Anais da Academia Brasileira de Ciências (2017) 89(1): 3-17
(Annals of the Brazilian Academy of Sciences)
Printed version ISSN 0001-3765 / Online version ISSN 1678-2690
http://dx.doi.org/10.1590/0001-3765201720150579
www.scielo.br/aabc

The Extended Log-Logistic Distribution: Properties and Application

STÊNIO R. LIMA and GAUSS M. CORDEIRO

Universidade Federal de Pernambuco,
Departamento de Estatística, Av. Prof. Moraes Rego, 1235, Cidade Universitária, 50740-540 Recife, PE, Brazil

Manuscript received on September 3, 2015; accepted for publication on October 7, 2016

ABSTRACT

We propose a new four-parameter lifetime model, called the extended log-logistic distribution, to generalize the two-parameter log-logistic model. The new model is quite flexible to analyze positive data. We provide some mathematical properties including explicit expressions for the ordinary and incomplete moments, probability weighted moments, mean deviations, quantile function and entropy measure. The estimation of the model parameters is performed by maximum likelihood using the BFGS algorithm. The flexibility of the new model is illustrated by means of an application to a real data set. We hope that the new distribution will serve as an alternative model to other useful distributions for modeling positive real data in many areas.

Key words: exponentiated generalize, generating function, log-logistic distribution, maximum likelihood estimation, moment.

INTRODUCTION

Numerous classical distributions have been extensively used over the past decades for modeling data in several areas. In fact, the statistics literature is filled with hundreds of continuous univariate distributions (see, e.g, Johnson et al. 1994, 1995). However, in many applied areas, such as lifetime analysis, finance and insurance, there is a clear need for extended forms of these distributions.

The *log-logistic* ("LL" for short) distribution is a very popular distribution pioneered to model population growth by Verhulst (1838). In income inequality literature, the LL model is well-known as the Fisk distribution due to Fisk (1961), and has also been widely used in many areas such as reliability, survival analysis, actuarial science, economics, engineering and hydrology. In some cases, the LL distribution is proved to be a good alternative to the log-normal distribution since it characterizes increasing hazard rate function (hrf). Further, its use is well appreciated in case of censored data usually common in reliability and life-testing experiments, Tahir et al. (2014).

Correspondence to: Stênio Rodrigues Lima
E-mail: stenioamil20@gmail.com

The cumulative distribution function (cdf) $G(x)$ and probability density function (pdf) $g(x)$ of the LL distribution are given by

$$G(x) = G(x;\alpha,\beta) = \frac{x^\beta}{\alpha^\beta + x^\beta}, \quad x > 0, \tag{1}$$

and

$$g(x) = g(x;\alpha,\beta) = \frac{\beta(x/\alpha)^{\beta-1}}{\alpha[1+(x/\alpha)^\beta]^2}, \tag{2}$$

respectively, where $\alpha > 0$ is the scale parameter and also the median of the distribution and $\beta > 0$ is the shape parameter. Some basic properties of the LL distribution are given, for example, by Kleiber and Kotz (2003), Lawless (2003) and Ashkar and Mahdi (2006). The moments (for $r < \beta$) are easily derived as, Tadikamalla (1980)

$$E(T^r) = \alpha^r B(1 - r\beta^{-1}, 1 + r\beta^{-1}) = \frac{r\pi\alpha^r\beta^{-1}}{\sin(r\pi\beta^{-1})}, \tag{3}$$

where $B(a,b) = \int_0^1 \omega^{a-1}(1-\omega)^{b-1}d\omega$ is the beta function. Hence,

$$E(T) = \frac{\pi\alpha\beta^{-1}}{\sin(\pi\beta^{-1})}, \; \beta > 1 \quad \text{and} \quad \mathrm{Var}(T) = \frac{2\pi\alpha^2\beta^{-1}}{\sin(2\pi\beta^{-1})} - \left[\frac{\pi\alpha\beta^{-1}}{\sin(\pi\beta^{-1})}\right]^2, \; \beta > 2.$$

If T has a LL distribution with scale parameter α and shape parameter β, LL(α,β) say, then $W = \log(T)$ has a logistic distribution with location parameter $\log(\alpha)$ and scale parameter $1/\beta$. It has attracted a wide applicability in survival and reliability over the last few decades, particularly for events whose failure rate increases initially and decreases later, for example, mortality from cancer following diagnosis or treatment, Gupta et al. (1999). It has also been used in hydrology to model stream flow and precipitation – see, Shoukri et al. (1988) and Ashkar and Mahdi (2006) – and for modeling flood frequency, see, Ahmad et al. (1988).

For a baseline continuous cdf $G(x)$, Cordeiro et al. (2013) defined the *exponentiated generalized* ("EG") class of distributions by the cdf

$$F(x) = \{1 - [1 - G(x)]^a\}^b, \qquad x \in \mathbb{R}, \tag{4}$$

where $a > 0$ and $b > 0$ are two extra parameters whose role is to govern the skewness and generate distributions with heavier/ligther tails. They are sought as a manner to furnish a more flexible distribution. Because of its tractable cdf (4), this class can be used quite effectively even if the data are censored. The EG class is suitable for modeling continuous univariate data that can be in any interval of the real line. The pdf corresponding to (4) is given by

$$f(x) = a\, b\, [1 - G(x)]^{a-1} \{1 - [1 - G(x)]^a\}^{b-1} g(x), \qquad x \in \mathbb{R}, \tag{5}$$

where $g(x) = \mathrm{d}G(x)/\mathrm{d}x$ is the baseline pdf. The two parameters in (5) can control both tail weights and possibly adding entropy to the center of the EG density. The baseline pdf $g(x)$ is a special case of (5) when $a = b = 1$. Setting $a = 1$ it gives to the exponentiated-G ("exp-G") class of distributions. If $b = 1$, we obtain Lehmann type II distribution. So, the distribution (5) generalizes both Lehmann types I and II alternative distributions; that is, this method can be interpreted as a double construction of Lehmann alternatives. Note that even if $g(x)$ is a symmetric density, the density $f(x)$ will not be symmetric.

In this paper, we propose and study some structural properties of a new four-parameter distribution with positive real support, called the *extended log-logistic* ("ELL") distribution, derived from equation (5) by taking $G(x)$ and $g(x)$ to be the cdf and pdf of the LL distribution, respectively. We also discuss maximum likelihood estimation of the model parameters of the proposed distribution. We adopt a different approach to much of the literature so far: rather than considering the beta and Kumaraswamy generators applied to a baseline distribution, we consider the EG generator applied to the LL distribution. The ELL distribution can be applied to positive real data in several areas in a similar manner of the beta log-logistic and Kumaraswamy log-logistic distributions.

