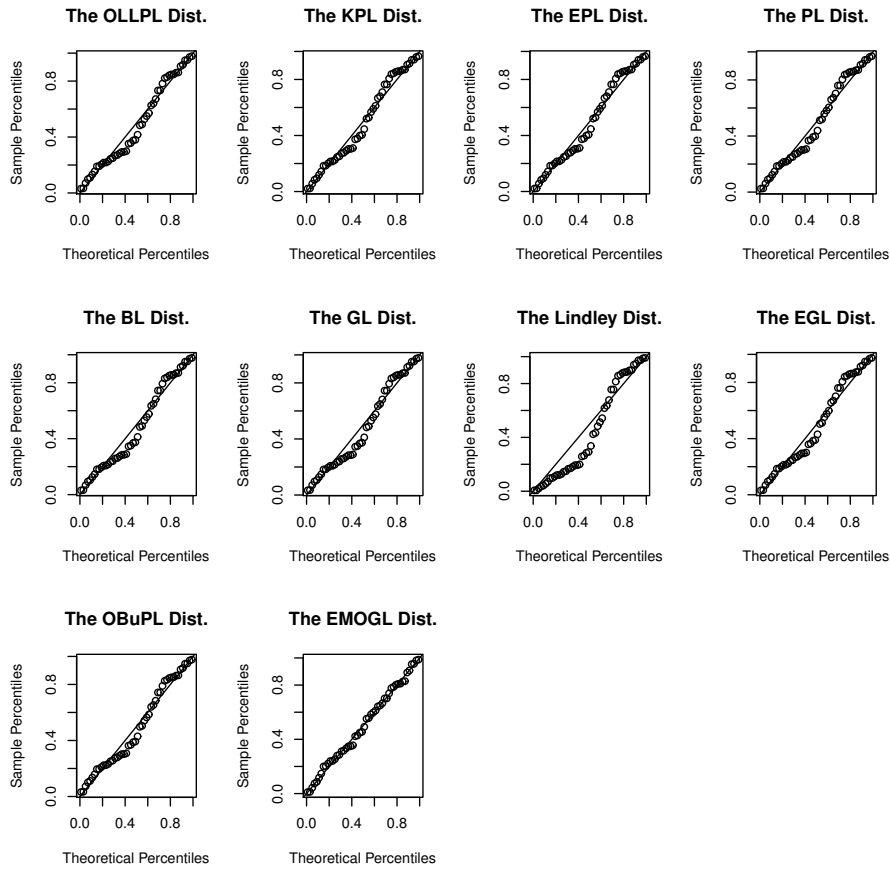
Figure 7. Probability plots for the second data set.

In the Bayesian estimation, according to that we do not know the actual value of the parameter, we may be adversely affected by loss when we choose an estimator. This loss can be measured by a function of the parameter and corresponding estimator.

Five well-known loss functions and associated Bayesian estimators and corresponding posterior risk are presented in Table 10. For more details see [15], [19] and [18]. In the next, we provide the posterior probability distributions for a complete

Table 10. Bayes estimator and posterior risk under different loss functions

loss function	Bayes estimator	Posterior risk
$L_1 = SELF = (\theta - d)^2$	$E(\theta\|x)$	$Var(\theta\|x)$
$L_2 = WSELF = \frac{(\theta-d)^2}{\theta}$	$(E(\theta^{-1}\|x))^{-1}$	$E(\theta\|x) - (E(\theta^{-1}\|x))^{-1}$
$L_3 = MSELF = \left(1 - \frac{d}{\theta}\right)^2$	$\frac{E(\theta^{-1}\|x)}{E(\theta^{-2}\|x)}$	$1 - \frac{E(\theta^{-1}\|x)^2}{E(\theta^{-2}\|x)}$
$L_4 = PLF = \frac{(\theta-d)^2}{d}$	$\sqrt{E(\theta^2\|x)}$	$2\left(\sqrt{E(\theta^2\|x)} - E(\theta\|x)\right)$
$L_5 = KLF = \left(\sqrt{\frac{d}{\theta}} - \sqrt{\frac{\theta}{d}}\right)$	$\sqrt{\frac{E(\theta\|x)}{E(\theta^{-1}\|x)}}$	$2\left(\sqrt{E(\theta\|x)E(\theta^{-1}\|x)} - 1\right)$

data set. Let we define the function φ as

$$\varphi(\alpha, \beta, \gamma, \lambda) = \alpha^{a-1}\beta^{h-1}\gamma^{c-1}\lambda^{e-1}e^{-(b\alpha+h\beta+d\gamma+f\lambda)}, \; \alpha > 0, \; \beta > 0, \; \gamma > 0, \; \lambda > 0.$$

