The Lindley Family of Distributions: Properties and Applications

Selen Cakmakyapan*† and Gamze Ozel‡

Abstract

In this paper, we propose a new class of distributions called the Lindley generator with one extra parameter to generate many continuous distributions. The new distribution contains several distributions as submodels, such as Lindley-Exponential, Lindley-Weibull, and Lindley-Lomax. Some mathematical properties of the new generator, including ordinary moments, quantile and generating functions, limiting behaviors, some entropy measures and order statistics, which hold for any baseline model, are presented. Then, we discuss the maximum likelihood method to estimate model parameters. The importance of the new generator is illustrated by means of three real data sets. Applications show that the new family of distributions can provide a better fit than several existing lifetime models.

Keywords: Lindley distribution, Lomax distribution, Weibull distribution, Estimation, Generating Function, Maximum Likelihood, Moments, Entropy.

2000 AMS Classification: 62E15, 60E05

1. Introduction

Statistical distributions are very useful in describing and predicting real world phenomena. Numerous classical distributions have been extensively used over the past decades for modeling data in several areas. Recent developments focus on definition of the new families of distributions that extend well-known distributions and at the same time provide great flexibility in modelling data. Hence, several classes of distributions have been introduced by adding one or more parameters to generate new distributions in the statistical literature. The well-known generators are Marshall-Olkin generated family (MO-G) by Marshall and Olkin [1], beta-G by Eugene et al. [2], Kumaraswamy-G (Kw-G) by Cordeiro and de Castro [3], McDonald-G (Mc-G) by Alexander et al. [4], transformed-transformer (T-X) family by Alzaatreh et al. [5], exponentiated T-X by Alzaghal et al. [6], Weibull-G by Bourguignon et al. [7], exponentiated half-logistic by Cordeiro et al. [8], Lomax-G by Cordeiro et al. [9], Zografos-Balakrishnan-G by Nadarajah et al. [10].

*Department of Statistics, Hacettepe University, Ankara, Turkey Email: selencakmakyapan@hacettepe.edu.tr

†Corresponding Author.

‡Department of Statistics, Hacettepe University, Ankara, Turkey Email: gamzeozl@hacettepe.edu.tr

The Lindley distribution was introduced by Lindley [11] to analyze failure time data, especially in applications modeling stress-strength reliability. The motivation of the Lindley distribution arises from its ability to model failure time data with increasing, decreasing, unimodal and bathtub shaped hazard rates. The Lindley distribution belongs to an exponential family and it can be written as a mixture of exponential and gamma distributions. The distribution represents a good alternative to the exponential failure time distributions that suffer from not exhibiting unimodal and bathtub shaped failure rates [12]. The properties and inferential procedure for the Lindley distribution were studied by Ghitany et al. [13, 14]. It is shown that the Lindley distribution is better than the exponential distribution when hazard rate is unimodal or bathtub shaped. Mazucheli and Achcar [15] also proposed the Lindley distribution as a possible alternative to exponential and Weibull distributions.

The probability density function (pdf) of a Lindley random variable X, with scale parameter θ is given by

$$h(x;\theta) = \frac{\theta^2}{1+\theta}\,(1+x)\,exp\,(-\theta x) \tag{1.1}$$

and the corresponding cumulative distribution function (cdf) of X is

$$H(x;\theta) = 1 - \frac{1+\theta+\theta x}{1+\theta}\exp(-\theta x) \tag{1.2}$$

The Lindley distribution does not provide enough flexibility for analyzing different types of lifetime data because of having only one parameter. To increase the flexibility for modelling purposes it will be useful to consider further alternatives of this distribution. Therefore, the aim of this study is to introduce a new family of distributions using the Lindley generator. The term *generator* means that we have a different distribution F for each baseline distribution G. Based on the transformer (T-X) generator of Alzaatreh et al. [5], we propose a new wider class of continuous distributions called Lindley-G family by integrating the Lindley density function having cdf given by

$$\begin{aligned} F_{Lindley-G}(x;\theta,\boldsymbol{\xi}) &= \int_0^{-log[-G(x;\boldsymbol{\xi})]} \frac{\theta^2}{1+\theta}(1+t)\exp(-\theta)dt \\ &= 1 - \left[1 - \frac{\theta}{\theta+1}\left[log\left(1-G(x;\boldsymbol{\xi})\right)\right]\right]\left[-G(x;\boldsymbol{\xi})\right]^{\theta} \end{aligned} \tag{1.3}$$

where G(x;$\boldsymbol{\xi}$) is a baseline cdf which depends on a $(rx1)$ parameter vector $\boldsymbol{\xi}$. The family pdf reduces to

$$f_{Lindley-G}(x;\theta,\xi) = g(x;\xi)\left[1 - log\left(1-G(x;\boldsymbol{\xi})\right)\right]\left[1-G(x;\xi)\right]^{\theta-1}\frac{\theta^2}{\theta+1} \tag{1.4}$$

where $g(x;\boldsymbol{\xi})$ is the baseline pdf. Henceforth, let G be a continuous baseline distribution. For each G distribution, we define the Lindley-G distribution with one extra parameter θ defined by the pdf in (1.4). A random variable X with pdf (1.4) is denoted by $X \sim Lindley - G(\theta,\xi)$.

We obtain the survival function corresponding to (1.3) as

$$
\begin{aligned}
S_{Lindley-G}(x;\theta,\xi) &= 1 - F(x;\theta,\xi) \\
&= \left[1 - \frac{\theta}{\theta+1} log\left[1 - G(x;\xi)\right]\right]\left[1 - G(x;\xi)\right]. \qquad (1.5)
\end{aligned}
$$

Then, the hazard rate function (hrf) of X is given by

$$
\begin{aligned}
\tau_{Lindley-G}(x;\theta,\boldsymbol{\xi}) &= \frac{f(x;\theta,\xi)}{S(x;\theta,\xi)} \\
&= \theta^2 \frac{g(x;\xi)\left[1 - log\left(1 - G(x;\xi)\right)\right]}{\left[\theta + 1 - \theta\left[log\left(1 - G(x;\xi)\right)\right]\right]\left[1 - G(x;\xi)\right]}. \qquad (1.6)
\end{aligned}
$$

The rest of the paper is organized as follows. In Section 2, we present three new generated distributions in the proposed family. We discuss the distributional properties of the proposed family, including quantile function, limiting behaviors, moments and generated functions in Section 3. Section 4 is devoted to the Renyi and Shannon entropies, reliability function and order statistics. Maximum likelihood estimation of the model parameters and the observed information matrix are presented in Section 5. In Section 6, applications to three real data sets are presented to illustrate the potentially of the new family. Conclusion is given in Section 7.

2. Special Lindley-G Distributions

The pdf in (1.4) allows greater flexibility of its tails and can be widely applied in many areas of statistics. Here, we present and study some special cases of this family because it extends several widely known distributions in the literature. The pdf is the most tractable when the cdf $G(x;\boldsymbol{\xi})$ and the pdf $g(x;\boldsymbol{\xi})$ have closed-forms.

2.1. Lindley-Weibull Distribution. Consider the Weibull distribution with density and distribution functions given by $g(x;a,b) = \frac{a}{b}\left(\frac{x}{b}\right)^{a-1}\exp\left[-\left(\frac{x}{b}\right)^{a}\right]$ and $G(x;a,b) = 1 - \exp\left[-\left(\frac{x}{b}\right)^{a}\right]$, respectively. Then, the Lindley-Weibull (LW) density function is given by

$$
f(x;\theta,a,b) = \frac{a\theta^2}{b(\theta+1)}\left(\frac{x}{b}\right)^{a-1}\left[1 + \left(\frac{x}{b}\right)^{a}\right]\exp\left[-\left(\frac{x}{b}\right)^{a}\theta\right] \qquad (2.1)
$$

where a is the shape parameter and b is the scale parameter. A random variable X with pdf (2.2) is denoted by $X \sim LW(\theta,a,b)$. For $a = 1$, it becomes the Lindley-Exponential (LE) distribution.

The corresponding cumulative density and hazard rate functions are, for $x \geq 0$, $a > 0$, $b > 0$, $\theta > 0$, respectively, given by

$$
F(x;\theta,a,b) = 1 - \left[1 + \frac{\theta}{\theta+1}\left(\frac{x}{b}\right)^{a}\right] exp\left[-\left(\frac{x}{b}\right)^{a}\theta\right], \qquad (2.2)
$$

$$\tau(x;\theta,a,b) = \frac{a\theta^2 x^{a-1}}{b^a(\theta+1)+\theta x^a}\left[1+\left(\frac{x}{b}\right)^a\right]. \tag{2.3}$$

Figure 1 displays plots for the probability density, cumulative distribution, survival and hazard rate functions of the LW distribution for several parameter values. Figure 1 indicates that the pdf of LW has various shapes. Both unimodal and monotonically decreasing shape appear possible. Monotonically decreasing shapes appear when a is small. Figure 1 also shows that the hrf of LW can have very flexible shapes, such as increasing, decreasing, upside-down bathtub.

