STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **11**, June 2023, pp 719–729.
Published online in International Academic Press (www.IAPress.org)

On Testing the Adequacy of the Lindley Model and Power Study

Hadi Alizadeh Noughabi [1,*], Mohammad Shafaei Noughabi [2]

[1] *Department of Statistics, University of Birjand, Birjand, Iran*
[2] *Department of Mathematics and Statistics, University of Gonabad, Gonabad, Iran*

Abstract The Lindley distribution may serve as a useful reliability model. Applications of this distribution are presented in statistical literature. In this article, goodness of fit tests for the Lindley distribution based on the empirical distribution function (EDF) are considered. In order to compute the test statistics, we use the maximum likelihood estimate (MLE) suggested by Ghitany et al. (2008), which is simple explicit estimator. Critical points of the proposed test statistics are obtained by Monte Carlo simulation. Power comparisons of the considered tests are carried out via simulations. Finally, two illustrative examples are presented and analyzed.

Keywords Empirical distribution function, Model validity, Goodness of fit tests, Lindley distribution, Monte Carlo simulation, Power study.

AMS 2010 subject classifications 62G10, 62G30

DOI: 10.19139/soic-2310-5070-1443

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), Ibrahim et al. (2019), Chesneau et al. (2021), and Tomy et al. (2021a,b) among others. Also, a retrospective study on Lindley distribution can be found in Tomy (2018).
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.

*Correspondence to: Hadi Alizadeh Noughabi (Email: alizadehhadi@birjand.ac.ir). Department of Statistics, University of Birjand, Birjand, Iran.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

Therefore, it is a clear need to check whether the Lindley model is a satisfactory model for the observations.
A Goodness of Fit (GOF) test determines how well your analysis data fits the calculated data model.
• If your data passes the GOF test, your data follows the data model closely, and you can rely on the predictions made by model.
• If your data fails the GOF test, it may not follow the model closely enough to confidently rely on model predictions.
A statistical hypothesis test uses to determine whether a variable is likely to come from a specified distribution or not. It is often used to evaluate whether sample data is representative of the full population.
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. These tests are valid when there are no unknown parameters in the hypothesized distribution. These tests become extremely conservative if they are used in case where unknown parameters must be estimated from the sample data.
Goodness of fit tests based on the empirical distribution function (EDF) are well-known in the literature and commonly used in practice and statistical Software. The known EDF-tests are Cramer-von Mises (W^2), Kolmogorov-Smirnov (D), Kuiper (V), Watson (U^2), and Anderson-Darling (A^2). For more details about these tests, see D'Agostino and Stephens (1986).
Many researchers have been interested in goodness of fit tests for different distributions and then different tests are developed in the literature. For example, see D'Agostino and Stephen (1986), Chen and Balakrishnan (1995), Huber-Carol et al. (2002), He and Xu (2013), Alizadeh and Chahkandi (2015), Alizadeh (2015,2016) and Jahanshahi et al. (2016). 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 and Balakrishnan (2015).
Moreover, Zhang (2002) introduced three goodness of fit test statistics based on the empirical distribution function and applied them for testing normality and showed that the new tests have higher power than the competing tests. In the present paper, we will apply these test statistics to test the hypothesis of the Lindley distribution and compare the power of these tests with the other tests.
The main contribution of the paper can express as follows. In this paper, we apply EDF-tests for the Lindley distribution. Moreover, the method of Zhang (2002) is stated and based on this method, we propose three goodness of fit tests for the Lindley distribution. Table of critical values and properties of the tests are presented. We show through extensive simulation studies that the proposed goodness-of-fit tests are 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 tests for the Lindley model with real data.
In Section 2, a summary of the Lindley distribution is presented. In Section 3, we consider goodness of fit test statistics based on the empirical distribution function and apply them for the Lindley distribution. In Section 4, the critical values of the test statistics are obtained by Monte Carlo simulations. Then power values of the tests are computed and then compared with each other. All simulations were carried out by using R 4.1.1 and with 100,000 replications. Section 5 contains applications of the tests in real examples.

2. A Summary of 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 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.

Since in EDF-based test statistics, we need to estimate the 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

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

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 $\sqrt{n}\left(\hat{\theta} - \theta\right) \to N(0, 1/\sigma^2)$.
We will use the ML estimator for the EDF-test statistics to test the goodness-of-fit for the Lindley distribution.
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.

