Lindley–Exponential distribution: properties and applications

Deepesh Bhati · Mohd. Aamir Malik · H. J. Vaman

Abstract In this paper, we introduce a new class of distributions generated by an integral transform of the probability density function of the Lindley distribution which results in a model that is more flexible in the sense that the derived model spans distributions with increasing failure rate, decreasing failure rate and upside down bathtub shaped hazard rate functions for different choices of parametric values. For this new model, various distributional properties including limiting distribution of extreme order statistics are established. Maximum likelihood estimators and the marginal confidence intervals of the parameters are obtained. The applicability of the proposed distribution is shown through application to real data sets. Through application to two real datasets, it is demonstrated that the proposed model fits better as compared to some other competing models. Further, the model is shown to be useful for analysing stress–strength model.

1 Introduction

Lifetime distribution represents an attempt to describe, statistically, the length of the life of a system, a device, and in general, time-to-event data. Lifetime distributions are frequently used in fields like medicine, biology, engineering, insurance etc. Many parametric models

D. Bhati (✉) · M. A. Malik · H. J. Vaman
Department of Statistics, Central University of Rajasthan, Kishangarh, India
e-mail: deepesh.bhati@curaj.ac.in

M. A. Malik
e-mail: aamirmalik.stats@gmail.com

H. J. Vaman
e-mail: vamanjalikop@curaj.ac.in

such as exponential, gamma, weibull have been frequently used in statistical literature to analyse lifetime data.

Recently, one parameter Lindley distribution has attracted researchers for its potential in modelling lifetime data, and it has been observed that this distribution has performed excellently in many applications. The Lindley distribution was originally proposed by Lindley [19] in the context of Bayesian statistics, as a counter example to fiducial statistics. The distributon can also be derived as a mixture of $\exp(\theta)$ and $\text{gamma}(2, \theta)$. More details on the Lindley distribution can be found in Ghitany et al. [7].

A random variable X is said to have the Lindley distribution with parameter θ if its probability density is defined as:

$$f_X(x; \theta) = \frac{\theta^2}{(\theta + 1)}(1 + x)e^{-\theta x}; x > 0, \theta > 0. \tag{1}$$

The corresponding cumulative distribution function is

$$F(x) = 1 - \frac{e^{-\theta x}(1 + \theta + \theta x)}{1 + \theta}; x \in \mathbb{R}^+, \theta > 0,$$

Ghitany et al. [6,7] have introduced a two-parameter weighted Lindley distribution and have pointed out its usefulness, in particular, in modelling biological data from mortality studies. Bakouch et al. [4] have come up with extended Lindley (EL) distribution, Adamidis and Loukas [1] have introduced a new lifetime distribution with decreasing failure rate. Shanker et al. [23] have introduced a two-parameter Lindley distribution. Zakerzadeh and Mahmoudi [24] have proposed a new two parameter lifetime distribution: model and properties. Hassan [11] has introduced convolution of Lindley distribution. Ghitany et al. [8] worked on the estimation of the reliability of a stress-strength system from power lindley distribution. Elbatal et al. [5] has proposed a new generalized Lindley distribution by considering the mixture of two gamma distribution.

Ristić and Balakrishnan [22] have introduced a new family of distributions generated by gamma pdf with survival function given by

$$\bar{F}_X(x) = \frac{1}{\Gamma(\alpha)} \int_0^{-\log(G(x))} t^{\alpha-1}e^{-t}dt, \qquad x \in \mathbb{R}^+, \alpha > 0,$$

where $G(x)$ is a chosen distribution function, called the parent distribution function (df).

In this paper, we introduce a new family of distribution generated by an integral transform of the pdf of a random variable T which follows one parameter Lindley distribution. The survival function of this new family is given as:

$$\bar{F}_X(x) = \frac{\theta^2}{1 + \theta} \int_0^{-\log(G(x))} (1 + t)e^{-\theta t}dt, \qquad x \in \mathbb{R}^+, \theta > 0 \tag{2}$$

where $\theta > 0$,the corresponding probability density function (pdf) is given by

$$f(x) = \frac{\theta^2}{1 + \theta}(1 - \log G(x)) G(x)^{\theta-1}g(x), \tag{3}$$

In this formation, we consider $G(x)$ corresponding to exponential distribution with cdf $(1 - e^{-\lambda x})$ which yields the survival function of the new distribution is as

$$\bar{F}(x) = 1 - \frac{\left(1 - e^{-\lambda x}\right)^{\theta} \left(1 + \theta - \theta \log\left(1 - e^{-\lambda x}\right)\right)}{1 + \theta}, \tag{4}$$

with corresponding density

$$f(x) = \frac{\theta^2 \lambda e^{-\lambda x} \left(1 - e^{-\lambda x}\right)^{\theta-1} \left(1 - \log\left(1 - e^{-\lambda x}\right)\right)}{(1 + \theta)}. \tag{5}$$

In this paper, we refer to random variable with survival function (4) as Lindley–Exponential (L–E) distribution with parameters θ and λ and denote it by L–E(θ, λ). Motivation for this paper arose from the following observation.

In modelling time-to-event data arising from a new application, the monotonicity nature of the hazard rate function is uncertain and hence an a priori choice of a model becomes challenging. In such cases, one usually resorts to nonparametric methods in general. However, if a parametric model fits the particular application well, it is preferable to choose one such model. There is thus a need to develop a distributional model which would span various monotonicity properties of the hazard rate function, and this paper is an attempt to develop a new model with two parameters model ensuring flexibility of hazard function to posses increasing, decreasing and upside down shapes. Having developed one such distribution, we observe the following properties that connect the new model to known distributions.

1. Let U be a random variable with the Log-lindley distribution $LL(\theta, 1)$ proposed by Goméz et al. [10]. Then the random variable $V = G^{-1}(U)$ has distribution with pdf as in (3).
2. If U be Lindley distributed random variable, then the random variable $X = G^{-1}(e^{-U})$ has pdf (3).

We now study properties of the L–E distribution and illustrate its applicability. The contents of the proposed work are organized as follows. Various distributional properties like shape of the pdf, quantile function, moment generating function, limiting distribution of sample statistics like maximum and minimum and entropy of L–E distribution are present in Sect. 2. The maximum likelihood estimation is presented in Sect. 3. Performance of the maximum likelihood estimators for small samples is assessed by simulation in Sect. 4. Applicability of the new distribution is shown in Sect. 5. Section 6, gives estimation of stress–strength parameter (R) by using maximum likelihood estimation method.

2 Distributional properties of L–E distribution

2.1 Shape of the density

Theorem 1 *The probability density function of the L–E distribution is decreasing for $0 < \theta < 1$ and uni-modal for $\theta > 1$. In the latter case, the mode is the root of the following equation:*

$$(1 - \theta e^{-\lambda x})\left(1 - \log(1 - e^{-\lambda x})\right) + e^{-\lambda x} = 0$$

Proof The first order derivative of $\log(f(x))$ is

$$\frac{d \log(f(x))}{dx} = \frac{\lambda\, r(x)}{\left(1 - e^{-\lambda x}\right)\left(1 - \log\left(1 - e^{-\lambda x}\right)\right)}. \tag{6}$$

where, $r(x) = \left(\theta e^{-\lambda x} + \left(1 - \theta e^{-\lambda x}\right) \log\left(1 - e^{-\lambda x}\right) - e^{-\lambda x} - 1\right)$. For $0 < \theta < 1$, the function $r(x)$ is negative. So $f'(x) < 0$ for all $x > 0$. This implies that f is decreasing for
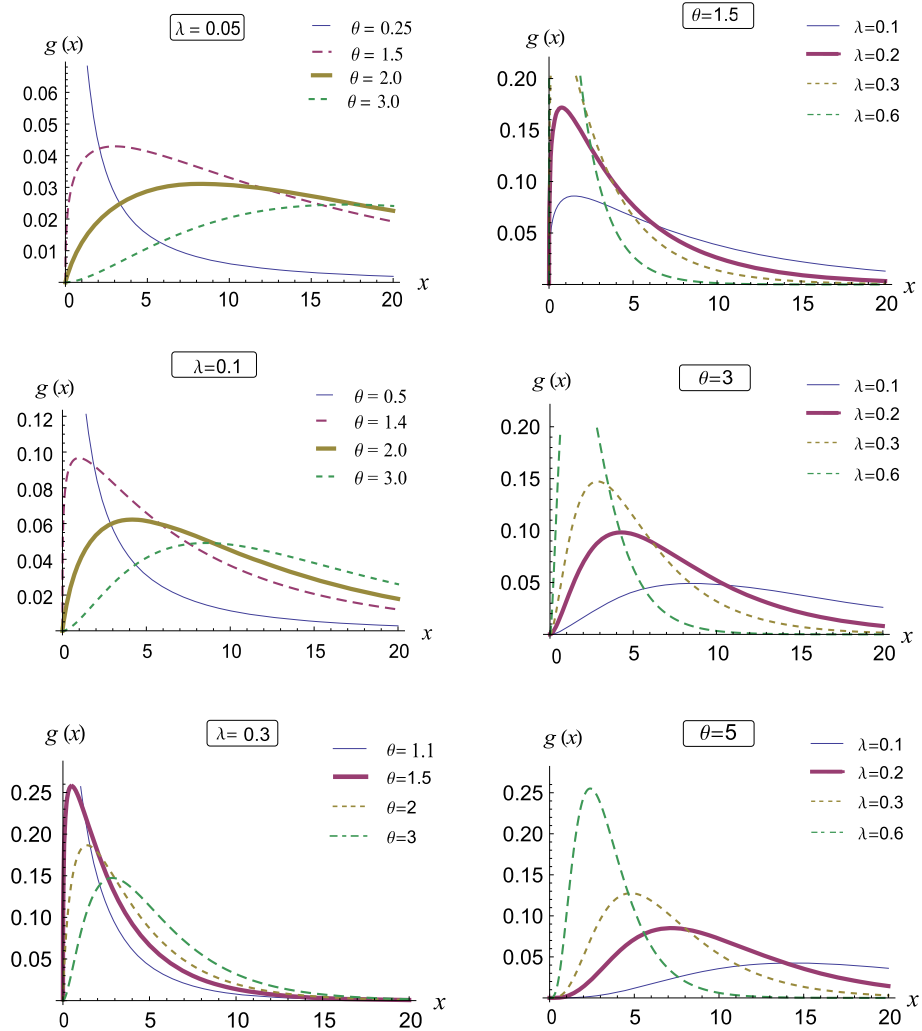
Fig. 1 PDF plot for various values of λ and θ

