Journal of Statistical Theory and Applications, Vol. 16, No. 4 (December 2017) 490–507

A new family of distributions to analyze lifetime data

M. Mansoor[1], M. H. Tahir[1,†], Gauss M. Cordeiro[2], Ayman Alzaatreh[3], M. Zubair[1,4]

[1]*Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur-63100, Pakistan*
[2]*Department of Statistics, Federal University of Pernambuco, Recife, PE, Brazil*
[3]*Department of Mathematics and Statistics, American University of Sharjah, UAE*
[4]*Department of Statistics, Govt. S.E College, Bahawalpur-63100, Pakistan*
Emails: mansoor.abbasi143@gmail.com; †mtahir.stat@gmail.com(corresponding author); gausscordeiro@gmail.com; aalzaatreh@aus.edu; zubair.stat@yahoo.com

Received 18 November 2016

Accepted 14 March 2017

In this paper, a new family of distributions is proposed by using quantile functions of known distributions. Some general properties of this family are studied. A special case of the proposed family is studied in detail, namely the Lomax-Weibull distribution. Some structural properties of the special model are established. This distribution has been applied to several censored and uncensored data sets with various shapes.

Keywords: Censoring; log-logistic distribution; Lomax distribution; quantile function; T-X family; Weibull distribution

2000 Mathematics Subject Classification: 60E05, 62E10, 62N05

1. Introduction

Adding extra parameter to an existing family of distributions is very common in the statistical distribution theory in order to introduce more flexibility. For example, Azzalini (1985) introduced the skew-normal distribution by introducing an extra parameter to the normal distribution. Azzalini (1985)'s skew normal distribution takes the following form (for $\lambda \in \Re$)

$$f(x;\lambda) = 2\phi(x)\Phi(\lambda x),$$

where $\phi(x)$ and $\Phi(x)$ are the probability density function (PDF) and cumulative distribution function (CDF) of a standard normal distribution, and λ is the skewness parameter. Although Azzalini introduced this method mainly for normal distribution, but it can be easily used for other symmetric distributions.

Marshall and Olkin (1997) proposed a general method for generating a new family of distributions in terms of the survival function as

$$\bar{G}(x;\alpha) = \frac{\alpha \bar{F}(x)}{1-\bar{\alpha}\bar{F}(x)}; \quad x \in \Re, \quad \alpha > 0,$$

where $\bar{\alpha} = 1-\alpha$ and $\bar{F}(x) = 1-F(x)$. For more details on lifetime distributions one may refer to Marshall and Olkin (2010) and Lai (2013).

Eugene et al. (2002) introduced beta generated family of distribution, where the beta distribution is used as a generator. The CDF of beta generated family is given by

$$G(x) = \int_0^{F(x)} b(t)dt,$$

where $F(x)$ is the CDF of any random variable.

Alzaatreh et al. (2013) introduced a new method for generating families of continuous distributions called the T-X family by replacing the beta PDF with a PDF, $r(t)$, of a any continuous random variable and applying a link function $W(\cdot)$ that satisfies some specific conditions.

Recently, Aljarrah et al. (2014) used the link function $W(\cdot)$ to be the quantile function (QF) of a random variable Y to generate the so-called *T-X*$\{Y\}$ family. Alzaatreh et al. (2014) unified the notations of the T-X{Y} family as follows: Let T, R and Y be the random variables with CDFs $F_T(x) = P(T \le x)$, $F_R(x) = P(R \le x)$ and $F_Y(x) = P(Y \le x)$, respectively. The corresponding QFs are $Q_T(p)$, $Q_R(p)$ and $Q_Y(p)$, where the QF is defined by $Q_Z(p) = \inf\{z : F_Z(z) \ge p\}$, $0 < p < 1$. If the densities exist, we denote them by $f_T(x)$, $f_R(x)$ and $f_Y(x)$. Further, we consider the random variables $T \in (a,b)$ and $Y \in (c,d)$ for $-\infty \le a < b \le \infty$ and $-\infty \le c < d \le \infty$. The CDF of the T-R{Y} class is defined by

$$F_X(x) = \int_a^{Q_Y(F_R(x))} f_T(t)\,dt = \mathbb{P}\Big[T \le Q_Y\big(F_R(x)\big)\Big] = F_T\Big(Q_Y\big(F_R(x)\big)\Big). \tag{1.1}$$

The PDF and hazard rate function (HRF) corresponding to Equation(1.1) are, respectively, given by

$$f_X(x) = f_R(x)\,\frac{f_T\Big(Q_Y\big(F_R(x)\big)\Big)}{f_Y\Big(Q_Y\big(F_R(x)\big)\Big)} \tag{1.2}$$

and

$$h_X(x) = h_R(x)\,\frac{h_T\Big(Q_Y\big(F_R(x)\big)\Big)}{h_Y\Big(Q_Y\big(F_R(x)\big)\Big)}. \tag{1.3}$$

Remark 1.1. If X follows the T-R{Y}, then

(i) $X \overset{d}{=} Q_R\big(F_Y(T)\big)$,

(ii) $Q_X(p) = Q_R\Big(F_Y\big(Q_T(p)\big)\Big)$,

(iii) if $T \overset{d}{=} Y$, then $X \overset{d}{=} R$, and

(iv) if $Y \overset{d}{=} R$, then $X \overset{d}{=} T$.

If T follows the Lomax distribution, then the T-R{Y} family in (1.1) reduces to the Lomax-R{Y} family. In this paper, we study some general properties of the Lomax-R{Y} family. The rest of the paper is organized as follows. In Section 2, we define this family. In Section 3, we study some of its general properties. In Section 4, we consider a member of the family, namely the Lomax-Weibull{log-logistic} distribution, and obtain some of its structural properties. Parameter estimation and simulation study are discussed in Section 5. In Section 6, we prove empirically the usefulness of this distribution to censored and uncensored real-life data sets. Finally, Section 7 offers some concluding remarks.

2. The Lomax-R{Y} family

Let T be a Lomax random variable with PDF $f_T(x) = k\,(1+x)^{-k-1}$ and CDF $F_T(x) = 1-(1+x)^{-k}$, then the CDF of the *Lomax-R*$\{Y\}$ family is defined from Equation(1.1) by

$$F_X(x) = 1 - \Big[1 + Q_Y\Big(F_R(x)\Big)\Big]^{-k}, \tag{2.1}$$

then the PDF corresponding to (2.1) is given by

$$f_X(x) = \frac{k\, f_R(x)\, \Big[1 + Q_Y\Big(F_R(x)\Big)\Big]^{-k-1}}{f_Y\Big[Q_Y\Big(F_R(x)\Big)\Big]}. \tag{2.2}$$

Equation (2.1) has a very easy mathematical interpretation, since it is just the Lomax cdf evaluated at the transformed point $Q_Y(F_R(x))$ for any positive random variable Y, holding for any other random variable R. Note that the CDF (2.1) and the PDF (2.2) have closed-forms when Q_Y and F_R have closed-forms. Further, the Lomax-R{Y} family has some advantage in terms of its applicability as shown in the data examples presented later.

The HRF of the Lomax-R{Y} family reduces to

$$h_X(x) = h_R(x) \times \frac{k\Big[1 + Q_Y\Big(F_R(x)\Big)\Big]^{-k}}{h_Y\Big[Q_Y\Big(F_R(x)\Big)\Big]}. \tag{2.3}$$

The Lomax-R{Y} family in Equation(2.2) can generate several extended Lomax classes. Table 1 gives some subclasses of the Lomax-R{Y} family.