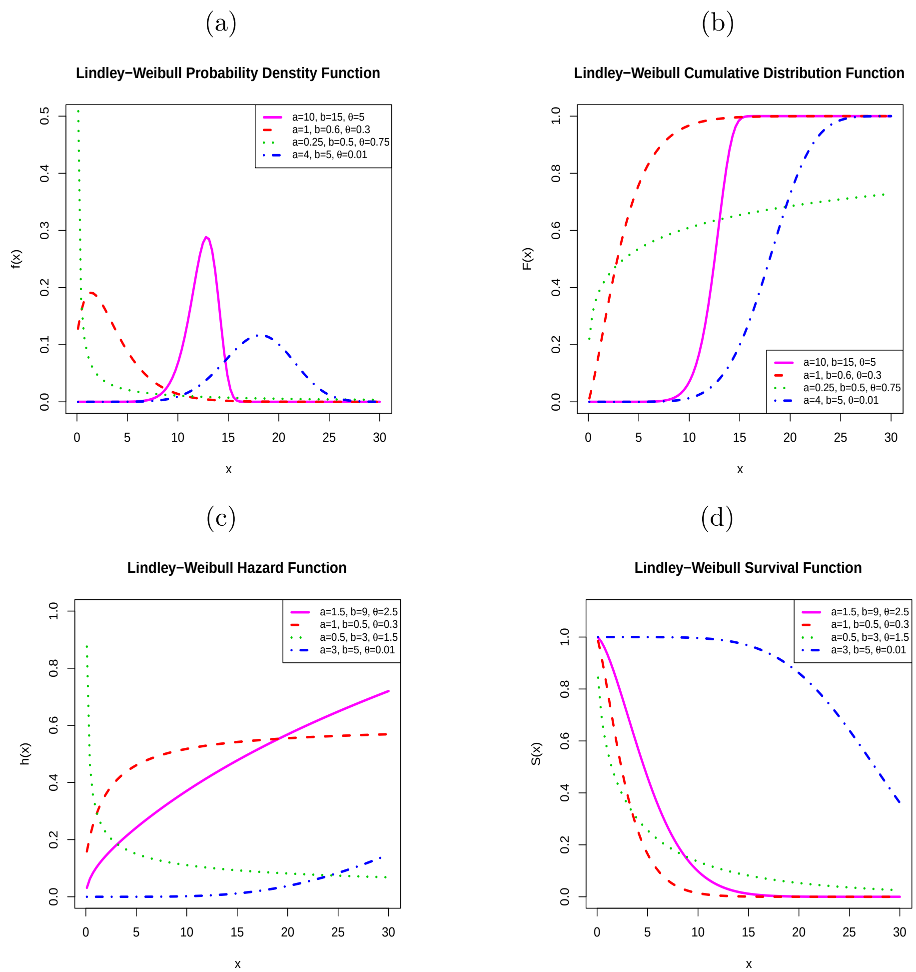

Figure 1. Probability density, cumulative density, hazard rate and survival functions of the LW distribution for some arbitrary parameters

2.2. Lindley-Lomax Distribution. Let X be a continuous random variable having a Lomax distribution with shape parameter $\alpha > 0$ and scale parameter

$\sigma > 0$. Then, the pdf and cdf of the Lomax distribution are, for $x \geq 0, \sigma > 0, \alpha > 0$, respectively, given by

$$g(x; \alpha, \sigma) = \frac{\alpha}{\sigma}\left(1 + \frac{x}{\sigma}\right)^{-(\alpha+1)} \tag{2.4}$$

$$G(x; \alpha, \sigma) = 1 - \left(1 + \frac{x}{\sigma}\right)^{-\alpha} \tag{2.5}$$

Note that standard Lomax distribution is obtained for $\sigma = 1$. Then, the cdf of Lindley-Lomax (LL) by inserting (2.5) in (1.3) for $x \geq 0$, $\sigma > 0$, $\alpha > 0$, $\theta > 0$ as

$$F(x; \theta, \alpha, \sigma) = 1 - \left[1 + \frac{\theta\alpha}{(\theta+1)} log\left(1 + \frac{x}{\sigma}\right)\right]\left(1 + \frac{x}{\sigma}\right)^{-\alpha\theta} \tag{2.6}$$

$$f(x; \theta, \alpha, \sigma) = \frac{\theta^2\alpha}{(\theta+1)\sigma}\left[1 + \alpha log\left(1 + \frac{x}{\sigma}\right)\right]\left(1 + \frac{x}{\sigma}\right)^{-(\alpha\theta+1)} \tag{2.7}$$

where θ, α are scale and α is shape parameters. A random variable X with pdf (2.7) is denoted by $X \sim LL(\theta, \alpha, \sigma)$. Note that the Lindley standard Lomax distribution is the special case of (2.7) for $\sigma = 1$.

The hrf of LL distribution is given by

$$\tau(x; \sigma, \alpha, \theta) = \frac{\theta^2\alpha\left[1 + \alpha log\left(1 + \frac{x}{\sigma}\right)\right]}{\left[(\theta + 1 + \theta\alpha) log\left(1 + \frac{x}{\sigma}\right)\right](\sigma + x)} \tag{2.8}$$

Plots for the probability density, cumulative density, hazard rate and survival functions of the LL distribution for several parameter values are displayed in Figure 2. The LL distribution given by (2.7) is much more flexible than the Lindley distribution and can allow for greater flexibility of the tails.

The pdf of the LL has unimodal and monotonically decreasing shapes. Figure 2 also shows that the LL distribution has decreasing hrf for small values of α and upside down bathtub hrf for the large values of α and θ.

3. Statistical Properties

In this section, we study the distributional properties of the Lindley-G. In particular, if $X \sim Lindley - G(\theta, \xi)$, then the shapes of the pdf, quantile function, moments, skewness, kurtosis are derived and studied in detail.

3.1. Useful Expansions. Despite the fact that the cdf and pdf of Lindley-G require mathematical functions that are widely available in modern statistical packages, frequently analytical and numerical derivations take advantage of power series for the pdf. We use the following expansion of Gradshteyn and Ryzhik [16] for a power series raised to any positive integer n.

$$\left(\sum_{i=0}^{\infty} a_i u^i\right)^n = \sum_{i=0}^{\infty} c_{n,i} u^i \tag{3.1}$$

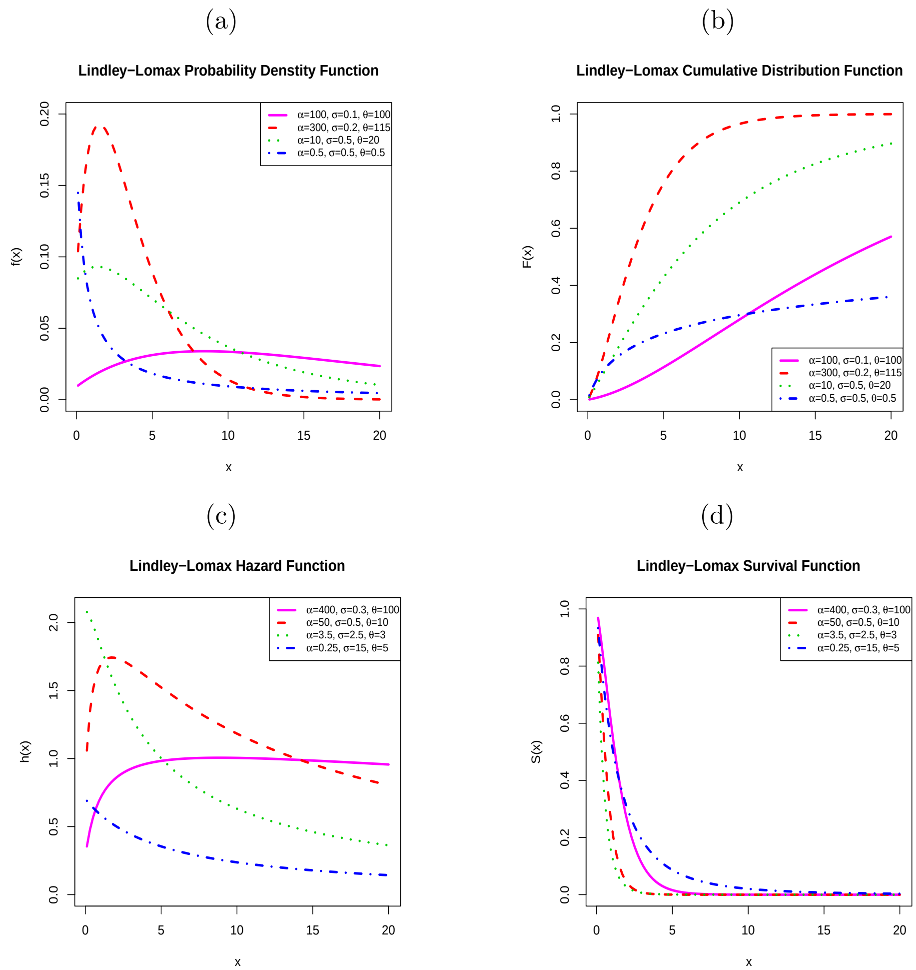

Figure 2. Probability density, cumulative density, hazard rate and survival functions of the LL distribution for some arbitrary parameters.

where $c_{n,i}$,$i = 1, 2, ...,$ for $c_{n,0} = a_0^n$, are easily obtained from the recurrence equation

$$c_n^i = (ia_0)^{-1} \sum_{m=1}^{i} [m(n+1) - i]\, a_m c_{n,i-m}$$

In this study, we also consider the following expansions:

$$(1 - z)^t = \sum_{i=0}^{\infty} (-1)^i \binom{t}{i} z^i, \ |z| < 1, \tag{3.2}$$

$$\text{(3.3)} \qquad log(1-z) = -\sum_{i=0}^{\infty} \frac{z^{i+1}}{i+1},\ |z|<1,$$

$$\text{(3.4)} \qquad log(1+z) = -\sum_{i=0}^{\infty} \frac{(-1)^{i+1} z^{i+1}}{i+1},\ |z|<1.$$

The mathematical relation given above will be useful to obtain moments and entropy function of the Lindley-G family.

3.2. Other representations. We now state some useful expansions for the pdf of Lindley-G family. If $X \sim Lindley - G(\theta, \xi)$, we obtain a double-mixture form of the Lindley-G family using expansions in (3.2) and (3.3) as

$$\text{(3.5)} \qquad f_{Lindley-G}(x;\theta,\xi) = \frac{\theta^2}{\theta+1} g(x;\xi) \sum_{k=0}^{\infty} w_k G^k(x;\xi) + \frac{\theta^2}{\theta+1} g(x;\xi) \sum_{k,j=0}^{\infty} w_k w_j G^{k+j+1}(x;\xi)$$

where the coefficient is given by $w_k = (-1)^k \binom{\theta-1}{k}$.