$0 < \theta < 1$. Also note that, $(\log f)'(0) = \infty$ and $(\log f)'(\infty) < 0$. This implies that for $\theta > 1$, $g(x)$ has a unique mode at x_0 such that $r(x) > 0$ for $x < x_0$ and $r(x) < 0$ for $x > x_0$. So, g is uni-modal function with mode at $x = x_0$. The pdf for various values of λ and θ are plotted in Fig. 1. □

Theorem 2 *The hazard function of L–E distribution is decreasing, increasing or upside down according as $0 < \theta \leq 1$, $\theta > 2$ and $1 < \theta < 2$ respectively.*

Proof Considering the hazard rate function (hrf) of the L–E distribution given by

$$h(x) = \frac{\theta^2 \lambda e^{\lambda - x} \left(1 - e^{-\lambda x}\right)^{\theta - 1} \left(1 - \log\left(1 - e^{-\lambda x}\right)\right)}{1 + \theta - \left(1 - e^{-\lambda x}\right)^{\theta} \left(1 + \theta - \theta \log\left(1 - e^{-\lambda x}\right)\right)}, \tag{7}$$
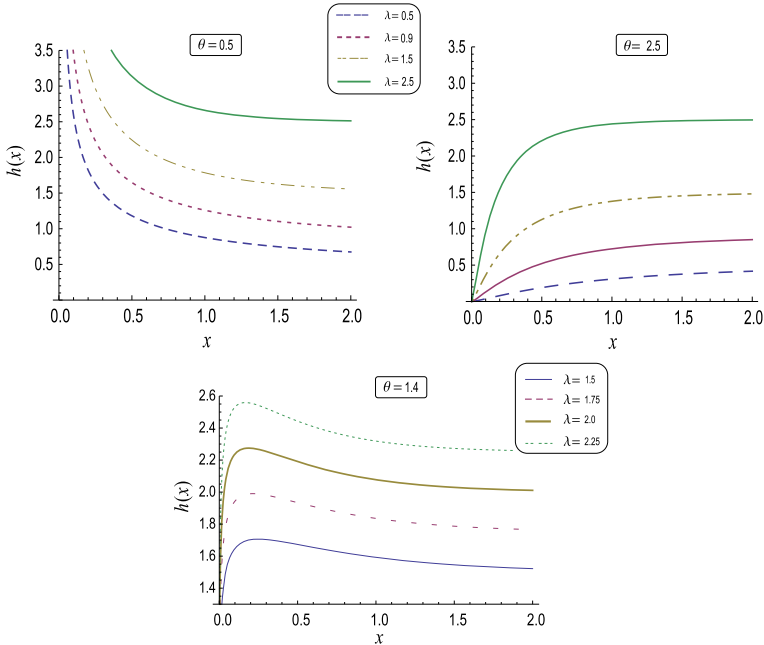
Fig. 2 Hazard function plot for various values of λ and θ

and using theorems of Glaser [9], we can discuss the shape characteristics of the hrf of L–E distribution. The function $\eta(x) = -f'(x)/f(x)$ for L–E distribution is given by

$$\eta(x) = \lambda \left(\frac{1 - \theta e^{-\lambda x}}{1 - e^{-\lambda x}} + \frac{e^{-\lambda x}}{(1 - e^{-\lambda x}) \log(1 - e^{-\lambda x})} \right)$$

and

$$\eta'(x) = \frac{\lambda^2 e^{-\lambda x} \left((\theta - 2) + (\theta - 1) \text{Log}^2 (1 - e^{-\lambda x}) - (2\theta - 3) \log(1 - e^{-\lambda x}) + e^{-\lambda x} \right)}{(1 - e^{-\lambda x})^2 (1 - \log(1 - e^{-\lambda x}))^2}.$$

If $0 < \theta \leq 1$, then $\eta'(x) < 0 \ \forall \ x > 0$. It follows from Theorem (b) of Glaser [9] that the failure rate is decreasing. Whereas, for $\theta > 2$, $\eta'(x) > 0 \forall x > 0$ which indicate increasing failure rate (see Theorem (b) of Glaser [9]. Furthermore, we note that the failure rate can be upside-down bathtub for $1 < \theta < 2$. Moreover, for $\theta > 0$ the hazard rate function follows relation $\lim_{x \to \infty} h(x) = \lambda$. These different failure rate for different values of parameters θ and λ are shown in Fig. 2. □

2.2 The quantile function of L–E distribution

The cdf, $F_X(x) = 1 - \bar{F}(x)$, can be obtained by using Eq. (4). Further, it can be noted that F_X is continuous and strictly increasing so that the quantile function of X is $Q_X(\gamma) = F_X^{-1}(\gamma)$, $0 < \gamma < 1$. In the following theorem, we give an explicit expression for Q_X in terms of the Lambert W function. For more details on Lambert W function we refer the reader to Jodŕa [12] and also to Nair et al. [20] for discussion on quantile functions.

Theorem 3 *For any $\theta, \lambda > 0$, the quantile function of the L–E distribution X is*

$$x_\gamma = F^{-1}(\gamma) = -\frac{1}{\lambda} \log \left(1 - \left(-\frac{W_{-1} \left(-e^{-\theta-1}(\theta+1)\gamma \right)}{(\theta+1)\gamma} \right)^{-1/\theta} \right), \qquad (8)$$

where W_{-1} denotes the negative branch of the Lambert W function.

Proof By assume $p = 1 - e^{-\lambda x}$ the cdf can be written as

$$F_X(x) = \frac{p^\theta(1 + \theta - \theta \log(p))}{1 + \theta}$$

for fixed $\theta, \lambda > 0$ and $\gamma \in (0, 1)$, the γth quantile function is obtained by solving $F_X(x) = \gamma$. By re-arranging the above, we obtain

$$\log(p^{-\theta}) - \gamma p^{-\theta}(1 + \theta) = -(1 + \theta)$$

it can be further written as

$$-p^{-\theta}(1+\theta)\gamma e^{-p^{-\theta}(1+\theta)\gamma} = -(1+\theta)\gamma e^{-(1+\theta)}. \qquad (9)$$

We see that $-p^{-\theta}(1 + \theta)\gamma$ is the Lambert-W function of real argument $-(1+\theta)\gamma e^{-(1+\theta)}$.
    Thus, we have

$$W \left(-(1+\theta)\gamma e^{-(1+\theta)} \right) = -p^{-\theta}(1+\theta)\gamma \qquad (10)$$

Moreover, for any θ and $\lambda > 0$ it is immediate that $p^{-\theta}(1 + \theta)\gamma > 1$, and it can also be checked that $(1+\theta)\gamma e^{-(1+\theta)} \in (-1/e, 0)$ since $\gamma \in (0, 1)$. Therefore, by taking into account the properties of the negative branch of the Lambert W function, we deduce the following.

$$W_{-1} \left(-(1+\theta)\gamma e^{-(1+\theta)} \right) = -p^{-\theta}(1+\theta)\gamma$$

Again, solving for x by using $p = 1 - e^{-\lambda x}$, we get

$$x_\gamma = F^{-1}(\gamma) = -\frac{1}{\lambda} \log \left(1 - \left(-\frac{W_{-1} \left(-e^{-\theta-1}(\theta+1)\gamma \right)}{(\theta+1)\gamma} \right)^{-1/\theta} \right). \qquad (11)$$

$\square$

Further the first three quantiles can be obtained by substituting $\gamma = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ in (11).

$$Q_1 = F^{-1}(1/4) = -\frac{1}{\lambda} \log \left(1 - \left(-\frac{4W_{-1} \left(\frac{-e^{-\theta-1}(1+\theta)}{4} \right)}{(1+\theta)} \right)^{-\frac{1}{\theta}} \right)$$

$$\text{Median}(M_d) = Q_2 = F^{-1}(1/2) = -\frac{1}{\lambda} \log \left(1 - \left(-\frac{2W_{-1} \left(\frac{-e^{-\theta-1}(1+\theta)}{2} \right)}{(1+\theta)} \right)^{-\frac{1}{\theta}} \right)$$

$$Q_3 = F^{-1}(3/4) = -\frac{1}{\lambda} \log \left(1 - \left(-\frac{4W_{-1} \left(\frac{-3e^{-\theta-1}(1+\theta)}{4} \right)}{3(1+\theta)} \right)^{-\frac{1}{\theta}} \right) \qquad (12)$$

2.3 Moments

The Moment Generating function of L–E random variable is given as

$$
\mathbb{M}_X(t) = \mathbb{E}(e^{Xt}) = \int_0^\infty \frac{\theta^2 \lambda e^{-x(\lambda - t)} \left(1 - e^{-\lambda x}\right)^{\theta - 1} \left(1 - \log\left(1 - e^{-\lambda x}\right)\right)}{\theta + 1} \, dx
$$
$$
= \frac{\theta \Gamma(\theta + 1)\Gamma\left(1 - \frac{t}{\lambda}\right)\left(-\psi^{(0)}(\theta) + \psi^{(0)}\left(-\frac{t}{\lambda} + \theta + 1\right) + 1\right)}{(\theta + 1)\Gamma\left(-\frac{t}{\lambda} + \theta + 1\right)} \quad (13)
$$

where, $\psi^{(n)}(z) = \frac{d^n \psi(z)}{dz^n}$ and $\psi(z) = \frac{\Gamma'(z)}{\Gamma(z)}$ known as digamma function.

