Lindley–exponential slash distribution

M Dwiki, S Nurrohmah and M Novita

Department of Mathematics, Faculty of Mathematics and Natural Sciences (FMIPA), Universitas Indonesia, Depok 16424, Indonesia

Corresponding author’s email: snurrohmah@sci.ui.ac.id

Abstract. There is a problem arising when Lindley distribution is used to model right–skewed and unimodal data, which lies in its inability to model data with its peak farther from 0. A modification is required to increase the flexibility of this distribution, with “transformed–transformer” being one of the proposed method. This method is done by making a composition of two random variables through their respective distribution functions, with random variable T being the “transformed”, and random variable X being the ”transformer”. In this paper, Lindley distribution was chosen to be the ”transformed” and exponential distribution was chosen to be the ”transformer”, constructing the Lindley–exponential distribution. However, there is a trouble using distribution if the data has a heavier tail. An alternative distribution is required while maintaining the properties of the Lindley-Exponential Distribution. Through transformation of variables method, a new distribution, Lindley-Exponential Slash Distribution is introduced, which is unimodal, right–skewed, heavier tailed. This paper covered some of its statistical characteristics, such as pdf, cdf, survival function, hazard rate, kth moment, mean, variance, skewness, and kurtosis. Parameter estimation was carried out with maximum likelihood estimation through numerical method. An application of distribution was illustrated on maximum annual precipitation data of Durham City.

Keywords: Heavy–tail, maximum likelihood estimation, slash distribution, transformation of variables, unimodal

1. Introduction

Lindley distribution was introduced by Lindley [1] from purely theoretical motivations. Ghitany et al. [2] explored the statistical properties of this distribution, and displayed some of its advantages over exponential distribution, specifically on its mode, coefficient of variation, skewness, kurtosis, and its hazard rate. It is also shown that the Lindley distribution has applicabilities in modeling lifetime data.

However, Lindley distribution has a limitation in modeling data with higher–valued peak or mode. A modification is required to increase this distribution’s flexibility. Ozel et al. [3] introduced a Lindley–X family of distributions using the “transformed-transformer” (T–X) method proposed by Alzaatreh et al. [4], and constructed the Lindley–exponential distribution as one of the examples. This distribution is constructed from a composition between Lindley distribution as the “transformed” and exponential distribution as the “transformer”.

The additional parameter added enables the Lindley–xponential distribution’s peak or mode to have a wider range of values compared to Lindley distribution. However, this distribution has its own limitation, and unable to model data that requires the distribution to have heavier right–tails. Rogers et al. [5] introduced a distribution, suitable to model a data with tails heavier than standard
normal distribution, called the Slash distribution, as one of the examples of long-tailed symmetrical distributions. The Slash distribution is constructed from a random variable \( Y = \frac{X}{U^{1/q}} \), where \( X \) is a random variable of standard normal distribution, and \( U \) is a random variable from uniform distribution.

Using the same method of constructing the Slash distribution, Gui [6] introduced a Lindley–exponential slash distribution, from random variable \( Y = \frac{X}{U^{1/q}} \), where \( X \) is a random variable Lindley–exponential distribution, and \( U \) is a random variable from uniform distribution. This new distribution will be able to model data with heavier tails than the Lindley–exponential distribution.

Therefore, we are motivated to discuss about Lindley–exponential slash distribution, as the generalization of Lindley–exponential distribution, its properties, mathematical derivations and its use in other data.

This paper is organized as the following. Section 2 discusses the basic statistical properties of the Lindley–exponential slash distribution, such as its density function, cumulative distribution function, survival function, hazard rate, its \( k \)th moments and maximum likelihood estimation for the parameters. Section 3 shows the application of this distribution on real data and this work is concluded in section 4.