Exponentiated-G (Exp-G) distribution is a very popular distribution family and have been studied by many authors in recent years, see Mudhokar et al. [17]-[19] for exponentiated Weibull, Gupta et al. [20] for exponentiated Pareto, Gupta and Kundu [21]-[23] for generalized exponential distributions and Nadarajah and Gupta [24] for exponential gamma distribution. Kumaraswamy-G (Kw-G) is another popular distribution family and also can be expressed as the Exp-G distribution. The various new distributions have been defined as a member of Kw-G family. Among these, Cordeiro et al. [25] investigated Kumaraswamy Gumbel distribution. Pascoa et al. [26] and Paraníba et al. [27] studied Kumaraswamy generalized gamma and Kumaraswamy Burr XII distributions, respectively. Further, Nadarajah et al. [28] studied several mathematical properties of Kw-G and Lemonte et al. [29] defined exponentiated Kumaraswamy distribution and its log-transform.

For an arbitrary baseline cdf $G(x)$, a random variable is said to have the Exp-G distribution with parameter $a > 0$, say $X \sim Exp - G(a, \xi)$, if its pdf is given by

$$\text{(3.6)} \qquad h_{Exp-G}(x,a,\xi) = ag(x,\xi)G^{a-1}(x,\xi)$$

Cordeiro and de Castro [3] introduced the Kw-G distribution with the pdf $f_{Kw-G}(x)$ given by

$$\text{(3.7)} \qquad f_{Kw-G}(x,a,\xi) = agb(x,\xi)G^{a-1}(x,\xi)\left[1-G^a(x,\xi)\right]^{b-1}$$

Nadarajah et al. [30] was expressed (3.7) in the form of the Exp-G distribution as

$$f_{Kw-G}(x,a,\xi) = a^{-1}\sum_{k=0}^{\infty}\frac{z_k}{k+1}h_{Exp-G}\left(x,a(k+1),\xi\right) \tag{3.8}$$

where $z_k = (-1)^k ab\binom{b-1}{k}$ and $h_{Exp-G}\left(x,a(k+1),\xi\right)$ is the pdf of the $Exp-G\left(x,a(k+1),\xi\right)$ distribution.

The pdf of the Lindley-G family can be derived using the concept of exponential and Kumaraswamy distributions. By this way, we can use the statistical properties of Exp-G and Kw-G distributions.

Using some series expansion, we obtain the pdf of Lindley-G family as a combination of Exp-G distribution which is given by

$$\begin{aligned} f_{Lindley-G}(x;\theta,\xi) &= \frac{\theta^2}{\theta+1}\sum_{k=0}^{\infty}\frac{h_{Exp-G}\left(x;(k+1),\xi\right)w_k}{k+1} \\ &\quad + \frac{\theta^2}{\theta+1}\sum_{k,j=0}^{\infty}\frac{h_{Exp-G}\left(x;(k+j+2),\xi\right)w_k w_j}{k+j+2} \end{aligned} \tag{3.9}$$

We also obtain the expression for the pdf of Lindley-G as a linear combination of Kw-G density function as

$$\begin{aligned} f_{Lindley-G}(x;\theta,\xi) &= \frac{\theta^2}{\theta+1}f_{Kw-G}(1,\theta,\omega) \\ &\quad + \frac{\theta^2}{\theta+1}\sum_{j=0}^{\infty}w_j f_{Kw-G}\left(\frac{j+1}{k+1}+1,\theta,\omega\right)\left(\frac{j+1}{k+1}+1\right)^{-1} \end{aligned} \tag{3.10}$$

3.3. Limiting Behaviors. We seek to investigate the behavior of the probability density, cumulative density, survival and hazard rate functions as $x \to 0$ and as $x \to \infty$.

Proposition 1. The limiting behaviors of (1.3), (1.4), (1.5) and (1.6) as $x \to 0$ are given by

$$f_{Lindley-G}\left(x;\theta,\xi\right) - \frac{\theta^2}{\theta+1}g\left(x;\xi\right) \ \ as\ x \to 0,$$

$$F_{Lindley-G}\left(x;\theta,\xi\right) - 0\ as\ x \to 0,$$

$$S_{Lindley-G}\left(x;\theta,\xi\right) - 1\ as\ x \to 0,$$

$$\tau_{Lindley-G}\left(x;\theta,\xi\right) - \frac{\theta^2}{\theta+1}g\left(x;\xi\right) \ \ as\ x \to 0.$$

Note that the asymptotes of (1.3), (1.4), (1.5) and (1.6) as $x \to \infty$ behave like Lindley distribution.

3.4. Shapes. The shapes of the pdf in (1.4) can be described analytically. The critical points of the pdf are the roots of (3.11)

$$\frac{g^{'}(x)}{g(x)}\left[1 - log\left(1 - G(x)\right)\right] + \frac{g(x)}{G(x)} = (\theta - 1)\frac{g(x)}{\left(1 - G(x)\right)}. \tag{3.11}$$

If is a root of (3.11), then it corresponds to a local maximum, a local minimum or a point of inflexion depending on whether $\lambda(x_0) < 0$, $\lambda(x_0) > 0$ or $\lambda(x_0) = 0$, where

$$\begin{aligned}\lambda(x_0) = {} & \frac{g^{''}(x)\left[1 - log\left(1 - G(x)\right)\right]}{g(x)} + \frac{g^{'}(x)}{\left(1 - G(x)\right)} - \frac{\left[g^{'}(x)\right]^2\left[1 - log\left(1 - G(x)\right)\right]}{g^2(x)} \\ & + \frac{g^{'}(x)}{G(x)} - \frac{g^2(x)}{G^2(x)} - (\theta - 1)\left[\frac{g^{'}(x)}{\left(1 - G(x)\right)} + \frac{g^2(x)}{\left(1 - G(x)\right)^2}\right].\end{aligned} \tag{3.12}$$

3.5. Quantile Function. Quantile functions are in widespread use in general statistics and often find representations in terms of lookup tables for key percentiles. Let X denote a Lindley-G random variable. The quantile function, $Q(u)$, $0 < u < 1$, for the T-X family of distributions is computed by using the formula of Alzaatreh et al. [5] as

$$Q(u) = F^{-1}\left[1 - exp\left(-H^{-1}(u)\right)\right], \tag{3.13}$$

$$H^{-1}(u) = -\frac{\theta + 1 + W\left[(u - 1)(\theta + 1)exp(-(\theta + 1))\right]}{\theta} \tag{3.14}$$

where $H^{-1}(u)$ is the inverse of the Lindley distribution function and $W(.)$ is Lambert function.

We can also use (3.14) for simulating the Lindley-G random variable. Let U be a uniform variable on the unit interval $(0, 1)$. Thus, by means of the inverse transformation method, we also consider the random variable X given by $X = F^{-1}\left[1 - exp\left(-H^{-1}(u)\right)\right]$. In particular, the median of the Lindley-G distribution can be written as $X = F^{-1}\left[1 - exp\left(-H^{-1}(0.5)\right)\right]$.

Skewness measures the degree of the long tail and kurtosis is a measure of the degree of tail heaviness. For the Lindley-G family, Bowley's skewness can be computed by using quantile function in (3.7) as

$$S = \frac{Q(3/4) - 2Q(1/2) + Q(1/4)}{Q(3/4) - Q(1/4)} \tag{3.15}$$

$$K = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(6/8) - Q(2/8)} \tag{3.16}$$

where $Q(.)$ represents the quantile function. When the distribution is symmetric, $S = 0$ and when the distribution is right (or left) skewed, $S > 0$ (or $S < 0$). As

K increases, the tail of the distribution becomes heavier. These measures are less sensitive to outliers and they exist even for distributions without moments.

Table 1. Kurtosis and skewness of the LW and LL distributions for various values of parameters.

	LW Distribution						**LL Distribution**				
	a	b	θ	K	S		α	σ	θ	K	S
Case 1	0.5	5	10	78.62528	181.4882	Case 4	0.5	5	10	87.62863	0.40513
	0.8	5	10	12.146	15.82786		0.8	5	10	20.65822	2.07005
	1	5	10	5.73755	6.18065		1	5	10	15.1531	4.40341
	2.5	5	10	-0.16491	0.36431		2.5	5	10	8.186753	8.4547
	4	5	10	-0.25576	-0.12587		4	5	10	7.140732	3.64231
	10	5	10	0.58162	-3.37595		10	5	10	6.255138	5.98153
	20	5	10	1.28433	-22.4142		20	5	10	5.98975	48644.42
Case 2	10	0.5	5	0.00686	-2733.558	Case 5	10	0.5	5	6.411015	595912.6
	10	0.8	5	0.60125	-686.1663		10	0.8	5	6.407388	145478
	10	1	5	0.61113	-351.3172		10	1	5	6.407437	74484.5
	10	2.5	5	0.61058	-22.4843		10	2.5	5	6.407439	4767.008
	10	4	5	0.61058	-5.48933		10	4	5	6.407439	1163.82
	10	10	5	0.61058	-0.35132		10	10	5	6.407439	74.4845
	10	20	5	0.61058	-0.04391		10	20	5	6.407439	9.31056
Case 3	5	10	0.5	0.2732	-0.01896	Case 6	5	10	0.5	*	*
	5	10	0.8	0.14226	-0.02122		5	10	0.8	*	*
	5	10	1	0.08343	-0.02226		5	10	1	80.94593	0.01918
	5	10	2.5	-0.07289	-0.02817		5	10	2.5	10.66335	0.71234
	5	10	4	-0.10265	-0.03374		5	10	4	8.97689	3.95326
	5	10	10	-0.11825	-0.05397		5	10	10	6.829499	90.4188
	5	10	20	-0.1196	-0.08121		5	10	20	6.430927	843.1556