3. GOF Tests for the Lindley Distribution

The GOF test checks whether our sample data is likely to be from a specific theoretical distribution. We have a set of data values, and an idea about how the data values are distributed. The test gives us a way to decide if the data values have a "good enough" fit to our idea, or if our idea is questionable. GOF tests are designed to measure how well the observed sample data fits some proposed model.
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

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

against the general alternative

$$H_1: \ f(x) \neq f_0(x;\theta), \qquad for\ any\ \theta,$$

where θ is specified or unspecified and $\Theta = R^+$.
Here, 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):

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

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

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

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, $F_0(x)$ is the cumulative distribution function of the Lindley distribution and $\hat{\theta}$ is the maximum likelihood estimate of the parameter θ.

Moreover, we consider the EDF-based tests proposed by Zhang (2002). Briefly, the approach of Zhang (2002) for the Lindely distribution is described as follows. Let

$$H_t: \ F(t) = F_0(t) = 1 - \frac{\theta + 1 + \theta x}{\theta + 1} e^{-\theta x},$$

and $\bar{H}_t: \ F(t) \neq F_0(t)$. According Zhang (2002), testing H vs $\bar{H}$ is equivalent to testing H_t vs $\bar{H}_t$ for every $t \in (0, \infty)$ in the sense that

$$H = \bigcap_{t \in (0,\infty)} H_t$$

and

$$\bar{H} = \bigcup_{t \in (0,\infty)} \bar{H}_t \ .$$

Now, define a binary random sample to test H_t vs $\bar{H}_t$ for each t;

$$X_{it} = I\left(X_i \leq t\right) \qquad i = 1, 2, ..., n,$$

where $P\left(X_{it} = 1\right) = F(t)$ and $P\left(X_{it} = 0\right) = 1 - F(t)$.

Let Z_t denotes a statistic based on X_{it} for testing H_t vs $\bar{H}_t$ where large values of Z_t reject H_t. For testing H_t vs $\bar{H}_t$, Zhang (2002) proposed two test statistics given by

$$Z = \int Z_t dw(t) \ \ and \ \ Z_{\max} = \sup_{t \in (0,\infty)} \left[Z_t w(t)\right],$$

where $w(t)$ is some weight function. Also, large values of these statistics reject H. Zhang (2002) for Z_t considered Pearson's Chi squared statistic

$$X_t^2 = \frac{n[F_n(t) - F_0(t)]^2}{F_0(t)\left[1 - F_0(t)\right]},$$

and the likelihood ratio statistic

$$G_t^2 = 2n \left\{ F_n(t) \log \frac{F_n(t)}{F_0(t)} + [1 - F_n(t)] \log \frac{1 - F_n(t)}{1 - F_0(t)} \right\},$$

where $F_n(t)$ is the empirical distribution function.
If we set $Z_t = X_t^2$ with

$$w(t) = n^{-1} F_0(t) \left[1 - F_0(t)\right], \ dw(t) = n^{-1} F_0(t) \left[1 - F_0(t)\right] dF_0(t),$$

and $w(t) = F_0(t)$, then the traditional Kolmogorov–Smirnov, Cramer–von Mises and Anderson–Darling statistics are obtained.
Moreover, if we consider $Z_t = G_t^2$ with $w(t) = 1$, $dw(t) = F_0(t)^{-1}[1 - F_0(t)]^{-1} dF_0(t)$ and $dw(t) = F_n(t)^{-1}[1 - F_n(t)]^{-1} dF_n(t)$, respectively, and further, $F_n(X_{(i)}) = \frac{i - 0.5}{n}$, then the test statistics proposed by Zhang (2002) are obtained. These test statistics for the Lindley distribution are as

$$Z_A = -\sum_{i=1}^{n} \left(\frac{\log F_0(X_{(i)}; \hat{\theta})}{n - i + 0.5} + \frac{\log \left[1 - F_0(X_{(i)}; \hat{\theta})\right]}{i - 0.5} \right),$$

$$Z_C = \sum_{i=1}^{n} \left(\log \left\{ \frac{F_0(X_{(i)}; \hat{\theta})^{-1} - 1}{(n - 0.5)/(i - 0.75) - 1} \right\} \right)^2,$$

$$Z_K = \max_{1 \leq i \leq n} \left((i - 0.5) \log \left\{ \frac{i - 0.5}{n F_0(X_{(i)}; \hat{\theta})} \right\} + (n - i + 0.5) \log \left\{ \frac{n - i + 0.5}{n(1 - F_0(X_{(i)}; \hat{\theta}))} \right\} \right).$$

It is obvious that for large values of the above test statistics the null hypothesis will be rejected. The test statistics are invariant under any affine transformation on the sample data. Therefore, they are distribution-free within the Lindley distribution family.