Table 1. Some Lomax-R{Y} classes based on different choices of the random variables R and Y.

S.No.	Y	$Q_Y(p)$	CDF of the Lomax-$R\{Y\}$
(a).	Log-logistic	$\alpha(\frac{p}{1-p})^{1/\beta}, \alpha, \beta > 0$	$1-\Big\{1+\alpha\Big[\frac{F_R(x)}{1-F_R(x)}\Big]^{1/\beta}\Big\}^{-k}$
(b).	Weibull	$\gamma\Big\{-\log(1-p)\Big\}^{1/c}, \quad \gamma, c > 0$	$1-\Big\{1+\gamma\Big[-\log(1-F_R(x))\Big]^{1/c}\Big\}^{-k}$
(c).	Exponentiated-exponential (EE)	$-\frac{1}{\theta}\log(1-p^{1/\alpha}), \quad \alpha, \theta > 0$	$1-\Big\{1-\frac{1}{\theta}\log\Big[1-(F_R(x))^{1/\alpha}\Big]\Big\}^{-k}$
(d)†.	Exponential	$-\log(1-p)$	$1-\Big\{1-\log[1-F_R(x)]\Big\}^{-k}$

Note: The Lomax class† in Table 1 has recently been studied by Cordeiro et al. (2014).

Theorem 2.1. *The PDF of the Lomax-R{Y} family can be expressed as*

$$f_X(x) = \sum_{i=0}^{\infty} b_{i+1}\, \pi_{R,i+1}(x), \tag{2.4}$$

where $\pi_{R,i+1}(x) = (i+1)\, F_R(x)^i\, f_R(x)$ *is the exponentiated-*F_R *(exp-*F_R *for short) density function with power parameter* $(i+1)$ *and the coefficients* b_{i+1}*'s depend on the parameters of the Lomax and the distribution of* Y.

Proof. Based on the generalized binomial expansion, we can write (2.1) as

$$\Big[1+Q_Y\Big(F_R(x)\Big)\Big]^{-k} = 1+\sum_{n=1}^{\infty}\frac{(-1)^n\,[k]_n}{n!}\,Q_Y\Big(F_R(x)\Big)^n,$$

where $[k]_n = k(k+1)\ldots(k+n-1)$ is the ascending factorial. Then,

$$F_X(x) = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}\,[k]_n}{n!}\,Q_Y\Big(F_R(x)\Big)^n. \tag{2.5}$$

If the QF, $Q_Y(u)$, does not have a closed-form expression, this function can usually be expressed as a power series of the form

$$Q_Y(u) = \sum_{i=0}^{\infty} a_i\, u^i, \tag{2.6}$$

where the coefficients a_i's are suitably chosen real numbers depending on the parameters of Y. For several important distributions such as the Weibull, log-logistic, exponentiated-exponential and exponential (listed in Table 1) and the normal, Student-t, gamma and beta distributions, among others, $Q_Y(u)$ can be expanded as in Equation (2.6).

By application of an equation in Section 0.314 of Gradshteyn and Ryzhik (2000) for a power series raised to a positive power, we can write from (2.6) for any n positive integer

$$Q_Y(u)^n = \left(\sum_{i=0}^{\infty} a_i\, u^i\right)^n = \sum_{i=0}^{\infty} c_{n,i}\, u^i, \tag{2.7}$$

where (for $n \geq 0$) $c_{n,0} = a_0^n$ and the coefficients $c_{n,i}'$s (for $i = 1,2,\ldots$) can be determined from the recurrence equation

$$c_{n,i} = (i a_0)^{-1}\sum_{m=1}^{i}[m(n+1)-i]\, a_m\, c_{n,i-m}.$$

The coefficient $c_{n,i}$ can be evaluated numerically in any algebraic or numerical software.

Combining equations (2.5) and (2.7), we obtain

$$F_X(x) = \sum_{i=0}^{\infty} b_i\, F_R(x)^i, \tag{2.8}$$

where (for $i \geq 0$) $b_i = b_i(k) = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}\,[k]_n}{n!}\, c_{n,i}$.

Now differentiating equation (2.8), we can write the PDF of T as

$$f_X(x) = \sum_{i=0}^{\infty} b_{i+1}\, \pi_{R,i+1}(x),$$

where $\pi_{R,i+1}(x)$ is defined above. □

Equation (2.4) reveals that the density function of the Lomax-R$\{Y\}$ family is a linear mixture of exp-F_R densities, where the coefficients are functions of the parameters of the Lomax and the distribution of Y. Thus, some mathematical properties of this class such as the ordinary and incomplete moments and generating function can be determined by knowing those of the exp-F_R distribution, which have been investigated by some authors for several baseline distributions.

3. Some properties

In this section, some general properties of the Lomax-R$\{$Y$\}$ family are investigated. The following Lemma gives the relationships between the random variables R and Y for some cases, which can be used to simulate the random variable X from the random variable T.

Lemma 3.1. *(Transformation): For Lomax random variable with PDF $f_T(x)$, the random variable:*
(i) $X = Q_R\Big(1+(\frac{T}{\alpha})^{-\beta}\Big)^{-1}$ follows the Lomax-R$\{$log-logistic$\}$ class in Table 1 (a),
(ii) $X = Q_R\Big(1-\mathrm{e}^{-(T/\gamma)^c}\Big)$ follows the Lomax-R$\{$Weibull$\}$ class in Table 1 (b),
(iii) $X = Q_R\Big(1-\mathrm{e}^{-\theta T}\Big)^{\alpha}$ follows the Lomax-R$\{$EE$\}$ class in Table 1 (c).

Remark 3.1. The following results are obtained from equations (2.1) and (2.2):

(a) The QFs for the (i) Lomax-R$\{$log-logistic$\}$, (ii) Lomax-R$\{$Weibull$\}$ and (ii) Lomax-R$\{$EE$\}$ classes are given by:
(i) $Q_X(p) = Q_R\Big(1+(\frac{Q_T(p)}{\alpha})^{-\beta}\Big)^{-1}$,
(ii) $Q_X(p) = Q_R\Big(1-\mathrm{e}^{-(Q_T(p)/\gamma)^c}\Big)$,
(iii) $Q_X(p) = Q_R\Big(1-\mathrm{e}^{-\theta Q_T(p)}\Big)^{\alpha}$,
respectively, where $Q_T(p) = (1-p)^{-1/k}-1$.

(b) The modes of the Lomax-R$\{$Y$\}$ classes are the solutions of the equation

$$x = \frac{f_R'(x)}{f_R(x)} - Q_Y'\Big(F_R(x)\Big)\left\{\frac{k+1}{1+Q_Y\Big(F_R(x)\Big)} + \frac{f'\Big(Q_Y\big(F_R(x)\big)\Big)}{f\Big(Q_Y\big(F_R(x)\big)\Big)}\right\} f_R(x).$$

(c)

$$\mathbb{E}(X^r) = \mathbb{E}_T\left\{\Big[Q_R\Big(F_Y(T)\Big)\Big]^r\right\}.$$

The Shannon's entropy of the Lomax-R$\{$Y$\}$ family is given in the following theorem.

