Pakistan Journal of Statistics and Operation Research

A powerful goodness-of-fit test for Lindley distribution with application to real data

Hadi Alizadeh Noughabi

Department of Statistics, University of Birjand, Birjand, Iran, alizadehhadi@birjand.ac.ir

Abstract

The Lindley distribution may serve as a useful reliability model. Applications of this distribution are presented in statistical literature. In this article, a powerful goodness of fit test for the Lindley distribution is proposed. In order to compute the proposed test statistic, we use the maximum likelihood estimate (MLE) suggested by Ghitany et al. (2008), which is simple explicit estimator. By Monte Carlo simulation, critical points of the proposed test statistic for different sample sizes are obtained. Power values of the proposed test are compared with the competing tests against various alternatives via simulations. Finally, two real data are presented and analyzed.

Key Words: Model validity; Lindley distribution; Goodness of fit tests; Kullback-Leibler information; Monte Carlo simulation; Power study.

Mathematical Subject Classification: 62G10; 62B10

1. Introduction

The modeling and analyzing lifetime data are crucial in many applied sciences including medicine, engineering, insurance and finance, amongst others. It is well known that the Lindley distribution is one of the fundamental models applied for reliability models. The Lindley distribution has been discussed by many authors in different practical cases, such as Bayesian estimation Ali et al. (2013), loading-sharing system mode Singh and Gupta (2012) and stress-strength reliability model Al-Mutairi el al. (2013). It deserves mentioning that the Lindley distribution provides a flexible shape to model the lifetime data. Moreover, Ghitany et al. (2008) presented a comprehensive study about its important mathematical and statistical properties, estimation of parameter and application showing the superiority of Lindley distribution over of the bank customers.

Since the distribution was proposed, it has been overlooked in the literature partly due to the popularity of the exponential distribution in the context of reliability analysis. Nonetheless, it has recently received considerable attention as a lifetime model to analyze survival data in the competing risks analysis and stress-strength reliability studies; see, for example, Ghitany et al. (2008), Mazucheli and Achcar (2011), Gupta and Singh (2013), Al-Mutairi el al. (2013), and Wang (2013), Valiollahi et al. (2017), Altun (2019), Kumar and Jose (2019), and Ibrahim et al. (2019), among others.

Ghitany et al. (2008) provide a nice overview of various statistical properties of the Lindley distribution. Furthermore, they argue that the Lindley distribution could be a better lifetime model than the exponential distribution using a real data set.

Therefore, it is a clear need to check whether the Lindley model is a satisfactory model for the observations.

Goodness-of-fit (GOF) tests are designed to measure how well the observed sample data fits some proposed model. One class of GOF tests that can be used consists of tests based on the distance between the empirical and hypothesized distribution functions. Five of the known tests in this class are Kolmogrov-Smirnov, Cramer-von-Mises, Anderson-Darling, Watson and Kuiper tests.

Many researchers have been interested in goodness of fit tests for different distributions and therefore different tests are developed in the literature. For example, see D'Agostino and Stephen (1986) and Huber-Carol et al. (2002). Moreover, goodness of fit tests based on censored samples are developed by some authors including Balakrishnan et

al. (2004), Balakrishnan et al. (2007), Lin et al. (2008), Habibi Rad et al. (2011), Pakyari and Balakrishnan (2012, 2013) and Alizadeh Noughabi and Balakrishnan (2015).

The main contribution of the paper can express as follows. In this paper, we propose a goodness of fit test statistic for the Lindley distribution based on an estimate of Kullback-Leibler divergence. The properties of the proposed test and critical points are presented. We show through extensive simulation studies that the proposed goodness-of-fit test is more powerful, or at least as good as the classical EDF-tests for different choices of sample sizes and alternatives. We also investigate the behavior of the proposed test for the Lindley model with real data.

In Section 2, a summary of the Lindley distribution is presented and then we introduce a goodness of fit test statistic based on the Kullback-Leibler divergence for the Lindley distribution. In Section 3, the critical values of the test statistic are obtained by Monte Carlo simulations. Then power values of the proposed test are computed and then compared with the competing tests. All simulations were carried out by using R 4.0.1 and with 100,000 replications. Section 4 contains applications of the suggested test in real examples.

