Journal of Mahani Mathematical Research Center

Shahid Bahonar University of Kerman

Print ISSN: 2251-7952 Online ISSN: 2645-4505

JMMRC 2012

TEST OF FIT FOR CAUCHY DISTRIBUTION BASED ON THE EMPIRICAL LIKELIHOOD RATIO WITH APPLICATION TO THE STOCK MARKET PRICE

H. Alizadeh Noughabi*

Article type: Research Article

(Received: 16 June 2020, Revised: 30 November 2021, Accepted: 12 December 2021)

(Available Online: 16 December 2021)

(Communicated by N. Shajareh Poursalavati)

Abstract. Recently, it has been shown that the density based empirical likelihood concept extends and standardizes these methods, presenting a powerful approach for approximating optimal parametric likelihood ratio test statistics. In this article, we propose a density based empirical likelihood goodness of fit test for the Cauchy distribution. The properties of the test statistic are stated and the critical points are obtained. Power comparisons of the proposed test with some known competing tests are carried out via simulations. Our study shows that the proposed test is superior to the competitors in most of the considered cases and can confidently apply in practice. Finally, a financial data set is presented and analyzed.

Keywords: Cauchy distribution, Empirical likelihood ratio, Goodness-of-fit test, Test power, Monte Carlo simulation.
2020 MSC: Primary 62G10, 62P20.

1. Introduction

A random variable X has a Cauchy distribution with parameters $\mu \in R$ and $\sigma > 0$, if its density function has the form:

$$f_0(x;\mu,\sigma) = \frac{1}{\pi\sigma\left[1+((x-\mu)/\sigma)^2\right]}, \qquad -\infty < x < \infty.$$

Here, σ is a positive scale parameter and μ is the location parameter. We henceforth denote this distribution by $C(\mu,\sigma)$. The corresponding cumulative distribution function is given by

$$F_0(x;\mu,\sigma) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{x-\mu}{\sigma}\right).$$

The Cauchy distribution can be considered as a model for describing data that arise as realizations of the ratio of two normal random variables. Min et al.

*Corresponding author, ORCID: 0000-0002-7515-1896
E-mail: alizadehhadi@birjand.ac.ir

DOI: 10.22103/jmmrc.2021.17747.1152

How to cite: H. Alizadeh Noughabi, *Test of fit for Cauchy distribution based on the empirical likelihood ratio with application to the stock market price*, J. Mahani Math. Res. Cent. 2022; 11(1): 79-94.

[13] found that Cauchy distribution describes the distribution of velocity differences induced by different vortex elements. Another application of the Cauchy distribution is presented by Stapf et al. [18]. They apply this distribution to study the polar and nonpolar liquids in porous glasses. Kagan [8] showed that the hypocenters on focal spheres of earthquakes is distributed as a Cauchy random variable. Winterton et al. [30] pointed out that the source of fluctuations in contact window dimensions is variation in contact resistivity, and the contact resistivity is distributed as a Cauchy random variable. Nolan [14] applied the Cauchy distribution to financial modeling. The Cauchy distribution is very extensively reviewed in Johnson et al. [7] and Kotz et al. [9]. Therefore, in practice, it is important to test whether the underlying distribution has a Cauchy form.

Many researchers have been interested in goodness of fit tests for different distributions and developed various tests in the literature. 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).

Recently, the density based empirical likelihood ratio goodness of fit tests are widely developed in statistical applications, see for example, Vexler et al. [24], [28], Vexler and Gurevich [21], Gurevich and Vexler [5], Shan et al. [15], Vexler and Yu [22], Yu et al. [32], Vexler et al. [26], [27], and Vexler et al. [23], Yu et al. [33], Zhao et al. [35], Gurevich and Vexler [6]. Also, there are packages in the STATA and R software for applying the EL approach to real data problems, see Tanajian et al. [19], Shepherd et al. [16] and Vexler et al. [23].