The rest of the paper is organized as follows. In Section "The ELL Distribution', we define the ELL distribution and provide plots of the density and hazard rate functions. In Section "useful representations", we derive a useful mixture representation for its density function in terms of LL densities. In Section "'Main Properties', we provide some mathematical properties of the proposed model. In fact, explicit expressions for the ordinary and incomplete moments, a power series expansion for the quantile function (qf), standard measures for the skewness and kurtosis, probability weighted moments, mean deviations and entropy measure. Estimation by the method of maximum likelihood is presented in Section "Maximum Likelihood Estimation". An application to a real data set illustrates the flexibility of the ELL distribution in Section "Application". Concluding remarks are given in Section "Conclusions".

THE ELL DISTRIBUTION

By inserting (1) and (2) in equation (5), we obtain the ELL density function (for $x > 0$) with positive parameters a, b, α and β, say ELL(a,b,α,β), given by

$$f(x) = \frac{ab\beta(x/\alpha)^{\beta-1}}{\alpha[1+(x/\alpha)^{\beta}]^2}\left(1-\frac{x^{\beta}}{\alpha^{\beta}+x^{\beta}}\right)^{a-1}\left[1-\left(1-\frac{x^{\beta}}{\alpha^{\beta}+x^{\beta}}\right)^{a}\right]^{b-1}. \tag{6}$$

The ELL density is straightforward to compute using any statistical software with numerical facilities. Also, there is no functional relationship between the parameters and they vary freely in the parameter space. The new distribution due to its flexibility in accommodating non-monotonic hrf may be an important distribution that can be used in a variety of problems in modeling survival data. The LL and Exp-LL (exponentiated log-logistic) distributions are clearly the most important sub-models of (6) for $a = b = 1$ and $a = 1$, respectively. If $b = 1$, we obtain Lehmann type II log-logistic distribution. The cdf and hrf corresponding to (6) are

$$F(x) = \left[1-\left(1-\frac{x^{\beta}}{\alpha^{\beta}+x^{\beta}}\right)^{a}\right]^{b} \tag{7}$$

and

$$h(x) = \frac{ab\beta(x/\alpha)^{\beta-1}\left(1-\frac{x^{\beta}}{\alpha^{\beta}+x^{\beta}}\right)^{a-1}\left[1-\left(1-\frac{x^{\beta}}{\alpha^{\beta}+x^{\beta}}\right)^{a}\right]^{b-1}}{\alpha[1+(x/\alpha)^{\beta}]^2\left\{1-\left[1-\left(1-\frac{x^{\beta}}{\alpha^{\beta}+x^{\beta}}\right)^{a}\right]^{b}\right\}}. \tag{8}$$

The qf of X, say $x_u = Q(u)$, follows easily by inverting (7) as

$$x_u = Q(u) = \frac{\alpha^{\beta}[1-(1-u^{1/\beta})^{1/\alpha}]}{(1-u^{1/\beta})^{1/\alpha}}, \quad u \in (0{,}1). \tag{9}$$

Quantiles of interest can be obtained from (9) by substituting appropriate values for u. In particular, the median of X is

$$Median(X) = Q(0.5) = \frac{\alpha^{\beta}[1-(1-0.5^{1/\beta})^{1/\alpha}]}{(1-0.5^{1/\beta})^{1/\alpha}}.$$

We can also use (9) for simulating ELL random variables: if U is a uniform random variable on the unit interval $(0{,}1)$, then

$$X = Q(U) = \frac{\alpha^{\beta}[1-(1-U^{1/\beta})^{1/\alpha}]}{(1-U^{1/\beta})^{1/\alpha}}$$

has the pdf (6).

We consider the generalized binomial expansion

$$(1-z)^b = \sum_{k=0}^{\infty} (-1)^k \binom{b}{k} z^k, \tag{10}$$

which holds for any real non-integer b and $|z| < 1$. Using (10) in equation (9), we can rewrite $Q(u)$ as

$$Q(u) = \sum_{k=1}^{\infty} a_k \, u^{k/\beta}, \tag{11}$$

where (for $k > 0$) $a_k = (-1)^k \alpha^{\beta} \binom{1/\alpha}{k}$.

Figures 1 and 2 display some shapes of (6) and (8), respectively, for specified parameter values. The pdf and hrf take various forms depending on the parameter values. The ELL distribution is much more flexible than the LL distribution, that is, the additional shape parameters (a and b) allow for a high degree of its flexibility. So, the new model can be very useful in many practical situations for modeling positive real data sets.

USEFUL REPRESENTATIONS

Some useful representations for (4) and (5) can be derived using the concept of exponentiated distributions. The properties of these distributions have been studied by many authors in recent years. See, for example, Mudholkar et al. (1995) and Mudholkar and Srivastava (1993) for exponentiated Weibull, Gupta et al. (1998) for exponentiated Pareto, Nadarajah (2005) for exponentiated Gumbel, Nadarajah and Gupta (2007) for exponentiated gamma, and Lemonte (2012) for exponentiated Kumaraswamy (Kw), among others.

Cordeiro and Lemonte (2015) proved, using (10) twice in equation (4), that the EG cdf can be expressed as $F(x) = \sum_{j=0}^{\infty} \omega_{j+1} H_{j+1}(x)$, where $\omega_{j+1} = \sum_{m=1}^{\infty} (-1)^{j+m+1} \binom{b}{m} \binom{ma}{j+1}$ and $H_{j+1}(x)$ is the exp-G cdf with power parameter $j+1$. By differentiating $F(x)$, we can write