2. The Lindley distribution and test Statistic

In this section, we express some properties of the Lindley distribution and then construct a goodness of fit test statistic for this distribution.

2.1 The Lindley Distribution

If the density function of the random variable X be as follows, then we say that X has a Lindley distribution.

$$f_0(x;\theta) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}, \qquad x > 0, \ \theta > 0.$$

Lindley distribution was proposed by Lindley (1958) in the context of Bayesian statistics, as a counter example of fiducial statistics. The cumulative distribution function of the Lindley distribution is as

$$F_0(x;\theta) = 1 - \frac{\theta+1+\theta x}{\theta+1}e^{-\theta x}.$$

The mean and variance of the distribution are

$$\mu = E(X) = \frac{\theta+2}{\theta(\theta+1)},$$

and

$$\sigma^2 = Var(X) = \frac{\theta^2 + 4\theta + 2}{\theta^2(\theta+1)^2}.$$

Ghitany et al. (2008) conducted a detailed study about various properties of Lindley distribution including skewness, kurtosis, hazard rate function, mean residual life function, stochastic ordering, stress-strength reliability, among other things; estimation of its parameter and application to model waiting time data in a bank.

In the literature of survival analysis and reliability theory, the exponential distribution is widely used as a model of lifetime data. However, the exponential distribution only provides a reasonable fit for modeling phenomenon with constant failure rates. Distributions like gamma, Weibull and lognormal have become suitable alternatives to the exponential distribution in many practical situations. Ghitany et al. (2008) found that the Lindley distribution can be a better model than one based on the exponential distribution.

The Lindley distribution belongs to an exponential family and it can be written as a mixture of an exponential with parameter θ and a gamma distribution with parameters $(2,\theta)$.

$$f_0(x;\theta) = pf_1(x) + (1-p)f_2(x) \qquad x > 0,$$

where $p = \theta/(1+\theta)$, $f_1(x) = \theta e^{-\theta x}$ and $f_2(x) = \theta^2 x e^{-\theta x}$.

Shanker et al. (2015) discussed a comparative study of Lindley and exponential distributions for modelling various lifetime data sets from biomedical science and engineering, and concluded that there are lifetime data where exponential distribution gives better fit than Lindley distribution and in majority of data sets Lindley distribution gives better fit than exponential distribution.

For computing the suggested test statistic, we need to estimate the unknown parameter θ. We apply the maximum likelihood estimate (MLE) approach to estimate the unknown parameter.

Suppose $X_1,...,X_n$ is a random sample from the Lindley distribution, the estimator for both maximum likelihood estimate (MLE) and method of moments estimate of the parameter θ is

$$\hat{\theta} = \frac{-(\bar{X}-1)+\sqrt{(\bar{X}-1)^2+8\bar{X}}}{2\bar{X}}, \qquad \bar{X} > 0.$$

Ghitany et al. (2008) showed that the estimator $\hat{\theta}$ of θ is positively biased: $E(\hat{\theta})-\theta > 0$, and it is consistent and asymptotically normal $\sqrt{n}(\hat{\theta}-\theta)\xrightarrow{D.} N(0,1/\sigma^2)$.

In complete sample case, Ghitany et al. (2008) developed different distributional properties, reliability characteristics and some inferential procedures for the Lindley distribution. Krishna and Kumar (2011) discussed reliability estimation in Lindley distribution with progressively type II right censored sample. Gupta and Singh (2013) gave parameter estimation of Lindley distribution with hybrid censored data. Also, Al-Mutairi et al. (2013) studied inferences on stress-strength reliability for Lindley distribution with complete sample information. Kumar et al. (2015) discussed estimation of stress-strength reliability using progressively first failure censoring. These studies suggest that in many real-life situations Lindley distribution serves as a better lifetime model than the so far popular distributions like exponential, gamma, Rayleigh, Weibull etc.