In parametric statistics, based on Neyman-Pearson lemma the likelihood ratio test is a uniformly most powerful test. Suppose that $X_1, ..., X_n$ are a random sample and we wish to test the hypothesis

$$H_0 : X_1, ..., X_n \sim f_0,$$

versus

$$H_1 : X_1, ..., X_n \sim f_1.$$

The most powerful test statistic for the above hypothesis is the likelihood ratio

$$\frac{\prod_{i=1}^{n} f_1(X_i)}{\prod_{i=1}^{n} f_0(X_i)},$$

where $f_0(x)$ and $f_1(x)$ are completely known. About the connection between NP lemma and likelihood ratio test, one can see Solomon [17], Berger and Wolpert [1], Lehmann [10], and Glover and Dixon [4].

As we know in nonparametric statistics, the alternative distribution is unknown and therefore, for goodness of fit tests based on EL ratio, we need to estimate the likelihood function $\prod_{i=1}^{n} f_1(X_i)$ and then we can use the likelihood ratio statistic. Vexler and Gurevich [20] estimated the likelihood ratio as

$$\prod_{i=1}^{n} \frac{2m}{n\left(X_{(i+m)} - X_{(i-m)}\right)},$$

and then proposed a test statistic for goodness of fit. Their test statistic is as

$$T_{mn} = \frac{\prod\limits_{i=1}^{n} \frac{2m}{n\left(X_{(i+m)} - X_{(i-m)}\right)}}{\prod\limits_{i=1}^{n} f_0(X_i;\hat{\theta})},$$

where $\hat{\theta}$ is the maximum likelihood estimator of θ , and $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ are order statistics obtained from $X_1, ..., X_n$ and also $X_{(i)} = X_{(1)}$ if $i < 1$, and $X_{(i)} = X_{(n)}$ if $i > n$.

Since T_{mn} depends on m, they proposed the following test statistic.

$$T_{mn} = \frac{\min\limits_{1 \leq m \leq n^{\delta}} \prod\limits_{i=1}^{n} \frac{2m}{n\left(X_{(i+m)} - X_{(i-m)}\right)}}{\prod\limits_{i=1}^{n} f_0(X_i;\hat{\theta})},$$

where $\delta \in (0, 1)$. They used their test statistic and proposed tests for the normal and uniform distributions. Moreover, Vexler et al. [25] applied the above test statistic and introduced a goodness of fit test for the inverse Gaussian distribution.

Recently, Mahdizadeh and Zamanzade [11-12] introduced some goodness of fit tests for Cauchy distribution and showed that their tests have a good performance in compared to the existing tests. Also, Ebner et al. [3] introduced a new characterization of the Cauchy distribution and proposed a class of goodness-of-fit tests to the Cauchy family. Villaseñor and González-Estrada [29] investigated goodness-of-fit tests for Cauchy distributions using data transformations.

The goal of this article is to propose a density based empirical likelihood ratio goodness of fit test for Cauchy distribution. In Section 2, we construct our test statistic and then its properties are stated. In Section 3, we obtain the critical values and the power values of the proposed test, and then power values are compared with those of the competing tests. Section 4 contains an illustrative example. The following section contains a brief conclusion.

2. The density based empirical likelihood ratio test statistic

Let $X_1, ..., X_n$ be an i.i.d. (independent identically distributed) sample from a population with unknown cumulative distribution function F and a probability density function f. We interest to test the null hypothesis

$$H_0: \{X_1, ..., X_n\} \text{ is a sample from Cauchy } C(\mu, \sigma),$$

where μ and σ are specified or unspecified. The alternative hypothesis is

$$H_1: \{X_1, ..., X_n\} \text{ is not a sample from Cauchy } C(\mu, \sigma).$$

If $f_0(x; \mu, \theta)$ denotes the density of Cauchy distribution, then the hypothesis of interest is

$$H_0: f(x) = f_0(x; \mu, \sigma) = \frac{1}{\pi\sigma\left[1 + ((x-\mu)/\sigma)^2\right]}, \qquad for\ some\ (\mu, \sigma) \in \Omega,$$

where $\Omega = R \times R^+$. The alternative to H_0 is

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

Here, we briefly describe the method of density based empirical likelihood ratio to construct a test statistic for the above hypothesis.
The likelihood ratio test statistic for the above hypothesis is defined as

$$LR = \frac{\prod_{i=1}^{n} f_{H_1}(X_i)}{\prod_{i=1}^{n} f_{H_0}(X_i; \theta)},$$

where $\theta = (\mu, \sigma)$.