4. Critical Points and Power Comparison

Because deriving the exact distribution of the test statistics are complicated, we obtain the critical values of the test statistics 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 statistics at level 0.05%

n	W^2	D	V	U^2	A^2	Z_A	Z_C	Z_k
10	0.1994	0.3140	0.4879	0.1560	1.203	3.778	13.057	1.993
20	0.2017	0.2270	0.3526	0.1568	1.224	3.599	15.527	2.463
30	0.2028	0.1876	0.2920	0.1576	1.238	3.522	17.016	2.729
40	0.2019	0.1627	0.2538	0.1569	1.242	3.477	18.031	2.880
50	0.2030	0.1464	0.2276	0.1569	1.245	3.449	18.924	3.018
60	0.2019	0.1335	0.2083	0.1572	1.236	3.427	19.426	3.105
70	0.2033	0.1244	0.1934	0.1578	1.243	3.412	20.034	3.191
80	0.2032	0.1165	0.1813	0.1573	1.246	3.399	20.422	3.280
90	0.2027	0.1098	0.1713	0.1574	1.243	3.389	20.780	3.312
100	0.2032	0.1042	0.1624	0.1573	1.242	3.382	21.179	3.361

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 $(\theta x)^{-1}(2\pi)^{-1/2}\exp\left(-(\log x)^2\Big/(2\theta^2)\right)$;
- the half-normal HN distribution with density $\Gamma(2/\pi)^{1/2}\exp\left(-x^2/2\right)$;
- the uniform distribution U with density $1, \quad 0 \le x \le 1$;
- 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 with distribution function $1-\exp\left(-(\log(x+1))^{\theta+1}\right)$;
- 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 $\alpha = 0.05$.
For each sample size and alternative, the bold type in these tables indicates the tests achieving the maximal power.