2. Lindley-exponential slash distribution

2.1. Lindley distribution

Let \( X \sim \text{Lindley}(\theta) \). Then, the density function of \( X \) is given by:

\[
    f(x) = \frac{\theta^2}{\theta + 1} (1 + x)e^{-\theta x}, \quad x > 0, \theta > 0.
\]  

(1)

It can be shown that when \( \theta \geq 1 \), the density function will be monotonically decreasing, and when \( 0 < \theta < 1 \), the density function will be increasing until \( \frac{1-\theta}{\theta} \), then it will monotonically decrease. Therefore, the mode of the Lindley distribution is given by:

\[
    \text{mode}(X) = \left\{ \begin{array}{ll}
        \frac{1-\theta}{\theta} & 0 < \theta < 1 \\
        0, & \theta \geq 1
    \end{array} \right.
\]  

(2)

2.2. Lindley exponential distribution

2.2.1. Transformed-Transformer (T-X) Method. Let \( r(t) \) be the pdf of a random variable \( T \), defined on \([0, \infty)\). Let \( F(x) \) be the cdf of a random variable \( X \). The cdf of the \( T - X \) family of distributions is given by:

\[
    G(x) = \int_0^{-\log(1-F(x))} r(t)dt
\]  

(3)

with its pdf given by:

\[
    g(x) = \frac{f(x)}{1 - F(x)} r(-\log(1 - F(x)))
\]  

(4)

which can be obtained by taking the derivative of equation 3.
2.2.2. **Lindley-X Family of Distributions.** Based on the “transformed-transformer” (T-X) method, the Lindley-X family of distributions has a pdf given by:

\[
g(x) = f(x)[1 - \log(1 - F(x))][1 - F(x)]^{\theta - 1} \frac{\theta^2}{\theta + 1}, x > 0, \theta > 0 \tag{5}
\]

where \( f(x) \) and \( F(x) \) are pdf and cdf of the “transformer” distribution, respectively.

2.2.3. **Lindley-Exponential Distribution.** Let \( X \) be a random variable of exponential distribution with pdf \( f(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \) and cdf \( F(x) = 1 - e^{-\frac{x}{\lambda}} \), as the “transformer”. Then, by equation 3, the Lindley-Exponential distribution density function is given by:

\[
f(x) = \frac{\theta^2}{(\theta + 1)\lambda} \left[ 1 + \frac{x}{\lambda} \right] e^{-\frac{x}{\lambda}}, x, \theta, \lambda > 0 \tag{6}
\]

here the Lindley distribution is a special case when \( \lambda = 1 \). It can be shown by taking the derivative of equation 6, that the mode of this distribution is given by:

\[
\text{mode}(X) = \begin{cases} 
\frac{(1 - \theta)\lambda}{\theta}, & 0 < \theta < 1, \lambda > 0 \\
0, & \theta \geq 1, \lambda > 0
\end{cases} \tag{7}
\]

The \( \lambda \) parameter allows the distribution to have a wider range of modes.

2.3. **Slash distribution**

Let \( X \) has a standard normal distribution, \( U \) has a uniform \( (0,1) \) distribution, where \( X \) and \( U \) are independent. Let a transformation be defined as:

\[
Y = \frac{X}{U^{1/q}}, q > 0 \tag{8}
\]

Then \( Y \) has a Slash distribution with parameter \( q \), with density function given by:

\[
f(y) = q \int_0^1 u^q \phi(yu)du, -\infty < y < \infty, q > 0 \tag{9}
\]

with \( \phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \), or the standard normal density function. This distribution has heavier tails than standard normal distribution. As the \( q \) gets lower the tails get heavier and the distribution gets flatter. For \( q \to \infty \), it will converge back to standard normal distribution. Figure 1 illustrates the behavior of the \( q \) on the density of Slash distribution:

2.4. **Lindley-exponential slash distribution**

Based on the construction of slash distribution, the construction of Lindley–exponential slash distribution is introduced.