$$f(x) = \sum_{j=0}^{\infty} w_{j+1} \, h_{j+1}(x), \tag{12}$$

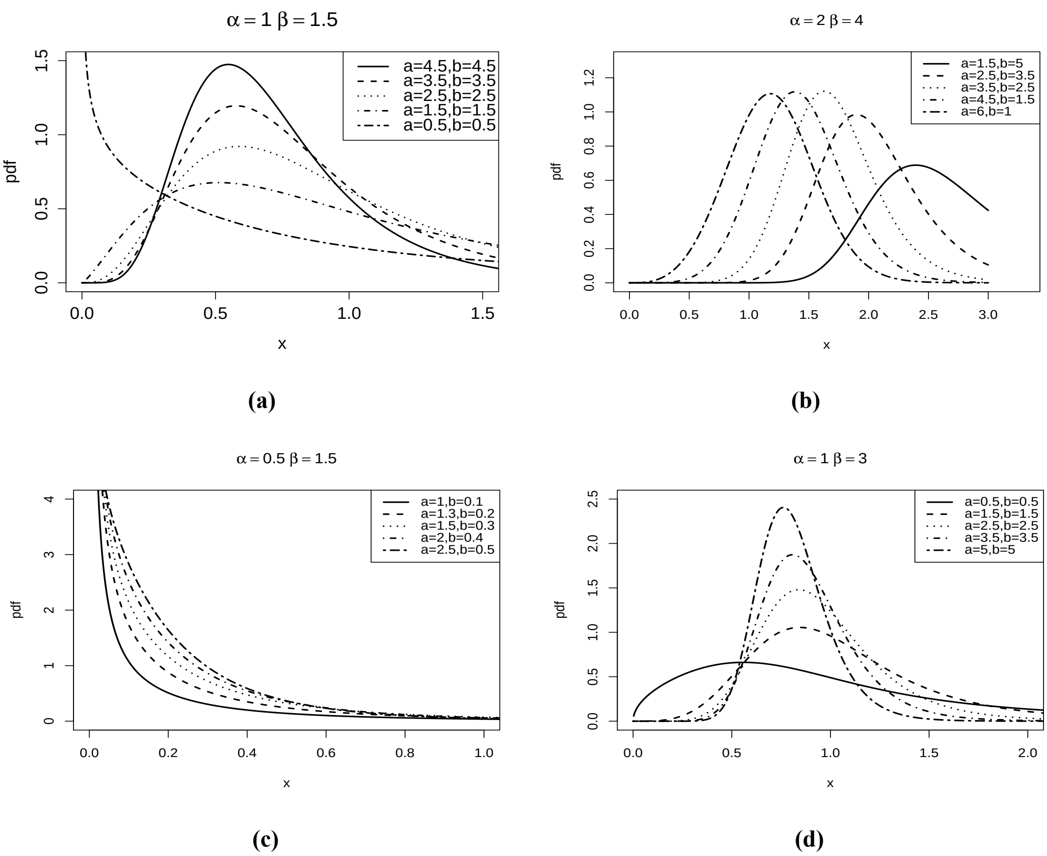


Figure 1 - Plots of the ELL pdf for some parameter ponto.

where $h_{j+1}(x)$ is the exp-G pdf with power parameter $j+1$. Thus, equation (12) reveals that the EG density function is a mixture of exp-G densities. This result is important to obtain some properties of the EG class from those exp-G properties.

An expansion for (6) can be derived by defining the exponentiated log-logistic (exp-LL) distribution. A random variable Z has the exp-LL distribution with power parameter $d > 0$, say $Z \sim$ exp-LL(d), if its cdf and pdf are given, respectively, by

$$H_d(x) = \left[\frac{x^\beta}{\alpha^\beta + x^\beta}\right]^d, \qquad h_d(x) = \frac{d\beta(x/\alpha)^{\beta-1}}{\alpha[1+(x/\alpha)^\beta]^2}\left[\frac{x^\beta}{\alpha^\beta + x^\beta}\right]^{d-1}.$$

First, the expansion holds

$$(1+z)^{-a} = \sum_{n=0}^{\infty} \binom{-a}{n} z^n. \tag{13}$$

Inserting this expansion in $h_d(x)$ gives

$$h_d(x) = \sum_{n=0}^{\infty} s_n\, x^{\beta(d+n)-1},$$

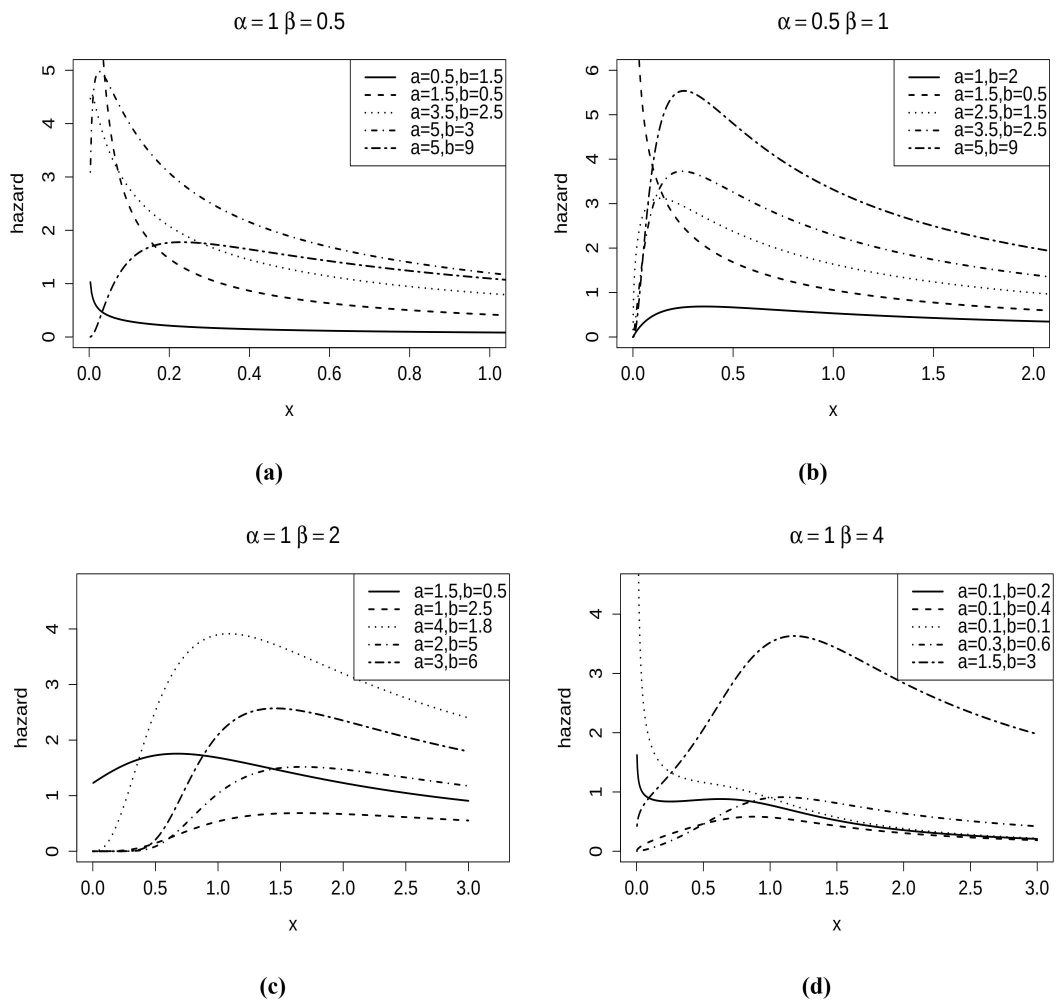


Figure 2 - Plots of the ELL hrf for some parameter ponto.

