STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **10**, September 2022, pp 1162–1167.
Published online in International Academic Press (www.IAPress.org)

Testing the Validity of Laplace Model Against Symmetric Models, Using Transformed Data

Hadi Alizadeh Noughabi*

Department of Statistics, University of Birjand, Birjand, Iran

Abstract In this paper, we first present three characterizations of Laplace distribution and then we introduce a goodness of fit test for Laplace distribution against symmetric distributions, based on one of the transformations. The power of the proposed test under various alternatives is compared with that of the existing tests, by simulation. To show the behavior of the proposed test in real cases, two real examples are presented.

Keywords Goodness of fit tests, Characterization, Laplace distribution, Monte Carlo simulation.

AMS 2010 subject classifications 62G10, 62G20

DOI:10.19139/soic-2310-5070-1030

1. Introduction

Laplace distribution $L(\mu, \theta)$ has the following density function:

$$f(x; \mu, \theta) = \frac{1}{2\theta} \exp\left(-|x-\mu| / \theta\right), \qquad -\infty < \mu < \infty\,, \;\; \theta > 0\,,$$

where θ is a positive scale parameter and μ is the location parameter. This distribution was discovered by Laplace in 1774 for describing the errors of measurement. It is also known as the double exponential distribution.

The Laplace distribution is widely used in many applications. It is therefore very important to test whether the underlying distribution has a Laplace form. Our goal in this article is constructing a new goodness of fit test for Laplace distribution using transformed data.

[1] proposed the following test statistic for testing uniformity.

$$TA = \frac{1}{n} \sum_{i=1}^{n} \left| u_i . \hat{f}(u_i) - F_0(u_i) \right|,$$

where $u_{(1)} \leq u_{(2)} \leq ... \leq u_{(n)}$ are order statistics from a random sample of size n, F_0 is the uniform distribution function and $|\,.\,|$ is the absolute value function. Also

$$\hat{f}(u_i) = \frac{1}{nh} \sum_{j=1}^{n} K(\frac{u_i - u_j}{h}),$$

is the kernel estimate of f, K is the standard normal density and the bandwidth h is obtained from the normal optimal smoothing formula, $h = 1.06 s n^{-\frac{1}{5}}$, where s is the sample standard deviation.

*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)

[1], [2] and [3] introduced some goodness of fit tests for exponentiality using transformed data. Also, they [4] introduced a goodness of fit test for normal distribution using transformed data. Then, [5] improved the exponentiality test using transformed data.

Main contribution of this paper is as follows. We first introduce a new goodness of fit test for Laplace distribution using transformed data. The properties of new test statistic are stated. Then, the critical values and power values of the proposed test are obtained and compared with the powers of the competing tests. We show that the proposed test has the higher power as compared with the other tests against the considered alternatives. Finally, two real examples to show the behavior of the proposed test in real cases are presented. Moreover, recently [6], [7] and [8] proposed new goodness of fit tests for some statistical distributions.

In Section 2, we introduce three characterizations of Laplace distribution and next, based on one of the transformations, we introduce a goodness of fit test for Laplace distribution against symmetric distributions. The power of the proposed test is obtained, by simulation, and the results are in Section 3 where the proposed test is compared with other tests. In Section 4, two real examples are presented and analyzed.

2. The test statistic

Given a random sample $X_1, ..., X_n$ from a continuous symmetric probability distribution F with a density $f(x)$, the hypothesis of interest is

$$H_0: \ f(x) = f_0(x; \mu, \theta) = \frac{1}{2\theta} \exp\left(-|x-\mu|/\theta\right), \qquad for\ some\ (\mu, \theta) \in \Theta,$$

where μ and θ are unspecified and $\Theta = \mathbb{R} \times \mathbb{R}^+$. The alternative to H_0 is

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

In order to obtain a test statistic, we use the following theorem.

Theorem 1. *Let X_1 and X_2 be two independent observations from a symmetric continuous distribution F. Then*
i) $W = \frac{|X_1|}{|X_1|+|X_2|}$ *is distributed as $U(0,1)$ if and only if F is Laplace with mean 0.*
ii) $Y = \frac{|X_1|}{|X_2|}$ is distributed as $F_{(2,2)}$ if and only if F is Laplace with mean 0.
iii) $Z = \frac{|X_1|-|X_2|}{|X_1|+|X_2|}$ is distributed as $U(-1,1)$ if and only if F is Laplace with mean 0.
($F_{(2,2)}$ is Fisher's distribution with 2 and 2 degrees of freedom and $|\,.\,|$ is the absolute value function.)