2.2 The Proposed Goodness-of-Fit Test

Assuming that $X_1,...,X_n$ is the sample from a distribution F, we wish to assess whether the unknown $F(x)$ can be satisfactorily approximated by a Lindley model $F_0(x)$.

The Kullback-Leibler (KL) discrimination has been widely studied in the literature as a central index for measuring quantitative similarity between two probability distributions. The KL discrimination of f from f_0 is defined by

$$D(f,f_0) = \int f(x)\log\frac{f(x)}{f_0(x)}\,dx. \tag{1}$$

Note that $D(f,f_0) = 0$ if and only if $f(x) = f_0(x)$ with probability 1.

Recently, Alizadeh Noughabi (2019) proposed a new estimate of (1) and then constructed a test statistic for testing the validity of a model. His test statistic is

$$T = -\frac{1}{n}\sum_{i=1}^{n}\log\left\{\frac{n}{2m}\Big(F_0(X_{(i+m)};\hat{\theta})-F_0(X_{(i-m)};\hat{\theta})\Big)\right\}, \tag{2}$$

where F_0 is the distribution function of f_0, m is a positive integer, $m \le n/2$, and $X_{(1)} \le X_{(2)} \le ... \le X_{(n)}$ are the order statistics and $X_{(i)} = X_{(1)}$ if $i < 1$, $X_{(i)} = X_{(n)}$ if $i > n$. Here, θ is a model parameter which is usually unknown, and $\hat{\theta}$ is a reasonable equivariant estimate of θ.

Alizadeh Noughabi (2019) showed that the test statistic is non-negative just like the Kullback-Leibler divergence, i.e., $T \ge 0$. Also, the test based on T is consistent. Then, he proposed tests for normal, exponential, Laplace and Weibull distributions and compared the power of these tests with the other existing tests and showed that his test has a good power against different alternatives. Moreover, Alizadeh Noughabi (2021) applied the above test statistic and proposed a new test for the logistic distribution. Here, we apply the Alizadeh Noughabi's test statistic and introduce a goodness of fit test for the Lindley distribution.

Suppose $X_1,...,X_n$ are a random sample from a continuous probability distribution F with density f. We are interested to test the hypothesis

$$H_0: f(x) = f_0(x;\theta) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}, \qquad \textit{for some } \theta \in \Theta,$$

against the general alternative

$$H_1: f(x) \ne f_0(x;\theta), \qquad \textit{for any } \theta,$$

where θ is specified or unspecified and $\Theta = \square^+$.

Based on the general statistic (2), we proposed the following test statistic for test of the Lindley distribution.

$$T = -\frac{1}{n}\sum_{i=1}^{n}\log\left\{\frac{n}{2m}\left(F_0(X_{(i+m)};\hat{\theta}) - F_0(X_{(i-m)};\hat{\theta})\right)\right\}, \tag{3}$$

where F_0 is the Lindley distribution function

$$F_0(x;\theta) = 1 - \frac{\theta+1+\theta x}{\theta+1}e^{-\theta x}, \tag{4}$$

and $\hat{\theta}$ is the maximum likelihood estimate (MLE) of the parameter θ, i.e.

$$\hat{\theta} = \frac{-(\bar{X}-1)+\sqrt{(\bar{X}-1)^2+8\bar{X}}}{2\bar{X}}. \tag{5}$$

We reject the null hypothesis for large values of (3). According to Alizadeh Noughabi (2019), the test statistic is non-negative, i.e., $T \geq 0$, and also the test based on T is consistent.

3. Critical points and power comparison
Because deriving the exact distribution of the test statistic is complicated, we obtain the critical values of the test statistic by Monte Carlo simulations. These values for different sample sizes are presented in Table 1.
It should be mentioned that the empirical percentiles given in Table 1 provides an excellent type I error control.