where $s_n = \frac{d\beta}{\alpha^{\beta(d+n)}}\binom{-d-1}{n}$. Thus,

$$h_{j+1}(x) = (j+1)\beta \sum_{n=0}^{\infty} \binom{-j-2}{n} \frac{x^{\beta(j+n+1)-1}}{\alpha^{\beta(j+n+1)}}. \tag{14}$$

Inserting the last expression in equation (12) gives

$$f(x) = \sum_{j,n=0}^{\infty} \omega_{j+1}\, s_n\, x^{\beta(j+n+1)-1}, \tag{15}$$

where $s_n = \frac{(j+1)\beta}{\alpha^{\beta(j+n+1)}}\binom{-j-2}{n}$.

In a more simplified form, the last equation reduces to

$$f(x) = \sum_{j,n=0}^{\infty} q_{j,n}\, g(x;\alpha,\beta(j+n+1)). \tag{16}$$

where $q_{j,n} = s_n\, w_{j+1}$. Equation (16) reveals that the ELL density function is a mixture of LL densities. The cdf corresponding (16) is evidently given by

$$F(x) = \sum_{j,n=0}^{\infty} q_{j,n}\, G(x;\alpha,\beta(j+n+1)). \tag{17}$$

where $G(x;\alpha,\beta(j+n+1))$ denotes the cdf of the LL distribution with shape parameter $\beta(j+n+1)$. So, several of the ELL mathematical properties can be obtained form those of the LL distribution. The coefficients $q_{j,n}$ depend only on the generator parameters. This equation is the main result of this section.

MAIN PROPERTIES

In this section, we obtain the ordinary and incomplete moments, moment generating function (mgf), qf, probability weighted moments (PWMs), entropy measure and mean deviations of the ELL distribution. The formulas derived throughout the paper can be easily handled in analytical softwares such as **MAPLE** and **MATHEMATICA** which have the ability to deal with analytic expressions of formidable size and complexity.

MOMENTS AND GENERATING FUNCTION

We derive a simple expression for the rth moment of X, $\mu_r' = \mathbb{E}(X^r) = \int_0^\infty x^r f(x)dx$. We have

$$\mu_r' = \sum_{j,n=0}^{\infty} q_{j,n} \int_0^\infty x^r g(x;\alpha,\beta(j+n+1))dx.$$

By using (3), we obtain

$$\mu_r' = \alpha^r\, B(1-r\beta(j+n+1)^{-1},1+r\beta(j+n+1)^{-1}) = \frac{r\pi\alpha^r\beta(j+n+1)^{-1}}{\sin(r\pi\beta(j+n+1)^{-1})}, \tag{18}$$

where $B(a,b) = \int_0^1 \omega^a (1-\omega)^{b-1} d\omega$ is the beta function. Setting $r=1$ in (18), it follows that the mean of X is

$$\mu_1' = \frac{\pi\alpha\beta(j+n+1)^{-1}}{\sin(\pi\beta(j+n+1)^{-1})}.$$

Further, the central moments (μ_n) and cumulants (κ_n) of X are obtained from (18) as

$$\mu_s = \sum_{k=0}^{p} \binom{s}{k} (-1)^k\, \mu_1'^s\, \mu_{s-k}' \qquad \textit{and} \qquad \kappa_s = \mu_s' - \sum_{k=1}^{s-1} \binom{s-1}{k-1} \kappa_k\, \mu_{s-k}',$$

respectively, where $\kappa_1 = \mu_1'$. The skewness $\gamma_1 = \kappa_3/\kappa_2^{3/2}$ and kurtosis $\gamma_2 = \kappa_4/\kappa_2^2$ can be calculated from the third and fourth standardized cumulants.

For empirical purposes, the shape of many distributions can be usefully described by the incomplete moments. These types of moments play an important role for measuring inequality, for example, income quantiles. The main application of the first incomplete moment refers to the Bonferroni and Lorenz curves. These curves are very useful in economics, reliability, demography, insurance and medicine.

The sth incomplete moment of X is defined by $m_s(y) = \int_0^y x^s f(x)\mathrm{d}x$. From equation (15), we obtain

$$m_s(y) = \sum_{j,n=0}^{\infty} q_{j,n} \frac{y^{\beta(j+n+1)+s}}{\beta(j+n+1)+s}. \tag{19}$$

Applications of these equations to obtain Bonferroni and Lorenz curves are important in several fields such as economics, reliability, demography, insurance and medicine. For a given probability π, these curves are given by $B(\pi) = m_1(q)/\pi\mu_1'$ and $L(\pi) = m_1(q)/\mu_1'$, respectively, where $\mu_1' = E(X)$, $q = \frac{\alpha^\beta[1-(1-u^{1/\beta})^{1/\alpha}]}{(1-u^{1/\beta})^{1/\alpha}}$ is the qf of X at π and $m_1(q)$ comes from (19) with $s = 1$.

QUANTILE FUNCTION

The effects of the shape parameters a and b on the skewness and kurtosis can be considered based on quantile measures. The shortcomings of the classical kurtosis measure are well-known. Bowley skewness, Kenney and Keeping (1962), and Moors kurtosis, Moors (1998), are based on quartiles and octiles given by

$$\mathcal{B} = \frac{Q(3/4) + Q(1/4) - 2Q(1/2)}{Q(3/4) - Q(1/4)}$$

and

$$\mathcal{M} = \frac{Q(3/8) - Q(1/8) + Q(7/8) - Q(5/8)}{Q(6/8) - Q(2/8)},$$

respectively. Clearly, $\mathcal{M} > 0$ and there is good concordance with the classical kurtosis measures for some distributions. These measures are less sensitive to outliers and they exist even for distributions without moments.