*The integral is probably divergent

We present skewness and kurtosis of the LW and LL distributions for various values of parameters in Table 1. Table 1 reveals that for fixed b and θ, the kurtosis initially decreases and thereafter increases in Case 1. Besides, the skewness decreases for fixed b and θ in Case 1 when a increases While a is from 0.5 to 1 or a is greater than 4, the LW distribution has positive kurtosis so it is called as leptokurtic distribution. For $a = 2.5$ or $a = 4$, we obtain negative kurtosis and platykurtic distribution. Table 1 also reveals that the skewness increases when b increases for fixed a and θ in Case 2. The kurtosis does not vary for $b \geqslant 2.5$ in Case 2 for fixed a and θ. It can be concluded that the parameter b does not effect on the kurtosis in Case 2. Note that we have leptokurtic and left skewed distribution in Case 2. Especially, the kurtosis is almost zero for and the distribution is called as mesokurtic. For fixed a and b, the skewness and kurtosis decrease when θ increases in Case 3. The LW distribution has more rounded peak and thinner tails while θ increases. In Case 4, the kurtosis decreases and the skewness increases when α increases for fixed σ and θ. The parameter σ does not effect on the kurtosis and while σ increases, the skewness decreases in Case 5. It can be noticed from Table 1 that the kurtosis decreases as θ increases. While θ increases, the effect on the kurtosis of the change in θ parameter decreases. Conversely, the effect on the skewness of the change in θ parameter decreases. Table 1 indicates that the LL distribution is right skewed and leptokurtic for all selected values of the parameters.

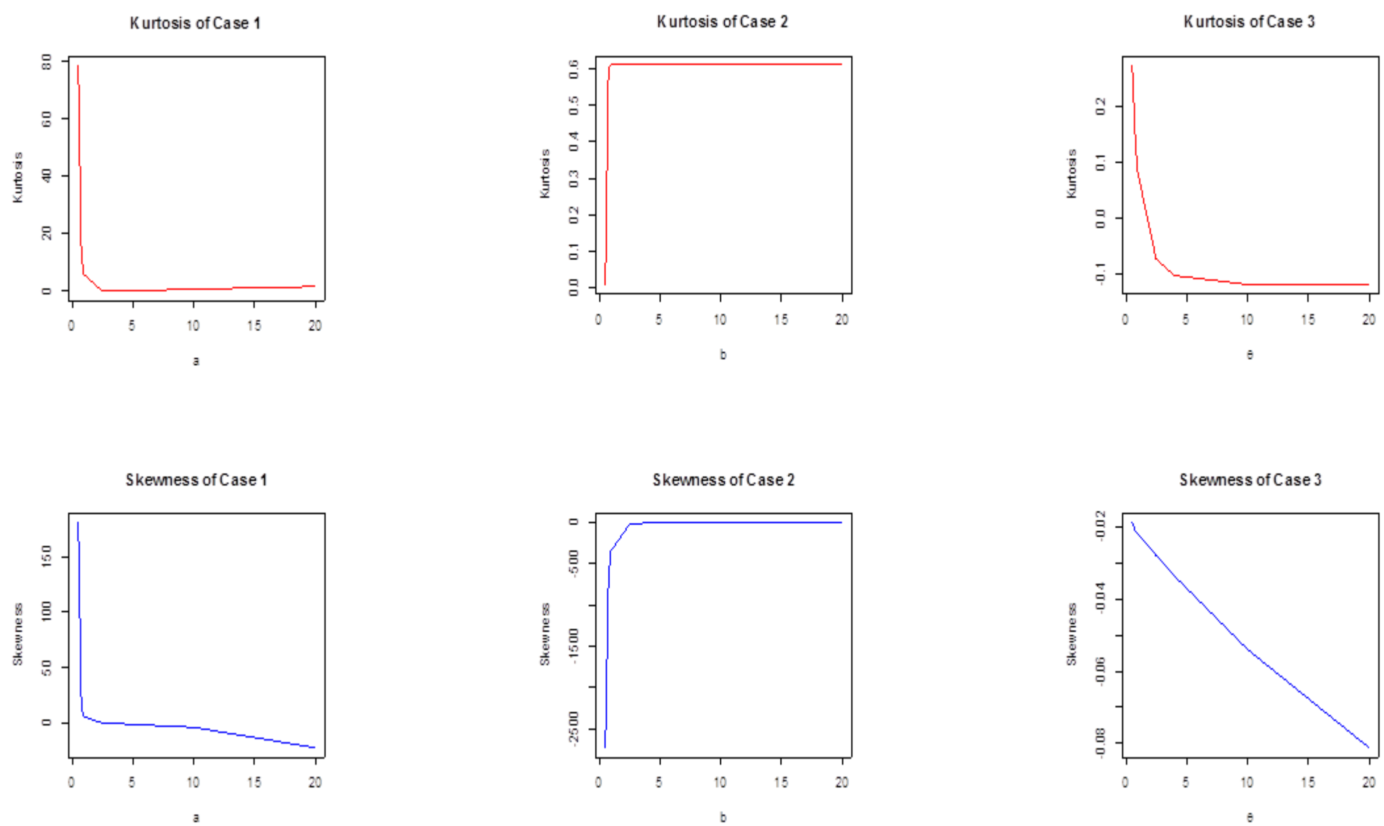

Figure 3. Various kurtosis and skewness shapes of the LW distribution for related parameters.

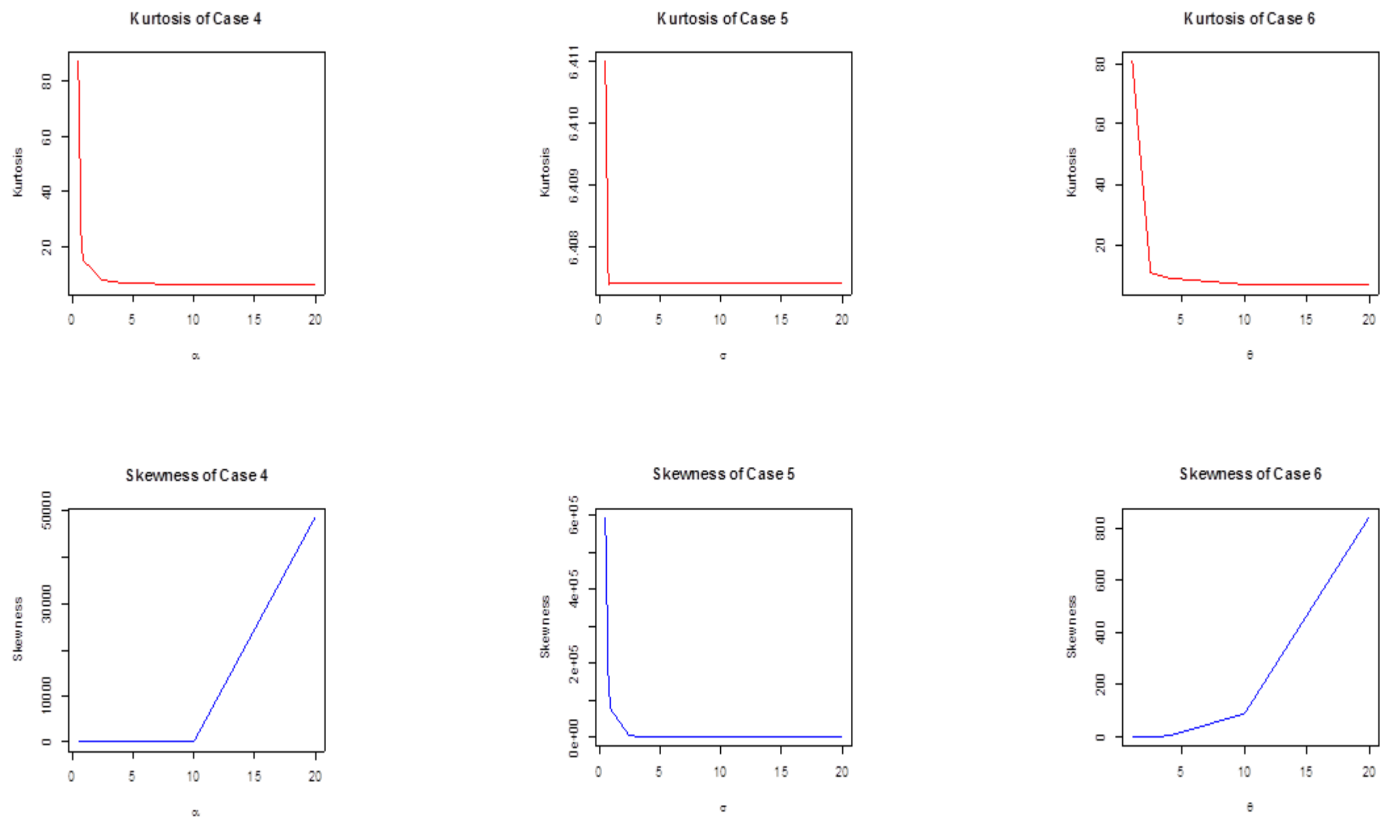

Figure 4. Various kurtosis and skewness shapes of the LL distribution for related parameters.