Table 1: Critical values of the proposed test statistic at level 5%

					n				
10	20	30	40	50	60	70	80	90	100
0.6945	0.4440	0.3780	0.3321	0.3046	0.2959	0.2807	0.2699	0.2684	0.2606

For power comparison purpose, we consider the popular and common tests which are used in practice and statistical software. The test statistics of these tests are briefly described as follows. For more details about these tests, see D'Agostino and Stephens (1986).
Let $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ are the order statistics based on the random sample $X_1,...,X_n$.

1. The Cramer-von Mises statistic (1931):

$$W^2 = \frac{1}{12n} + \sum_{i=1}^{n}\left(\frac{2i-1}{2n} - F_0(X_{(i)};\hat{\theta})\right)^2.$$

2. The Watson statistic (1961):

$$U^2 = W^2 - n(\bar{P}-0.5)^2,$$

    where $\bar{P}$ is the mean of $F_0(X_{(i)};\hat{\theta}),\ i=1,...,n$.

3. The Kolmogorov-Smirnov statistic (1933):

$$D = \max(D^+, D^-).$$

where

$$D^+ = \max_{1\leq i\leq n}\left\{\frac{i}{n} - F_0(X_{(i)};\hat{\theta})\right\};\ D^- = \max_{1\leq i\leq n}\left\{F_0(X_{(i)};\hat{\theta}) - \frac{i-1}{n}\right\}.$$

4. The Kuiper statistic (1960):

$$V = D^+ + D^-.$$

5. The Anderson-Darling statistic (1952):

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left\{\log F_0(X_{(i)};\hat{\theta}) + \log\left[1 - F_0(X_{(n-i+1)};\hat{\theta})\right]\right\}.$$

In the above test statistics, $F_0(\mathrm{x})$ is the cumulative distribution function of the Lindley distribution and $\hat{\theta}$ is the maximum likelihood estimate of the parameter θ, which are presented in equations (4) and (5), respectively. It is obvious that for large values of the above test statistics the null hypothesis H_0 will be rejected.

Table 2: Empirical powers of the tests against IFR alternatives at significance level 5%.

Alternative	n	W^2	D	V	U^2	A^2	T
$W(1.4)$	10	0.1303	0.1174	0.1104	0.1170	0.0894	0.1882
	20	0.2258	0.1966	0.1761	0.1884	0.1917	0.3261
	30	0.3237	0.2691	0.2330	0.2635	0.2967	0.4687
	50	0.5098	0.4231	0.3736	0.4167	0.5036	0.6651
$\Gamma(2)$	10	0.1175	0.1028	0.1101	0.1188	0.0810	0.1987
	20	0.2011	0.1754	0.1772	0.1935	0.1800	0.3590
	30	0.2879	0.2412	0.2369	0.2687	0.2827	0.5195
	50	0.4745	0.4014	0.3875	0.4408	0.5104	0.7378
HN	10	0.0952	0.0887	0.0844	0.0875	0.0678	0.1187
	20	0.1364	0.1234	0.1084	0.1149	0.1076	0.1745
	30	0.1835	0.1552	0.1340	0.1446	0.1492	0.2360
	50	0.2839	0.2321	0.1960	0.2139	0.2445	0.3486
U	10	0.3386	0.2647	0.3088	0.2957	0.2615	0.3944
	20	0.6318	0.4888	0.6071	0.5477	0.5793	0.7617
	30	0.8309	0.6764	0.8143	0.7416	0.8056	0.9273
	50	0.9756	0.9000	0.9777	0.9417	0.9756	0.9969
$CH(1)$	10	0.0937	0.0868	0.0772	0.0789	0.0673	0.1093
	20	0.1364	0.1220	0.0998	0.1061	0.1074	0.1604
	30	0.1826	0.1557	0.1230	0.1332	0.1477	0.2205
	50	0.2796	0.2301	0.1810	0.1933	0.2379	0.3316
$CH(1.5)$	10	0.4268	0.3505	0.3359	0.3553	0.3348	0.4992
	20	0.7600	0.6343	0.6239	0.6480	0.7160	0.8351
	30	0.9200	0.8205	0.8176	0.8370	0.9071	0.9622
	50	0.9943	0.9684	0.9736	0.9763	0.9943	0.9985
$LF(2)$	10	0.1386	0.1235	0.1113	0.1187	0.0972	0.1699
	20	0.2282	0.1943	0.1706	0.1802	0.1851	0.2706
	30	0.3292	0.2723	0.2327	0.2527	0.2828	0.3855
	50	0.5133	0.4204	0.3663	0.3955	0.4662	0.5687
$LF(4)$	10	0.2056	0.1790	0.1594	0.1700	0.1469	0.2437
	20	0.3777	0.3160	0.2752	0.2980	0.3192	0.4164
	30	0.5308	0.4386	0.3864	0.4204	0.4758	0.5783
	50	0.7680	0.6595	0.6067	0.6401	0.7313	0.7976
$EV(0.5)$	10	0.0923	0.0861	0.0749	0.0782	0.0670	0.1113
	20	0.1384	0.1221	0.1020	0.1074	0.1068	0.1607
	30	0.1833	0.1557	0.1242	0.1345	0.1467	0.2210
	50	0.2779	0.2262	0.1803	0.1933	0.2378	0.3273
$EV(1.5)$	10	0.0923	0.0861	0.0749	0.0782	0.0670	0.2200
	20	0.1384	0.1221	0.1020	0.1074	0.1068	0.3937
	30	0.1833	0.1557	0.1242	0.1345	0.1467	0.5611
	50	0.2779	0.2262	0.1803	0.1933	0.2378	0.8023