Theorem 3.1. *The Shannon's entropy of the Lomax-R$\{Y\}$ family can be expressed as*

$$\eta_X = \log\left(\frac{1}{k}\right) + \eta_R + (k+1)\mathbb{E}\Big[\log\Big\{1+Q_Y\Big(F_R(x)\Big)\Big\}\Big] + \mathbb{E}\Big[\log\Big\{f_Y\Big(Q_Y\big(F_R(x)\big)\Big)\Big\}\Big]. \quad (3.1)$$

Corollary 3.1. *The Shannon's entropies for the (i) Lomax-R{log-logistic}, (ii) Lomax-R{Weibull}, and (iii) Lomax-R{EE} classes are, respectively, given by*

$$(i)\ \eta_X = \log\left(\frac{\beta}{k}\right) + \eta_R - \left(\frac{1-\beta}{\beta}\right)\mathbb{E}\Big[\log F_R(x)\Big] - \left(\frac{1+\beta}{\beta}\right)\mathbb{E}\Big[-\log\Big(1-F_R(x)\Big)\Big]$$
$$+(k+1)\,\mathbb{E}\Big\{\log\Big[1+\Big(\frac{F_R(x)}{1-F_R(x)}\Big)^{\frac{1}{\beta}}\Big]\Big\},$$
$$(ii)\ \eta_X = \log\left(\frac{c}{\gamma^c k}\right) + \eta_R + (k+1)\,\mathbb{E}\Big\{\log\Big[1+\gamma\Big(-\log\big(1-F_R(x)\big)\Big)^{\frac{1}{c}}\Big]\Big\}$$
$$+(c-1)\,\mathbb{E}\Big[\log\Big\{\gamma\Big(-\log\big[1-F_R(x)\big]\Big)^{\frac{1}{c}}\Big\}\Big] - \mathbb{E}\Big[-\log\big(1-F_R(x)\big)\Big],$$
$$(iii)\ \eta_X = \log\left(\frac{\theta\alpha}{k}\right) + \eta_R + +(k+1)\,\mathbb{E}\Big\{\log\Big[1+\log\Big(1-[F_R(x)]^{\frac{1}{\alpha}}\Big)^{-\frac{1}{\theta}}\Big]\Big\}$$
$$+\mathbb{E}\Big\{\log\Big[1-\big(F_R(x)\big)^{\frac{1}{\alpha}}\Big]\Big\} + \left(\frac{\alpha-1}{\alpha}\right)\mathbb{E}\Big[\log\big(F_R(x)\big)\Big].$$

4. The Lomax-Weibull{Log-logistic} distribution

Based on Table 1 (a), the PDF of the Lomax-R{log-logistic} class is

$$f_X(x) = \frac{k\alpha}{\beta}\,\frac{f_R(x)}{[1-F_R(x)]^2}\Big[\frac{F_R(x)}{1-F_R(x)}\Big]^{\frac{1}{\beta}-1}\left\{1+\alpha\Big[\frac{F_R(x)}{1-F_R(x)}\Big]^{\frac{1}{\beta}}\right\}^{-k-1}. \tag{4.1}$$

Remark 4.1. From (4.1), we have:

(i) When $x \to -\infty$, $f_X(x) \sim \frac{k\alpha}{\beta}\, f_R(x)\, [F_R(x)]^{\frac{1}{\beta}-1}$,

(ii) When $x \to \infty$, $f_X(x) \sim \frac{k}{\alpha^k\beta}\, f_R(x)\, [1-F_R(x)]^{\frac{k}{\beta}-1}$.

Next, we provide some properties of a special case of (4.1), the Lomax-Weibull{Log-logistic} (for short, LW{LL}) distribution. We eliminate the redundancy of the scale and shape parameters by setting $\alpha = 1$.

If a random variable R has the Weibull distribution, then, the PDF and CDF of the LW{LL} distribution are, respectively, given by

$$f(x) = \frac{k\,c}{\beta\,\gamma}\left(\frac{x}{\gamma}\right)^{c-1} \mathrm{e}^{(x/\gamma)^c}\Big[\mathrm{e}^{(x/\gamma)^c}-1\Big]^{\frac{1}{\beta}-1}\left\{1+\Big[\mathrm{e}^{(x/\gamma)^c}-1\Big]^{\frac{1}{\beta}}\right\}^{-k-1} \tag{4.2}$$

and

$$F(x) = 1-\left\{1+\Big[\mathrm{e}^{(x/\gamma)^c}-1\Big]^{\frac{1}{\beta}}\right\}^{-k}. \tag{4.3}$$

Equation (4.3) is a generalization of the Weibull distribution. Clearly, this CDF reduces to the Weibull CDF when $\beta = k = 1$.

Remark 4.2. For the LW{LL} distribution, we have:

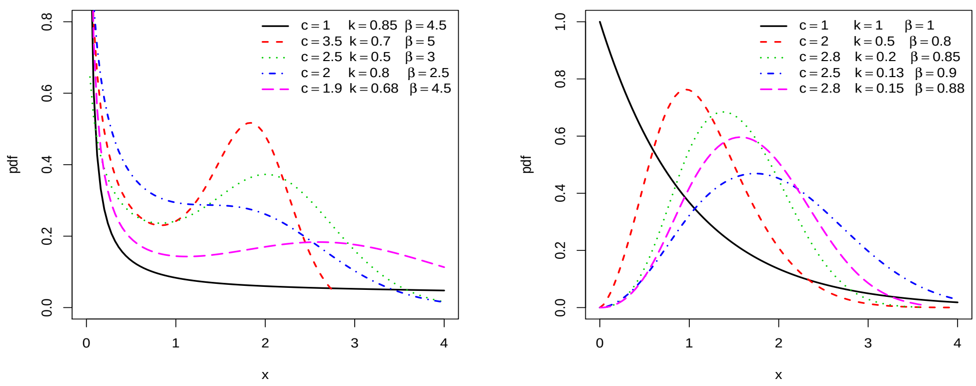
(i) When $x \to 0$, $f_X(x) \sim \frac{kc}{\beta\gamma}\left(\frac{x}{\gamma}\right)^{c-1}\left[1-\mathrm{e}^{-\left(\frac{x}{\gamma}\right)^c}\right]^{\frac{1}{\beta}-1}$,

(ii) When $x \to \infty$, $f_X(x) \sim \frac{kc}{\beta\gamma}\left(\frac{x}{\gamma}\right)^{c-1}\mathrm{e}^{-\frac{k}{\beta}\left(\frac{x}{\gamma}\right)^c}$.