The joint posterior distribution in terms of a given likelihood function $L(data)$ and joint prior distribution $\pi(\alpha, \beta, \gamma, \lambda)$ defined as

$$\pi^*(\alpha, \beta, \gamma, \lambda | data) \propto \pi(\alpha, \beta, \gamma, \lambda)L(data). \tag{25}$$

Hence, we get joint posterior density of parameters α, β, γ and λ for complete sample data by combining the likelihood function and joint prior density (24). Therefore, the joint posterior density function is given by

$$\pi^*(\alpha, \beta, \gamma, \lambda | \underline{x}) = K\varphi(\alpha, \beta, \gamma, \lambda)L(\underline{x}, \boldsymbol{\xi}) \tag{26}$$

where

$$L(\underline{x}; \boldsymbol{\xi}) = \prod_{i=1}^{n} \frac{\beta\lambda^2(1+x_i)e^{-\lambda x_i}\left[1 - (1 + \frac{\lambda x_i}{1+\lambda})e^{-\lambda x_i}\right]^{\alpha-1}\left\{\alpha + (\gamma-\alpha)\left[1 - (1 + \frac{\lambda x_i}{1+\lambda})e^{-\lambda x_i}\right]^{\gamma}\right\}}{\left\{\left[1 - (1 + \frac{\lambda x_i}{1+\lambda})e^{-\lambda x_i}\right]^{\alpha} + \beta\left[1 - (1 - (1 + \frac{\lambda x_i}{1+\lambda})e^{-\lambda x_i})^{\gamma}\right]\right\}^2}. \tag{27}$$

and K is given as

$$K^{-1} = \int_0^{\infty}\int_0^{\infty}\int_0^{\infty}\int_0^{\infty} \varphi(\alpha, \beta, \gamma, \lambda)L(\underline{x}, \boldsymbol{\xi})d\alpha d\beta d\gamma d\lambda.$$

It is clear from the equation (26) that there is no closed form for the Bayesian estimators under the Five loss functions described in Table 10, so we suggest using an $MCMC$ procedure based on 10000 replicates to compute Bayesian estimators. The corresponding Bayesian point and interval estimation and posterior risk are provided in Tables 11 and 12 for flood peaks data set. Table 12 provides 95% credible and HPD intervals for each parameter of the $EMOGL$ distribution. The posterior samples extracted by using Gibbs sampling technique. Moreover, we provide the posterior summary plots in Figures 8 and 9. These plots confirm that the sampling process is of prime quality and convergence is occurred.

Table 11. Bayesian estimates $\widehat{\theta}$ and their posterior risks $r_{\widehat{\theta}}$ of the parameters under different loss functions based on the flood peaks data.

Data	Flood peaks			
Bayesian estimation				
Loss function	$\widehat{\alpha}\;(r_{\widehat{\alpha}})$	$\widehat{\beta}\;(r_{\widehat{\beta}})$	$\widehat{\gamma}\;(r_{\widehat{\gamma}})$	$\widehat{\lambda}\;(r_{\widehat{\lambda}})$
SELF	0.3019 (0.0067)	3.7853 (0.6092)	0.0415 (0.0003)	0.0886 (0.0003)
WSELF	0.2804 (0.0216)	3.6237 (0.1616)	0.0346 (0.0068)	0.0858 (0.0028)
MSELF	0.2588 (0.0769)	3.4609 (0.0449)	0.0277 (0.2012)	0.0829 (0.0342)
PLF	0.3128 (0.0218)	3.8649 (0.1593)	0.0449 (0.0068)	0.0900 (0.0027)
KLF	0.2909 (0.0755)	3.7036 (0.0441)	0.0379 (0.1886)	0.0872 (0.0325)

Analogous Bayesian results for Failure times data set are provided in Tables 13 and 14. Moreover, the corresponding associated Figures are provided by Figures 10 and 11.