Figures 3 and 4 are given to show which parameters lead to a particular properties of the distributions. Figure 3 shows that kurtosis and skewness decrease and an exponential decay shapes occurs when increases for fixed other parameters (Case 1). On the other hand, exponential growths occur for kurtosis and skewness as increases (Case 2). If increases while other parameters are fixed (Case 3), kurtosis has an exponential decay shape, whereas skewness increases almost linearly. The shape changes for the LL pdf can be seen in Figure 4. The increase of parameter causes an exponential decay on kurtosis. Skewness is effected slightly by the increase of parameter to a point, then skewness increases almost linearly (Case 4). When increases for fixed other parameters (Case 5), kurtosis and skewness decrease. The increase of parameter in Case 6 causes decreasing kurtosis but increasing skewness.

3.6. Moments. Some of the most important characteristics of a distribution can be studied through moments. Let $G(x;\xi) = u$ and $G^{-1}(x;\xi) = Q(u) = x$, then the nth moment $\mu_n^{'} = E(x^n)$, $n = 1, 2, ...$, can be obtained as

$$\begin{aligned}
\mu_n^{'} &= E(X^n) = \int_0^\infty x^n \frac{g(x;\xi)}{1-G(x;\xi)} h(-log\left[1-G(x;\xi)\right])dx \\
&= \frac{\theta^2}{\theta+1}\int_0^1 Q^n(u)(1-u)^{\theta-1}\left[1-log\,(1-u)\right]du \\
&= \frac{\theta^2}{\theta+1}\int_0^1 Q^n(u)\left[\sum_{k=0}^{\infty}(-1)^k\binom{\theta-1}{k}u^k\right]\left[1+u\sum_{i=0}^{\infty}\frac{u^i}{i+1}\right] \\
&= \frac{\theta^2}{\theta+1}\left\{\sum_{i=0}^{\infty}(-1)^k\binom{\theta-1}{k}\int_0^1 Q^n(u)u^k du + \sum_{k=0}^{\infty}\sum_{i=0}^{\infty}\frac{(-1)^k}{i+1}\binom{\theta-1}{k}\int_0^1 Q^n(u)u^{k+i+1}du\right\} \\
&= \frac{\theta^2}{\theta+1}\left\{\sum_{i=0}^{\infty}(-1)^k\binom{\theta-1}{k}I(n,k) + \sum_{k=0}^{\infty}\sum_{i=0}^{\infty}\frac{(-1)^k}{i+1}\binom{\theta-1}{k}I(n,k+i+1)\right\} \qquad (3.17)
\end{aligned}$$

where $I(n,k) = \int_0^1 Q^n(u)u^k du$. Further, the central moments (μ_n) and cumulants (κ_n) of the X can be obtained, respectively, as

$$\mu_n = \sum_{k=0}^{r}(-1)^k\binom{n}{k}\mu_1^{'k}\mu_{n-k}^{'} \text{ and } \kappa_n = \mu_n^{'} - \sum_{k=0}^{r}(-1)^k\binom{n}{k}\mu_1^{'k}\mu_{n-k}^{'} \qquad (3.18)$$

where $\kappa_1 = \mu_1^{'}$, $\kappa_2 = \mu_2^{'} - \mu_1^{'2}$, and $\kappa_3 = \mu_3^{'} - 3\mu_2^{'}\mu_1^{'} + 2\mu_1^{'3}$, etc. The skewness $\gamma_1 = \kappa_3/\kappa_2^{3/2}$ and kurtosis $\gamma_2 = \kappa_4/\kappa_2^2$ can also been computed from the second, third and fourth cumulants.

First four ordinary moments of the LW and LL distributions for various values of parameters presented in Tables 2 and 3, respectively.

3.7. Moment Generating Function. The moment generating function (mgf) is widely used as an alternative way to analytical results compared with working directly with the pdf and cdf. Let $G(x;\xi) = u$ and $G(x;\xi)^{-1} = Q(u) = x$, then we give a formula for the mgf $M(t) = E(e^{tX})$ of X as

Table 2. First four moments of the LW distribution for various values of a, b and θ.

a	b	θ	$\mu_1^{'}$	$\mu_2^{'}$	$\mu_3^{'}$	$\mu_4^{'}$
0.5	5	10	0.118182	0.081818	0.139091	0.435273
5	5	10	2.949293	9.151767	29.58539	98.96873
20	5	10	4.357902	19.06387	83.69372	368.6616
10	0.5	5	0.411711	0.171913	0.007268	0.003107
10	5	5	4.117108	17.19131	72.68229	310.7231
10	20	5	16.46843	275.0609	4651.666	79545.11
5	10	0.5	11.95325	148.2952	1896.042	24865.04
5	10	5	6.876524	49.71569	374.2077	2912.798
5	10	20	5.091351	27.27943	152.3065	880.1200

Table 3. First four moments of the LL distribution for various values of α, σ and θ.

α	σ	θ	$\mu_1^{'}$	$\mu_2^{'}$	$\mu_3^{'}$	$\mu_4^{'}$
0.5	5	10	1.392045	5.271464	46.63826	1216.856
5	5	10	0.111507	0.025244	0.000871	0.00408
20	5	10	0.027421	0.001501	0.000123	0.000014
10	0.5	5	0.011939	0.000286	0.00001	0.000005
10	5	5	0.119395	0.028565	0.010319	0.005021
10	20	5	0.477579	0.457034	0.60436	1.285388
5	10	1	4.0625	40.97222	927.0833	62083.33
5	10	5	0.489005	0.490859	0.762927	1.639918
5	10	20	0.105869	0.022607	0.007305	0.003176

$$
\begin{aligned}
(3.19)\quad \mu_n^{'} &= E(e^{tX}) = \int_0^{\infty} e^{tX}\frac{g(x;\xi)}{1-G(x;\xi)}h(-log\,[1-G(x;\xi)]) \\
&= \frac{\theta^2}{\theta+1}\int_0^{\infty} e^{tQ(u)}(1-u)^{\theta-1}\,[1-log(1-u)]\,du \\
&= \frac{\theta^2}{\theta+1}\int_0^{\infty} e^{tQ(u)}\left[\sum_{k=0}^{\infty}(-1)^k\binom{\theta-1}{k}u^k\right]\left[1+u\sum_{i=0}^{\infty}\frac{u^i}{i+1}\right] \\
&= \frac{\theta^2}{\theta+1}\left\{\sum_{k=0}^{\infty}(-1)^k\binom{\theta-1}{k}I_e(t,k)+\sum_{k=0}^{\infty}\sum_{i=0}^{\infty}\frac{(-1)^k}{i+1}\binom{\theta-1}{k}I_e(t,k+i+1)\right\}
\end{aligned}
$$

where $I_e(t,k) = \int_0^{\infty} e^{tQ(u)}u^k$

4. Other Measures

4.1. Entropies. The entropy of a random variable X with density function $f(x)$ is a measure of variation of the uncertainty. Two popular entropy measures are the Renyi and Shannon entropies [31],[32]. Here, we derive expressions for the

Renyi and Shannon entropies when X is a Lindley-G random variable. The Renyi entropy of a random variable with pdf $f(x)$ is defined as

$$I_R(\gamma) = \frac{1}{1-\gamma} log \int_0^\infty f^\gamma(x)dx, \tag{4.1}$$

for $\gamma > 0$ and $\gamma \neq 1$. Using the power series in (3.1) and also the generalized binomial expansions in (3.2) and (3.3), we obtain

$$\begin{aligned}\int_0^\infty f^\gamma_{Lindley-G}(x;\theta,\xi)dx &= \int_0^\infty \frac{\theta^{2\gamma}}{(\theta+1)^\gamma} \left[1 - log\left(1-G(x;\xi)\right)\right]^\gamma \\ &\times \left[1-G(x;\xi)\right]^{\gamma\theta-\gamma} g^\gamma(x;\xi)dx \\ &= \frac{\gamma\theta^{2\gamma}}{(\theta+1)^\gamma} D_{j,k,i} I_{i,k}\end{aligned}$$

where $D_{j,k,i} = \sum_{j=0}^\infty \sum_{k=0}^\infty \sum_{i=0}^\infty \binom{\gamma}{j} c_{j,k} (-1)^i \binom{\gamma\theta-\gamma}{i}$, $I_{i,k} = \int_0^\infty G^{i+k+\gamma}(x) g^\gamma(x)dx$ and $c_{j,k} = (ka_0)^{-1} \sum_{m=1}^i \left[k - m(j+1)\right] a_m c_{j,k-m}$ for $a_k = (k+1)^{-1}$. Then, the Renyi entropy of the Lindley-G distribution is given by

$$I_R(\gamma) = \frac{1}{1-\gamma} log \left[\frac{\gamma\theta^{2\gamma}}{(\theta+1)^\gamma} D_{j,k,i} I_{i,k}\right] \tag{4.2}$$

The Shannon entropy of a random variable X is defined by $E[-logf(X)]$. It is the special case of Renyi entropy when $\gamma > 1$. Using the pdf of Lindley-G family, we obtain $-logf_{Lindley-G}(x;\theta,\xi) = -log\left[1 - log\left(1-G(x;\xi)\right)\right] - (\theta - 1)log\left(1-G(x;\xi)\right) - log\theta^2 + log\theta^2 + log(\theta+1) - log\left(g(x;\xi)\right)$ and for the Lindley-G family direct calculation yields

$$\begin{aligned}E\left[-logf_{Lindley-G}(X;\theta,\xi)\right] &= \sum_{i=0}^\infty \sum_{k=0}^\infty \frac{(-1)^{i+1}}{i+1} c_{i+1,k} E\left[(X;\xi)\right] \\ &+ (\theta-1) \sum_{k=0}^\infty \frac{1}{k+1} E\left[G^{k+1}(X;\xi)\right] \\ &- log\theta^2 + log\left(\theta+1\right) - E\left[log\left(g(X;\xi)\right)\right]\end{aligned} \tag{4.3}$$

where $c_{i+1,k} = (ka_0)^{-1} \sum_{m=1}^k \left[m(i+2) - k\right] a_m c_{i+1,k-m}$ and $a_k = (k+1)^{-1}$.