Henceforth, a random variable having the PDF (4.2) is denoted by $X \sim \text{LW}\{\text{LL}\}(k,\beta,c,\gamma)$. From (4.2) and (4.3), the HRF of X, say $h(x)$, is given by

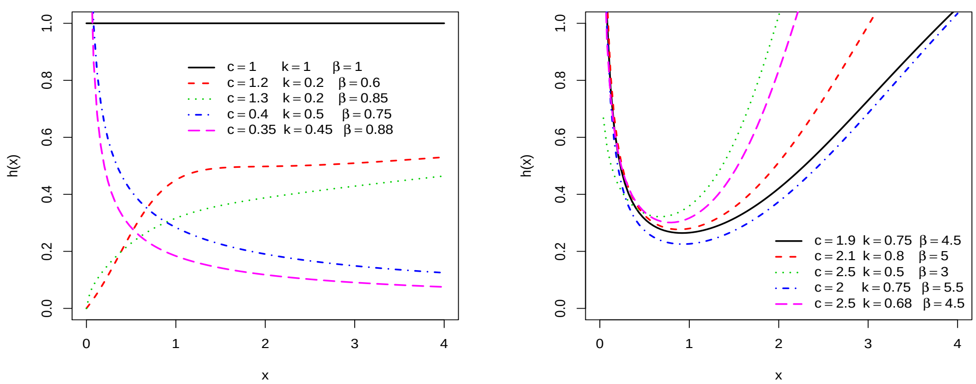
$$h(x) = \frac{kc}{\beta\gamma}\left(\frac{x}{\gamma}\right)^{c-1} \mathrm{e}^{(x/\gamma)^c}\left[\mathrm{e}^{(x/\gamma)^c}-1\right]^{\frac{1}{\beta}-1}\left\{1+\left[\mathrm{e}^{(x/\gamma)^c}-1\right]^{\frac{1}{\beta}}\right\}^{-k-1}.$$

In Figures 1 and 2 some plots of the PDF and HRF of the LW$\{$LL$\}$ model are displayed for some parameter values. Figure 1 reveals that the LW$\{$LL$\}$ density has various shapes such as bimodal, approximately symmetric, right-skewed, left-skewed and reversed-J. Also, Figure 2 shows that the LW$\{$LL$\}$ HRF can have constant, increasing, decreasing, upside-down bathtub and bathtub shapes.



(a) (b)

Fig. 1. Plots of the LW$\{$LL$\}$ densities for selected parameters.



(a) (b)

Fig. 2. Plots of the LW$\{$LL$\}$ hazard rates for selected parameters.

Remark 4.3. From Lemma 1 (i), we have:
(a) If a random variable T follows the Lomax distribution with parameter k, then using Lemma 1 (i)

$$X = \gamma\left[\log\left(1+T^{\beta}\right)\right]^{\frac{1}{c}} \tag{4.4}$$

has the LW{LL} distribution with parameters k, β, c and γ.(b) The QF, say $Q(u)$, of the LW{LL} distribution is given by

$$Q(u) = \gamma\left\{\log\left[1+\left((1-u)^{-1/k}-1\right)^{\beta}\right]\right\}^{1/c}, \quad 0<u<1. \tag{4.5}$$

(c) By using the result $\eta_R = 1-\log(\gamma)+\xi(1-\gamma^{-1})$, the Shannon entropy of the LW{LL} distribution is given by

$$\eta_X = 1+\log\left(\frac{\beta\gamma}{kc}\right)+\xi\,\Gamma\left(1-\frac{1}{c}\right)-\left(\frac{1}{\beta}+1\right)\gamma^{1-c}+(k+1)\sum_{j=0}^{\infty}\frac{(-1)^j}{(1-j)}\left[\frac{\iota\,\pi}{\beta}+\xi\right],$$

where $\xi \approx 0.577216$ is the Euler constant.

Lemma 4.1. *The rth moment of the LW{LL} distribution is given by*

$$\mu_r' = \mathbb{E}(X^r) = \frac{kr}{c}\gamma^r\sum_{\ell=0}^{\infty}\sum_{j=0}^{\ell}\frac{(-1)^{\ell+j}}{(r/c-\ell)}\binom{\ell}{j}\binom{\ell-r/c}{\ell}p_{j,\ell}\,\frac{\Gamma\Big(\beta(\ell+r/c)+1\Big)\Gamma\Big(k-\beta(\ell+r/c)\Big)}{\Gamma(k+1)}, \tag{4.6}$$

where $\beta(\ell+r/c)$ is any non-integer real and the constants $p_{j,\ell}$ can be determined recursively (for $\ell \geq 1$) by (with $p_{j,0}=1$)

$$p_{j,\ell} = \ell^{-1}\sum_{m=1}^{\ell}\left[m(j+1)-\ell\right]p_{j,\ell-m}.$$

Proof. Using (4.4), the rth moment of X is given by

$$\mathbb{E}(X^r) = k\gamma^r\int_0^{\infty}(1+x)^{-k-1}\left[\log\left(1+x^{\beta}\right)\right]^{r/c}dx. \tag{4.7}$$

By using the power series http://functions.wolfram.com/ElementaryFunctions/Log/06/01/04/03/, we obtain

$$\left[\log\left(1+x^{\beta}\right)\right]^{r/c} = r/c\sum_{\ell=0}^{\infty}\binom{\ell-r/c}{\ell}\sum_{j=0}^{\ell}\frac{(-1)^{\ell+j}\binom{\ell}{j}}{(r/c-\ell)}p_{j,\ell}\,x^{\beta\left(\ell+r/c\right)}. \tag{4.8}$$

Equation (4.6) follows by substituting (4.8) in (4.7) and noting that

$$\int_0^{\infty}x^{\beta\left(\ell+r/c\right)}(1+x)^{-k-1}dx = \frac{\Gamma\Big(\beta(\ell+r/c)+1\Big)\Gamma\Big(k-\beta(\ell+r/c)\Big)}{\Gamma(k+1)},$$

where $\beta(\ell+r/c)$ is any non-integer real. □

Further, the central moments (μ_r) and cumulants (κ_r) of X are obtained from the ordinary moments by

$$\mu_r = \sum_{k=0}^{r} (-1)^k \binom{r}{k} \mu_1'^k \mu_{r-k}' \qquad \text{and} \qquad \kappa_r = \mu_r' - \sum_{k=1}^{r-1} \binom{r-1}{k-1} \kappa_k \mu_{r-k}',$$

respectively, where $\kappa_1 = \mu_1'$. The skewness and kurtosis of X can be determined from the ordinary moments using well-known relationships.

Lemma 4.2. *Let X be a random variable having the Lomax-R$\{$log-logistic$\}$ and R be non-negative. If $\mathbb{E}(X^r) < \infty$ and $k > \beta$, then $\mathbb{E}(X^r) \leq \mathbb{E}(R^r)\left[1 + k\,\alpha^{-\beta} B(k-\beta, \beta+1)\right]$.*

Proof. If the random variable R is non-negative and X follows the T-R{Y} class in (1.2) with $\mathbb{E}(X^r) < \infty$, one can check that $\mathbb{E}(X^r) \leq \mathbb{E}(R^r)\,\mathbb{E}[1/(1-F_Y(T))]$ (see Theorem 2.1, Aljarrah et al., 2014). Further, if Y follows the log-logistic distribution with parameters α and β, then $\mathbb{E}[1/(1-F_Y(T))] = 1 + \alpha^{-\beta}\,\mathbb{E}(T^\beta)$, where $T \sim \text{Lomax}(k)$. The result follows by noting that $\mathbb{E}(T^\beta) = k\,B(k-\beta, \beta+1)\,,\, k > \beta$. □

Theorem 4.1. *If X has the LW$\{$LL$\}$ distribution, $\mathbb{E}(X^r) < \infty$ and if $k > \beta$, then*

$$\mathbb{E}(X^r) \leq \gamma^r\,\Gamma\Big(1 + r/c\Big)\left[1 + k\,\alpha^{-\beta} B(k-\beta, \beta+1)\right].$$