Proof. [2] showed that the above transformations (i), (ii) and (iii) (with absolute value signs removed) characterize exponential distribution. It is also mentioned in [9]. We know that if X has a symmetric distribution, then X is distributed as Laplace$(0,\ \theta)$ if and only if $|X|$ is distributed as exponential(θ). Thus the results follow. □

To construct a test using the mentioned characterizations, we transform the sample data $(X_1, X_2, ..., X_n)$ to

$$Y_i = X_i - \mu,$$

where $\mu = E(X)$. Note that if X_i's constitute an i.i.d. sample from a Laplace distribution $L(\mu, \theta)$, then Y_i's are $L(0, \theta)$.

Now, we transform Y_i's to

$$W_{ij} = \frac{|Y_i|}{|Y_i| + |Y_j|}, \qquad i \neq j\ ,\ i, j = 1, 2, ..., n.$$

By the above theorem, under the null hypothesis, each W_{ij} has a uniform distribution, and it seems to be appropriate to use our proposed uniformity test (described in the Introduction) to test the uniformity of the distribution of W_{ij}'s and thus to test that the distribution of X_i's is Laplace.

Briefly, the proposed procedure is as follows.
Suppose the order statistics of the sample are $x_{(1)}, x_{(2)}, ..., x_{(n)}$.

1. Find the maximum likelihood estimate of the location parameter (under the assumption that the observations come from Laplace distribution) as follows. The median of the sample is the estimate of μ. Thus

$$\hat{\mu} = \begin{cases} x_{(\frac{n+1}{2})} & if\ n\ is\ odd. \\ (x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)})/2 & if\ n\ is\ even. \end{cases}$$

2. Make the transformation $y_i = x_i - \hat{\mu}$ for $i = 1, 2, ..., n$. Next transform the y_i's to

$$w_{ij} = \frac{|y_i|}{|y_i| + |y_j|}, \qquad i \neq j\ ,\ i, j = 1, 2, ..., n.$$

3. Compute the test statistic as

$$DA = \frac{1}{n'} \sum_{i=1}^{n'} \left| w_i \hat{f}(w_i) - F_0(w_i) \right| ,$$

where $n' = n(n-1)$ and $\hat{f}(.)$ and $F_0(.)$ are as introduced in preceding Section.
4. Find the critical point C^* at a given significance level (α), for sample size n. If DA is greater than C^*, the null hypothesis is rejected at level α.

The test statistic is invariant with respect to location and scale transformations. To see this, note that if x_i's are multiplied by a constant $c > 0$ and are added by a constant d then $\hat{\mu}$ is multiplied by c and is then added by d. Thus

$$w_{ij} = \frac{|cx_i + d - (c\hat{\mu} + d)|}{|cx_i + d - (c\hat{\mu} + d)| + |cx_j + d - (c\hat{\mu} + d)|} = \frac{|x_i - \hat{\mu}|}{|x_i - \hat{\mu}| + |x_j - \hat{\mu}|} = \frac{|y_i|}{|y_i| + |y_j|},$$

that is, w_{ij} is not changed. Thus the test statistic remains invariant.

Remark 1. Since the proposed test statistic is invariant to transformations of location and scale and the parameter space for Laplace distribution is transitive, the distribution of the proposed test statistic does not depend on the unknown parameters and the test is exact. Consequently, we can conclude that the actual size of the proposed test is equal to the nominal size. Therefore, we don't need to provide the size of the test for comparison. In our power comparisons, the actual size of the tests are equal to the nominal size i.e., 0.05.

We think that the transformed data approach is useful because

1. The estimation of the scale parameter is not necessary, in fact by transformation the scale parameter is eliminated and the test statistic does not depend on it. For example in testing exponentiality using transformed data the scale parameter (λ) is eliminated and the test statistic does not depend on λ, thus we don't need to estimate it. Here, for Laplace distribution the scale parameter (θ) is eliminated by transformation and the test statistic does not depend on θ, then it is not necessary to estimate scale parameter for computing the proposed test statistic. Note that the statistics based on empirical distribution function (EDF) depend on both parameters and they must be estimated.
2. The data that we obtain by transformation, although not independent, are much more numerous than the original data. For example for $n = 10$ the number of transformed data is $n' = n(n-1) = 90$ and we deal with a larger sample.