4.2. Reliability. In the context of reliability, the stress-strength model describes the life of a component which has a random strength X_1 that is subjected to a random stress X_2. The component fails at the instant that the stress applied to it exceeds the strength, and the component will function satisfactorily whenever $X_1 > X_2$. Hence, $R = Pr(X_1 > X_2)$ is a measure of component reliability. Here, we obtain the reliability function R when $X_1 \sim Lindley - G(\theta_1, \xi)$ and $X_2 \sim Lindley - G(\theta_1, \xi)$ are independent random variables. Probabilities of this form have many applications especially in engineering concepts.

Let f_i denote the pdf of X_i and F_i denote the cdf of X_i for $i = 1, 2$, then we obtain

$$
\begin{aligned}
f_1(x)F_2(x) = \quad & \frac{\theta_1^2}{\theta_1+1}\left\{g(x;\xi)\left[1-G(x;\xi)\right]^{\theta_1-1} - g(x;\xi)log\left[1-G(x;\xi)\right]\left[1-G(x;\xi)\right]^{\theta_1-1}\right\} \\
& -\frac{\theta_1^2}{\theta_1+1}\left\{g(x;\xi)\left[1-G(x;\xi)\right]^{\theta_1+\theta_2-2} - \left(\frac{\theta_2}{\theta_2+1}-1\right)g(x;\xi)log\left[1-G(x;\xi)\right]\right. \\
& \left.\times\left[1-G(x;\xi)\right]^{\theta_1+\theta_2-2}\right\} \\
& +\frac{\theta_1^2}{\theta_1+1}\left\{\frac{\theta_2}{\theta_2+1}g(x;\xi)log\left[1-G(x;\xi)\right]^2\left[1-G(x;\xi)\right]^{\theta_1+\theta_2-2}\right\}
\end{aligned} \tag{4.4}
$$

From (4.4), the reliability function for the Lindley-G family is given by

$$
\begin{aligned}
R &= \int_0^\infty f_1(x)F_2(x)dx \\
&= \frac{\theta_1^2}{\theta_1+1}\left\{-\frac{1}{\theta_1} - \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{j+1}}{(ji+j+2)(i+1)}\binom{\theta_1-1}{j}\right\} \\
&- \frac{\theta_1^2}{\theta_1+1}\left\{-\frac{1}{\theta_1+\theta_2-2}\left(\frac{\theta_2}{\theta_1+1}-1\right)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{j+1}}{(ji+j+2)(i+1)}\binom{\theta_1+\theta_2-2}{j}\right\} \\
&+ \frac{\theta_1^2}{\theta_1+1}\left\{\frac{\theta_2}{\theta_2+1}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i\binom{k}{i}\frac{c_{2,j}}{i+j+3}\right\}
\end{aligned} \tag{4.5}
$$

where $c_{2,j} = (ja_0)^{-1}\sum_{m=1}^{j}(3m-j)a_m c_{2,j-m}$ and $a_j = (j+1)^{-1}$.

4.3. Order Statistics. Order statistics make their appearance in many areas of statistical theory and practice. They enter in the problems of estimation and hypothesis tests in a variety of ways. Therefore, we now discuss some properties of the order statistics for the proposed class of distributions.

Let $X_{i:n}$ denote the ith order statistic. Then, the pdf $f_{i:n}(x)$ of the ith order statistic for a random sample $X_1, X_2, ..., X_n$ from $F(x)$ distribution is given by

$$
\begin{aligned}
f_{i:n}(x) &= \frac{n!}{(i-1)!(n-i)!}f(x)F(x)^{i-1}\left[1-F(x)\right]^{n-i} \\
&= \frac{n!}{(i-1)!(n-i)!}\sum_{j=1}^{n-i}\binom{n-i}{j}f(x)F^{j+i-1}(x)
\end{aligned} \tag{4.6}
$$

Using (3.2) and (3.3), the pdf of $X_{i:n}$ for the Lindley-G family can be expressed as

$$
\begin{aligned}
f_{i:n}(x) &= \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}\left(1+\sum_{k=0}^{\infty}\frac{G^{k+1}(x)}{k+1}\sum_{l=0}^{\theta-1}G^l(x)\frac{\theta^2}{\theta+1}g(x)\right) \\
&= \sum_{z=0}^{i+j-1}(-1)^z\binom{i+j-1}{z}\left\{\left[\sum_{w=0}^{\theta}(-1)^w\binom{\theta}{w}G^w(x)\right]^z\left(1+\frac{\theta}{\theta+1}\sum_{s=0}^{\infty}\frac{G^{s+1}(x)}{s+1}\right)^z\right\}
\end{aligned} \tag{4.7}
$$

where $f(.)$ and $F(.)$ are the probability density and cumulative density functions of the Lindley-G distribution, respectively.

5. Maximum Likelihood Estimation

Several approaches for parameter point estimation have been proposed in the literature but the maximum likelihood method is the most commonly employed. The maximum likelihood estimates (MLEs) enjoy desirable properties and can be used for constructing confidence intervals. Large sample theory for these estimates delivers simple approximations that work well infinite samples. Statisticians often seek to approximate quantities such as the density of a test statistic that depends on the sample size in order to obtain better approximate distributions. The resulting approximation for the MLEs in distribution theory is easily handled either analytically or numerically.

Let $x_1, x_2, ..., x_n$ be observed values from the Lindley-G distribution with parameters θ and ξ. The likelihood function for (θ, ξ) is given by

$$L(\theta;\xi) = \prod_{i=1}^{n} \left\{ \frac{\theta^2}{\theta+1} \left[1 - log\left(1 - G(x_i;\xi)\right)\right] \left(1 - G(x_i;\xi)\right)^{\theta-1} g(x_i;\xi) \right\}.$$

The log-likelihood function of the parameters (θ, ξ) can be expressed as

$$\text{(5.1)} \qquad logL = n\left[2log\theta - log(\theta+1)\right] + \sum_{i=1}^{n} log\left[1 - log\left(1 - G(x_i;\xi)\right)\right]$$
$$+ (\theta-1)\sum_{i=1}^{n} log\left[1 - G(x_i;\xi)\right] + \sum_{i=1}^{n} log\left[g(x_i;\xi)\right]$$

The log-likelihood function can be maximized either directly by using SAS (PROC NLMIXED) or Ox program (sub-routine MaxBFGS) [33] or by solving the nonlinear likelihood equations obtained by differentiating (5.1). The first derivatives of $logL$ with respect to parameters θ and ξ are

$$\begin{aligned} \tfrac{\partial logL}{\partial \theta} &= \frac{2n}{\theta} - \frac{n}{\theta+1} + \sum_{i=1}^{n} log\left[1 - F(x_i;\xi)\right], \\ \tfrac{\partial logL}{\partial \xi} &= \sum_{i=1}^{n} \frac{1}{\left[1 - log\left(1 - G(x_i;\xi)\right)\right]\left[1 - G(x_i;\xi)\right]} \frac{\partial G(x_i;\xi)}{\partial \xi} \\ &\quad -(\theta-1)\sum_{i=1}^{n} \frac{1}{\left[1 - G(x_i;\xi)\right]} \frac{\partial G(x_i;\xi)}{\partial \xi}. \end{aligned}$$

The MLEs of θ and ξ, say $\hat{\theta}$ and $\hat{\xi}$, are the simultaneous solutions of the equations $\frac{\partial logL}{\partial \theta} = 0, \frac{\partial logL}{\partial \xi} = 0$. Maximization of (5.1) can be performed by using *nlm, adequacymodel* or *optimize* in R statistical package. For interval estimation of (θ, ξ) and hypothesis tests, we require the observed information matrix. The observed information matrix for (θ, ξ) can be determined as

$$I = \begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix}$$

where

$$I_{11} = \frac{\partial logL}{\partial \theta^2} = \frac{-2n}{\theta^2} + \frac{n}{(\theta+1)^2},$$

$$I_{22} = \frac{\partial logL}{\partial \xi^2} = \sum_{i=1}^{n} \frac{log\,[1-G(x_i;\xi)]}{[1-log\,(1-G(x_i;\xi))]^2\,[1-G(x_i;\xi)]^2} \left(\frac{\partial G(x_i;\xi)}{\partial \xi}\right)$$

$$-\,(\theta-1)\sum_{i=1}^{n}\left[\frac{1}{[1-G(x_i;\xi)^2]}\frac{\partial G(x_i;\xi)}{\partial \xi} + \frac{\partial^2 G(x_i;\xi)}{\partial \xi^2}\right]$$

$$I_{12} = \frac{\partial^2 logL}{\partial \theta \partial \xi} = -\sum_{i=1}^{n} \frac{1}{[1-F(x_i;\xi)]}\frac{\partial F(x_i;\xi)}{\partial \xi}.$$

For large n, distribution of $(\theta-\hat{\theta}, \xi-\hat{\xi})$ can be approximated by a $(r+1)$ multivariate normal distribution with zero means and variance-covariance matrix I^{-1}. Some statistical properties of $(\hat{\theta}, \hat{\xi})$ can be derived based on this normal approximation.

6. Application

In this section, we analyze three real data sets to demonstrate the performance of the LW and LL distributions in practice. We obtained the data sets from the Turkish State Meteorological Service (http://www.mgm.gov.tr/en-us/forecast-5days.aspx). First, we describe the data sets. Then, we fit some distributions to the data sets using MLE and the aim is to compare proposed distributions with several kind of distributions.