Proof. The result follows from Lemma 4.2. □

5. Estimation and Simulations

Several approaches for parameter estimation were proposed in the literature but the maximum likelihood method is the most commonly employed. The maximum likelihood estimators (MLEs) enjoy desirable properties and can be used when constructing confidence intervals for the model parameters and also in test statistics. The normal approximation for these estimators in large sample distribution theory is easily handled either analytically or numerically. So, we consider the estimation of the unknown parameters for the LW$\{$LL$\}$ distribution from complete samples only by maximum likelihood. Let $x_1, \ldots, x_n$ be a sample of size n from this distribution given by (4.2). We consider the estimation of the unknown parameters by the maximum likelihood method. The log-likelihood function for the vector of parameters $\theta = (k, \beta, c, \gamma)^\top$ can be expressed as

$$\begin{aligned}\ell = n\log\left(\frac{kc}{\beta\gamma}\right) + \left(\frac{c-1}{\gamma}\right)\sum_{i=1}^{n}\log(x_i) + \sum_{i=1}^{n}\left(\frac{x_i}{\gamma}\right)^c + \left(\frac{1-\beta}{\beta}\right)\sum_{i=1}^{n}\log\left[\mathrm{e}^{\left(\frac{x_i}{\gamma}\right)^c} - 1\right] \\ -\,(k+1)\sum_{i=1}^{n}\log\left\{1 + \left[\mathrm{e}^{\left(\frac{x_i}{\gamma}\right)^c} - 1\right]^{\frac{1}{\beta}}\right\}. \qquad (5.1)\end{aligned}$$

Equation (5.1) can be maximized directly by using the `R` (`optim` function), `SAS` (`PROC NLMIXED`) or `Ox` program (sub-routine `MaxBFGS`).

5.1. *Simulation study*

In this section, we evaluate the performance of the MLEs of the model parameters for the LW{LL} distribution using Monte Carlo simulation varying the sample size and for selected parameter values. The simulation study is repeated 5,000 times each with sample sizes $n = 25, 50, 100, 200, 400, 600$ and parameter values: I: $c = 0.8$, $k = 0.5$, $\beta = 2$, $\gamma = 1$ and II: $c = 1.5$, $k = 1.1$, $\beta = 0.7$, $\gamma = 1$. The MLEs are determined by maximizing the log-likelihood function in Equation (5.1) using the `optim` routine in the R software. Table 2 provides the average bias (Bias), mean square error (MSE), coverage probability (CP), average lower bound (LB) and average upper bound (UB) values for the parameters c, k, β and γ under different sample sizes. From the results of the simulations, we can verify that the biases and MSEs decrease in general when the sample size n increases. The CP of the confidence intervals are quite close to the 95% nominal level. Therefore, the MLEs and their asymptotic results can be used for estimating and constructing confidence intervals even for reasonably small sample sizes.

Table 2. Monte Carlo simulation results: Bias, MSE, CP, LB and UB.

Parameter	n	Bias	MSE	CP	LB	UB
			I			
c	25	0.210	0.526	0.90	0.830	2.364
	50	0.085	0.130	0.91	0.651	1.763
	100	0.048	0.073	0.92	0.530	1.463
	200	0.026	0.035	0.93	0.508	1.262
	400	0.009	0.021	0.95	0.531	1.127
	600	−0.002	0.014	0.96	0.555	1.056
k	25	0.106	1.316	0.88	1.862	3.068
	50	0.037	0.389	0.90	0.988	2.061
	100	0.009	0.208	0.91	0.424	1.436
	200	−0.012	0.079	0.93	0.159	1.094
	400	0.003	0.044	0.95	0.117	0.939
	600	0.011	0.028	0.95	0.168	0.862
β	25	0.708	3.786	0.92	3.471	8.815
	50	0.554	2.122	0.96	2.652	7.164
	100	0.327	1.070	0.96	1.680	5.047
	200	0.186	0.474	0.97	1.432	3.871
	400	0.091	0.194	0.96	1.390	3.103
	600	0.067	0.112	0.95	1.448	2.786
γ	25	0.192	2.521	0.84	4.374	6.679
	50	0.240	1.632	0.89	5.539	7.946
	100	0.112	0.895	0.93	1.927	4.029
	200	0.060	0.541	0.93	1.096	2.957
	400	0.052	0.242	0.95	0.596	2.154
	600	0.042	0.123	0.96	0.490	1.783

Table 2 (Continued).

Parameter	n	Bias	MSE	CP	LB	UB
			II			
c	25	0.405	1.334	0.90	1.196	4.175
	50	0.332	0.810	0.92	1.067	3.703
	100	0.311	0.549	0.93	0.903	3.339
	200	0.250	0.323	0.95	0.877	2.830
	400	0.168	0.167	0.96	0.969	2.401
	600	0.117	0.110	0.96	1.062	2.178
k	25	−0.083	1.311	0.82	2.938	4.943
	50	−0.001	1.052	0.84	2.209	4.341
	100	0.184	1.186	0.89	1.440	3.875
	200	0.168	0.788	0.91	0.755	3.028
	400	0.167	0.522	0.95	0.484	2.611
	600	0.122	0.363	0.97	0.385	2.302
β	25	0.133	0.657	0.83	0.755	2.317
	50	0.006	0.238	0.86	0.649	1.737
	100	−0.039	0.126	0.90	0.458	1.315
	200	−0.071	0.035	0.92	0.358	0.967
	400	−0.051	0.015	0.94	0.419	0.887
	600	−0.035	0.010	0.96	0.474	0.856
γ	25	0.016	2.773	0.84	4.291	6.172
	50	0.011	1.700	0.85	2.914	4.746
	100	0.121	1.090	0.90	1.654	3.618
	200	0.073	0.452	0.91	0.636	2.369
	400	0.084	0.260	0.94	0.428	2.030
	600	0.061	0.181	0.96	0.400	1.845

6. Applications

In this section, five applications of the LW$\{$LL$\}$ model are presented to illustrate its flexibility to fit data sets having various shapes. In the applications, the model parameters are estimated by the method of maximum likelihood. The Akaike information criterion (AIC), Bayesian information criterion (BIC) and Kolmogrove-Smirnov (K-S) statistic are calculated to compare the LW$\{$LL$\}$ model with other models.

6.1. *Uncensored data sets*

6.1.1. *Data set 1: Cancer data.*

The first data set represents the remission times (in months) of 128 bladder cancer patients studied by Lee and Wang (2003). These data were used by Zea et al. (2012) and Lemonte and Cordeiro (2013) for the beta exponentiated-Pareto and extended Lomax models, respectively. The MLEs, AICs and BICs for the fitted McDonald-Lomax (McL), beta-Lomax (BL) and Kumaraswamy-Lomax (KwL) are taken from Lemonte and Cordeiro (2013) while the KS values reported in Table 3 are obtained using R software. The estimates and goodness-of-fit statistics of the exponentiated-Weibull (EW) (Mudholkar and Srivastava, 1993) and Weibull models are also reported in Table 3. The figures in Table 3 indicate that the McL, BL, KwL and EW models provide adequate fits, but the LW$\{$LL$\}$ model provides the best fit with lowest AIC and K-S values. The distribution of these data is highly skewed to the right (skewness = 3.29). This application suggests that the LW$\{$LL$\}$ model has the ability to fit right-skewed data sets. For a visual comparison, we provide the PP-plots of the fitted models in Figure 3. Clearly, the LW$\{$LL$\}$ model fits the data more closely.