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

$Alternative$	n	W^2	D	V	U^2	A^2	Z_A	Z_C	Z_K
$W(1.4)$	10	0.1303	0.1174	0.1104	0.1170	0.0894	**0.1421**	0.1278	0.0926
	20	0.2258	0.1966	0.1761	0.1884	0.1917	**0.2755**	0.2431	0.1739
	30	0.3237	0.2691	0.2330	0.2635	0.2967	**0.3891**	0.3491	0.2491
	50	0.5098	0.4231	0.3736	0.4167	0.5036	**0.5846**	0.5404	0.4088
$\Gamma(2)$	10	0.1175	0.1028	0.1101	0.1188	0.0810	**0.1474**	0.1280	0.0864
	20	0.2011	0.1754	0.1772	0.1935	0.1800	**0.3073**	0.2592	0.1889
	30	0.2879	0.2412	0.2369	0.2687	0.2827	**0.4540**	0.3873	0.2909
	50	0.4745	0.4014	0.3875	0.4408	0.5104	**0.6815**	0.6155	0.5063
HN	10	**0.0952**	0.0887	0.0844	0.0875	0.0678	0.0919	0.892	0.0701
	20	0.1364	0.1234	0.1084	0.1149	0.1076	**0.1371**	0.1268	0.0885
	30	**0.1835**	0.1552	0.1340	0.1446	0.1492	0.1787	0.1608	0.1054
	50	**0.2839**	0.2321	0.1960	0.2139	0.2445	0.2777	0.2389	0.1531
U	10	**0.3386**	0.2647	0.3088	0.2957	0.2615	0.2876	0.2893	0.1707
	20	0.6318	0.4888	0.6071	0.5477	0.5793	**0.6444**	0.5826	0.3495
	30	0.8309	0.6764	0.8143	0.7416	0.8056	**0.9004**	0.8174	0.6838
	50	0.9756	0.9000	0.9777	0.9417	0.9756	**0.9986**	0.9867	0.9898
$CH(1)$	10	**0.0937**	0.0868	0.0772	0.0789	0.0673	0.0867	0.0835	0.0681
	20	**0.1364**	0.1220	0.0998	0.1061	0.1074	0.1294	0.1209	0.0873
	30	**0.1826**	0.1557	0.1230	0.1332	0.1477	0.1723	0.1546	0.1010
	50	**0.2796**	0.2301	0.1810	0.1933	0.2379	0.2659	0.2278	0.1451
$CH(1.5)$	10	0.4268	0.3505	0.3359	0.3553	0.3348	0.4089	0.4035	0.2708
	20	0.7600	0.6343	0.6239	0.6480	0.7160	**0.7607**	0.7502	0.5319
	30	0.9200	0.8205	0.8176	0.8370	0.9071	**0.9287**	0.9180	0.7391
	50	0.9943	0.9684	0.9736	0.9763	0.9943	**0.9968**	0.9950	0.9521
$LF(2)$	10	**0.1386**	0.1235	0.1113	0.1187	0.0972	0.1309	0.1255	0.0913
	20	**0.2282**	0.1943	0.1706	0.1802	0.1851	0.2192	0.2039	0.1393
	30	**0.3292**	0.2723	0.2327	0.2527	0.2828	0.3091	0.2830	0.1892
	50	**0.5133**	0.4204	0.3663	0.3955	0.4662	0.4748	0.4285	0.2968
$LF(4)$	10	**0.2056**	0.1790	0.1594	0.1700	0.1469	0.1922	0.1845	0.1309
	20	**0.3777**	0.3160	0.2752	0.2980	0.3192	0.3479	0.3302	0.2320
	30	**0.5308**	0.4386	0.3864	0.4204	0.4758	0.4877	0.4567	0.3225
	50	**0.7680**	0.6595	0.6067	0.6401	0.7313	0.7172	0.6766	0.5190
$EV(0.5)$	10	**0.0923**	0.0861	0.0749	0.0782	0.0670	0.0872	0.0837	0.0688
	20	**0.1384**	0.1221	0.1020	0.1074	0.1068	0.1291	0.1198	0.0881
	30	**0.1833**	0.1557	0.1242	0.1345	0.1467	0.1706	0.1544	0.1040
	50	**0.2779**	0.2262	0.1803	0.1933	0.2378	0.2650	0.2289	0.1437
$EV(1.5)$	10	**0.1681**	0.1456	0.1420	0.1547	0.1170	0.1488	0.1472	0.1024
	20	**0.3359**	0.2706	0.2529	0.2634	0.2658	0.3038	0.2932	0.1825
	30	0.4612	0.3645	0.3618	0.3805	0.4152	**0.4725**	0.4380	0.2624
	50	0.7218	0.5906	0.5811	0.5943	0.6880	**0.7300**	0.6495	0.4412