2.4.1. **Pdf and Cdf.** Let \( X \) be a random variable from Lindley–exponential distribution with parameters \((\theta, \lambda)\), and \( U \) a random variable from uniform\((0,1)\) distribution, where \( X \) and \( U \) are independent. Let a transformation be defined as:

\[
Y = \frac{X}{U^{1/q}}, q > 0. \tag{10}
\]
Then the random variable $Y$ has a Lindley–exponential Slash distribution with parameters $(\theta, \lambda, q)$, with the density function given by:

$$g(y) = \frac{q \theta^2}{(\theta + 1)\lambda} \int_0^1 \left[ 1 + \frac{yt}{\lambda} \right] e^{-\frac{\theta yt}{\lambda}} t^{1/q} dt, y, 0, \lambda, q > 0$$  \hspace{1cm} (11)

**Proof:** From the transformation in equation 10, the Jacobian determinant is as the following:

$$J = \begin{vmatrix} \frac{\partial x}{\partial y} & \frac{\partial x}{\partial w} \\ \frac{\partial u}{\partial y} & \frac{\partial u}{\partial w} \end{vmatrix} = \begin{vmatrix} w^{1/q} & \left(\frac{1}{q}\right)y^{w^{1/q}-1} \\ 0 & 1 \end{vmatrix} = w^{1/q}$$

Therefore, the joint pdf is as following:

$$g_{Y,W}(y,w) = f_{X,U}(yw^{1/q},w)|J| = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2w^{2/q}}{2}\right)w^{1/q}$$  \hspace{1cm} (12)

The marginal pdf of $Y$ is obtained by integrating equation 12.

For $q \to \infty$, this distribution converges back to Lindley–exponential distribution. The plot of the density function will be shown. Through integration by parts, the distribution function of Lindley–exponential slash distribution is given by:

$$G(y) = \frac{1}{\lambda(\theta + 1)} - e^{-\frac{\theta yt}{\lambda}} \frac{(\lambda \theta + \theta yt)}{\lambda(\theta + 1)} t^{q-1} dt, y, \theta, \lambda, q > 0$$  \hspace{1cm} (13)
Figure 2. Plot of density function of Lindley–exponential slash distribution with various parameters.

Figure 2 shows the density function of Lindley–exponential Slash distribution for various cases. It can be seen that for $\theta \geq 1$, the density function will be monotonically decreasing, and when $0 < \theta < 1$, the density function will be increasing, then it will monotonically decrease, similar to the behavior in Lindley distribution. The lesser the $q$, the heavier the tail and the distribution gets flatter. For the higher $q$, the distribution converges to Lindley–exponential distribution. For higher $\lambda$, the peak has higher value, and the peak has lesser probability.

2.4.2. Survival Function and Hazard Rate. The survival function of Lindley–exponential slash distribution is as following:

$$S(y) = 1 - q \int_0^y \frac{\lambda(t + 1) - e^{-\theta yt}}{\lambda(t + 1)} t^{q-1} dt, y, \theta, \lambda, q > 0$$  (14)

while the hazard rate is as follows:

$$h(y) = \frac{q \theta^2}{(\theta + 1) \lambda} \int_0^y \left[ 1 + \frac{yt}{\lambda} \right] e^{-\theta yt} t^{q-1} dt, y, \theta, \lambda, q > 0$$  (15)
Figure 3 shows the hazard rate of Lindley–exponential slash distribution for various cases. It can be seen that the hazard rate can be bathtub shaped, while the hazard rate of Lindley and Lindley–exponential distribution is limited only to increasing function.