Hence the first and second raw moments can be obtained by differentiating the MGF $\left(\frac{d\mathbb{M}_X(t)}{dt}\right)_{t=0}$ and $\left(\frac{d^2\mathbb{M}_X(t)}{dt^2}\right)_{t=0}$ respectively.

$$
\mathbb{E}(X) = \frac{\gamma\theta + \frac{1}{\theta} + \theta(-\psi^{(1)}(\theta + 1)) + (\theta + 1)\psi^{(0)}(\theta) + \gamma + 1}{(\theta + 1)\lambda} \quad (14)
$$

$$
\mathbb{E}(X^2) = \frac{1}{\theta^2(\theta + 1)\lambda^2} \left((\theta + 1)\left(\left(\gamma^2 + \frac{\pi^2}{6}\right)\theta^2 + 2\gamma\theta + 1\right)\right.
$$
$$
+ \theta(\theta + 1)\psi^{(0)}(\theta)(2\gamma\theta + \theta\psi^{(0)}(\theta) + 2)
$$
$$
\left. - \theta^2(\theta + 2\theta(\psi^{(0)}(\theta) + \gamma + 3)\psi^{(1)}(\theta + 1) + \theta^3\psi^{(2)}(\theta + 1)\right) \quad (15)
$$

where γ is Eulergamma constant $= 0.577216$.

Table 1 displays the mode, mean and median for L–E distribution for different choices of parameter λ and θ. It can be noted from the table that all three measures of central tendency decrease as λ increases, and the measures increase as θ increase. Also for any choice of λ and θ Mean > Median > Mode which is an indication of positive skewness.

2.4 Limiting distribution of sample minimum and maximum

We can derive the asymptotic distribution of the sample minima $X_{1:n}$ by using Theorem 8.3.6 of Arnold et al. [3], it follows that the asymptotic distribution of $X_{1:n}$ is Weibull type with shape parameter $\theta > 0$ if

$$
\lim_{t \to 0^+} \frac{F(tx)}{F(t)} = x^\theta, \quad \forall x > 0
$$

From (4), $F(X) = 1 - \bar{F}(X)$, using the identities

$$
\lim_{t \to 0^+} \frac{(1 - e^{-\lambda tx})^{\theta - 1}}{(1 - e^{-\lambda t})^{\theta - 1}} = x^{\theta - 1}
$$

and

$$
\lim_{t \to 0^+} \frac{(1 - \log(1 - e^{-\lambda tx}))}{(1 - \log(1 - e^{-\lambda t}))} = 1
$$

Thus, we get

$$
\lim_{t \to 0^+} \frac{F(tx)}{F(t)} = x \lim_{t \to 0^+} \frac{f(tx)}{f(t)} = x \lim_{t \to 0^+} \frac{(1 - e^{-\lambda tx})^{\theta - 1}(1 - \log(1 - e^{-\lambda tx}))}{(1 - e^{-\lambda t})^{\theta - 1}(1 - \log(1 - e^{-\lambda t}))} = x^\theta
$$

Table 1 Mode, mean and median for various value of parameter

$\theta \downarrow$	$\lambda \rightarrow$	0.1	0.5	1	1.5	2	2.5
1.1	Mode	0.001219	0.000244	0.000122	0.000081	0.000061	0.000049
	Mean	7.446760	1.489350	0.744676	0.496451	0.372330	0.297871
	Median	4.478034	0.895607	0.447803	0.298536	0.223902	0.179121
1.5	Mode	1.508590	0.301719	0.150850	0.100573	0.075429	0.060344
	Mean	9.861580	1.972300	0.986158	0.657438	0.493079	0.394463
	Median	6.920488	1.384098	0.692048	0.461366	0.346024	0.276820
2	Mode	4.174510	0.834902	0.417451	0.278300	0.208725	0.166980
	Mean	12.367100	2.473421	1.236710	0.824470	0.618350	0.494680
	Median	9.528006	1.905601	0.952801	0.635200	0.476400	0.381120
2.5	Mode	6.540190	1.308040	0.654019	0.436013	0.327010	0.261608
	Mean	14.444000	2.888800	1.444400	0.962930	0.722200	0.577761
	Median	11.704978	2.340996	1.170498	0.780332	0.585249	0.468199
3	Mode	8.569080	1.713820	0.856908	0.571272	0.428454	0.342763
	Mean	16.204700	3.240930	1.620470	1.080310	0.810233	0.648186
	Median	13.549240	2.709848	1.354924	0.903283	0.677462	0.541970
3.5	Mode	10.317200	2.063400	1.031720	0.687811	0.515858	0.412687
	Mean	17.726300	3.545200	1.772630	1.181760	0.886310	0.709053
	Median	15.138649	3.027730	1.513865	1.009243	0.756932	0.605546
4	Mode	11.840500	2.368100	1.184050	0.789366	0.592025	0.473620
	Mean	19.062700	3.812500	1.906270	1.270850	0.953130	0.762510
	Median	16.529903	3.305981	1.652990	1.101994	0.826495	0.661196

Similarly, using the following identity

$$\lim_{t\to\infty} \frac{\left(1 - e^{-\lambda(t+x)}\right)^{\theta-1}}{\left(1 - e^{-\lambda t}\right)^{\theta-1}} = e^{-\lambda x(\theta-1)}$$

and

$$\lim_{t\to\infty} \frac{1 - \log\left(1 - e^{-\lambda(t+x)}\right)}{1 - \log\left(1 - e^{-\lambda t}\right)} = 1$$

It can be seen that

$$\lim_{t\to\infty} \frac{1 - F(t+x)}{1 - F(t)} = \lim_{t\to\infty} \frac{(1+\theta) - (1 - e^{-\lambda(x+t)})^{\theta}(1+\theta - \theta\log(1 - e^{-\lambda(x+t)}))}{(1+\theta) - (1 - e^{-\lambda t})^{\theta}(1+\theta - \theta\log(1 - e^{-\lambda t}))}$$

by using L-Hópitals rule,

$$\lim_{t\to\infty} \frac{e^{-\lambda x}\left(1 - e^{-\lambda(t+x)}\right)^{\theta-1}\left(1 - \log\left(1 - e^{-\lambda(t+x)}\right)\right)}{\left(1 - e^{-\lambda t}\right)^{\theta-1}\left(1 - \log\left(1 - e^{-\lambda t}\right)\right)} = e^{-\lambda\theta x}$$

Hence, it follows from Theorem 1.6.2 in Leadbetter et al. [17] that there must be norming constants $a_n > 0$, b_n, $c_n > 0$ and d_n such that

$$\Pr\{a_n(M_n - b_n) \le x\} \to e^{-e^{-\theta\lambda x}} \tag{16}$$

and

$$\Pr\{c_n(m_n - d_n) \leq x\} \to 1 - e^{-x^\theta} \tag{17}$$

as $n \to \infty$. By following Corollary 1.6.3 in Leadbetter et al. [17], we can determine the form of the norming constants. As an illustration, one can see that $a_n = \theta$ and $b_n = F^{-1}(1 - 1/n)$, where $F^{-1}(.)$ denotes the inverse function of $F(.)$.

2.5 Entropy

In many fields of science such as communication, physics and probability, entropy is an important concept to measure the amount of uncertainty associated with a random variable X. Several entropy measures and information indices are available but among them the most popular entropy called Rényi entropy defined as

$$\mathfrak{J}(\zeta) = \frac{1}{1 - \zeta} \log \left(\int_{\mathbb{R}^+} f^\zeta(x) dx \right), \quad \text{for} \quad \zeta > 1 \quad \text{and} \quad \zeta \neq 1 \tag{18}$$

In our case

$$\int_{\mathbb{R}^+} f^\zeta dx = \int_{\mathbb{R}^+} \left(\frac{\theta^2 \lambda}{\theta + 1} \right)^\zeta e^{-x\lambda\zeta} \left(1 - e^{-x\lambda} \right)^{\zeta(\theta-1)} \left(1 - \log \left(1 - e^{-x\lambda} \right) \right)^\zeta dx$$

substituting $x = -\frac{1}{\lambda} \log(1 - e^{-u})$ and using power series expansion $(1 - z)^\alpha = \sum_{j=0}^\infty (-1)^j \binom{\alpha}{j} z^j$, the above expression reduces to

$$\int_{\mathbb{R}^+} f^\zeta dx = \frac{\theta^{2\zeta} \lambda^{\zeta-1}}{(\theta + 1)^\zeta} \int_0^\infty (u + 1)^\zeta \left(1 - e^{-u} \right)^{\zeta-1} e^{u(-(\zeta(\theta-1)+1))} du$$

$$= \frac{\theta^{2\zeta} \lambda^{\zeta-1}}{(\theta + 1)^\zeta} \int_0^\infty \sum_{j=0}^\infty (-1)^j \binom{\zeta - 1}{j} e^{u(-(\zeta(\theta-1)+j+1))} (u + 1)^\zeta du$$

$$= \frac{\theta^{2\zeta} \lambda^{\zeta-1}}{(\theta + 1)^\zeta} \sum_{j=0}^\infty (-1)^j \binom{\zeta - 1}{j} \int_0^\infty (u + 1)^\zeta e^{u(-(\zeta(\theta-1)+j+1))} du$$

$$= \frac{\theta^{2\zeta} \lambda^{\zeta-1}}{(\theta + 1)^\zeta} \sum_{j=0}^\infty (-1)^j e^{\zeta(\theta-1)+j+1} E_{-\zeta} [(j + (\theta - 1)\zeta + 1)] \tag{19}$$

where $E_n(z) = \int_1^\infty e^{-zt} t^{-n} dt$ known as exponential integral function. For more details see http://functions.wolfram.com/06.34.02.0001.01.