The model selection is carried out using the Akaike information criterion (AIC), Consistent Akaike information criteria (CAIC), Bayesian information criterion (BIC), and Hannan-Quinn information criterion (HQIC) given by

$$\begin{aligned} AIC &= -2logL + 2p, \\ CAIC &= -2logL + \frac{2pn}{n-p-1}, \\ BIC &= -2logL + plogn, \\ HQIC &= -2logL + 2plog(logn), \end{aligned} \tag{6.1}$$

where p is the number of the model parameters and n is the sample size. The model with minimum AIC (or CAIC, BIC, and HQIC) value is chosen as the best model to fit the data.

Finally, we give the histograms of the data sets and plot the fitted density functions to obtain a visual comparison of the adjustments of the models.

6.1. Particulate Matter Data. The considered first data set is corresponding to daily atmospheric particulate matter (PM_{10}) observations. PM_{10} is microscopic solid or liquid matter suspended in the Earth's atmosphere. PM_{10} is formed by the mixture of oil, gasoline, and diesel fuel combustions. This pollutant is analyzed in this study because it may indicate a much higher health risk despite its low representation when compared to gas pollutants. It enters the body exclusively through the respiratory system and its effects depend on whether or not it enters the respiratory tract, with the degree of penetration depending on particle size [34]. Kocaeli is one of the most industrialized cities of Turkey. Many industrial facilities in terms of air pollution, constitute a risk. Because of Kocaeli's location, which is on the junction of Turkey's whole motorway transport, increases the importance of this issue. Hence, the analysis of PM10 is important for Kocaeli.

The daily PM10 values are measured and 683 observations are recorded for Kocaeli-Dilovasi station. The period of the data set is between 2012 and 2015.

The descriptive statistics of the PM10 data is given in Table 3. Table 3 indicates that the data has positive skewness and kurtosis. Note that the right tail is longer, the mass of the distribution is concentrated on the left of the figure and it has a peaked distribution.

Table 4. Descriptive statistics of the PM10 data.

Mean	89.34261
Standard Deviation	44.81
Median	76
Trimmed Mean	82.74
Median Absolute Deviation	31.13
Minimum	23
Maximum	390
Range	367
Skewness	1.72478
Kurtosis	4.54262
Standard Error	1.71

We fit the LW distribution to the data sets using MLE and compared the proposed distribution with W (Weibull), L (Lomax), SL (Standard Lomax), E (Exponential) and Lindley distributions. We present the results of AIC, CAIC, BIC, and HQIC statistics for the models in Table 4. These results show that the LW distribution has the lowest AIC, CAIC, BIC and HQIC values among all the fitted models, and so it could be chosen as the best model.

Table 5. The measures AIC, CAIC, BIC, and HQIC for PM10 data.

Distribution	AIC	CAIC	BIC	HQIC
LW	**6946.458**	**6946.493**	**6960.037**	**6951.713**
W	7003.099	7003.116	7012.152	7006.602
Lindley	7130.332	7130.337	7134.858	7132.083
E	7504.726	7504.731	7509.252	7506.477
L	7506.757	7506.775	7515.81	7510.261
SL	9406.102	9406.108	9410.629	9407.854

We obtain the MLEs of the model parameters for PM10 data in Table 5.

Table 6. The MLEs of the models for PM10 data.

Distribution	Estimated Parameters
LW(a, b, θ)	(1.54074828, 6.53244086, 0.03192352)
W(a, b)	(2.119833, 101.312713)
Lindley(θ)	(0.02215732)
E(λ)	(89.34252)
L(α, σ)	(16164.71, 1444141.14)
SL(α)	(0.2271563)

Finally, we obtain a density plot in Figure 5 to compare the fitted densities of the models with the empirical histogram of the observed data. Figure 5 shows that the fitted density for the LW distribution (the black one) is closer to the empirical histogram than the fits of the other distributions.

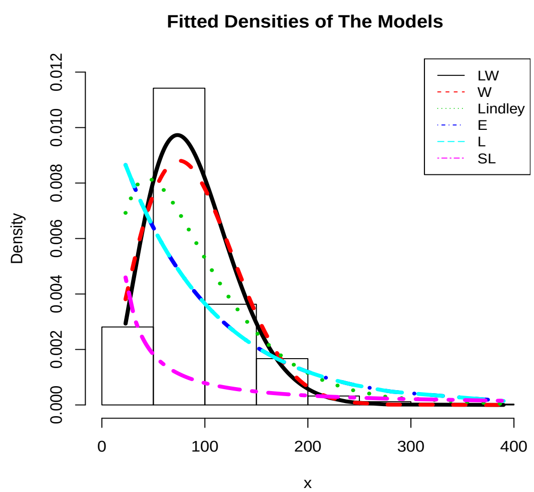

Figure 5. Fitted densities of the distributions for the PM10 data.

6.2. Sulfur Dioxide Data. Sulfur dioxide (SO2) is known to be one of the combustion end products of sulfur containing fossil fuels. The major health impact of SO2 include effects on breathing, respiratory illness, weakness of lung defenses, increase in the effects of existing respiratory and cardiovascular disease, and death [35]. Explanations of effects of pressure, temperature, and wind speed on the samplers for SO2 has been reported in many studies [36],[37]. In particular, cities with heavy industrial activities have high levels of SO2 concentrations. With a population of over one million, Bursa is one of the most crowded cities in Turkey and has heavy industry consisting of automotive, textile, and food industries. Due to insufficient ventilation and high population and industrial densities, Bursa has a potential for serious air pollution problems. Hence, the estimation of SO2 measures is important for Bursa. For the second application, we consider a real data set corresponding to sulfur dioxide (SO_2) measure in Bursa (in days). The recorded 652 observations are between 2012 and 2015. The descriptive statistics of the SO_2 data is given in Table 6. As seen in Table 6, the data is skewed to the right and positive kurtosis indicates a peaked distribution.

Table 7. Descriptive statistics of SO_2 data.

Mean	7.07
Standard Deviation	7
Median	4
Trimmed Mean	5.66
Median Absolute Deviation	2.97
Minimum	0
Maximum	48
Range	48
Skewness	2.09
Kurtosis	5.15
Standard Error	0.27

We fit the SO_2 data with proposed LL and LSL (Lindley-Standard Lomax), EL (Extended Lomax), L (Lomax), SL (Standard Lomax), Lindley and E (Exponential) distributions.

We obtained AIC, CAIC, BIC, and HQIC statistics to compare models in Table 7. These results show that the LL distribution has the lowest AIC, CAIC, BIC and HQIC values among all the fitted models, and so it could be chosen as the best model.

Table 8. The measures AIC, CAIC, BIC, and HQIC for SO_2 data.

Distribution	AIC	CAIC	BIC	HQIC
LL	**3775.281**	**3775.318**	**3788.721**	**3780.493**
LSL	4064.214	4064.2324	4073.174	4067.689
EL	3791.675	3791.694	3800.635	3795.150
L	3859.221	3859.240	3868.182	3862.696
SL	4432.379	4432.385	4436.859	4434.116
Lindley	3858.727	3858.733	3863.207	3860.464
E	3857.110	3857.116	3861.590	3858.847

Here, for more discussion, we obtain the MLEs of parameters for SO_2 data in Table 8.

The histogram of SO_2 data and plots of the fitted distributions are shown in Figure 6. We conclude from Figure 6 that the LL distribution yield the best fit and hence can be adequate for the data.

6.3. Ozone Data. The serious air quality problems, specifically inverse health effects, have been experienced in megacities of both developing and developed

Table 9. The MLEs of the models for SO2 data.

Distribution	Estimated Parameters
LL(α, σ, θ)	(158.680729,10.673049,0.027630)
LSL(α, θ)	(64.487807, 0.017045)
EL(α, λ)	(50.051436, 2.245115)
L(α, σ)	(163.9941, 1153.1473)
SL(α)	(0.553628)
Lindley(θ)	(0.2540985)
E(λ)	(7.073619)

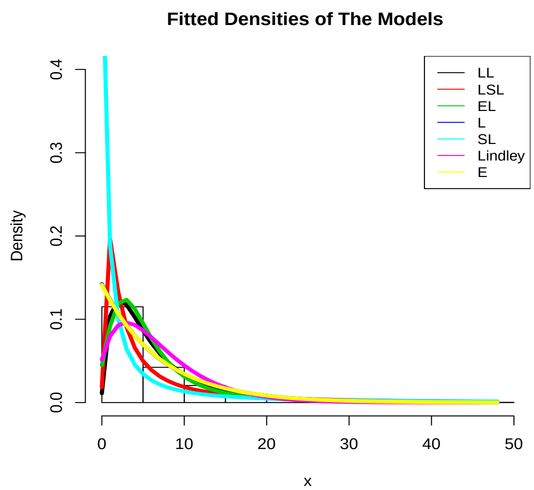

Figure 6. Fitted densities of the distributions for the SO2 data.

countries due to the exposure to high concentrations of ozone (O3). Istanbul is the center of industry, economics, finance and culture in Turkey. It has a serious air pollution problem due to domestic heating, industry and traffic. O3 is an important pollutant produced by a series of complicated photochemical reactions in Istanbul. Therefore, the estimation of O3 levels is vital for Istanbul.

The third data set represents 592 daily ozone (O3) measures from 2012 to 2015 in Kadikoy, Istanbul.

Table 9 shows the descriptive statistics of the O3 data. Table 9 shows that the data is skewed to the right and positive kurtosis indicates a peaked distribution.

The values of AIC, CAIC, BIC, and HQIC are presented in Table 10. Based on Table 10, we obtain that the LL model gives the lowest values for the AIC, CAIC, BIC, and HQIC for O3 data.

It is clear that LL distribution provides the overall best fit and therefore could be chosen as the more adequate model for explaining O3 data set. Table 11 lists the MLEs of the parameters for O3 data.

Table 10. Descriptive statistics of O3 data.