Table 3: Empirical powers of the tests against UFR, DFR and BFR alternatives at level 5%.

Alternative	n	W^2	D	V	U^2	A^2	T
$LN(0.8)$	10	0.1413	0.1302	0.1279	0.1403	0.1068	0.2004
	20	0.2221	0.1968	0.2204	0.2448	0.2110	0.4034
	30	0.3180	0.2720	0.3268	0.3652	0.3440	0.6224
	50	0.5147	0.4436	0.5541	0.6054	0.6131	0.8784
$LN(1.5)$	10	0.5140	0.4823	0.3849	0.4001	0.5544	0.0643
	20	0.8027	0.7664	0.6690	0.6869	0.8197	0.2057
	30	0.9257	0.9020	0.8342	0.8489	0.9306	0.3079
	50	0.9900	0.9842	0.9642	0.9697	0.9905	0.5571
$DL(1)$	10	0.0877	0.0813	0.0809	0.0862	0.0629	0.1228
	20	0.1185	0.1064	0.1139	0.1236	0.1041	0.1941
	30	0.1486	0.1274	0.1445	0.1619	0.1445	0.2826
	50	0.2123	0.1771	0.2245	0.2533	0.2394	0.4268
$DL(1.5)$	10	0.1999	0.1735	0.1751	0.1937	0.1462	0.3207
	20	0.3844	0.3271	0.3228	0.3634	0.3601	0.5843
	30	0.5568	0.4783	0.4598	0.5241	0.5677	0.7873
	50	0.8123	0.7363	0.7129	0.7832	0.8509	0.9473
$W(0.8)$	10	0.1960	0.1750	0.1288	0.1366	0.2748	0.0077
	20	0.3570	0.3095	0.2295	0.2438	0.4417	0.0040
	30	0.4933	0.4319	0.3201	0.3476	0.5752	0.0024
	50	0.7062	0.6330	0.5093	0.5395	0.7720	0.0022
$\Gamma(0.4)$	10	0.5137	0.4712	0.3701	0.3914	0.7163	0.0106
	20	0.8109	0.7663	0.6579	0.6850	0.9222	0.0374
	30	0.9354	0.9074	0.8310	0.8551	0.9810	0.0427
	50	0.9943	0.9894	0.9697	0.9762	0.9990	0.1099
$CH(0.5)$	10	0.3912	0.3546	0.2711	0.2860	0.5728	0.0051
	20	0.6670	0.6127	0.4979	0.5281	0.8141	0.0105
	30	0.8331	0.7839	0.6733	0.7102	0.9251	0.0096
	50	0.9669	0.9464	0.8924	0.9137	0.9903	0.0234

Table 4: Powerful tests against different alternatives

IFR	UFR	DFR-BFR
T	T & A^2	A^2

By Monte Carlo simulations, power of the tests against various alternatives are evaluated. The following alternatives are considered in power comparison.