6.1.2. *Data set 2: Carbone Fibre Data.*

The second data set has recently been used by Cordeiro and Lemonte (2011) to illustrate the applicability of the beta-Birnbaum-Saunders (BBS) distribution. The MLEs, AICs and BICs for the fitted BBS and Birnbaum-Saunders (BS) distributions are taken from Cordeiro and Lemonte (2011), and the KS values in Table 4 are calculated by using the R software. We also provide estimates and goodness-of-fit statistics of the EW and Weibull models in Table 4. The results in Table 4 indicate that the LW$\{$LL$\}$ model provides the best fit with the lowest AIC, BIC and K-S values. The distribution of these data is slightly skewed to the left (skewness=-0.13). This application reveals that the LW$\{$LL$\}$ distribution has the ability to fit left-skewed data sets. Further, the PP-plots in Figure 4 also support the results in Table 4.

6.1.3. *Data set 3: Aarset Data.*

The third data set is taken from Aarset (1987) which represents the lifetimes of 50 devices. Recently, Silva et al. (2010) fitted the beta modified-Weibull (BMW) distribution to these data to illustrate its potentiality. The MLEs, AICs and BICs values for the fitted BMW, modified-Weibull (MW), beta-Weibull (BW) (Lee et al., 2007) and EW models are taken from Silva et al. (2010) and the other results in Table 5 are obtained by using the R software. The figures in Table 5 indicate that LW$\{$LL$\}$ model provides the best fit as compared to those of the BMW, MW, BW, EW and Weibull distributions. This application suggests that the LW$\{$LL$\}$ model can be used to fit bathtub density shaped data sets. Furthermore, the PP-plots in Figure 5 also support the results in Table 5.

Table 3. MLEs (standard errors in parentheses) and the AIC, BIC and K-S statistics for data set 1.

Distribution	Estimates					AIC	BIC	K-S
LW$\{$LL$\}$$(k, \beta, c, \gamma)$	0.8811 (0.746)	0.4528 (0.294)	0.5703 (0.426)	10.7816 (12.2413)		826.93	838.336	0.0315
McL$(\alpha, \beta, a, \eta, c)$	0.8085 (3.364)	11.2929 (15.818)	1.5060 (0.243)	4.1886 (25.029)	2.1046 (3.079)	829.82	844.09	0.0391
BL(α, β, a, η)	3.9191 (18.192)	23.9281 (27.338)	1.5853 (0.280)	0.1572 (5.024)		828.14	839.29	0.0406
KwL(α, β, a, η)	0.3911 (2.386)	12.2973 (17.316)	1.5162 (0.228)	11.0323 (87.144)		827.88	839.29	0.0389
EW(a, α, γ)	0.6544 (0.1346)	2.7957 (1.2626)	3.3461 (1.8890)			827.36	835.91	0.0450
W(α, γ)	1.0478 (0.0675)	9.5606 (0.8529)				832.17	837.87	0.0700

Table 4. MLEs (standard errors in parentheses) and the AIC, BIC and K-S statistics for data set 2.

Distribution	Estimates				AIC	BIC	K-S
LW{LL}(k, β, c, γ)	0.2163 (0.0506)	1.3772 (0.2872)	3.8199 (0.0026)	1.9127 (0.2413)	178.273	186.99	0.0806
BBS(α, β, a, b)	1.045 (0.004)	57.600 (0.331)	0.1923 (0.026)	1876 (605.0)	190.71	200.40	0.1380
EW(a, α, γ)	3.9097 (1.069)	0.8009 (0.353)	3.2304 (0.345)		179.88	186.45	0.0849
BS(α, β)	0.3911 (2.386)	12.29733 (17.316)			204.38	208.75	0.184
W(α, γ)	3.4411 (0.3309)	3.0622 (0.1149)			186.13	190.51	0.082

Table 5. MLEs (standard errors in parentheses) and the AIC, BIC and K-S statistics for data set 3.

Distribution	Estimates					AIC	BIC	K-S
LW{LL}(k, β, c, γ)	0.9288 (0.1448)	8.9661 (1.1608)	5.8394 (0.0025)	50.3961 (0.0196)		440.56	448.21	0.1112
BMW(a, b, α, λ, γ)	0.1975 (0.0462)	0.1647 (0.0830)	0.0002 (0.0001)	0.0541 (0.0157)	1.3771 (0.3387)	451.60	461.20	0.1304
BW(a, b, α, γ)	0.1835 (0.0509)	0.0748 (0.0353)	0.0007 (0.0004)	2.3615 (0.1715)		463.90	471.90	0.1657
MW(α, λ, γ)	0.0624 (0.0267)	0.0233 (0.0048)	0.3548 (0.1127)			460.30	466.00	0.1337
EW(a, α, γ)	0.0011 (0.0010)	0.4668 (0.0889)	1.5936 (0.1858)			480.50	486.2	0.2237
W(α, γ)	0.9488 (0.1195)	44.8551 (6.9333)				486.00	489.82	0.1932

6.2. *Censored data sets*

In this section, we provide applications of the LW{LL} model for two censored data sets. The LW{LL} survival function has closed-form expression and therefore can be used effectively for lifetime data in presence of censoring. We adopt the AIC and BIC statistics to compare the fits of the LW{LL} distribution with other models such as the beta-Weibull (BW), beta-Lomax (BL) (Lemonte and Cordeiro, 2013), Weibull-Lomax (WL) (Tahir et al., 2015), EW and Weibull.

Consider a data set $D = (x; r)$, where $x = (x_1 \dots x_n)^T$ are the observed failure times and $r = (r_1, \dots r_n)^T$ are the censored failure times. The indicator r_i is equal to 1 if a failure is observed and 0 otherwise. Suppose that the data are independently and identically distributed and come from a distribution with PDF given by Equation (4.2). Let $\theta = (k, \beta, c, \gamma)^T$ denotes a vector of parameters. The log-likelihood of θ can be written as

$$L(D;\theta) \propto \prod_{i=1}^{n} [f(x_i;\theta)]^{r_i} [S(x_i;\theta)]^{1-r_i}$$

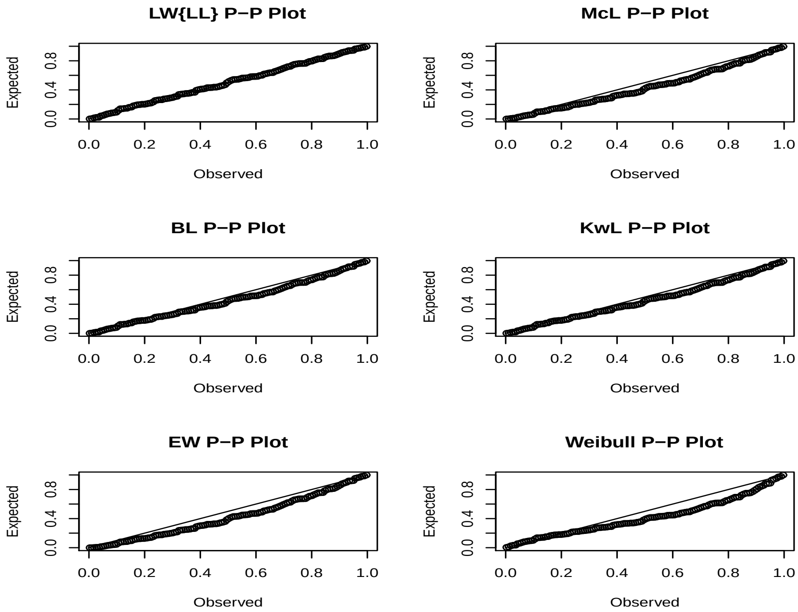


Fig. 3. PP-plots for data set 1.