Thus according to (18) the Rényi entropy of L–E(θ, λ) distribution given as

$$\mathfrak{J}(\zeta) = \frac{1}{1 - \zeta} \log \left(\frac{\theta^{2\zeta} \lambda^{\zeta-1}}{(\theta + 1)^\zeta} \sum_{j=0}^\infty (-1)^j e^{\zeta(\theta-1)+j+1} E_{-\zeta} [(j + (\theta - 1)\zeta + 1)] \right). \tag{20}$$

Moreover, The Shannon entropy defined by $E[\log(f(x))]$ is a special case derived from $\lim_{\zeta \to 1} \mathfrak{J}(\zeta)$

3 Maximum likelihood estimators

In this section we shall discuss the point and interval estimation of the parameters of L–E(θ, λ). The log-likelihood function $l(\Theta)$ of single observation (say x_i) for the vector of parameter $\Theta = (\theta, \lambda)^\top$ is

$$l_n(x, \Theta) = 2\log(\theta) + \log(\lambda) - \log(\theta + 1) + (\theta - 1)\log(1 - e^{-\lambda x})$$
$$+ \log(1 - \log(1 - e^{-\lambda x})) - \lambda x$$

The associated score function is given by $U_n = \left(\frac{\partial l_n}{\partial \theta}, \frac{\partial l_n}{\partial \lambda} \right)^\top$, where

$$\frac{\partial l_n}{\partial \theta} = \frac{2}{\theta} - \frac{1}{\theta + 1} + \log\left(1 - e^{-\lambda x}\right) \tag{21}$$

$$\frac{\partial l_n}{\partial \lambda} = \frac{1}{\lambda} - x + (\theta - 1)\frac{x e^{-\lambda x}}{1 - e^{-\lambda x}} + \frac{x e^{-\lambda x}}{\left(1 - e^{-\lambda x}\right)\left(1 - \log\left(1 - e^{-\lambda x}\right)\right)}. \tag{22}$$

As we know th expected value of score function equals zero, i.e. $\mathbb{E}(U(\Theta))$, which implies $\mathbb{E}\left(\log\left(1 - e^{-\lambda x}\right)\right) = \frac{1}{\theta+1} - \frac{2}{\theta}$.

The total log-likelihood of the random sample $x = (x_1, \ldots, x_n)^\top$ of size n from X is given by $l_n = \sum_1^n l^{(i)}$ and th total score function is given by $U_n = \sum_{i=1}^n U^{(i)}$, where $l^{(i)}$ is the log-likelihood of ith observation and $U^{(i)}$ as given above. The maximum likelihood estimator $\hat{\Theta}$ of Θ is obtained by solving Eqs. (21) and (22) numerically or this can also obtained easily by using nlm() function in R. The initial guess for the estimators were obtained from the inner region of 3D contour plot of log-likelihood function for a given sample. For example, in Fig. 3, the contour plot of log-likelihood function for different θ and λ, the initial estimates were taken from interior. The associated Fisher information matrix is given by

$$K = K_n(\Theta) = n \begin{bmatrix} \kappa_{\theta,\theta} & \kappa_{\theta,\lambda} \\ \kappa_{\lambda,\theta} & \kappa_{\lambda,\lambda} \end{bmatrix} \tag{23}$$
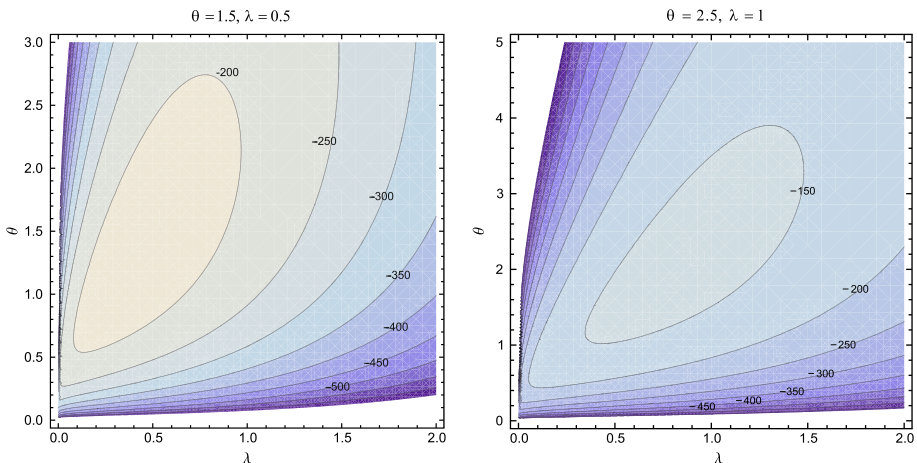

Fig. 3 Contour plot of log-likelihood surface for different values of θ and λ

where

$$\kappa_{\theta,\theta} = \frac{n}{(\theta+1)^2} - \frac{2n}{\theta^2}$$

$$\kappa_{\lambda,\theta} = \kappa_{\theta,\lambda} = \mathbb{E}\left(\frac{Xe^{-\lambda X}}{1-e^{-\lambda X}}\right)$$

$$\kappa_{\lambda,\lambda} = \frac{1}{\lambda^2} + \mathbb{E}\left(\frac{(\theta-1)\left(X^2 e^{-\lambda X}\right)}{\left(1-e^{-\lambda X}\right)^2}\right) - \mathbb{E}\left(\frac{X^2 e^{-\lambda X}\left(-e^{-\lambda X} - \log\left(1-e^{-\lambda X}\right)+1\right)}{\left(1-e^{-\lambda X}\right)^2\left(1-\log\left(1-e^{-\lambda X}\right)\right)^2}\right) \tag{24}$$

The above expressions depend on some expectations which can be easily computed using numerical integration. Under the usual regularity conditions, the asymptotic distribution of

$$\sqrt{n}\left(\hat{\Theta} - \Theta\right) \text{ is } N_2(0, K(\Theta)^{-1}) \tag{25}$$

where $\lim_{n\to\infty} K_n(\Theta)^{-1} = K(\Theta)^{-1}$. The asymptotic multivariate normal $N_2(0, K(\Theta)^{-1})$ distribution of $\hat{\Theta}$ can be used to construct approximate confidence intervals. An asymptotic confidence interval with significance level α for each parameter θ and λ is

$$\text{ACI}(\theta, 100(1-\alpha)\%) \equiv (\hat{\theta} - z_{\alpha/2}\sqrt{\kappa_{(\theta,\theta)}}, \hat{\theta} + z_{\alpha/2}\sqrt{\kappa_{(\theta,\theta)}})$$

$$\text{ACI}(\lambda, 100(1-\alpha)\%) \equiv (\hat{\lambda} - z_{\alpha/2}\sqrt{\kappa_{(\lambda,\lambda)}}, \hat{\lambda} + z_{\alpha/2}\sqrt{\kappa_{(\lambda,\lambda)}}) \tag{26}$$

where $z_{1-\alpha/2}$ denotes $1 - \alpha/2$ quantile of standard normal random variable.

4 Simulation

In this section, we investigate the behaviour of the ML estimators in finite sample sizes through a simulation study based on different L–E(θ, λ). The observations are generated using cdf technique presented in Sect. 4 from L–E(θ, λ) and in the following Sects. 4.1 and 4.2, we investigate the performance of maximum likelihood estimators $(\hat{\theta}, \hat{\lambda})$ for various combinations of parameters (θ, λ) and also with respect to sample size n respectively.

4.1 Performance of estimators for different parametric values

A simulation study consisting of following steps is carried out for each triplet (θ, λ, n), where $\theta = 0.5, 1, 2, \lambda = 0.5, 1, 2, 3$ and $n = 25, 50, 75, 100$.

1. Choose the values $\theta_\circ, \lambda_\circ$ for the corresponding elements of the parameter vector $\Theta = (\theta, \lambda)$, to specify L–E$(\theta, \lambda)$ distribution;
2. choose sample size n;
3. generate N independent samples of size n from L–E(θ, λ);
4. compute the ML estimate $\hat{\Theta}_n$ of $\Theta_\circ$ for each of the N samples;
5. compute the average bias $= \frac{1}{N}\sum_{i=1}^{N}(\Theta_i - \Theta_\circ)$, the average mean square error *MSE* $(\Theta) = \frac{1}{N}\sum_{i=1}^{N}(\Theta_i - \Theta_\circ)^2$ and average width (AW) of 95% of confidence limit of the obtained estimates over all N samples.

In all three Tables, Tables 2, 3 and 4, it can be clear seen that the value of average bias, MSE and average width decreases with increase in sample size n.