- the Weibull distribution with density $\theta x^{\theta-1}\exp\left(-x^\theta\right)$, denoted by $W(\theta)$,

- the gamma distribution with density $\Gamma(\theta)^{-1}x^{\theta-1}\exp(-x)$, denoted by $\Gamma(\theta)$,

- the lognormal distribution $LN(\theta)$ with density
  $(\theta x)^{-1}(2\pi)^{-1/2}\exp\left(-(\log x)^2/(2\theta^2)\right)$,

- the half-normal HN distribution with density $\Gamma(2/\pi)^{1/2}\exp\left(-x^2/2\right)$,

- the uniform distribution U with density $1,\ \ 0\le x\le 1$,

- the modified extreme value $EV(\theta)$, with distribution function

$$1 - \exp\left(\theta^{-1}(1 - e^x)\right),$$

- the linear increasing failure rate law $LF(\theta)$ with density

$$(1 + \theta x)\exp\left(-x - \theta x^2/2\right),$$

- Dhillon's (1981) distribution $DL(\theta)$ with distribution function

$$1 - \exp\left(-\left(\log(x+1)\right)^{\theta+1}\right),$$

- Chen's (2000) distribution $CH(\theta)$, with distribution function $1 - \exp\left(2\left(1 - e^{x^\theta}\right)\right).$

These alternatives include densities f with decreasing failure rates (DFR), increasing failure rates (IFR) as well as models with unimodal failure rate (UFR) functions and bathtub failure rate (BFR) functions.

To assess the power values of the tests, we generate 100,000 random samples from the alternative hypothesis for different choices of sample sizes and then the test statistics are calculated. Then power of the corresponding test is computed by the frequency of the event ''the statistic is in the critical region''. Tables 2 and 3 display and compares the power values of the tests for sample sizes $n = 10, 20, 30, 50$ at the significance level $\alpha = 0.05$.

Tables 2 reveals a superiority of the proposed test based on T statistic to all competing tests as we can say that this test outperforms all other tests against IFR alternatives. The power differences between the proposed test and the other tests are substantial.

Based on the power values in Table 3, it is seen that the proposed test based on T statistic has the most power against UFR alternatives and just for the $LN(1.5)$ alternative, the test A^2 has the most power. The power differences between these tests and the other tests are substantial.

From Table 3, it is evident that the test based on A^2 statistic has the most power against DFR and BFR alternatives and power differences between this test and the other tests are substantial.

Although there is no uniformly most powerful test against all alternatives, the tests based on T and A^2 statistics can be recommended in practice. In general, we can conclude that the proposed test T and also the Anderson-Darling test A^2 have a good performance and therefore can be used in practice. Finally, we summarized the results in Table 4. This table presents the best test in terms of power against different alternatives.

4. Illustration with real data

We illustrate, by two real examples, how the proposed test can be applied to test the goodness-of-fit for the Lindley distribution when a random sample is available.

Example 1. We use the data set of waiting times (in minutes) before service of 100 bank customers as discussed by Ghitany et al. (2008). The waiting times (in minutes) are as follows:
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, 4.0, 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, 8.0, 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, 23.0, 27.0, 31.6, 33.1, 38.5.

Krishna and Kumar (2011) considered four reliability models, namely exponential, Lindley, gamma, and lognormal. According to Bayesian information criterion (BIC), they found that the Lindley model is the best fit for these data. Thus, Lindley distribution is fitting the above data quite satisfactorily. The main advantage of using Lindley distribution over gamma and lognormal distributions is that it involves only one parameter. Hence, maximum likelihood and other inferential procedures become simple to deal with, especially from computational point of view.