3. Simulation study

In this section, the performance of the proposed test is investigated using Monte Carlo simulations.
For small to moderate sample sizes 5, 10, 15, 20, 25, 30 and 50, we used Monte Carlo methods with 10,000

Table 1. Critical values of the proposed test statistic

		α		
n	n'	0.01	0.05	0.10
5	20	0.6804	0.3314	0.2145
10	90	0.3174	0.1943	0.1466
15	210	0.2269	0.1532	0.1229
20	380	0.1964	0.1379	0.1081
25	600	0.1844	0.1263	0.1013
30	870	0.1563	0.1132	0.0951
50	2450	0.1251	0.0904	0.0764

replicates from standard Laplace distribution to obtain critical values of our procedure. These values are reported in Table 1.

Since the tests based on the empirical distribution function are important tests which are commonly used in practice, we compare the power of the proposed test with the power of these tests. These tests are the Cramer-von Mises W^2, the Watson U^2, the Anderson-Darling A^2, the Kolmogorov-Smirnov D, and the Kuiper V. For further information about these tests see [10].

The following symmetric distributions were considered as alternative distributions:

- standard normal distribution, denoted by $N(0, 1)$,
- student's t distribution with 3, 5, and 10 degrees of freedom, denoted by $t(3)$, $t(5)$, $t(10)$, respectively,
- standard Cauchy distribution, denoted by $C(0, 1)$,
- standard logistic distribution, denoted by $Lo(0, 1)$,
- uniform distribution, denoted by $U(0, 1)$,
- beta distribution with parameters 2 and 2, denoted by $Beta(2, 2)$,
- Normal Inverse Gaussian distribution, denoted by $NIG(\alpha, \beta, \mu, \delta)$.

The NIG distribution is completely determined by four parameters, α, β, μ and δ. For an NIG distribution α, β, μ and δ are shape, skewness, location and scale parameters respectively. The mean, variance, skewness and kurtosis of NIG are defined respectively by $\mu + \beta\delta/\gamma$, $\delta\alpha^2/\gamma^3$, $3\beta/\alpha\sqrt{\delta\gamma}$ and $3(1 + 4\beta^2/\alpha^2)/\delta\gamma$,where $\gamma = \sqrt{\alpha^2 - \beta^2}$.

The estimated powers of the present test, based on 10,000 samples of size n equal to 10 and 20, are given in Table 2. The maximum power in each row is bold faced.

We see that the proposed test has the highest power against all considered alternatives distributions (with exception of Cauchy distribution), as compared with the other tests.
Generally, we observe that the proposed test performs quite well, as compared with the other tests. Thus, in applications, the use of the proposed test is recommended for achieving a higher gain in power.

4. Real data examples

Two real examples to show the behavior of the proposed test in real cases is presented in this section.

Table 2. Empirical powers of the tests against symmetric distributions at level 5%.

$altern.$	n	W^2	U^2	A^2	D	V	DA
$N(0,1)$	10	0.046	0.056	0.041	0.043	0.053	**0.096**
	20	0.075	0.114	0.067	0.086	0.106	**0.196**
$Lo(0,1)$	10	0.049	0.049	0.043	0.046	0.049	**0.078**
	20	0.060	0.072	0.054	0.062	0.070	**0.132**
$C(0,1)$	10	0.347	**0.397**	0.360	0.332	0.383	0.021
	20	0.536	**0.633**	0.556	0.515	0.611	0.010
$t(3)$	10	0.059	**0.065**	0.060	0.061	0.063	0.059
	20	0.069	0.081	0.073	0.071	0.080	**0.081**
$t(5)$	10	0.045	0.053	0.045	0.045	0.049	**0.077**
	20	0.060	0.072	0.056	0.064	0.068	**0.119**
$t(10)$	10	0.050	0.052	0.044	0.049	0.053	**0.088**
	20	0.067	0.086	0.060	0.073	0.085	**0.158**
$U(0,1)$	10	0.112	0.156	0.097	0.100	0.132	**0.191**
	20	0.245	0.445	0.242	0.242	0.371	**0.478**
$Beta(2,2)$	10	0.071	0.092	0.061	0.064	0.087	**0.144**
	20	0.124	0.232	0.116	0.151	0.209	**0.322**
$NIG(0.5,0,1,0)$	10	0.054	**0.061**	0.054	0.057	0.060	0.056
	20	0.052	0.060	0.057	0.053	0.061	**0.061**
$NIG(1,0,1,0)$	10	0.046	0.045	0.043	0.045	0.045	**0.061**
	20	0.051	0.058	0.051	0.057	0.057	**0.098**
$NIG(2,0,1,0)$	10	0.044	0.047	0.040	0.044	0.048	**0.074**
	20	0.054	0.063	0.044	0.053	0.063	**0.118**

Example 1. The following data are 100 breaking strengths of yarn presented by [11].