Mean	13.72
Standard Deviation	11.59
Median	10
Trimmed Mean	11.97
Median Absolute Deviation	8.9
Minimum	0
Maximum	59
Range	59
Skewness	1.32
Kurtosis	1.41
Standard Error	0.48

Table 11. The measures AIC, CAIC, BIC, and HQIC for O3 data.

Distribution	AIC	CAIC	BIC	HQIC
LL	**4252.959**	**4253.000**	**4266.110**	**4258.082**
LSL	4671.969	4671.989	4680.736	4675.384
EL	4284.424	4284.444	4293.191	4287.839
L	4288.543	4288.564	4297.31	4291.958
SL	5032.572	5032.579	5036.956	5034.280
Lindley	4266.889	4266.896	4271.272	4268.596
E	4286.543	4286.550	4290.927	4288.251

Table 12. The MLEs of the models for O3 data.

Distribution	Estimated Parameters
LL(α, σ, θ)	(155.62618025, 36.53180222,0.04260437)
LSL(α, θ)	(62.72699803, 0.01332717)
EL(α, λ)	(123.785305, 2.022086)
L(α, σ)	(4405680, 60469137)
SL(α)	(0.4199579)
Lindley(θ)	(0.1370131)
E(λ)	(13.71791)

Table 11 shows that the LL distribution can fit the current data better than other models. Then, the histogram of O3 data and plots of the fitted distributions are shown in Figure 7. We also conclude from Figure 7 that the fitted LL distribution yield the best fits and hence can be adequate for the data.

7. Concluding Remarks

The idea of generating new extended models from classic ones has been of great interest among researchers in the past decade. In this paper, we introduce a new class of models called the 'Lindley-G' family of distributions which can generate all classical continuous distributions. For any parent continuous distribution G, we

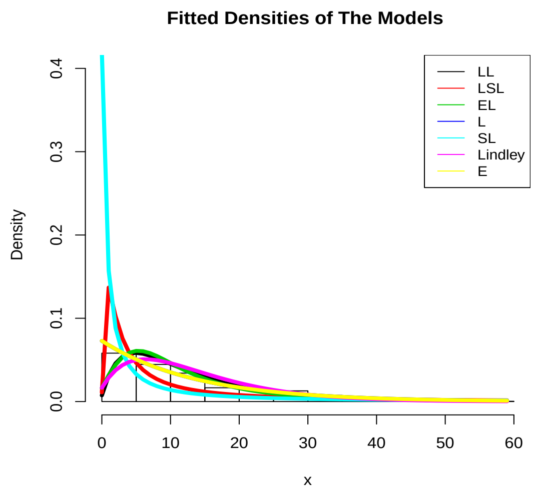

Figure 7. Fitted densities of the distributions for the O3 data.

define corresponding Lindley-G distribution. Hence, the new class extends several common distributions, such as the Exponential, Weibull and Lomax distributions. We study some of statistical and mathematical properties of the new generator, such as ordinary moments, cumulants, generating and quantile functions, Shannon entropy, Renyi entropy, and order statistics. We discuss maximum likelihood estimation and inference on the model parameters. Three applications of the new family demonstrate its usefulness and potentiality to analysis of real data.

References

[1] A.N. Marshall and I. Olkin, *A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families*, Biometrika **84**,641-552, (1997).

[2] N. Eugene, C. Lee and F. Famoye, *Beta-normal distribution and its applications*, Commun Stat Theory **31**,497-512,(2002).

[3] G.M. Cordeiro and M.D. Castro, *A new family of generalized distributions*, J Statist Comput Simulation, **81**, 883-898, (2011).

[4] C. Alexander, G.M. Cordeiro, E.M.M Ortega and J.M. Sarabia, *Generalized beta-generated distributions*, Comput Stat Data An **56**, 1880-1897, (2012).

[5] A. Alzaatreh, C. Lee and F. Famoye, *A new method for generating families of distributions*, Metron **71**, 63-79, (2013).

[6] A. Alzaghal, F. Famoye ,C. Lee, *Exponentiated T-X family of distributions with some applications*, Int J Prob Stat **2**, 31ÔÇô49,(2013).

[7] M. Bourguignon, R.B. Silva, G.M. Cordeiro, , *The Weibull-G family of probability distributions*, J Data Sci **12**, 53-68,(2014).

[8] G.M. Cordeiro, M. Alizadehand, E.M.M. Ortega,*The exponentiated half-logistic family of distributions: properties and applications*, J Prob Stat doi:10.1155/2014/864396 (2014).

[9] G.M. Cordeiro, E.M.M. Ortega., B. Popovic, R.R. Pescim, *The Lomax generator of distributions: Properties, minification process and regression model*, Appl Math Comput **247**, 465-486, (2014).

[10] S. Nadarajah, G.M. Cordeiro, E.M.M. Ortega, *The ZografosÔÇôBalakrishnan-G family of distributions: mathematical properties and applications*, Commun Stat Theory **44**, 186-215,(2015).
[11] D.V. Lindley, *Fiducial distributions and Bayes' theorem*, J Roy Stat Soc B **20**, 102-107,(1958).
[12] H.S. Bakouch, M.A. Bander, A.A. Al-Shaomrani, V.A.A. Marchi and F. Louzada, *An extended Lindley distribution*, J Korean Stat Soc **41**, 75-85, (2012).
[13] M.E. Ghitany, B. Atieh and S. Nadarajah, *Lindley distribution and its application*, Math Comput Simulat **78** 493-506, (2008).
[14] M.E. Ghitany, F. Alqallaf, D. K. Al-Mutairi and H.A. Husain, *A two-parameter Lindley distribution and its applications to survival data*, Math Comput Simulat **81**, 1190ÔÇô1201, (2011).
[15] J. Mazucheli and J.A. Achcar, *The Lindley distribution applied to competing risks lifetime data*, Comput Meth Prog Bio **104**, 188-92, (2011).
[16] I.S. Gradshteyn and I.M. Ryzhik,*Table of Integrals, Series and Products*, Academic Press, New York, 2007.
[17] G.S. Mudholkar and A.D. Hutson, *The exponentiated Weibull family: some properties and a flood data application*, Commun Stat Theory Meth **25**, 3059-3083, (1996).
[18] G.S. Mudholkar and D.K. Srivastava, , *Exponentiated Weibull family for analyzing bathtub failure-rate data*, IEEE Trans Rel **42**, 299-302, (1993).
[19] G.S. Mudholkar, D.K. Srivastava and M. Freimer, *The exponentiated Weibull family; a reanalysis of the bus motor failure data*, Technometrics **37**, 436-445, (1995).
[20] R.C. Gupta, P. L. Gupta, and R. D. Gupta, *Modeling failure time data by Lehmann alternatives*, Commun Stat Theory **27**, 887-904, (1998).
[21] R D. Gupta and D. Kundu, *Generalized Exponential Distributions* , Aust NZ J Stat **41**, 173-188, (1999).
[22] R.D. Gupta and D. Kundu, *Exponentiated Exponential Family; An Alternative to Gamma and Weibull*, Biometrical J **33**, 117-130, (2001).
[23] R.D. Gupta and D. Kundu, *Generalized Exponential Distributions: Different Methods of Estimation*, J Stat Comput Sim **69**, 315-338, (2001).
[24] S. Nadarajah and A.K. Gupta, *A generalized gamma distribution with application to drought data*, Math Comput Simulat **74**, 1-7, (2007).
[25] G.M. Cordeiro, S. Nadarajah and E.M.M. Ortega, *The Kumaraswamy Gumbel distribution*, Statist Meth Applic **21**, 139-168, (2012).
[26] M.A.R. Pascoa, E.M.M. Ortega, and G.M. Cordeiro, *The Kumaraswamy generalized gamma distribution with application in survival analysis*, Statist Methodol **8**, 411ÔÇô433, (2011).
[27] P.F. Paranaba, E.M.M. Ortega, G.M. Cordeiro and M.A.R. Pascoa, *The Kumaraswamy Burr XII distribution: theory and practice*, J Stat Comput Sim **82**, 1ÔÇô27, (2012).
[28] S. Nadarajah, G.M. Cordeiro and E.M.M. Ortega, *General results for the Kumaraswamy-G distribution*, J Stat Comput Sim, 951-972, (2012).
[29] A.J. Lemonte, W. Barreto-Souza and G.M. Cordeiro, *The exponentiated Kumaraswamy distribution and its log-transform*, Braz J Probab Statist **27** 31-53, (2013).
[30] S. Nadarajah, G.M. Cordeiro, and E. M. Ortega, *General results for the Kumaraswamy-G distribution*, J Stat Comput Sim **82** 951-979, (2012).
[31] A. Renyi, *On measures of entropy and information*, In: Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, 547-561. University of California Press, Berkeley, 1961.
[32] C.E. Shannon, *Prediction and entropy of printed English*, Bell Syst Tech J **30**, 50-64, (1951).
[33] J.A. Doornik, *Ox 5: An Object-Oriented Matrix Programming Language*, Timberlake Consultants, London, 2007.
[34] G. Ozel and S. Cakmakyapan, *A new approach to the prediction of PM10 concentrations in Central Anatolia Region*, Turkey, Atmos Pollut Res **6**, 735-741, (2015).
[35] A. Akdemir, *The creation of pollution mapping and measurement of ambient concentration of sulfur dioxide and nitrogen dioxide with passive sampler*, J Environ Health Sci Eng **12**, (2014).

[36] E.E. Ukpebor, S.I. Ahonkhal and H. Heydtman, *NO2 measurement with passive sampler: assessment of the sensitivity of two types of palmes diffusion tubes for NO2*, Intern J Environ Studies **61**, 67-71, (2004).

[37] T. Banerjee, S.B. Singh and R.K. Srivastava, *Development and performance evaluation of statistical models correlating air pollutants and meteorological variables at Pantnagar*, India Atmos Res **99** 505-517, (2011).