Now, consider an equation for a power series raised to a positive integer r

$$\left(\sum_{m=0}^{\infty} w_m\, x^m\right)^r = \sum_{m=0}^{\infty} p_{r,m}\, x^m, \tag{20}$$

given in Gradshteyn and Ryzhik (2007), where the coefficients $p_{r,m}$ (for $m = 1,2,\ldots$) can be determined from $w_0,\ldots,w_m$ based on the recurrence equation $p_{r,m} = (m\, w_0)^{-1} \sum_{k=1}^{m} [(r+1)k - m]\, w_k\, p_{r,m-k}$, with $p_{r,0} = w_0^r$. Clearly, $p_{r,m}$ can be computed numerically in any algebraic or numerical software.

Using (10) in equation (9), we can rewrite $Q(u)$ as

$$Q(u) = \sum_{j=1}^{\infty} s_j\, u^{j/\beta}, \quad u \in (0,1), \tag{21}$$

where $s_j = \alpha^\beta (-1)^j \binom{-1/\alpha}{j}$.

Let $W(\cdot)$ be any integrable function in the positive real time. We can write

$$\int_0^\infty W(x) f(x) dx = \int_0^1 W\left(\sum_{j=1}^{\infty} s_j u^{j/\beta}\right) du. \tag{22}$$

Equations (21) and (22) are the main results of this section since various ELL mathematical quantities can follow from them. For example, $\mu_n' = \int_0^1 (\sum_{j=1}^\infty s_j u^{j/\beta})^n \mathrm{d}u = \sum_{j=1}^\infty h_{n,j} \int_0^1 u^{j/b} \mathrm{d}u = \sum_{j=1}^\infty h_{n,j}/(j/b+1)$, where $h_{n,j}$ can be computed from the quantities s_j based on equation (20).

PROBABILITY WEIGHTED MOMENTS

The PWMs of a random variable T with cdf $G(X)$ are defined by $\delta_{s,r} = \mathbb{E}[X^s\, G(X)^r]$ for s and r positive integers. Here, the PWMs of the LL distribution are used to compute the ordinary moments of the ELL distribution. If $T \sim LL(\alpha,\beta)$, for $s < \beta$, we obtain

$$\begin{aligned}\delta_{s,r} &= \frac{\beta}{\alpha^{\beta}} \int_0^{\infty} x^{s+\beta-1} \left[1+\left(\frac{x}{\alpha}\right)^{\beta}\right]^{-2} \left\{1-\left[1+\left(\frac{x}{\alpha}\right)^{\beta}\right]^{-1}\right\}^{r} dx \\ &= \alpha^{s} \int_0^{1} \omega^{-s\beta^{-1}} (1-\omega)^{r+s\beta^{-1}} d\omega = \alpha^{s}\, B(r+1+s\beta^{-1}, 1-s\beta^{-1}). \end{aligned} \quad (23)$$

The ordinary moments of X follow as $\delta_{s,0} = \mathbb{E}(X^s)$. Based on (21), we can write

$$\delta_{s,r} = \int_0^1 \left(\sum_{j=1}^{\infty} s_j u^{j/\beta}\right)^{s} u^r \mathrm{d}u.$$

Then, using (20), we obtain

$$\delta_{s,r} = \sum_{j=1}^{\infty} t_{s,j} \int_0^1 u^{r+j/b} \mathrm{d}u = \sum_{j=1}^{\infty} \frac{t_{s,j}}{r+j/b+1}, \quad (24)$$

where $t_{s,j} = (j\,\nu_0)^{-1} \sum_{k=1}^{j} [(s+1)k - j]\, \nu_k\, h_{s,j-k}$ and $t_{s,0} = \nu_0^s$. Equation (24) is very simple to be calculated analytically or numerically using symbolic computer systems.

MEAN DEVIATIONS

The amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. The mean deviations about the mean ($\delta_1 = E(|X - \mu_1'|)$) and about the median ($\delta_2 = E(|X - M|)$) of X can be expressed as $\delta_1 = 2\mu_1'\, F(\mu_1') - 2m_1\,(\mu_1')$ and $\delta_2(X) = \mu_1' - 2m_1(M)$, respectively, where $\mu_1' = E(X)$, $M = Median(X)$ is the median, and $m_1(z)$ is the first incomplete moment obtained from (19) with $s = 1$.

ENTROPY MEASURE

Entropy has been used in various situations in science as a measure of variation of the uncertainty. Numerous measures of entropy have been studied and compared in the literature. Let $X \sim ELL$(a,b,α,β). Here, we shall derive an explicit expression for the Shannon entropy. We consider this entropy since it plays a similar role as the kurtosis measure in comparing the shapes of various densities and measuring heaviness of the tails. It is defined by $\mathcal{I}_S = \mathbb{E}\{-\log[f(X)]\}$, which implies

$$\begin{aligned}\mathcal{I}_S \;&= -\log\left(\tfrac{ab\beta}{\alpha}\right) - (a-1)\mathbb{E}\left[\log\left(\tfrac{\alpha^{\beta}}{\alpha^{\beta}+x^{\beta}}\right)\right] - (b-1)\mathbb{E}\left\{\log\left[1-\left(\tfrac{\alpha^{\beta}}{\alpha^{\beta}+x^{\beta}}\right)\right]\right\} \\ &\quad -(\beta-1)\mathbb{E}\left[\log\left(\tfrac{x}{\alpha}\right)\right] + 2\mathbb{E}\left\{\log\left[1+\left(\tfrac{x}{\alpha}\right)\right]\right\}.\end{aligned}$$

We can prove that

$$
\begin{aligned}
\mathbb{E}\left[\log\left(\frac{x}{\alpha}\right)\right] &= \frac{ba^b\alpha^{\beta(a+2b-ab-1)}}{\beta\Gamma(ab+1)}\Gamma(b)\Gamma(b(a-1)+1)\times \\
&\quad \left\{\beta\log(1/\alpha)-\log(\alpha^{-\beta})+\psi(b)-\psi[b(a-1)+1]\right\} \\
&- \frac{(b-1)\alpha^{\beta a(1-b)}[\gamma-1+\beta\log(1/\alpha)+\log(\alpha^{-\beta})+\psi(ab-1)]}{\beta(ab-1)} \\
&+ \frac{ab\alpha^{\beta(ab-2)+1}}{\beta\Gamma(ab+1)}\Gamma(a-1/\beta+2)\Gamma(a(b-1)+1/\beta-1)\times \\
&\quad \left\{\beta\log(1/\alpha)-\log(\alpha^{\beta})-\psi(a-1/\beta+2)+\psi(a(b-1)+1/\beta-1)\right\}.
\end{aligned}
$$

Some other expected values can be determined analytically using the **MATHEMATICA** software and they can be obtained from the authors upon request.