Here, we apply the proposed procedure to this data set. First, the ML estimator of θ is computed as:

$$\hat{\theta} = 0.1866.$$

Then, the value of test statistic is computed and also the critical value of the test at the significance level 0.05 is obtained from Table 1. Moreover, the values of competing test statistics are computed. Results are summarized in Table 5.

Table 5: The value of the test statistics and critical values at 5% level.

Test	Value of the test statistic	Critical value	Decision
W^2	0.0582	0.2032	Not reject H_0
D	0.0677	0.1042	Not reject H_0
V	0.1094	0.1624	Not reject H_0
U^2	0.0551	0.1573	Not reject H_0
A^2	0.4863	1.242	Not reject H_0
T	0.2546	0.2606	Not reject H_0

Because the value of each test statistic is smaller than the corresponding critical value, the Lindley hypothesis is accepted for these data at the significance level of 0.05. Therefore, we can conclude that the underlying distribution of these data is a Lindley distribution.

Example 2. The following data are 15 electronic components in an accelerated life test, presented by Lawless (1982): 1.4, 5.1, 6.3, 10.8, 12.1, 18.5, 19.7, 22.2, 23.0, 30.6, 37.3, 46.3, 53.9, 59.8, 66.2.

The proposed test can be used to investigate whether the data come from a Lindley distribution. The values of the proposed test statistic and also other test statistics are as

$$W^2 = 0.0375, \ D = 0.1103, \ V = 0.2184,$$

$$U^2 = 0.0371, A^2 = 0.3187, T = 0.2523,$$

and at 5% significance level, the critical values of the tests are 0.2006, 0.2596, 0.4038, 0.1561, 1.2158, 3.6671, 14.535 and 0.5443, respectively. Therefore, the tests accept the null hypothesis that the electronic components follow a Lindley model at significance level of 0.05.

5. Conclusions

In this paper, we have proposed a powerful goodness-of-fit tests for the Lindley distribution and have shown that the suggested test has a good performance. Through Monte Carlo simulations, we have carried out an extensive power study on the suggested test with the competing tests. It is shown that our test outperforms in most cases all other tests. Finally, we have used two real data sets and have illustrated how the proposed test can be applied to test the goodness-of-fit for the Lindley distribution when a random sample is available.

References

1. Ali, S., Aslam, M., Kazmi, S. (2013). A study of the effect of the loss function on Bayes Estimate, posterior risk and hazard function for Lindley distribution. Applied Mathematical Modelling, 37, 6068–6078.
2. Alizadeh Noughabi, H. (2021). A new goodness of fit test for the logistic distribution. Sankhya B, https://doi.org/10.1007/s13571-021-00262-1.
3. Alizadeh Noughabi, H. (2019). A new estimator of Kullback–Leibler information and its application in goodness of fit tests. Journal of Statistical Computation and Simulation, 89, 1914-1934.
4. Alizadeh Noughabi, H. and Balakrishnan, N. (2015). Goodness of fit using a new estimate of Kullback-Leibler information based on type II censored data. IEEE Transactions on Reliability, 64, 627-635.
5. Al-Mutairi, D.K., Ghitany, M.E., Kundu, D. (2013). Inferences on the stress-strength reliability from Lindley distributions. Communications in Statistics-Theory and Methods, 42, 1443–1463.
6. Altun, E. (2019). Two-sided Lindley distribution with inference and applications. Journal of the Indian Society for Probability and Statistics, 20, 255–279.
7. Anderson, T.W. and Darling, D.A. (1954). A test of goodness of fit. Journal of American Statistical Association, 49, 765-769.

8.   Balakrishnan, N., Habibi Rad, A. and Arghami, N.R. (2007). Testing exponentiality based on Kullback–Leibler information with progressively type-II censored data. IEEE Transactions on Reliability, 56, 301–307.

9.   Balakrishnan, N., Ng, H.K.T. and Kannan, N. (2004). Goodness-of-fit tests based on spacings for progressively type-II censored data from a general location-scale distribution. IEEE Transactions on Reliability, 53, 349–356.