9. Conclusion

In present paper, a new distribution which is called extended Marshal-Olkin generalized Lindley (EMOGL) was introduced. This distribution was included some important distributions as special cases such as extended generalized Lindley, Marshall-
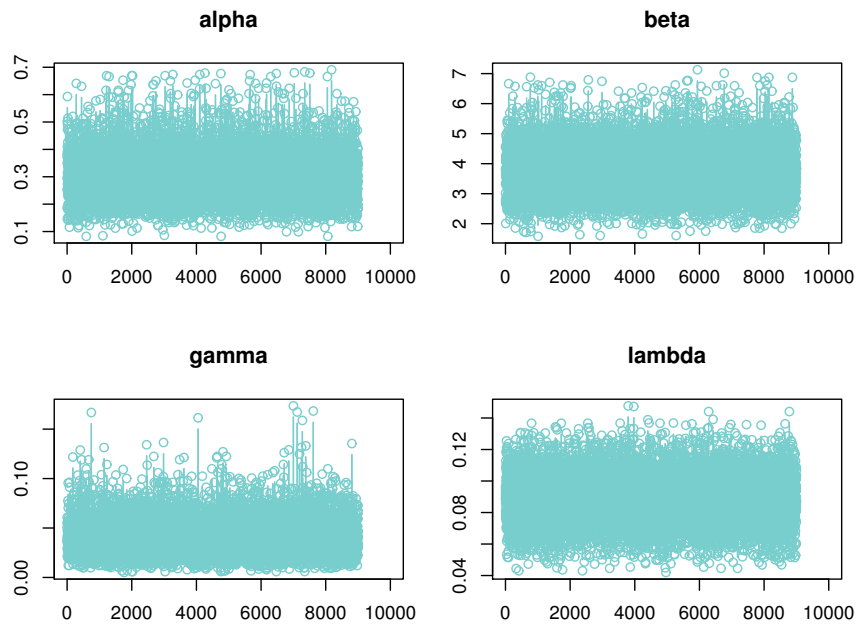
Figure 8. Plots of Bayesian analysis and performance of Gibbs sampling for the flood peaks data set. Trace plots of each parameter of $EMOGL$ distribution.
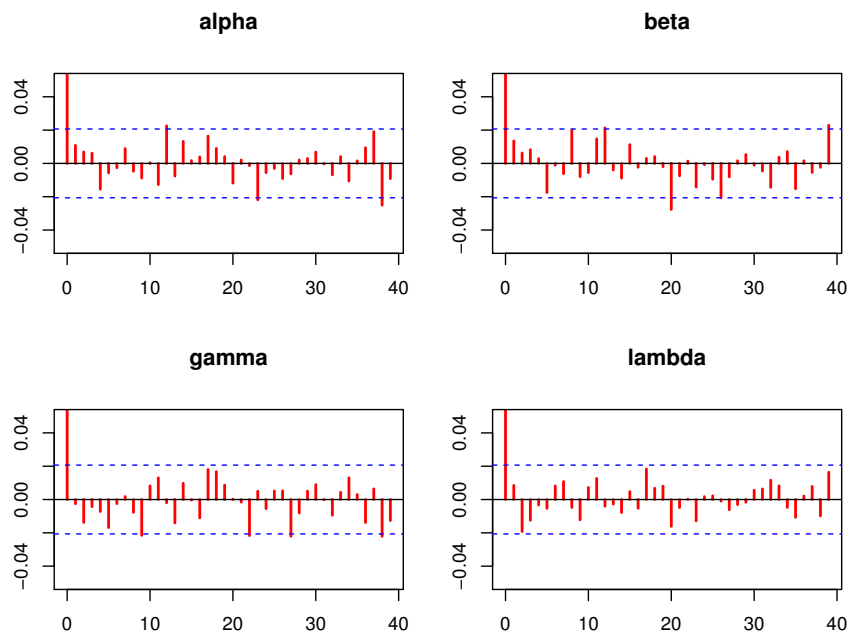

Figure 9. Plots of Bayesian analysis and performance of Gibbs sampling for the flood peaks data set. Autocorrelation plots of each parameter of $EMOGL$ distribution.