2.4.3. Moments. The \( k \)^{th} moment of the Lindley–exponential slash distribution is obtained through integration by substitution and uses of gamma function:

\[
E(Y^k) = \frac{q^k \Gamma(k+\theta + \lambda)}{(q-k)(\theta+1)\theta^k}, q > k
\]  

(16)

based from equation 12, the mean and the variance are as following:

\[
\mu = \frac{q(\theta+2)\lambda}{(q-1)(\theta+1)}, q > 1
\]

(17)

\[
\sigma^2 = \frac{q^2(\theta^2+4\theta+2) + q^2(\theta^2+4\theta+2) + 2(\theta^2+4\theta+3)}{(q-1)^2(q-2)(\theta+1)^2}, q > 2
\]

(18)

while the skewness and kurtosis are as following:

\[
y_1 = \frac{2\sqrt{q-2} \left[ 6(\theta^3+6\theta^2+9\theta+4) + q^4(\theta^3+6\theta^2+6\theta+2) - 5q^3(\theta^3+6\theta^2+6\theta+2) 
+ 6q^2(2\theta^3+12\theta^2+14\theta+5) - 6q(2\theta^3+12\theta^2+15\theta+5) \right]}{\sqrt{q(q-3)(-2q(\theta^2+4\theta+2) + q^2(\theta^2+4\theta+2) + 2(\theta^2+4\theta+3))}}, q > 3
\]

(19)

\[
y_2 = \frac{3(q-2) \left[ 4\theta(\theta^3+8\theta^2+18\theta^2+16\theta+5) + q^3(3\theta^3+24\theta^2+44\theta^2+32\theta+8) - 9q^2(3\theta^3+24\theta^2+44\theta^2+32\theta+8) 
+ 24q^4(53\theta^4+249\theta^3+799\theta^2+592\theta+152) - 4q^4(55\theta^4+440\theta^3+851\theta^2+640\theta+168) 
+ 24q^5(110\theta^5+880\theta^4+1780\theta^3+1496\theta^2+399\theta+23) - 24q(70\theta^5+56\theta^4+118\theta^3+96\theta^2+27) \right]}{q(q-3)(q-4)(-2q(\theta^2+4\theta+2) + q^2(\theta^2+4\theta+2) + 2(\theta^2+4\theta+3))}, q > 4
\]

(20)

It can be seen that the parameter \( \lambda \) does not affect the skewness and kurtosis. For a high \( q \), the moments converge to the Lindley–exponential distribution.

2.4.4. Maximum likelihood estimation. Let \( y_1, y_2, \ldots, y_n \) be a random sample with size \( n \) from the Lindley–exponential slash distribution with parameter \( (\theta, \lambda, q) \). Then, the likelihood function is given by:

\[
L(\theta, \lambda, q; y_1, y_2, \ldots, y_n) = \left( \frac{q\theta^2}{(\theta+1)\lambda} \right)^n \prod_{i=1}^{n} \int_{0}^{1} \left[ 1 + \frac{y_i t}{\lambda} \right] e^{-\frac{\theta y_i t}{\lambda}} t^q dt
\]

(21)

Taking the logarithm of the likelihood function, the log–likelihood function is given by:

\[
\log(L) = n \log(q) + 2n \log(\theta) - n \log(\theta+1) - n \log(\lambda) + \sum_{i=1}^{n} \log \left[ \int_{0}^{1} \left( 1 + \frac{y_i t}{\lambda} \right) e^{-\frac{\theta y_i t}{\lambda}} t^q dt \right]
\]

(22)

Finding the \( \theta, \lambda, q \) that maximizes the log-likelihood function requires taking the partial derivatives of the function with respect to \( \theta, \lambda, q \) respectively and equate them to zero. However, it is very difficult to obtain the analytical solution to this system of equations, so numerical method is required. The minimize routine of the scipy package from python 3 is used in this research.
3. Data application

We consider a data set representing a maximum annual precipitation (in 0.1 mm) in Durham city, recorded from year 1973 to 2012, taken from the National Centers for Environmental Information, NOAA. The data are as follows: 5.08, 2.06, 2.73, 0.74, 0.53, 1.26, 2.3, 0.51, 2.51, 1.05, 1.27, 0.98, 1.48, 1.15, 1.17, 1.55, 3.45, 0.22, 3.5, 1.76, 2.15, 1.02, 2.3, 1.5, 4.03, 1.48, 2.25, 0.93, 4.26, 0.59, 0.73, 1.7, 1.36, 2.12, 1.1, 5.1, 1.87, 6.5, 3.28, 0.84.