Table 2 Average bias of the simulated estimates

	n	$\lambda = 0.5$		$\lambda = 1$		$\lambda = 2$		$\lambda = 3$	
		bais (θ)	bais (λ)	bais (θ)	bais (λ)	bais (θ)	bais (λ)	bais (θ)	bais (λ)
$\theta = 0.5$	25	0.0448	0.1836	0.0476	0.4173	0.0519	0.8777	0.0534	1.1682
	50	0.0187	0.0673	0.0196	0.1596	0.0164	0.2694	0.0148	0.3427
	75	0.0092	0.0403	0.0131	0.0792	0.0136	0.2166	0.0120	0.2608
	100	0.0067	0.0267	0.0091	0.0728	0.0074	0.1318	0.0095	0.1988
$\theta = 1$	25	0.1200	0.1012	0.1073	0.1873	0.1158	0.3884	0.1085	0.5350
	50	0.0478	0.0405	0.0412	0.0655	0.0507	0.1723	0.0522	0.2444
	75	0.0364	0.0268	0.0265	0.0370	0.0286	0.1091	0.0285	0.1415
	100	0.0150	0.0136	0.0198	0.0315	0.0183	0.0548	0.0244	0.1266
$\theta = 2$	25	0.3599	0.0650	0.3270	0.1268	0.3628	0.2334	0.3401	0.3451
	50	0.1003	0.0204	0.1100	0.0460	0.1200	0.1021	0.1209	0.1324
	75	0.0686	0.0136	0.0654	0.0312	0.0803	0.0653	0.0955	0.1173
	100	0.0562	0.0136	0.0457	0.0234	0.0511	0.0417	0.0588	0.0922

Table 3 Average MSE of the simulated estimates

|	n	$\lambda = 0.5$		$\lambda = 1$		$\lambda = 2$		$\lambda = 3$	
		MSE (θ)	MSE (λ)	MSE (θ)	MSE (λ)	MSE (θ)	MSE (λ)	MSE (θ)	MSE (λ)
$\theta = 0.5$	25	0.0164	0.1988	0.0177	1.0549	0.0204	5.3140	0.0200	8.3846
	50	0.0051	0.0459	0.0055	0.2089	0.0053	0.7757	0.0051	1.5790
	75	0.0032	0.0270	0.0031	0.1047	0.0033	0.4656	0.0032	0.9631
	100	0.0024	0.0161	0.0022	0.0763	0.0022	0.2870	0.0024	0.7081
$\theta = 1$	25	0.1313	0.0802	0.0998	0.2457	0.1012	1.1482	0.1067	2.2693
	50	0.0314	0.0200	0.0283	0.0716	0.0302	0.3081	0.0316	0.7259
	75	0.0195	0.0119	0.0163	0.0431	0.0178	0.1875	0.0171	0.3556
	100	0.0112	0.0078	0.0128	0.0331	0.0111	0.1131	0.0125	0.3082
$\theta = 2$	25	0.8890	0.0344	0.7999	0.1393	0.9890	0.5046	0.9142	1.1439
	50	0.1733	0.0097	0.1815	0.0425	0.1633	0.1597	0.1857	0.3790
	75	0.1131	0.0063	0.1008	0.0244	0.1075	0.1069	0.1028	0.2132
	100	0.0730	0.0046	0.0712	0.0185	0.0730	0.0785	0.0625	0.1586

4.2 Performance of estimators with respect to sample size n

In this subsection, we assess the performance of ML estimators of $\hat{\theta}$ and $\hat{\lambda}$ as sample size n, varies from $n = 10, 11, \ldots, 100$, for $\theta = 1$ and $\alpha = 0.5$. For each of these sample sizes, we generate one thousand samples by using inversion method discussed in Sect. 4 and obtain maximum likelihood estimators and standard errors of ML estimates, $(\hat{\theta}, \hat{\lambda})$ and $(s_{i,\hat{\theta}}, s_{i,\hat{\lambda}})$ for $i = 1, 2, \ldots, 1{,}000$ respectively. For each repetition we comput bias, mean squared error and Coverage Lengths (CL):

Table 4 Average Width of 95 % confidence limit

	n	$\lambda = 0.5$		$\lambda = 1$		$\lambda = 2$	
		AW_λ	AW_θ	AW_λ	AW_θ	AW_λ	AW_θ
$\theta = 0.5$	25	1.1446	0.3930	2.3347	0.3926	4.6154	0.3934
	50	0.7317	0.2663	1.4689	0.2648	2.9775	0.2661
	75	0.5770	0.2143	1.1683	0.2143	2.3471	0.2156
	100	0.4947	0.1837	0.9945	0.1831	1.9738	0.1834
$\theta = 1$	25	0.7434	0.9113	1.4803	0.9132	2.9303	0.9259
	50	0.4935	0.6117	0.9907	0.6234	1.9692	0.6125
	75	0.3971	0.4899	0.7893	0.4921	1.5998	0.4913
	100	0.3404	0.4199	0.6846	0.4230	1.3461	0.4207
$\theta = 2$	25	0.5391	2.2947	1.0864	2.2898	2.1495	2.2439
	50	0.3675	1.4861	0.7424	1.4952	1.4827	1.4946
	75	0.2983	1.1865	0.5941	1.1697	1.1866	1.1877
	100	0.2585	1.0143	0.5129	1.0119	1.0383	1.0147
$\theta = 3$	25	0.4773	4.2754	0.9515	4.0969	1.8912	3.9875
	50	0.3237	2.5275	0.6533	2.5749	1.3011	2.5913
	75	0.2637	1.9924	0.5269	1.9904	1.0539	2.0121
	100	0.2268	1.7051	0.4549	1.7273	0.9148	1.7245

Fig. 4 Bias for $\hat{\theta}$ and $\hat{\lambda}$ versus n
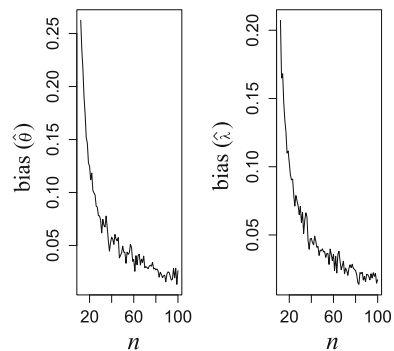
$$bias_\Theta(n) = \frac{1}{1,000} \sum_{i=1}^{1,000} \left(\hat{\Theta}_i - \Theta\right), \; MSE_\Theta(n) = \frac{1}{1,000} \sum_{i=1}^{1,000} \left(\hat{\Theta}_i - \Theta\right)^2,$$

$$\mathrm{CL}_\Theta(n) = \frac{2 \cdot 1.956}{1,000} \sum_{i=1}^{1,000} s_{i,\hat{\Theta}}$$

The coverage probabilities (CP) are given by

$$\mathrm{CP}_\Theta(n) = \frac{1}{1,000} \sum_{i=1}^{1,000} \mathbb{I}\left(\hat{\Theta} - 1.96 s_{i,\hat{\Theta}} < \Theta < \hat{\Theta} - 1.96 s_{i,\hat{\Theta}}\right)$$

for $\Theta = (\theta, \lambda)$ and $\mathbb{I}$ is an indicator function.

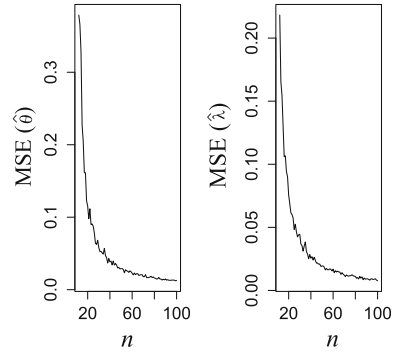**Fig. 5** MSE for $\hat{\theta}$ and $\hat{\lambda}$ versus n

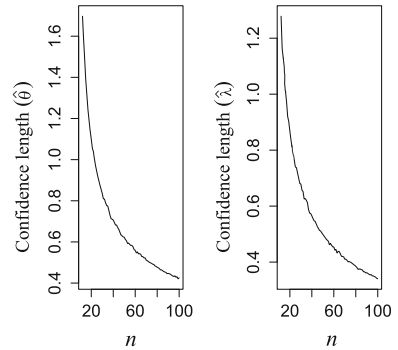**Fig. 6** Converge length for $\hat{\theta}$ and $\hat{\lambda}$ versus n
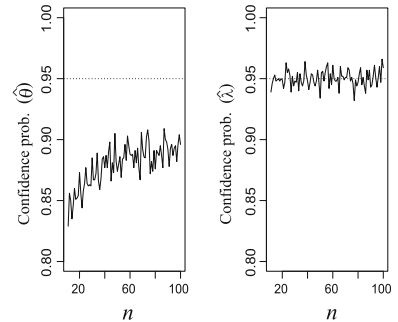
Fig. 7 Converge probabilities for $\hat{\theta}$ and $\hat{\lambda}$ versus n

Figures 4, 5, 6 and 7 shows behaviour of bias, mean squared error, coverage length and coverage probability of parameter (θ and λ) as one varies vary n. The horizontal lines in Fig. 7 correspond to the coverage probabilities being 0.95.

The following observations can be drawn from the figures: the biases for both the parameter appear generally positive and approach to zero with increasing n; the mean squared errors and coverage lengths of (θ and λ) decrease with n. The coverage probabilities for λ appear generally less than the nominal level but reach the nominal level as sample size increases whereas the coverage probabilities for θ appear generally closer than the nominal level. These observations are for $\theta = 1$ and $\alpha = 0.5$ and similar observations were noted for other values of parameters.