66, 117, 132, 111, 107, 85, 89, 79, 91, 97, 138, 103, 111, 86, 78, 96, 93, 101, 102, 110, 95, 96, 88, 122, 115, 92, 137, 91, 84, 96, 97, 100, 105, 104, 137, 80, 104, 104, 106, 84, 92, 86, 104, 132, 94, 99, 102, 101, 104, 107, 99, 85, 95, 89, 102, 100, 98, 97, 104, 114, 111, 98, 99, 102, 91, 95, 111, 104, 97, 98, 102, 109, 88, 91, 103, 94, 105, 103, 96, 100, 101, 98, 97, 97, 101, 102, 98, 94, 100, 98, 99, 92, 102, 87, 99, 62, 92, 100, 96, 98.

[12] concluded that the Laplace assumption is not rejected for any of EDF tests, even at a significance level of 0.10. Our procedure can be used to investigate the above conclusion. The value of the proposed test statistic is $T = 0.0416$ and the critical value is 0.061 at 0.05 level. Since the value of the proposed test statistic is smaller than the critical value, the Laplace assumption is accepted and our procedure confirms the results obtained by EDF methods.

Example 2. [13] considered the following dataset, consisting of 33 differences in flood levels between stations on a river.

1.96, 1.97, 3.60, 3.80, 4.79, 5.66, 5.76, 5.78, 6.27, 6.30,6.76, 7.65, 7.84, 7.99, 8.51, 9.18, 10.13, 10.24, 10.25, 10.43, 11.45, 11.48, 11.75, 11.81, 12.34, 12.78, 13.06, 13.29, 13.98, 14.18, 14.40, 16.22, 17.06.

They suggested that the Laplace distribution might provide a good fit. Puig and Stephens (2000) concluded that D, W^2, and U^2 just reject the Laplace distribution for the data at 0.05 level.
For this example we find $\hat{\mu} = 10.13$, $T = 0.1414$ and the critical value is 0.1072 at 0.05 level. Then the Laplace assumption is rejected as it is by D, W^2, and U^2 tests.

Acknowledgement

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

REFERENCES

1. H. Alizadeh Noughabi, and N.R. Arghami, *Testing exponentiality using transformed data*, Journal of Statistical Computation and Simulation, vol. 81, pp. 511–516, 2011a.
2. H. Alizadeh Noughabi, and N.R. Arghami, *Testing exponentiality based on characterizations of the exponential distribution*, Journal of Statistical Computation and Simulation, vol. 81, no. 2, pp. 1641–1651, 2011b.
3. H. Alizadeh Noughabi, *Testing The Validity of the Exponential Model Based on Type II Censored Data Using Transformed Sample Data*, STATISTICA, vol. 76, pp. 221–232, 2016.
4. H. Alizadeh Noughabi and N.R. Arghami, *Testing normality using transformed data*, Communications in Statistics – Theory and Methods, vol. 42, pp. 3065–3075, 2013.
5. B.R. Dhumal, and D.T. Shirke, , *A modified test for testing exponentiality using transformed data*, Journal of Statistical Computation and Simulation, vol. 84, pp. 397–403, 2014.
6. V. Zardasht, *Results on relative mean residual life and relative cumulative residual entropy*, Statistics, Optimization and Information Computing, vol. 7, pp. 150–159, 2019.
7. H. Alizadeh Noughabi, H. Alizadeh Noughabi, and J. Jarrahiferiz, *Informational Energy and Entropy Applied to Testing Exponentiality*, Statistics, Optimization and Information Computing, vol. 8, pp. 220-228, 2020.
8. V. Fakoor, M. Ajami, S.M.A. Jahanshahi, A. Shariati, *A Density-Based Empirical Likelihood Ratio Approach for Goodness-of-fit Tests in Decreasing Densities*, Statistics, Optimization and Information Computing, vol. 8, pp. 66–79, 2020.
9. N. Balakrishnan, and A.P. Basu, *The Exponential Distribution: Theory, Methods and Applications*, Amsterdam: Gordon and Breach. Publishers, 1995.
10. R.B. D'Agostino, and M.A. Stephens, *Goodness-of-Fit Techniques*, New York: Marcel Dekker, Inc., 1986.
11. A. J. Duncan, *Quality Control and Industrial Statistics*, IL: Irwin, Homewood, 1974.
12. P. Puig, and M.A. Stephens, *Tests of fit for the Laplace distribution with applications*, Technometrics, vol. 42, pp. 417–424, 2000.
13. L.J. Bain, and M. Englehardt, *Interval Estimation for the Two Parameter Double Exponential Distribution*, Technometrics, vol. 15, pp. 875–887, 1973.