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

$Alternative$	n	W^2	D	V	U^2	A^2	Z_A	Z_C	Z_K
$LN(0.8)$	10	0.1413	0.1302	0.1279	0.1403	0.1068	**0.1726**	0.1430	0.1126
	20	0.2221	0.1968	0.2204	0.2448	0.2110	**0.4129**	0.3128	0.2564
	30	0.3180	0.2720	0.3268	0.3652	0.3440	**0.6510**	0.5054	0.4371
	50	0.5147	0.4436	0.5541	0.6054	0.6131	**0.9127**	0.8086	0.7546
$LN(1.5)$	10	0.5140	0.4823	0.3849	0.4001	**0.5544**	0.4511	0.4371	0.4706
	20	0.8027	0.7664	0.6690	0.6869	**0.8197**	0.7253	0.7356	0.7405
	30	0.9257	0.9020	0.8342	0.8489	**0.9306**	0.8747	0.8856	0.8820
	50	0.9900	0.9842	0.9642	0.9697	**0.9905**	0.9756	0.9793	0.9764
$DL(1)$	10	0.0877	0.0813	0.0809	0.0862	0.0629	**0.0982**	0.0835	0.0630
	20	0.1185	0.1064	0.1139	0.1236	0.1041	**0.1945**	0.1539	0.1152
	30	0.1486	0.1274	0.1445	0.1619	0.1445	**0.2845**	0.2241	0.1712
	50	0.2123	0.1771	0.2245	0.2533	0.2394	**0.4588**	0.3649	0.3051
$DL(1.5)$	10	0.1999	0.1735	0.1751	0.1937	0.1462	**0.2491**	0.2201	0.1558
	20	0.3844	0.3271	0.3228	0.3634	0.3601	**0.5408**	0.2201	0.1558
	30	0.5568	0.4783	0.4598	0.5241	0.5677	**0.7416**	0.6787	0.5611
	50	0.8123	0.7363	0.7129	0.7832	0.8509	**0.9333**	0.9031	0.8335
$W(0.8)$	10	0.1960	0.1750	0.1288	0.1366	**0.2748**	0.1959	0.1975	0.2262
	20	0.3570	0.3095	0.2295	0.2438	**0.4417**	0.3197	0.3401	0.3610
	30	0.4933	0.4319	0.3201	0.3476	**0.5752**	0.4330	0.4650	0.4724
	50	0.7062	0.6330	0.5093	0.5395	**0.7720**	0.6341	0.6630	0.6632
$\Gamma(0.4)$	10	0.5137	0.4712	0.3701	0.3914	**0.7163**	0.6698	0.6776	0.6941
	20	0.8109	0.7663	0.6579	0.6850	**0.9222**	0.8973	0.9060	0.9055
	30	0.9354	0.9074	0.8310	0.8551	**0.9810**	0.9722	0.9748	0.9727
	50	0.9943	0.9894	0.9697	0.9762	**0.9990**	0.9984	0.9986	0.9980
$CH(0.5)$	10	0.3912	0.3546	0.2711	0.2860	**0.5728**	0.5048	0.5141	0.5454
	20	0.6670	0.6127	0.4979	0.5281	**0.8141**	0.7534	0.7701	0.7835
	30	0.8331	0.7839	0.6733	0.7102	**0.9251**	0.8868	0.8982	0.9012
	50	0.9669	0.9464	0.8924	0.9137	**0.9903**	0.9814	0.9835	0.9836

Table 4. Powerful tests against different alternatives

IFR	UFR	DFR-BFR
W^2 & Z_A	A^2 & Z_A	A^2

Based on the power values in Table 2, it is seen that the tests based on W^2 and Z_A statistics have the most power against IFR alternatives. The power differences between these tests and the other tests are substantial. Although for this type of alternatives the tests W^2 and Z_A have the most power but the power differences of these tests with each other are small and we can select one of the tests based on W^2 or Z_A statistic as a powerful test.

From Table 3, it is evident that the tests based on A^2 and Z_A statistics have the most power against UFR alternatives and power differences between these tests and the other tests are substantial.

Tables 3 reveals a superiority of the test based on A^2 statistic to all other tests as we can say that this test outperforms

all other tests against DFR and BFR alternatives.
Although there is no uniformly most powerful test against all alternatives, the tests based on W^2, A^2, Z_A statistics can be recommended in practice. In general, we can conclude that the proposed tests W^2, A^2 and Z_A 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.

5. Illustration with real data

We illustrate, by two real examples, how the tests 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 EDF procedures to this data set. First, the ML estimator of is computed as:

$$\hat{\theta} = \frac{-\left(\bar{X}-1\right)+\sqrt{\left(\bar{X}-1\right)^2+8\bar{X}}}{2\bar{X}} = 0.1866\,.$$

Then, the value of each test statistic is computed and also the critical value of each test at the significance level 0.05 is obtained from Table 1. Results are summarized in Table 3. 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 tests can be used to investigate whether the data come from a Lindley distribution. Based on the test statistics described in Section 3, the values of the proposed statistics are as

$$W^2 = 0.0375,\ D = 0.1103,\ V = 0.2184,\ U^2 = 0.0371,$$
$$A^2 = 0.3187,\ Z_A = 3.335,\ Z_C = 2.666,\ Z_K = 0.4140,$$

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 2.2734, respectively. Therefore, the proposed tests accept the null hypothesis that the electronic components follow a Lindley model at significance level of 0.05.