Table 5 The estimates of parameters and goodness-of-fit statistics for Illustration 1

Model	Parameter	−LL	AIC	BIC	K–S statistic
L–E	$\hat{\lambda} = 0.0962, \hat{\theta} = 1.229$	401.78	807.564	807.780	0.0454
PL	$\hat{\theta} = 0.385, \hat{\beta} = 0.744$	402.24	808.474	808.688	0.0446
L	$\hat{\theta} = 0.196$	419.52	841.040	843.892	0.0740
NGLD	$\hat{\theta} = 0.180, \hat{\alpha} = 4.679, \hat{\beta} = 1.324$	412.75	831.501	840.057	0.1160

5 Application to real datasets

In this section, we illustrate, the applicability of L–E Distribution by considering two different datasets used by different researchers. We also fit L–E distribution, Power–Lindley distribution [8], New Generalized Lindley Distribution [5], Lindley Distribution [19], and Exponential distribution. Namely
(i) Power–Lindley distribution (PL(α, β)):

$$f_1(x) = \frac{\alpha\beta^2}{1+\beta}(1+x^\alpha)x^{\alpha-1}e^{-\beta x^\alpha}, \qquad x, \alpha, \beta > 0.$$

(ii) New Generalized Lindley distribution (NGLD(α, β, θ)):

$$f_2(x) = \frac{e^{-\theta x}}{1+\theta}\left(\frac{\theta^{\alpha+1}x^{\alpha-1}}{\Gamma(\alpha)} + \frac{\theta^\beta x^{\beta-1}}{\Gamma(\beta)}\right), \qquad x, \alpha, \theta, \beta > 0.$$

(iii) Lindley Distribution (L(θ))

$$f_3(x) = \frac{\theta^2}{1+\theta}(1+x)e^{-\theta x}, \qquad x, \alpha, \beta > 0.$$

For each of these distributions, the parameters are estimated by using the maximum likelihood method, and for comparison we use Negative LogLikelihood values ($-LL$), the Akaike information criterion (AIC) and Bayesian information criterion (BIC) which are defined by $-2LL + 2q$ and $-2LL + q\log(n)$, respectively, where q is the number of parameters estimated and n is the sample size. Further Kolmogorov–Smirnov test (K–S) $= \sup_x |F_n(x) - F(x)|$, where $F_n(x) = \frac{1}{n}\sum_{i=1}^n \mathbb{I}_{x_i \leq x}$ is empirical distribution function and $F(x)$ is cumulative distribution function is calculated and shown for all the datasets.

5.1 Illustration 1: Application to bladder cancer patients

We consider an uncensored data set corresponding to remission times (in months) of a random sample of 128 bladder cancer patients (Lee and Wang [18]) as presented in Appendix A.1 in Table 9. The results for these data are presented in Table 5. We observe that the L–E distribution is a competitive distribution compared with other distributions. In fact, based on the values of the AIC and BIC criteria and K–S test statistic, we observe that the L–E distribution provides the best for these data among all the models considered. The probability density function and empirical distribution function are presented in Fig. 8 for all considered distributions for these data.
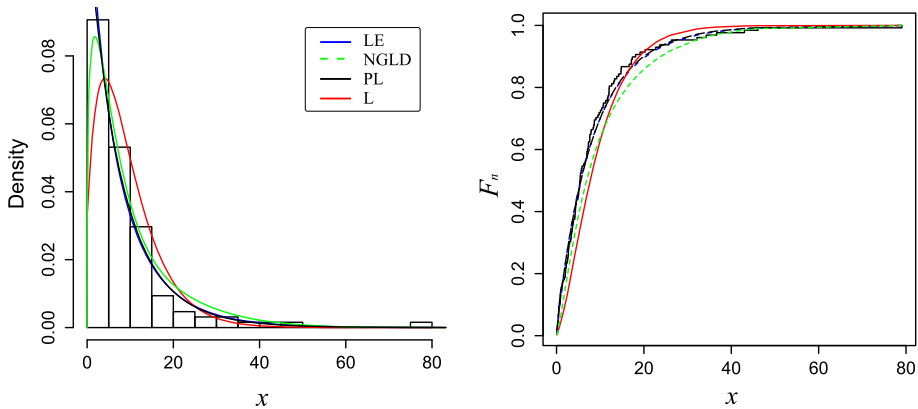
Fig. 8 PDF plot for various values of λ and θ

Table 6 The estimates of parameters and goodness-of-fit statistics for Illustration 2

Model	Parameter	LL	AIC	BIC	K–S
L–E	$\hat{\theta} = 2.650, \hat{\lambda} = 0.1520$	317.005	638.01	638.1337	0.0360
PL	$\hat{\theta} = 0.1530; \hat{\beta} = 1.0832$	318.319	640.64	640.64	0.0520
L	$\hat{\theta} = 0.187$	319.00	640.00	640.00	0.0680
E	$\hat{\theta} = 0.101$	329.00	660.00	660.00	0.1624
NGLD	$\hat{\theta} = 0.2033; \hat{\beta} = 2.008; \hat{\alpha} = 2.008$	317.3	640.60	640.60	0.0425

5.2 Illustration 2: Application to waiting times in a queue

As second example, we consider 100 observations on waiting time (in minutes) before the customer received service in a bank (see Ghitany et al. [7]). The data sets are presented in Appendix A.2 as Table 10. The results for these data are presented in Table 6. From these results we can observe that L–E distribution provide smallest AIC and BIC values as compare to Power lindley, new generalized Lindley distribution, Lindley and exponential and hence best fits the data among all the models considered. The results are presented in Table 6 and probability density function and empirical distribution function are shown in Fig. 9.

6 Estimation of the reliability of a stress–strength system

The stress–strength parameter (R), defined as $R = P(X > Y)$, plays an important role in the reliability analysis as it measures the system performance. Some of the significant work on the stress–strength model can be seen in Kundu and Gupta [14,15], Raqab and Kundu [21], Kundu and Raqab [16], Krishnamoorthy et al. [13], and the references cited therein. Recently Al-Mutairi et al. [2] have presented the estimation of the stress–strength parameter where both strength and stress of the system follow lindley distribution with different shape parameters. In this section, we consider a situation where $X \sim \text{L–E}(\theta_1, \lambda)$ and $Y \sim \text{L–E}(\theta_2, \lambda)$ and are independent of each other. In our case, the stress–strength parameter R is given by
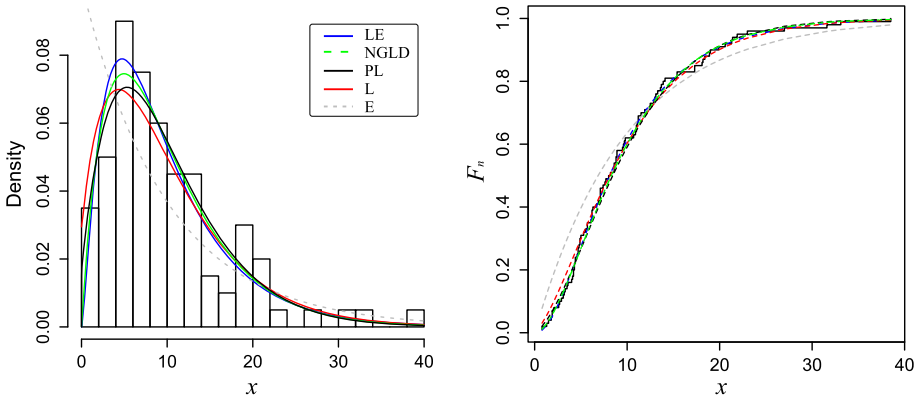
Fig. 9 PDF plot for various values of λ and θ

$$R = P(X > Y) = \int_0^\infty P(X > Y | Y = y) \, f_Y(y) dy$$

$$= \int_0^\infty S_X(y) \, f_Y(y) dy$$

$$= 1 - \int_0^\infty \frac{\theta_2^2 \lambda \left(1 - e^{-\lambda y}\right)^{\theta_1 + \theta_2 - 1} e^{-\lambda y} \left(1 - \log\left(1 - e^{-\lambda y}\right)\right) \left(1 + \theta_1 \left(1 - \log\left(1 - e^{-\lambda y}\right)\right)\right)}{(1 + \theta_1)(1 + \theta_2)} dy$$

$$= 1 - \frac{\theta_2^2 \left(\theta_1^3 + (2\theta_2 + 3)\theta_1^2 + (\theta_2 + 1)(\theta_2 + 3)\theta_1 + \theta_2^2 + \theta_2\right)}{(\theta_1 + 1)(\theta_2 + 1)(\theta_1 + \theta_2)^3} \tag{27}$$

Remarks:

(i) R is independent of λ

(ii) When $\theta_1 = \theta_2$, R = 0.5. This is intuitive that X and Y are i.i.d. and there is an equal chance that X is bigger than Y.

Since R in Eq. (27) is a function of stress–strength parameters θ_1 and θ_2, we need to obtain the maximum likelihood estimators (MLEs) of θ_1 and θ_2 to compute the MLE of R under invariance property of the MLE. Suppose that $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ are independent random samples from L–E (θ_1, λ) and L–E (θ_2, λ) respectively, then likelihood function is given by

$$L(\theta_1, \theta_2) = \frac{\theta_2^{2m} \theta_1^{2n} \lambda^{m+n}}{(\theta_1 + 1)(\theta_2 + 1)} \prod_{i=1}^n e^{-\lambda x_i} \left(1 - e^{-\lambda x_i}\right)^{\theta_1 - 1} \left(1 - \log\left(1 - e^{-\lambda x_i}\right)\right)$$

$$\cdot \prod_{j=1}^m e^{-\lambda y_j} \left(1 - e^{-\lambda y_j}\right)^{\theta_2 - 1} \left(1 - \log\left(1 - e^{-\lambda y_j}\right)\right).$$

The log-likelihood function is given by

$$\log L(\theta_1, \theta_2) = 2n \log (\theta_1) + 2m \log (\theta_2) + (m + n) \log(\lambda) - n \log (\theta_1 + 1) - m \log (\theta_2 + 1)$$

$$- (\theta_1 - 1) s_1 - \lambda \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \log(1 - \log(1 - e^{-\lambda x_i})) - (\theta_2 - 1) s_2$$

$$- \lambda \sum_{j=1}^{m} y_j + \sum_{j=1}^{m} \log(1 - \log(1 - e^{-\lambda y_j})),$$

where $s_1(\lambda) = -\sum_{i=1}^{n} \log \left(1 - e^{-\lambda x_i}\right)$ and $s_2(\lambda) = -\sum_{j=1}^{m} \log \left(1 - e^{-\lambda y_j}\right)$.