MAXIMUM LIKELIHOOD ESTIMATION

We consider the estimation of the unknown parameters of the ELL distribution by the method of maximum likelihood. Let $x_1,\ldots,x_n$ be a random sample from (6). Let $\Theta=(a,b,\alpha,\beta)^\top$ be the vector of model parameters. The log-likelihood function $\log L=\log L(\boldsymbol{\theta})$ for θ is given by

$$
\begin{aligned}
\log L \quad &= n\log\left(\tfrac{ab\beta}{\alpha}\right)+(a-1)\textstyle\sum_{i=1}^n\log\left(\tfrac{\alpha^\beta}{\alpha^\beta+x_i^\beta}\right)+(b-1)\textstyle\sum_{i=1}^n\log\left[1-\left(\tfrac{\alpha^\beta}{\alpha^\beta+x_i^\beta}\right)^a\right] \\
&\quad +(\beta-1)\textstyle\sum_{i=0}^n\log\left(\tfrac{x}{\alpha}\right)-2\textstyle\sum_{i=0}^n\log\left[1+\left(\tfrac{x}{\alpha}\right)^\beta\right]. \qquad (25)
\end{aligned}
$$

Maximization of (25) can be performed by using well-established routines like NLM or OPTIMIZE in the R statistical package, NLMIXED procedure in the SAS or MaxBFGS in the Ox program or, alternatively, by solving the nonlinear likelihood equations obtained by differentiating (25).

The components of the score vector $\boldsymbol{U}=\boldsymbol{U}(\boldsymbol{\theta})=(\partial\log L/\partial a,\partial\log L/\partial b,\partial\log L/\partial\alpha,\partial\log L/\partial\beta)^\top$ are given by

$$
\begin{aligned}
\frac{\partial\log L}{\partial a} &= \frac{n}{a}+n\log(\alpha^\beta)-n\log(\alpha^\beta+x^\beta)-\sum_{i=0}^n\left[a(b-1)\frac{\alpha^{\beta a}(\alpha^\beta+x^\beta)}{\alpha[a(\alpha x)^\beta+x^{\beta a}]}\right], \\
\frac{\partial\log L}{\partial b} &= \frac{n}{b}+\sum_{i=0}^n\log\left[1-\left(\frac{\alpha^\beta}{\alpha^\beta+x^\beta}\right)^a\right], \\
\frac{\partial\log L}{\partial \alpha} &= -\frac{n}{\alpha}-\frac{n(\beta-1)}{\alpha}+(a-1)\sum_{i=0}^n\frac{\beta\alpha^{\beta-1}x^\beta}{\alpha^\beta(\alpha^\beta+x^\beta)}+\sum_{i=0}^n\frac{a(b-1)\beta\alpha^{\beta a-1}x^\beta}{a(\alpha x)^\beta+x^{\beta a}}+ \\
&\quad 2\sum_{i=0}^n\frac{\alpha^\beta\beta x^\beta}{\alpha^\beta(\alpha^\beta+x^\beta)}, \\
\frac{\partial\log L}{\partial \beta} &= \frac{n}{\beta}+(a-1)\sum_{i=0}^n\frac{(\alpha x)^\beta\log(\alpha x)}{(\alpha^\beta+x^\beta)^2}+\sum_{i=0}^n\frac{a(b-1)\alpha^{\beta a}x^\beta\log(\alpha/x)}{a(\alpha x)^\beta+x^{\beta a}}+\sum_{i=0}^n\log\left(\frac{x}{\alpha}\right)- \\
&\quad 2\sum_{i=0}^n\frac{\alpha\beta x^{\beta-1}}{\alpha^\beta+x^\beta}.
\end{aligned}
$$

The maximum likelihood estimates (MLEs) in $(\boldsymbol{\theta})$, say $(\hat{\boldsymbol{\theta}})$, are the simultaneous solutions of the equations $\partial L/\partial\boldsymbol{\theta} = \mathbf{0}$. For interval estimation and tests of hypotheses on the model parameters, we require the 4×4 observed information matrix $\boldsymbol{J}(\boldsymbol{\theta})$, whose elements can be obtained from the authors under request.

Upon conditions that are fulfilled for parameters in the interior of the parameter space, the asymptotic distribution of $(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is $\mathcal{N}_4(\mathbf{0}, \boldsymbol{J}(\boldsymbol{\theta})^{-1})$, which can be used to construct approximate confidence intervals and confidence regions for the parameters. We believe that the standard likelihood regularity conditions are satisfied for the ELL model. These conditions are: i) the support of the distribution does not depend on unknown parameters; ii) the parameter space is open and the log-likelihood function has a global maximum in it; iii) the third order log likelihood derivatives have finite expected values; iv) the fourth order log likelihood derivatives exist almost everywhere, and are continuous in an open neighborhood that contains the true parameter value; v) the expected information matrix is positive definite and finite. These regularity conditions hold for almost distributions which satisfy i). So, they are not restrictive.

APPLICATION

In this section, we compare the fits of the ELL distribution defined in (6), the exponentiated log-logistic (ELLog) Rosaiah et al. (2006), McDonald log-logistic (McLL) Tahir et al. (2014), beta log-logistic (BeLL) Lemonte (2012), Kumaraswamy log-logistic (KwLL) de Santana et al. (2012), Marshal-Olkin log-logistica (MoLL) Gui (2013) and log-logistic (LL) to a real data set. In many applications there is qualitative information about the hrf, which can help with selecting a particular model. In this context, a device called the total time on test (**TTT**) plot Aarset (1987) is useful. The **TTT** plot is obtained by plotting $G(r/n) = [(\sum_{i=1}^{r} y_{i:n}) + (n-r)y_{r:n}]/\sum_{i=1}^{n} y_{i:n}$, where $r = 1, \ldots, n$ and $y_{i:n}$ $(i = 1, \ldots, n)$ are the order statistics of the sample, against r/n. It is a straight diagonal for constant failure rates, it is convex for decreasing failure rates and concave for increasing failure rates. It is first convex and then concave if the failure rate is bathtub-shaped. It is first concave and then convex if the failure rate is upside-down bathtub.