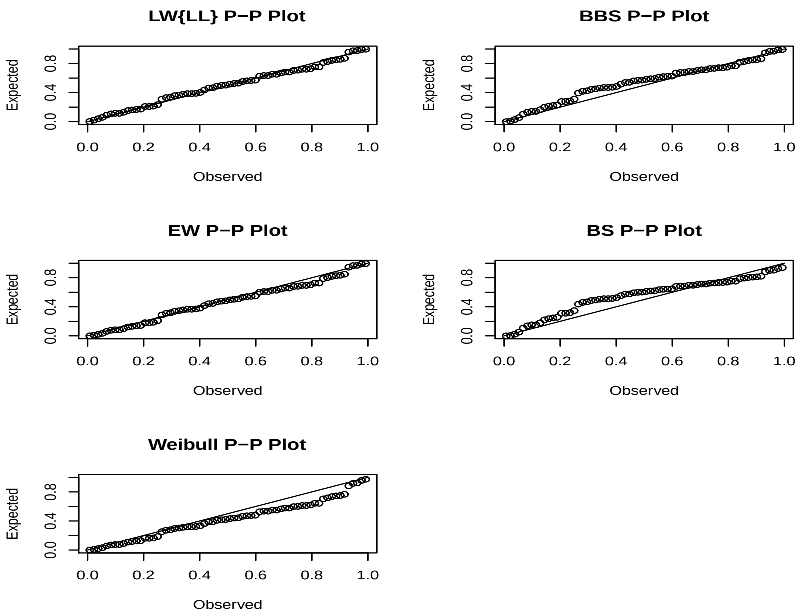


Fig. 4. PP-plots for data set 2.

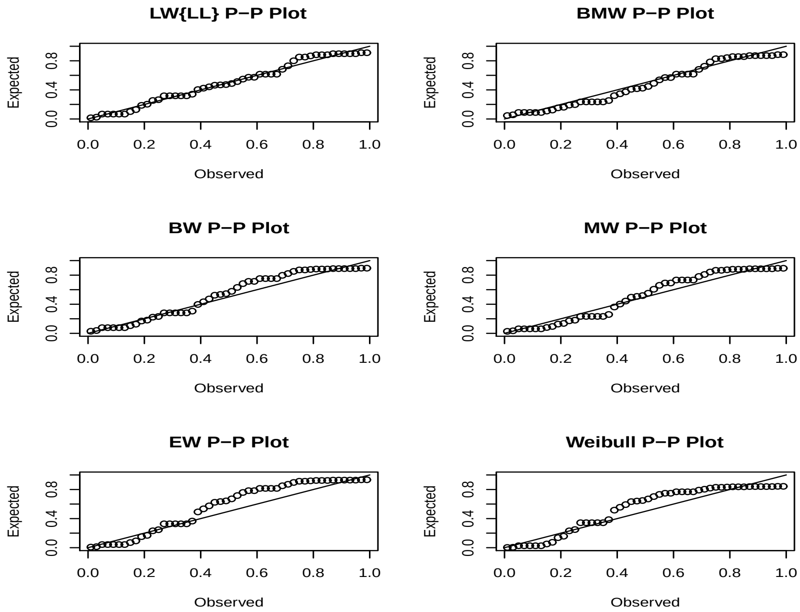


Fig. 5. PP-plots for data set 3.

and then the log-likelihood reduces to

$$\ell = r_i \sum_{i=1}^{n} \left[\log\left(\frac{k\,c}{\beta\,\gamma}\right) + (c-1)\log\left(\frac{x_i}{\gamma}\right) + \left(\frac{x_i}{\gamma}\right)^c + \left(\frac{1}{\beta}-1\right)\log\left[\mathrm{e}^{\left(\frac{x_i}{\gamma}\right)^c} - 1\right] \right.$$
$$\left. -(k+1)\log\left\{1+\left[\mathrm{e}^{\left(\frac{x_i}{\gamma}\right)^c}-1\right]^{\frac{1}{\beta}}\right\}\right]$$
$$+k\,(1-r_i)\sum_{i=1}^{n}\left[\log\left\{1+\left[\mathrm{e}^{\left(\frac{x_i}{\gamma}\right)^c}-1\right]^{\frac{1}{\beta}}\right\}\right]. \tag{6.1}$$

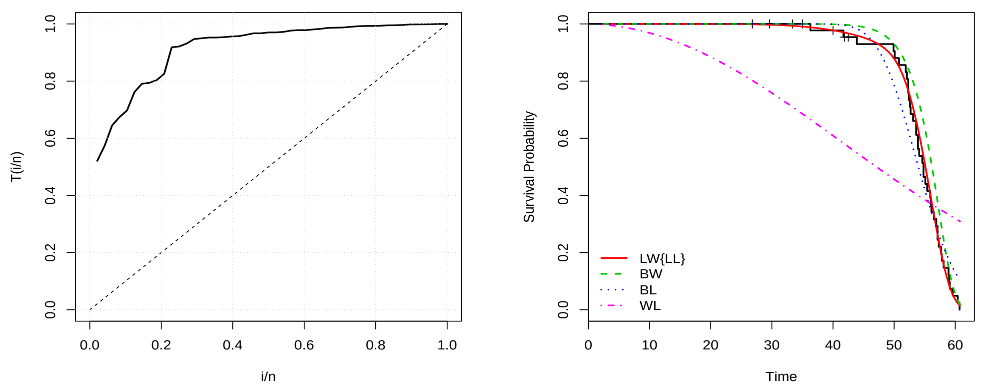
The log-likelihood can be maximized numerically to obtain the MLEs. There are various routines available for numerical maximization of ℓ. We use the routine optim in the R software.

6.2.1. *Data set 4: Cord failure data.*

These data represent strengths in coded units of 48 pieces of weathered braided cord. The data set has 14.5% of censored observations (7 in total). The detailed description of the data is given in Crowder et al. (1991). The TTT plot (due to Aarset, 1987) for these data, given in Figure 6(a), has a concave shape which suggests an increasing hazard shape. Therefore, the LW{LL} distribution is an appropriate model for fitting these data. The Kaplan-Meier and the survival curves of the fitted models are displayed in Figure 6(b). The MLEs are listed in Table 6.

Table 6. MLEs (standard errors in parentheses) and the AIC and BIC statistics for data set 4.

Distribution	Estimates				AIC	BIC
LW{LL}(k, β, c, γ)	0.0791 (0.0519)	0.3912 (0.1990)	18.6020 (2.3954)	46.6390 (2.0771)	234.6417	242.6390
BW(a, b, α, γ)	0.6450 (0.3094)	1.2787 (5.8190)	27.3528 (14.4741)	58.5659 (11.7415)	237.8018	245.2866
BL(a, b, α, γ)	3.1113 (4.7821)	190.8120 (5.9527)	147.5627 (5.5005)	126.1514 (2.5780)	258.4820	265.9668
WL$(\alpha, \lambda, \gamma)$	0.0039 (0.0005)	1.6460 (0.2689)	1.3141 (0.1992)	4.5885 (2.5780)	364.7034	372.1882
EW(a, α, γ)	0.4549 (0.2798)	27.1854 (11.9894)	57.9520 (1.3058)		237.8087	243.4223
W(α, λ)	16.3088 (2.0411)	56.0281 (0.5598)			239.5294	243.2718



(a) (b)

Fig. 6. (a) TTT plot (b) Kaplan-Meier and estimated survival curves for data set 4.