10. D'Agostino, R.B. and Stephens, M.A. (Eds.) (1986). Goodness-of-fit Techniques. New York: Marcel Dekker.

11. Ghitany, M.E., Atieh, B. and Nadarajah, S. (2008). Lindley distribution and its application. Mathematics Computing and Simulation, 78, 493-506.

12. Gupta, PK, Singh, B. (2013). Parameter estimation of Lindley distribution with hybrid censored data. International Journal of System Assurance Engineering and Management, 4, 378–385.

13. Habibi Rad, A., Yousefzadeh, F. and Balakrishnan, N. (2011). Goodness-of-fit test based on Kullback–Leibler information for progressively Type-II censored data. IEEE Transactions on Reliability, 60, 570–579.

14. Huber-Carol, C., Balakrishnan, N., Nikulin, M. S. and Mesbah, M. (2002). Goodness-of-fit tests and model validity. Boston, Basel, Berlin: Birkhäuser.

15. Ibrahim M., Yadav, A.S., Yousof, H.M., Goual, H., Hamedani, G. (2019). A new extension of Lindley distribution: modified validation test, characterizations and different methods of estimation. Communications for Statistical Applications and Methods, 26:473-495.

16. Kolmogorov, A.N. (1933). Sulla Determinazione Empirica di une legge di Distribuzione. Giornale dell'Intituto Italiano degli Attuari, 4, 83-91.

17. Krishna, H. and Kumar, K. (2011). Reliability estimation in Lindley distribution with progressively type II right censored sample. Mathematics and Computers in Simulation, 82, 281–294.

18. Kuiper, N.H. (1960). Tests concerning random points on a circle. Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Series A, 63, 38-47.

19. Kumar, C.S. and Jose, R. (2019). On double Lindley distribution and some of its properties. American Journal of Mathematical and Management Sciences, 38, 23-43.

20. Kumar, K., Krishna, H. and Garg, R. (2015). Estimation of P(Y < X) in Lindley distribution using progressively first failure censoring. International Journal of System Assurance Engineering and Management, 6, 330–341.

21. Lawless, J.F. (1982). Statistical models and methods for lifetime data. In: John Wiley & Sons (Ed.), New York, USA.

22. Lin, C-T., Huang, Y-L. and Balakrishnan, N. (2008). A new method for goodness-of-fit testing based on Type-II right censored samples. IEEE Transactions on Reliability, 57, 633–642.

23. Lindley, D.V. (1958). Fiducial distribution and Bayes' theorem. Journal of the Royal Statistical Society, 20, 102–107.

24. Mazucheli, J. and Achcar, J.A. (2011). The Lindley distribution applied to competing risks lifetime data. Computer Methods and Programs in Biomedicine, 2011, 104, 188- 192.

25. Pakyari, R. and Balakrishnan, N. (2012). A general purpose approximate goodness-of-fit test for progressively Type-II censored data. IEEE Transactions on Reliability, 61, 238–244.

26. Pakyari, R. and Balakrishnan, N. (2013). Goodness-of-fit tests for progressively Type-II censored data from location-scale distributions. Journal of Statistical Computation and Simulation, 83, 167-178.

27. Shanker, R., Hagos, F., Sujatha, S. (2015). On modeling of Lifetimes data using exponential and Lindley distributions. Biometrics & Biostatistics International Journal, 2, 1–9.

28. Singh, B. and Gupta, P.K. (2012). Load-sharing system model and its application to the real data set. Mathematics and Computers in Simulation, 82, 1615–1629.

29. Valiollahi, R., Asgharzadeh, A. and Ng, H.K.T. (2017). Prediction for Lindley distribution based on type-II right censored samples. Journal of The Iranian Statistical Society, 16, 1-19.

30. von Mises, R. (1931). Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik. Leipzig and Vienna: Deuticke.

31. Wang, M. (2013). A new three-parameter lifetime distribution and associated inference. arXiv:1308.4128 [stat.ME].

32. Watson, G.S. (1961). Goodness of fit tests on a circle. Biometrika, 48, 109-114.