The MLE of $\hat{\Theta}$ of Θ is solution of the following non linear equations:

$$\frac{\partial \log L}{\partial \alpha} = \frac{n + m}{\lambda} + (\theta_1 - 1) \sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{(1 - e^{-\lambda x_i})} + (\theta_2 - 1) \sum_{j=1}^{m} \frac{y_j e^{-\lambda y_j}}{(1 - e^{-\lambda y_j})} - \sum_{i=1}^{n} x_i - \sum_{j=1}^{m} y_j$$

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

$$- \sum_{j=1}^{m} \frac{y_j e^{-y_j \lambda}}{(1 - e^{-y_j \lambda}) (1 - \log (1 - e^{-y_j \lambda}))}$$

$$\frac{\partial \log L}{\partial \theta_1} = \frac{2n}{\theta_1} - \frac{n}{\theta_1 + 1} - s_1$$

$$\frac{\partial \log L}{\partial \theta_2} = \frac{2m}{\theta_2} - \frac{m}{\theta_2 + 1} - s_2 \tag{28}$$

from above equations, it follows that

$$\hat{\theta}_1 = \theta_1(\hat{\lambda}) = \frac{\left(s_1(\hat{\lambda}) - n\right) + \sqrt{\left(s_1(\hat{\lambda}) - n\right)^2 + 8n s_1(\hat{\lambda})}}{2 s_1(\hat{\lambda})}$$

$$\hat{\theta}_2 = \theta_2(\hat{\lambda}) = \frac{\left(s_2(\hat{\lambda}) - m\right) + \sqrt{\left(s_2(\hat{\lambda}) - m\right)^2 + 8m s_2(\hat{\lambda})}}{2 s_2(\hat{\lambda})} \tag{29}$$

where $\hat{\lambda}$ is the solution of the non-linear equation:

$$H(\lambda) = \frac{n + m}{\lambda} + (\theta_1(\lambda) - 1) \sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{(1 - e^{-\lambda x_i})} + (\theta_2(\lambda) - 1)$$

$$\times \sum_{j=1}^{m} \frac{y_j e^{-\lambda y_j}}{(1 - e^{-\lambda y_j})} - \sum_{i=1}^{n} x_i - \sum_{j=1}^{m} y_j$$

$$- \sum_{i=1}^{n} \frac{x_i e^{-x_i \lambda}}{(1 - e^{-x_i \lambda}) (1 - \log (1 - e^{-x_i \lambda}))} - \sum_{j=1}^{m} \frac{y_j e^{-y_j \lambda}}{(1 - e^{-y_j \lambda}) (1 - \log (1 - e^{-y_j \lambda}))}. \tag{30}$$

The value of $\hat{\lambda}$ can be substituted in (29) to get $\hat{\theta}_1$ and $\hat{\theta}_2$.

Hence, using the invariance property of the MLE, the maximum likelihood estimator $\hat{R}_{mle}$ of R can be obtained by substituting $\hat{\theta}_k$ for $k = 1,2$ in Eq. (27).

$$\hat{R}_{mle} = 1 - \frac{\theta_2^2 \left(\theta_1^3 + (2\theta_2 + 3)\,\theta_1^2 + (\theta_2 + 1)\,(\theta_2 + 3)\,\theta_1 + \theta_2^2 + \theta_2 \right)}{(\theta_1 + 1)\,(\theta_2 + 1)\,(\theta_1 + \theta_2)^3} \Bigg|_{\theta_1 = \hat{\theta}_1, \theta_2 = \hat{\theta}_2}. \quad (31)$$

6.1 Asymptotic confidence interval

For an estimator $\hat{\theta}_k$ to be asymptotically efficient for estimating θ_k for large sample, we should have $\sqrt{n_k}\big(\hat{\theta}_k - \theta_k\big) \xrightarrow{D} N\big(0, I(\theta_k)^{-1}\big)$

$$I(\theta_k) = -E\left(\frac{\partial^2 \log L}{\partial \theta_k^2} \right) = -E\left(\frac{n_k}{(\theta_k + 1)^2} - \frac{2n_k}{\theta_k^2} \right) = \frac{2n_k}{\theta_1^2} - \frac{n_k}{(\theta_1 + 1)^2}$$

$$I_1 = I(\theta_1) = -E\left(\frac{n}{(\theta_1 + 1)^2} - \frac{2n}{\theta_1^2} \right) = \frac{2n}{\theta_1^2} - \frac{n}{(\theta_1 + 1)^2}$$

$$I_2 = I(\theta_2) = -E\left(\frac{m}{(\theta_2 + 1)^2} - \frac{2m}{\theta_2^2} \right) = \frac{2m}{\theta_2^2} - \frac{m}{(\theta_2 + 1)^2}$$

where $n_1 = n$ and $n_2 = m$.

Therefore, as easily follows it $n \to \infty$ and $m \to \infty$

$$\frac{\hat{R} - R}{\sqrt{\dfrac{d_2^2}{mI_2} + \dfrac{d_1^2}{nI_1}}} \xrightarrow{D} N(0, 1)$$

where

$$d_1 = \frac{\partial R}{\partial \theta_1} = \frac{\theta_1 \theta_2^2 \left(\theta_1^3 + 2\,(\theta_2 + 3)\,\theta_1^2 + \left(\theta_2^2 + 8\theta_2 + 12\right)\theta_1 + 2\left(\theta_2^2 + 3\theta_2 + 3\right) \right)}{(\theta_1 + 1)^2\,(\theta_2 + 1)\,(\theta_1 + \theta_2)^4}$$

$$d_2 = \frac{\partial R}{\partial \theta_2} = -\frac{\theta_1^2 \theta_2 \left(\theta_2^3 + 6\theta_2^2 + 12\theta_2 + \theta_1^2\,(\theta_2 + 2) + 2\theta_1\left(\theta_2^2 + 4\theta_2 + 3\right) + 6 \right)}{(\theta_1 + 1)\,(\theta_2 + 1)^2\,(\theta_1 + \theta_2)^4}$$

Interval Using the asymptotic distribution of $\hat{R}$, $100(1 - \alpha)\,\%$ confidence interval for R can be easily obtained as

$$\hat{R} \mp Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{d_2^2}}{m\hat{I_2}} + \frac{\hat{d_1^2}}{n\hat{I_1}}}.$$

6.2 Illustration 3: Application to stress–strength data

In illustration 3, we present the results obtained from estimation methods of R proposed in Sect. 6, we consider two datasets, reported in Ghitany et al. [7], Al-Mutairi et al. [2], represents the waiting times (in minutes) before customer service in two different banks having 100(n) and 60(m) number of observations respectively. Q–Q plot in Fig. 10 and KS value in Table 7 indicate that LE Distribution fits better than Lindley Distribution (L). Thus, $\hat{R}_{MLE}$ and 95 %
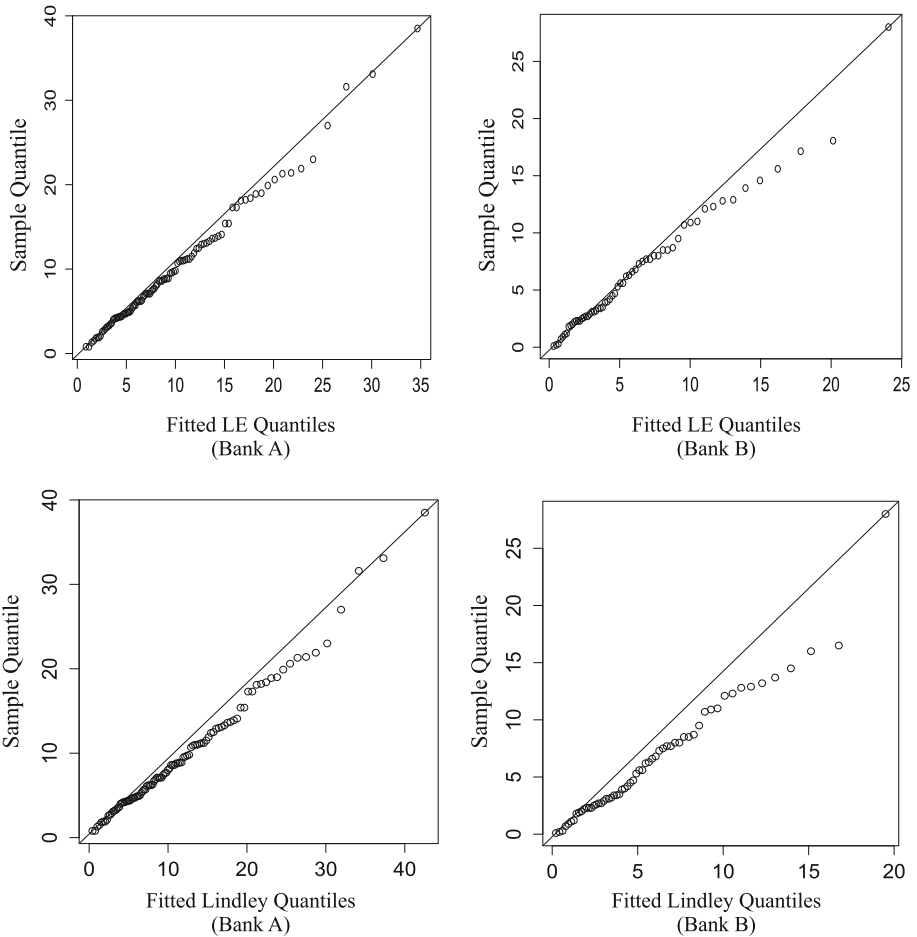
Fig. 10 Q–Q plot of the fitted LE and Lindley distribution for Bank A and Bank B data set