Olkin Lindley and Marshall- Olkin generalized Lindley. Some mathematical properties and asymptotic behavior of the new

Table 12. Credible and HPD intervals of the parameters α, β, γ and λ for the flood peaks.

	Credible interval	HPD interval
α	(0.243, 0.350)	(0.149, 0.451)
β	(3.254, 4.282)	(2.298, 2.298)
γ	(0.030, 0.051)	(0.012, 0.074)
λ	(0.078, 0.099)	(0.058, 0.117)

Table 13. Bayesian estimates $\widehat{\theta}$ and their posterior risks $r_{\widehat{\theta}}$ of the parameters under different loss functions based on the failure times.

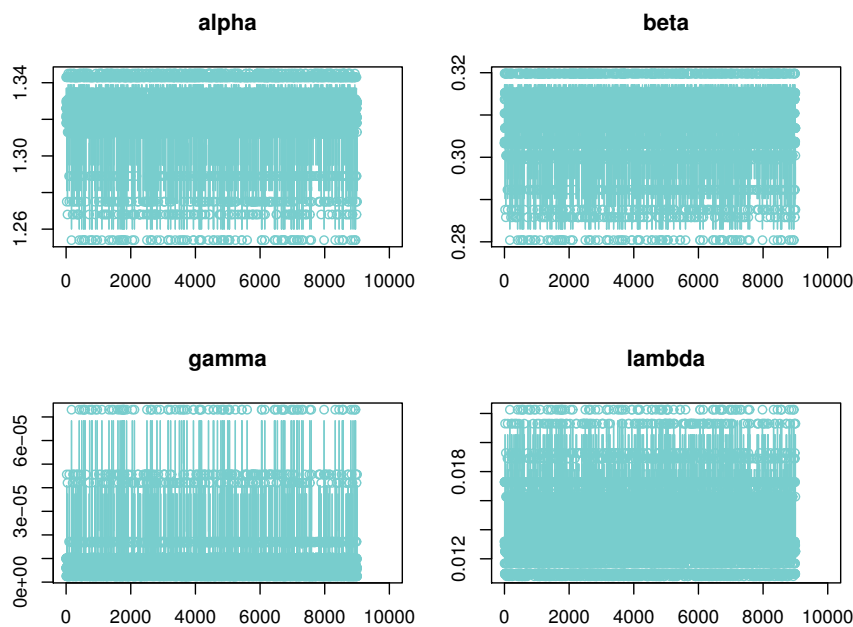| Data	Failure times			
Bayesian estimation				
Loss function	$\widehat{\alpha}\ (r_{\widehat{\alpha}})$	$\widehat{\beta}\ (r_{\widehat{\beta}})$	$\widehat{\gamma}\ (r_{\widehat{\gamma}})$	$\widehat{\lambda}\ (r_{\widehat{\lambda}})$
SELF	1.3235 (0.00022)	0.30888 (0.00005)	8.5471e-06 (9.07e-11)	0.013459 (7.06e-06)
WSELF	1.3233 (0.00017)	0.30869 (0.00019)	5.6692e-06 (2.87e-06)	0.013024 (0.00043)
MSELF	1.3231 (0.00013)	0.30849 (0.00064)	4.2564e-06 (0.2492)	0.012672 (0.02702)
PLF	1.3236 (0.00017)	0.30898 (0.00019)	1.2801e-05 (8.50e-06)	0.013719 (0.00052)
KLF	1.3234 (0.00013)	0.30879 (0.00063)	6.9609e-06 (0.45571)	0.013239 (0.03312)