We present the descriptive statistics of the data at table 1. Then we fit the data set using Lindley, Lindley–exponential and the Lindley–exponential slash distribution, respectively, using maximum likelihood estimation. The results are shown in table 2.

Figure 4 illustrates the fitted models with the MLE estimates. From table 2, it can be seen that the Lindley distribution has the lowest log-likelihood value, also from figure 4, the Lindley distribution did not model the peak properly. Visually, the Lindley–exponential and Lindley–exponential slash distribution seem to be able to model the data properly. The Lindley–exponential and Lindley exponential slash distribution yield the same result, which is shown from the same $\widehat{\theta}$ and $\widehat{\lambda}$ values. The high value of $\overline{\theta}$ shows the convergence of the Lindley–exponential slash distribution to Lindley–exponential distribution. Unfortunately, we have not been able to find an appropriate data to showcase the advantage Lindley–exponential slash distribution has over the Lindley–exponential distribution, due to the light–tailed nature of the data.

To examine if the distributions have fitted the data, a Kolmogorov–smirnov test is carried out, with its result shown on table 3. It can be seen that all 3 distributions could fit the data. However, the Lindley distribution performed less well compared to the other distributions due to its highest test statistic value. It can also be seen that the Lindley–exponential and Lindley–exponential slash distribution has the same performance, indicated by the same test statistic value.
Table 1. Summary for maximum annual precipitation data set.

| Sample Size | Mean | Median | Std. Dev | Skewness | Kurtosis | Min | Max |
|-------------|------|--------|----------|----------|----------|-----|-----|
| 40          | 2.01 | 1.525  | 1.42     | 1.348    | 4.393    | 0.22| 6.5 |

Table 2. MLE parameter estimates of the Lindley, Lindley–exponential and Lindley–exponential slash distribution for the data set.

| Distribution                  | $\theta$ | $\lambda$ | $\hat{q}$ | Loglik |
|-------------------------------|----------|-----------|-----------|--------|
| Lindley                       | 0.777    | -         | -         | -65.283|
| Lindley–exponential           | 0.01     | 0.01      | -         | -61.835|
| Lindley–exponential slash     | 0.01     | 0.01      | 134.731   | -61.835|

Figure 4. Plot of density functions from MLE estimates of the Lindley, Lindley–exponential, and Lindley–exponential slash distribution with the data set.

Table 3. Kolmogorov–smirnov test result of the Lindley, Lindley–exponential and Lindley–exponential slash distribution for the data set.

| Distribution                  | Critical value ($\alpha = 0.05$) | Test statistic | Null hypothesis |
|-------------------------------|----------------------------------|----------------|----------------|
| Lindley                       | 0.215                            | 0.127          | Supported      |
| Lindley–exponential           | 0.215                            | 0.0834         | Supported      |
| Lindley–exponential slash     | 0.215                            | 0.0834         | Supported      |

4. Conclusion
We introduced the Lindley–exponential slash distribution, which is the quotient of two independent random variables, a Lindley–exponential random variable and a power of uniform random variable, and the Lindley–exponential distribution is a special case. The resulting distribution can model data with heavier right–tail, while Lindley and Lindley–exponential distributions are limited to data with light tail.
We derived its statistical and probabilistic properties, computed the MLE through a numerical method. We then present an application of the distribution to a real data set.

References
[1] Lindley D V 1958 J. R. Stat. Soc. B 20 102-7
[2] Ghitany M E, Atieh B and Nadarajah S 2008 Math. Comput. Simulat. 78 493-506
[3] Cakmakyapan S and Ozel G 2016 Hacettepe. J. Math. Stat. 46 1113-37
[4] Alzaatreh A, Lee C and Famoye F 2013 Metron 71 63-79
[5] Rogers W H and Tukey J W 1972 Stat. Neerl. 26 211-26
[6] Gui W 2012 J. Appl. Stat. Sci. 20 283-98