6. Conclusions

In this paper, we have evaluated the empirical distribution function-based goodness-of-fit tests for the Lindley distribution, and have shown that the considered tests have a good performance. Through Monte Carlo simulations, we have carried out an extensive power study on the considered tests. It is shown that some of the tests outperform in most cases all other tests. Finally, we have used two real data sets and have illustrated how the considered tests can be applied to test the goodness-of-fit for the Lindley distribution when a random sample is available.

Acknowledgement

The authors are grateful to anonymous referees and the associate editor for providing some useful comments on an earlier version of this manuscript.

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. (2015), Testing exponentiality based on the likelihood ratio and power comparison, Annals of Data Science, 2, 195–204.
3. Alizadeh Noughabi, H. (2016), Two powerful tests for normality, Annals of Data Science, 3, 225–234.
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. Alizadeh Noughabi, H. and Chahkandi, M. (2015), Informational energy and its application in testing normality, Annals of Data Science, 2, 391–401.
6. 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.
7. Altun, E. (2019), Two-sided Lindley distribution with inference and applications, Journal of the Indian Society for Probability and Statistics, 20, 255–279.
8. Anderson, T.W. and Darling, D.A. (1954), A test of goodness of fit, Journal of American Statistical Association, 49, 765-769.
9. 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.
10. 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.
11. Chen G. and Balakrishnan, N. (1995), A general purpose approximate goodness-of-fit test, Journal of Quality Technology, 27, 154–161.
12. Chesneau, C., Tomy, L., Gillariose, J. (2021), A new modified Lindley distribution with properties and applications, Journal of Statistics and Management Systems, 24(7), 1383-1403.
13. Chesneau, C., Tomy, L., Jose, M. (2021a), Wrapped modified Lindley distribution, Journal of Statistics and Management Systems, 24 (5), 1025-1040.
14. D'Agostino, R.B. and Stephens, M.A. (Eds.) (1986), Goodness-of-fit Techniques, New York: Marcel Dekker.
15. Ghitany, M.E., Atieh, B. and Nadarajah, S. (2008), Lindley distribution and its Application, Mathematics Computing and Simulation, 78, 493-506.
16. 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.
17. 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.
18. He, D.J. and Xu, X.Z. (2013), A goodness-of-fit testing approach for normality based on the posterior predictive distribution, Test, 22, 1-18.
19. Huber-Carol, C., Balakrishnan, N., Nikulin, M. S. and Mesbah, M. (2002), Goodness-of-fit tests and model validity, Boston, Basel, Berlin: Birkhäuser.
20. 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.
21. Jahanshahi, S. M. A., Habibi Rad, A., Fakoor, V. (2016), A goodness-of-fit test for Rayleigh distribution based on Hellinger distance, Annals of Data Science, 3, 401–411.
22. Kolmogorov, A.N. (1933), Sulla Determinazione Empirica di une legge di Distribuzione. Giornale dell'Intituto Italiano degli Attuari, 4, 83-91.
23. 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.
24. Kuiper, N.H. (1960), Tests concerning random points on a circle, Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Series A, 63, 38-47.

25. 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.
26. 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.
27. Lawless, J.F. (1982), Statistical models and methods for lifetime data. In: John Wiley & Sons (Ed.), New York, USA.
28. 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.
29. Lindley, D.V. (1958), Fiducial distribution and Bayes' theorem, Journal of the Royal Statistical Society, 20, 102–107.
30. 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.
31. 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.
32. 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.
33. Shanker, R., Hagos, F., Sujatha, S. (2015), On modeling of Lifetimes data using exponential and Lindley distributions, Biometrics & Biostatistics International Journal, 2, 1–9.
34. 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.
35. Tomy, L. (2018), A retrospective study on Lindley distribution, Biometrics & Biostatistics International Journal, 7(3), 163-169.
36. Tomy, L., Veena, G., Chesneau, C. (2021b), The sine modified Lindley distribution, Mathematical and Computational Applications, 26 (81), 1-15.
37. 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.
38. von Mises, R. (1931), Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik, Leipzig and Vienna: Deuticke.
39. Wang, M. (2013), A new three-parameter lifetime distribution and associated inference. arXiv:1308.4128 [stat.ME].
40. Watson, G.S. (1961), Goodness of fit tests on a circle, Biometrika, 48, 109-114.
41. Zhang, J. (2002), Powerful goodness-of-fit tests based on the likelihood ratio, Journal of Royal Statistical Society, Series B, 64, 281-294.