Figure 10. Plots of Bayesian analysis and performance of Gibbs sampling for the Failure times data set. Trace plots of each parameter of $EMOGL$ distribution.
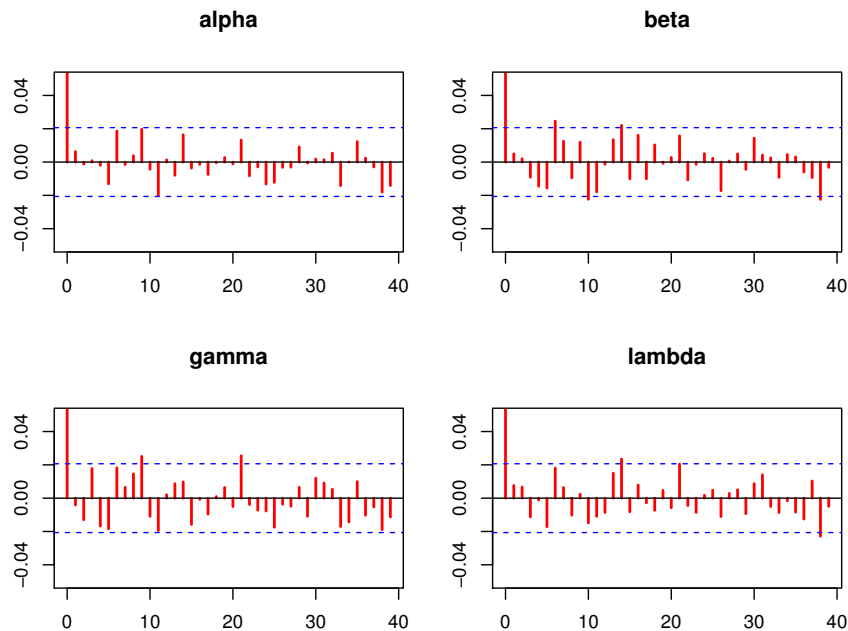
Figure 11. Plots of Bayesian analysis and performance of Gibbs sampling for the Failure times data set. Autocorrelation plots of each parameter of $EMOGL$ distribution.

Table 14. Credible and HPD intervals of the parameters α, β, γ and λ for the Failure times.

	Credible interval	HPD interval
α	(1.321, 1.329)	(1.289, 1.345)
β	(0.304, 0.315)	(0.2923, 0.3198)
γ	(5.8e-06, 9.8e-06)	(2.486e-06, 1.713e-05)
λ	(0.011, 0.013)	(0.0108, 0.0189)

distribution were provided. These properties were shown the EMOGL distribution has considerable flexibility. Moreover, the MLEs of EMOGL parameters and asymptotic distribution based on the MLEs were presented using Fisher information matrix. Furthermore, a simulation study was carried out to evaluate the efficiency of obtained MLEs and it was illustrated the good performance of these estimations. In addition, the EMOGL distribution was compared with some other extended distributions of Lindley such as generalized Lindley, Beta Lindley, exponentiated power Lindley and odd log-logistic power Lindley using two real data sets. Using these real data sets, it was shown numerically that the EMOGL has better performance and high flexibility in comparing to other extended distributions of Lindley which were considered. Finally, the Bayesian estimation and Gibbs sampling procedure for the considered data sets were discussed.

REFERENCES

1. AM Abouammoh, Arwa M Alshangiti, and IE Ragab. A new generalized lindley distribution. *Journal of Statistical Computation and Simulation*, 85(18):3662–3678, 2015.
2. Alfred Akinsete, Felix Famoye, and Carl Lee. The beta-pareto distribution. *Statistics*, 42(6):547–563, 2008.
3. M. Alizadeh, M. Afshari, B. Hosseini, , and T. G Ramires. Extended exp-g family of distributions: Properties and applications. *Communication in Statistics-Simulation and Computation*, 49(7):1730–1745, 2020.
4. M. Alizadeh, E. Altunb, G. Ozelc, M. Afshari, and A. Eftekharian. A new odd log-logistic lindley distribution with properties and applications. *Sankhya Ser A*, 81(2).
5. Morad Alizadeh, SMT K MirMostafaee, Emrah Altun, Gamze Ozel, and Maryam Khan Ahmadi. The odd log-logistic marshall-olkin power lindley distribution: Properties and applications. *Journal of Statistics and Management Systems*, 20(6):1065–1093, 2017.