We consider an uncensored data set from Nichols and Padgett (2006) consisting of 100 observations on breaking stress of carbon fibres (in Gba). The **TTT** plot for the exceedances of flood peaks data in Figure 3 indicates a bathtub-shaped hrf and, therefore, the appropriateness of the ELL distribution to fit these data.

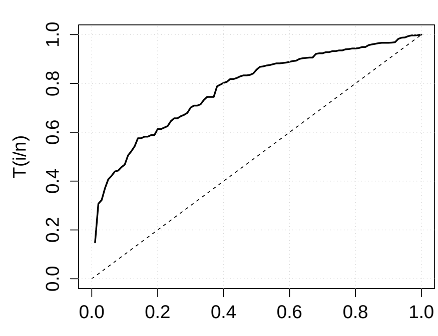


Figure 3 - The TTT-plot for the breaking stress of carbon fibres (in Gba) data.

Further, we fit the ELL, ELLog, McLL, BeLL, KwLL, MoLL and LL models (for $x > 0$) with corresponding densities:

$$
\begin{aligned}
\text{ELLog} \quad &: \quad f_{ELLog}(x) = \frac{\alpha a}{\beta^{\alpha a}} x^{\alpha a - 1} \left[1 + \left(\frac{x}{\beta}\right)^{\alpha}\right]^{-(a+1)}, \\
\text{McLL} \quad &: \quad f_{McLL}(x) = \frac{c\alpha}{B(ac^{-1},b)\, \beta^{a\alpha-1}} x^{\alpha a - 1} \left[1 + \left(\frac{x}{\beta}\right)^{\alpha}\right]^{-(a+1)} \times \\
&\qquad \left[1 - \left\{1 - \left[1 + \left(\frac{x}{\beta}\right)^{\alpha}\right]^{-1}\right\}^{c}\right]^{b-1}, \\
\text{BeLL} \quad &: \quad f_{BeLL}(x) = \frac{\alpha}{B(a,b)\, \beta^{a\alpha-1}} x^{a\alpha - 1} \left[1 + \left(\frac{x}{\beta}\right)^{\alpha}\right]^{-(\alpha+\beta)}, \\
\text{KwLL} \quad &: \quad f_{KwLL}(x) = \left(\frac{ab\alpha}{\beta}\right) \left(\frac{x}{\beta}\right)^{a\alpha - 1} \left[1 + \left(\frac{x}{\beta}\right)^{\alpha}\right]^{-(a+1)} \left\{1 - \left[1 - \frac{1}{1 + \left(\frac{x}{\beta}\right)^{\alpha}}\right]^{a}\right\}^{b-1}, \\
\text{MoLL} \quad &: \quad f_{MoLL}(x) = \frac{\alpha^{\beta} \beta a x^{\beta - 1}}{(x^{\beta} + \alpha^{\beta} a)^2}, \\
\text{LL} \quad &: \quad f_{LL}(x) = \frac{\alpha}{\beta^{\alpha}} x^{\alpha - 1} \left[1 + \left(\frac{x}{\beta}\right)^{\alpha}\right]^{-2},
\end{aligned}
$$

where $a, b, c > 0$, $\alpha > 0$ and $\beta > 0$.

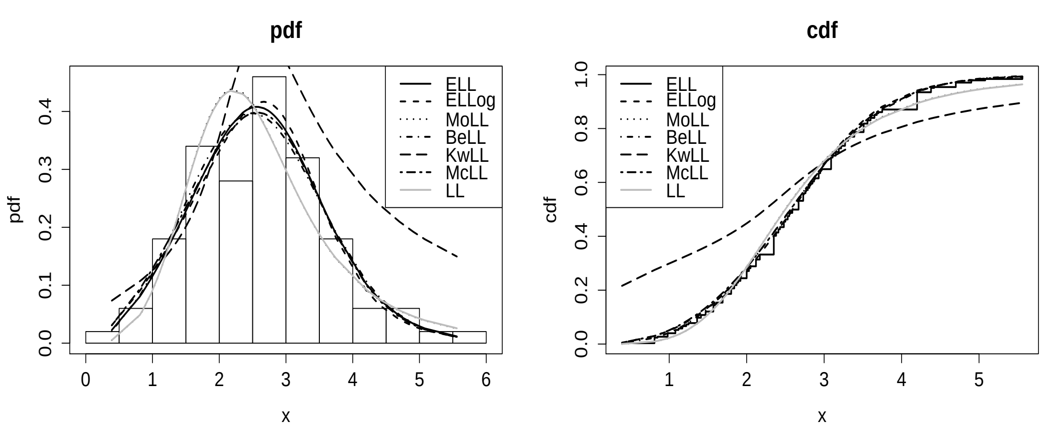
Table I lists the MLEs of the parameters (their standard errors are given in parentheses) for the ELL, ELLog, McLL, BeLL, KwLL, MoLL and LL models fitted to the exceedances of flood peak data. We estimate the unknown parameters of each model by maximum likelihood. There exists many maximization methods in R packages like NR (Newton-Raphson), BFGS (Broyden-FletcherGoldfarb-Shanno), BHHH (Berndt-Hall-Hall-Hausman), SANN (Simulated-Annealing) and NM (Nelder-Mead). The MLEs are calculated using the Limited Memory quasi-Newton code for Bound-constrained optimization (L-BFGS-B). Further, the Anderson-Darling (A^*) and Cramér-Von Mises (W^*) statistics are computed to compare the fitted models. The computations are carried out using the R-package *AdequacyModel* given freely from *http://cran.r-project.org/web/packages/AdequacyModel/AdequacyModel.pdf.*

Plots of the estimated pdfs and cdfs of the fitted models are displayed in Figure 4. It is clear that the ELL, BeLL and McLL distributions present practically the same fits, i.e., there is significant difference among the curves. Additionally, these plots indicate that these models provide better fits than the other models.

The statistics W^* and A^* are described in Chen and Balakrishnan (1995). In general, the smaller the values of these measures, the better the fit to the data. The statistics W^* and A^* for all models fitted to the current data are given in Table II. From the figures in this table, we conclude that the ELL model fits these data better than the other models. Therefore, the ELL distribution may be an interesting alternative to other models available in the literature for modeling positive real data with bathtub-shaped hrf. In summary, the proposed distribution produces a better fit for the current data than other known lifetime models and it could be chosen since has fewer parameters to be estimated.