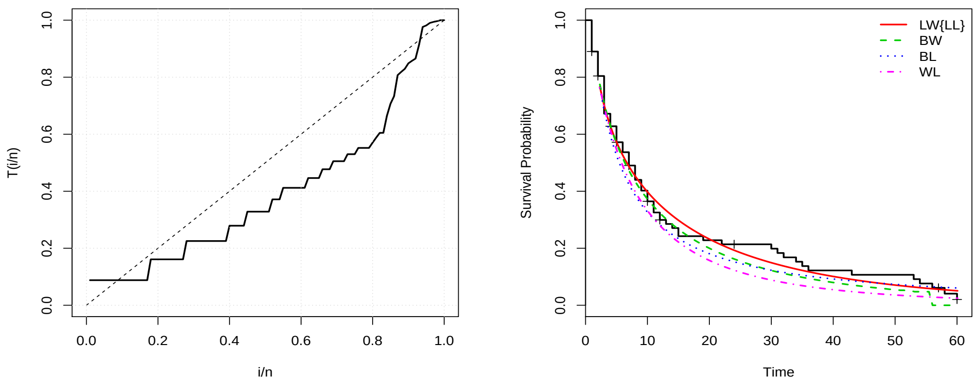
The results in Table 6 reveal that the LW{LL} distributions provides a better fit than those of the other models. Also, it is evident from Figure 6(b) that the LW{LL} model captures the pattern of the Kaplan-Meier curve better than the other models.

6.2.2. *Data set 5. HIV data.*

These data represent the survival times of HIV+ individuals using a follow up time. Subjects were enrolled in the study from January 1, 1989 to December 31, 1995. The data set consists of 100 observations with 20% of censored elements. More details about the data can be found in Hosmer and Lemeshow (1989). The TTT plot for the data is given in Figure 7(a), which suggests an upside down bathtub shape. Therefore, the LW{LL} distribution could be an appropriate model for the data. The Kaplan-Meier and survival curves for the fitted models are displayed in Figure 7(b). The MLEs are listed in Table 7. The smallest values of AIC and BIC statistics for the LW{LL} model suggest that it provides the best fit. Also, from Figure 7(b), it is evident that LW{LL} model captures the pattern of the Kaplan-Meier curve better than the other models.

Table 7. MLEs (standard errors in parentheses) and the AIC and BIC statistics for data set 5.

Distribution	Estimates				AIC	BIC
LW{LL}(k, β, c, γ)	0.0072 (0.0030)	39.8580 (1.5715)	0.6475 (0.0876)	1.6208 (0.1482)	549.9042	560.3248
BW(a, b, α, γ)	19.7146 (2.1398)	0.0917 (0.0111)	0.5494 (0.0121)	0.0859 (0.0111)	567.2130	577.6337
BL(a, b, α, γ)	10.8779 (0.2505)	4.5514 (0.1543)	5.1465 (1.6517)	0.095 (0.0114)	569.4952	579.9159
WL(α, λ, γ)	8.8564 (1.2966)	2.1932 (0.7224)	0.1002 (0.0579)	0.5165 (0.6813)	572.7972	583.2179
EW(a, α, γ)	4.7343 (1.5040)	0.4178 (0.0586)	1.3779 (0.7973)		573.1632	580.9787
W(α, λ)	0.8321 (0.0687)	13.2791 (1.8147)			583.0585	588.2688



(a) (b)

Fig. 7. (a) TTT Plot (b) Kaplan-Meier and estimated survival curves for data set 5.

7. Concluding remarks

The literature, several ways of extending well-known distributions are been proposed. Consequently, a significant progress has been made toward the generalization of existing distributions. In this context, we define new family, the *Lomax-R*$\{Y\}$ family of distributions. We obtain explicit expressions for the quantile functions, Shannon entropy and ordinary and incomplete moments. We consider a special model, namely the Lomax-Weibull{log-logistic} distribution, which is a generalization of Weibull distribution. Some applications show that the Lomax-Weibull{log-logistic} distribution can be used effectively in modeling censored and uncensored data sets with various shapes.

Acknowledgements

The authors wish to thank the Editor and the two anonymous referees for comments which greatly improved the earlier version of the paper.

References

[1] Aarset, M.V. (1987). How to identify bathtub hazard rate. *IEEE Transaction on Reliability*, **36**, 106–108.

[2] Aljarrah, M.A., Lee, C. and Famoye, F. (2014). A method of generating T-X family of distributions using quantile functions. *Journal of Statistical Distributions and Applications*, **1**, Art. 2.

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

[4] Alzaatreh, A., Lee, C. and Famoye, F. (2014). T-normal family of distribution: A new approach to generalize the normal distribution. *Journal of Statistical Distributions and Applications* **1**, Art. 16.

[5] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandavian Journal of Statistics*, **12**, 171–178.

[6] Cordeiro, G.M.and Lemonte, A.J. (2011). The beta-Birnbaum–Saunders distribution: An improved distribution for fatigue life modeling. *Compuational Statisics and Data Anaysis*. **55**, 1445–1461.

[7] Cordeiro, G.M., Ortega, E.M.M., Popović, B.V. and Pescim, R.R. (2014). The Lomax generator of distributions: Properties, minification process and regression model. *Appied Mathematics and Computations*, **247**, 465–486.

[8] Crowder, M.J., Kimber, A.C., Smith, R.L. and Sweeting, T.J. (1991). *The Statistical Analysis of Reliability Data*. Chapman and Hall: London.

[9] Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics–Theory and Methods*. **31**, 497–512.

[10] Gradshteyn, I.S. and Ryzhik, I.M. (2000). *Table of Integrals, Series, and Products.* Sixth edition. Academic Press: San Diego.

[11] Hosmer, D W. and Lemeshow, S. (1999). *Applied Logistic Regression*. Wiley: New York.

[12] Lee C, Famoye. F. and Olumolade, O. (2007). Beta-Weibull distribution: some properties and applications to censored data. *Journal of Modern and Applied Statistical Methods*, **6**, 173–186.

[13] Lee, E.T. and Wang, J.W. (2003). *Statistical Methods for Survival Data Analysis*. Third edition. Wiley: New York.

[14] Lemonte, A.J. and Cordeiro, G.M. (2013). An extended Lomax distribution. *Statistics*, **47**, 800–816.

[15] Marshall, A.N. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. *Biometrika*, **84**, 641–652.

[16] Mudholkar, G.S. and Srivastava, D.K. (1993). Exponentiated Weibull family for analyzing bathtub failure data. *IEEE Transaction on Reliability*, **42**, 299–302.

[17] Silva, G.O., Ortega, E.M.M. and Cordeiro, G.M. (2010). The beta modified Weibull distribution. *Lifetime Data Analysis*, **16**, 409–430.

[18] Tahir, M.H., Cordeiro, G.M., Mansoor, M. and Zubair, M. (2015). The Weibull-Lomax distribution: properties and applications. *Hacettepe Journal of Mathematics and Statistics*, **44**, 455–474.

[19] Zea, L.M., Silva, R.B., Bourguignon, M., Santos, A.M. and Cordeiro, G.M. (2012). The beta exponentiated Pareto distribution with application to bladder cancer susceptibility. *International Journal of Statistics and Probability*, **1**, 8–19.