6. Morad Alizadeh, Gamze Ozel, Emrah Altun, Mousa Abdi, and GG Hamedani. The odd log-logistic marshall-olkin lindley model for lifetime data. *Journal of Statistical Theory and Applications*, 16(3):382–400, 2017.
7. Gokcen Altun, Morad Alizadeh, Emrah Altun, and Gamze Ozel. Odd burr lindley distribution with properties and applications. *Hacettepe Journal of Mathematics and Statistics*, 46(2):255–276, 2017.
8. Ayman Alzaatreh, Carl Lee, and Felix Famoye. A new method for generating families of continuous distributions. *Metron*, 71(1):63–79, 2013.
9. A Asgharzadeh, Hassan S Bakouch, S Nadarajah, F Sharafi, et al. A new weighted lindley distribution with application. *Brazilian Journal of Probability and Statistics*, 30(1):1–27, 2016.
10. A Asgharzadeh, S Nadarajah, and F Sharafi. Weibull lindley distribution. *REVSTAT–Statistical Journal*, 16(1):87–113, 2018.
11. Samir K Ashour and Mahmoud A Eltehiwy. Exponentiated power lindley distribution. *Journal of Advanced Research*, 6(6):895–905, 2015.
12. Hassan S Bakouch, Bander M Al-Zahrani, Ali A Al-Shomrani, Vitor AA Marchi, and Francisco Louzada. An extended lindley distribution. *Journal of the Korean Statistical Society*, 41(1):75–85, 2012.
13. C. E. Bonferroni. Elmenti di statistica generale. *Libreria Seber, Firenze*, 1930.
14. Selen Cakmakyapan and Gamze Ozel. The lindley family of distributions: Properties and applications. *Hacettepe Journal of Mathematics and Statistics*, 46:1–27, 2016.
15. R Calabria and G Pulcini. Point estimation under asymmetric loss functions for left-truncated exponential samples. *Communications in Statistics-Theory and Methods*, 25(3):585–600, 1996.
16. Frank S Gomes-Silva, Ana Percontini, Edleide de Brito, Manoel W Ramos, Ronaldo Venâncio, and Gauss Moutinho Cordeiro. The odd lindley-g family of distributions. *Austrian Journal of Statistics*, 46(1):65–87, 2017.
17. I.S. Gradshteyn and I.M. Ryzhik. *Table of Integrals, Series, and Products*. Elsevier/Academic Press, 2014.
18. O. Kharazmi and S. Jahangard. A new family of lifetime distributions in terms of cumulative hazard rate function. *Communications in Statistics-Theory and Methods*, 65(1):1–22, 2020.
19. Saadatinik A. Kharazmi, O. and S. Jahangard. Odd hyperbolic cosine exponential-exponential (ohc-ee) distribution. *Annals of Data Science*, 6(4):765–785, 2019.
20. Malcolm R Leadbetter, Georg Lindgren, and Holger Rootzén. *Extremes and related properties of random sequences and processes*. Springer Science & Business Media, 2012.
21. Erich L Lehmann and George Casella. *Theory of Point Estimation*. Springer Science & Business Media, 2006.
22. Dennis V Lindley. Fiducial distributions and bayes' theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*, 20(1):102–107, 1958.
23. Max O Lorenz. Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, 9(70):209–219, 1905.
24. Pararai Mavis, Warahena-Liyanage Gayan, and Broderick O. Oluyede. A new class of generalized power lindley distribution with applications to lifetime data. *Asian Journal of Mathematics and Applications*, 5(1):53–96, 2016.
25. Faton Merovci and Vikas Kumar Sharma. The beta-lindley distribution: properties and applications. *Journal of Applied Mathematics*, http://dx.doi.org/10.1155/2014/198951, 2014.
26. DN Prabhakar Murthy, Min Xie, and Renyan Jiang. *Weibull models*, volume 505. John Wiley & Sons, 2004.
27. Saralees Nadarajah, Hassan S Bakouch, and Rasool Tahmasbi. A generalized lindley distribution. *Sankhya B*, 73(2):331–359, 2011.
28. V. Ranjbar, M. Alizadeh, and E. Altun. Extended generalized lindley distribution: Properties and applications. *Journal of Mathematical Extension*, 13:117–142, 2019.
29. Hojatollah Zakerzadeh and Ali Dolati. Generalized lindley distribution. *Journal of Mathematical Extension*, 3(2):1–17, 2009.