Table 7 Kolmogorov–Smirnov distances value for Bank A and Bank B data set

Data set	Model	Parameter	KS value
Bank A	LE	$\hat{\lambda} = 0.152, \hat{\theta} = 2.650$	0.0360
	L	$\hat{\theta} = 0.187$	0.0680
Bank B	LE	$\hat{\lambda} = 0.1758, \hat{\theta} = 1.767$	0.0496
	L	$\hat{\theta} = 0.2797$	0.0797

confidence interval were computed and presented in Table 8. It is clear that, from the Table 8 that the estimated value of stress–strength parameter obtained from L–E distribution is lower than the one obtained from Lindley distribution. Moreover, standard error of R is also lower for LE distribution compared to Lindley distribution. Hence we recommend, LE distribution for stress-strength analysis also.

Table 8 The maximum likelihood estimator of R for LE and Lindley model

Model	Parameter	$\hat{R}_{MLE}$	95 % confidence limit
LE	$\hat{\theta}_1 = 1.3982$, $\hat{\theta}_2 = 1.5023$, $\hat{\lambda} = 0.1096$	0.476	(0.4684, 0.4840)
L	$\hat{\theta}_1 = 0.187$, $\hat{\theta}_2 = 0.280$	0.645	(0.565, 0.720)

Conclusion

We have proposed a new two parameter class of distribution called Lindley–Exponential (L–E) distribution generated by Lindley distribution possess increasing, decreasing or upside down hazard function for different choices of the parameters. We have derived important properties of the L–E distribution like moments, entropy, asymptotic distribution of sample maximum and sample Minimum. Maximum likelihood of the parameters are obtained which can be used to get asymptotic confidence intervals. We have also illustrated the application of L–E distribution to two real data sets used by researchers earlier and compare it with other popular models. Further the stress–strength analysis were carried out and compared with that of Lindley distribution. Our application to real data set indicate that L–E distribution performs satisfactorily or better than its competitors and can be recommended for lifetime modelling the encountered in engineering, medical science, biological science and other applied sciences.

Appendix

Appendix A.1: Dataset used in illustration 1

See Table 9.

Table 9 The remission times (in months) of bladder cancer patients

0.08	2.09	3.48	4.87	6.94	8.66	13.11	23.63	0.2	2.23	0.26	0.31	0.73
0.52	4.98	6.97	9.02	13.29	0.4	2.26	3.57	5.06	7.09	11.98	4.51	2.07
0.22	13.8	25.74	0.5	2.46	3.64	5.09	7.26	9.47	14.24	19.13	6.54	3.36
0.82	0.51	2.54	3.7	5.17	7.28	9.74	14.76	26.31	0.81	1.76	8.53	6.93
0.62	3.82	5.32	7.32	10.06	14.77	32.15	2.64	3.88	5.32	3.25	12.03	8.65
0.39	10.34	14.83	34.26	0.9	2.69	4.18	5.34	7.59	10.66	4.5	20.28	12.63
0.96	36.66	1.05	2.69	4.23	5.41	7.62	10.75	16.62	43.01	6.25	2.02	22.69
0.19	2.75	4.26	5.41	7.63	17.12	46.12	1.26	2.83	4.33	8.37	3.36	5.49
0.66	11.25	17.14	79.05	1.35	2.87	5.62	7.87	11.64	17.36	12.02	6.76	
0.4	3.02	4.34	5.71	7.93	11.79	18.1	1.46	4.4	5.85	2.02	12.07	

Appendix A.2: Dataset used in illustration 2

See Table 10.

Table 10 Waiting times (min) of 100 bank customers

0.8	0.8	1.3	1.5	1.8	1.9	1.9	2.1	2.6	2.7
2.9	3.1	3.2	3.3	3.5	3.6	4	4.1	4.2	4.2
4.3	4.3	4.4	4.4	4.6	4.7	4.7	4.8	4.9	4.9
5.0	5.3	5.5	5.7	5.7	6.1	6.2	6.2	6.2	6.3
6.7	6.9	7.1	7.1	7.1	7.1	7.4	7.6	7.7	8
8.2	8.6	8.6	8.6	8.8	8.8	8.9	8.9	9.5	9.6
9.7	9.8	10.7	10.9	11.0	11.0	11.1	11.2	11.2	11.5
11.9	12.4	12.5	12.9	13.0	13.1	13.3	13.6	13.7	13.9
14.1	15.4	15.4	17.3	17.3	18.1	18.2	18.4	18.9	19.0
19.9	20.6	21.3	21.4	21.9	23	27	31.6	33.1	38.5

References

1. Adamidis, K., Loukas, S.: A lifetime distribution with decreasing failure rate. Stat. Probab. Lett. **39**, 35–42 (1998)
2. Al-Mutairi, D.K., Ghitany, M.E., Kundu, D.: Inference on stress–strength reliability from Lindley distribution. Commun. Stat. Theory Methods **42**, 1443–1463 (2013)
3. Arnold, B.C., Balakrishnan, N., Nagaraja, H.N.: A First Course in Order Statistics. Wiley, New York (2013)
4. Bakouch, H.S., Al-Zahrani, B.M., Al-Shomrani, A.A., Marchi, V.A., Louzada, F.: An extended Lindley distribution. J. Korean Stat. Soc. **41**, 75–85 (2012)
5. Elbatal I., Merovci F., Elgarhy M.: A new generalized Lindley distribution. Math. Theory Model. **3**(13) (2013)
6. Ghitany, M.E., Alqallaf, F., Al-Mutairi, D.K., Husain, H.A.: A two-parameter weighted Lindley distribution and its applications to survival data. Math. Comput. Simul. **81**(6), 1190–1201 (2011)
7. Ghitany, M.E., Atieh, B., Nadarajah, S.: Lindley distribution and its application. Math. Comput. Simul. **78**, 493–506 (2008)
8. Ghitany, M.E., Al-Mutairi, D.K., Aboukhamseen, S.M.: Estimation of the reliability of a stress–strength system from power lindley distributions. Commun. Stat.-Simul. Comput. **44**(1), 118–136 (2015)
9. Glaser, R.E.: Bathtub and related failure rate characterizations. J. Am. Stat. Assoc. **75**, 667–672 (1980)
10. Gómez, E.D., Sordo, M.A., Calderín, E.O.: The LogLindley distribution as an alternative to the beta regression model with applications in insurance. Insur. Math. Econ. **54**, 49–57 (2014)
11. Hassan, M.K.: On the convolution of Lindley distribution. Columbia Int. Publ. Contemp. Math. Stat. **2**(1), 47–54 (2014)
12. Jorda, P.: Computer generation of random variables with Lindley or Poisson–Lindley distribution via the Lambert W function. Math. Comput. Simul. **81**, 851–859 (2010)
13. Krishnamoorthy, K., Mukherjee, S., Guo, H.: Inference on reliability in two parameter exponential stress–strength model. Metrika **65**, 261–273 (2007)
14. Kundu, D., Gupta, R.D.: Estimation of $R = P(Y < X)$ for generalized exponential distributions. Metrika **61**, 291–380 (2005)
15. Kundu, D., Gupta, R.D.: Estimation of $R = P(Y < X)$ for Weibull distributions. IEEE Trans. Reliab. **55**, 270–280 (2006)
16. Kundu, D., Raqab, M.Z.: Estimation of $R = P(Y < X)$ for three-parameter Weibull distribution. Stat. Probab. Lett. **79**, 1839–1846 (2009)
17. Leadbetter, M.R., Lindgren, G., Rootzn, H.: Extremes and Related Properties of Random Sequences and Processes. Springer Statist. Ser. Springer, Berlin (1983)

18. Lee, E.T., Wang, J.W.: Statistical Methods for Survival Data Analysis, 3rd edn. Wiley, Hoboken (2003)
19. Lindley, D.V.: Fiducial distributions and Bayes theorem. J. R. Stat. Soc. Ser. B (Methodological). **20**(1), 102–107 (1958)
20. Nair, N., Nair, N.U., Sankaran, P.G., Balakrishnan, N.: Quantile-Based Reliability Analysis. Springer, Berlin (2013)
21. Raqab, M.Z., Kundu, D.: Comparison of different estimators of $P(Y < X)$ for a scaled Burr type X distribution. Commun. Stat. Simul. Comput. **34**, 465–483 (2005)
22. Ristić, M.M., Balakrishnan, N.: The gamma exponentiated exponential distribution. J. Stat. Comput. Simul. **82**(8), 1191–1206 (2012)
23. Shanker, R., Sharma, S., Shanker, R.: A two-parameter Lindley distribution for modeling waiting and survival times data. Appl. Math. **4**, 363–368 (2013)
24. Zakerzadeh, H., Mahmoudi, E.: A new two parameter lifetime distribution: model and properties. arXiv:1204.4248v1 [stat.CO], 2012 (1998)