TABLE I
The MLEs (standard errors in parentheses) of the model parameters.

Model	Estimates					AIC
ELL(a,b,α,β)	3.607	4.989	1.885	0.548		**288.80**
	(0.176)	(0.751)	(0.137)	(0.032)		
ELLog(a,α,β)	3.369	7.345	0.336			288.92
	(0.219)	(1.494)	(0.098)			
McLL(a,b,c,α,β)	0.506	1.996	1.214	4.985	3.729	292.34
	(0.162)	(0.429)	(3.439)	(33.868)	(0.966)	
BeLL(a,b,α,β)	3.424	4.891	0.851	5.359		290.38
	(39.488)	(0.171)	(0.148)	(0.381)		
KwLL(a,b,α,β)	2.448	10.821	0.014	0.172		296.81
	(0.026)	(0.055)	(0.001)	(0.012)		
MoLL(a,α,β)	1.800	4.125	3.843			298.55
	(30.412)	(0.344)	(267.146)			
LL(α,β)	2.499	4.109				296.55
	(0.105)	(0.343)				



(a) **(b)**

Figure 4 - Estimated pdfs and cdfs of the ELL,ELLog, McLL, BeLL, KwLL, MoLL and LL distributions fitted to the data from Nichols and Padgett (2006).

TABLE II
The W^* and A^* statistics.

Model	W^*	A^*
ELL	**0.040**	**0.279**
ELLog	0.046	0.302
McLL	0.052	0.331
BeLL	0.058	0.356
KwLL	0.137	0.821
MoLL	0.239	1.241
LL	0.239	1.243

CONCLUSIONS

In this paper, we propose a new four-parameter distribution, called the extended log-logistic (ELL) distribution, and study some of its general structural properties. This distribution has the support in the positive real interval. Further, the new distribution includes as special models other known distributions. We provide explicit expressions for the ordinary and incomplete moments, probability weighted moments, quantile function, mean deviations and entropy measure. The model parameters are estimated by maximum likelihood. The usefulness of the new model is illustrated by means of one application to real data. For these data the new model provides a consistently better fit than other known lifetime models. We hope that the proposed model may attract wider applications for modeling positive real data sets in many areas such as engineering, survival analysis, hydrology, economics, among others.

REFERENCES

AARSET MV. 1987. How to identify bathtub hazard rate. IEEE T Rehabil Eng 36: 106-108.

AHMAD MI, SINCLAIR CD AND WERRITTY A. 1988. Log-logistic ood frequency analysis. J Hydrol 98: 205-224.

ASHKAR F AND MAHDI S. 2006. Fitting the log-logistic distribution by generalized moments. J Hydrol 328: 694-703.

CHEN G AND BALAKRISHNAN N. 1995. A general purpose approximate goodness-of-fit test. J Qual Technol 27: 154-161.

CORDEIRO GM, LEMONTE AL AND CAMPELO AK. 2016. Extended Arcsine Distribution to Proportional Data: Properties and Applications. Stud Sci Math Hung 53.

CORDEIRO GM, ORTEGAM EMM AND DA CUNHA DCC. 2013. The exponentiated generalized class of distributions. JDS 11: 1-27

DE SANTANA TVF, ORTEGA EMM, CORDEIRO GM AND SILVA GO. 2012. The Kumaraswamy-log-logistic distribution. J Stat Theor Appl 11: 265-291.

FISK PR. 1961. The graduation of income distribution. Econometrica 29: 171-185.

GRADSHTEYN IS AND RYZHIK IM. 2007. Table of Integrals, Series, and Products. Academic Press, New York.

GUI W. 2013. Marshall-Olkin extended log-logistic distribution and its application in minification processes. Appl Math Sci 7: 3947-3961.

GUPTA RC, GUPTA PL AND GUPTA RD. 1998. Modeling failure time data by Lehman alternatives. Commun Stat Theor Meth 27: 887-904.

GUPTA RD, AHMAN O AND LVIN S. 1999. A study of log-logistic model in survival analysis. Biometrical J 41: 431-443.

JOHNSON N, KOTZ S AND BALAKRISHNAN N. 1994. Continuous Univariate Distributions -Volume 1, 2nd ed., New York: J Wiley, p. 284-285.

KENNEY JF AND KEEPING ES. 1962. Mathematics of Statistics, Part 1, 3rd ed., Princeton, NJ, p. 101-102.

KLEIBER C AND KOTZ S. 2003. Statistical Size Distributions in Economics and Actuarial Sciences. J Wiley: New York.

LAWLESS JF. 2003. Statistical Models and Methods for Lifetime Data. J Wiley: New York.

LEMOENTE AL. 2015. The beta log-logistic distribution. Braz J Probab Stat 29: 172-173.

MOORS JJA. 1998. A quantile alternative for kurtosis. J Roy Stat Soc D 37: 25-32.

MUDHOLKAR GS AND SRIVASTAVA DK. 1993. Exponentiated Weibull family for analyzing bathtub failure-rate data. IEEE T Rehabil Eng 42: 299-302.

MUDHOLKAR GS, SRIVASTAVA DK AND FREIMER M. 1995. The exponentiated Weibull family: A reanalysis of the bus-motor-failure data. Technometrics 37: 436-445.

NADARAJAH S. 2005. The exponentiated Gumbel distribution with climate application. Environmetrics 17: 13-23.

NADARAJAH S AND GUPTA AK. 2007. The exponentiated gamma distribution with application to drought data. Calcutta Statist Assoc 59: 29-54.

NICHOLS MD AND PADGETT WJ. 2006. A Bootstrap control chart for Weibull percentiles. Qual Reliab Eng Int 22: 141-151.

ROSAIAH K, KANTAM RRL AND KUMAR CHS. 2006. Reliability test plans for exponentiated log-logistic distribution. Econ Qual Control 21: 279-289.

SHOUKRI MM, MIAN IUH AND TRACY DS. 1988. Sampling properties of estimators of the log-logistic distribution with application to canadian precipitation data. Can J Stat 16: 223-236.

TADIKAMALLA PR. 1980. A look at the Burr and related distributions. Int Stat Rev 48: 337-344.

TAHIR MH, MANSOOR M, ZUBAIR M AND HAMEDANI GG. 2014. McDonald Log-Logistic distribution. J Stat Theory Appl 13: 65-82.

VERHULST PF. 1838. Notice sur la loi que la population suit dans son accroissement. Corresp Math Phys 10: 113-121.