When density function under H_1 is known (f_{H_1}), Neyman-Pearson lemma guarantees that the LR test is the MP test. If it is unknown, we will use the maximum empirical likelihood method to estimate the numerator. Also, we use the maximum likelihood estimators for the unknown parameters. Since for Cauchy distribution these estimators do not have a close form, we obtain them by Newton-Raphson method. As we know Newton-Raphson method needs the starting values and here we set starting values for the unknown parameters μ and σ the median and the half-interquartile range. Suppose ξ_p is the sample pth quantile. Then, the starting values are

$$\mu_0 = Median\,(X_i) \quad ; \quad \sigma_0 = (\xi_{0.75} - \xi_{0.25})/2.$$

Therefore, we report our results based on the starting points mentioned in above.
Consider

$$L_f = \prod_{i=1}^{n} f_{H_1}(X_i) = \prod_{i=1}^{n} f_{H_1}(X_{(i)}) = \prod_{i=1}^{n} f_i\,,$$

where $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ are the order statistics of the observations and $f(X_{(i)}) = f_i$. We apply the empirical likelihood method to derive the values of f_i that maximize L_f with the constraint $\int f(s)ds = 1$ under the alternative hypothesis. The following proposition, proved by Vexler and Gurevich [20], express this constraint more explicitly.

Proposition 2.1. *Let $f(x)$ be a density function. Then*

$$\sum_{j=1}^{n} \int_{X_{(j-m)}}^{X_{(j+m)}} f(x)dx = 2m \int_{X_{(1)}}^{X_{(n)}} f(x)dx - \sum_{k=1}^{m-1} (m-k) \int_{X_{(n-k)}}^{X_{(n-k+1)}} f(x)dx - \sum_{k=1}^{m-1} (m-k) \int_{X_{(k)}}^{X_{(k+1)}} f(x)dx \,,$$

where $X_{(j)} = X_{(1)}$ if $j \leq 1$ and $X_{(j)} = X_{(n)}$ if $j \geq n$.

Let

$$\Delta_m = \frac{1}{2m} \sum_{j=1}^{n} \int_{X_{(j-m)}}^{X_{(j+m)}} f(x)dx,$$

and since $\int_{X_{(1)}}^{X_{(n)}} f(x)dx \leq \int_{-\infty}^{\infty} f(x)dx = 1$, from Lemma 1,

$$\Delta_m \leq 1.$$

When $m/n \to 0$ as $m, n \to \infty$, we can expect that $\Delta_m \approx 1$. The integration $\int_{X_{(j-m)}}^{X_{(j+m)}} f(x)dx$ can be approximated by $\left(X_{(j+m)} - X_{(j-m)}\right) f_j$ and thus

$$\sum_{j=1}^{n} \int_{X_{(j-m)}}^{X_{(j+m)}} f(x)dx \approx \sum_{j=1}^{n} \left(X_{(j+m)} - X_{(j-m)}\right) f_j.$$

Therefore, Δ_m can be approximated by

$$\hat{\Delta}_m = \frac{1}{2m} \sum_{j=1}^{n} \left(X_{(j+m)} - X_{(j-m)}\right) f_j.$$

Now, by using the Lagrange multiplier method to maximize $l = \log(L_f) = \sum_{j=1}^{n} \log f_j$, under the constrain $\hat{\Delta}_m \leq 1$, we have

$$l(f_1, f_2, ..., f_n, \eta) = \sum_{j=1}^{n} \log f_j + \eta \left(\frac{1}{2m} \sum_{j=1}^{n} \left(X_{(j+m)} - X_{(j-m)}\right) f_j - 1 \right),$$

where η is a Lagrange multiplier. By taking the derivative of the above equation respect to each f_j and η, we obtain the values of $f_1, f_2, ..., f_n$. The form of values is as

$$f_j = \frac{2m}{n\left(X_{(j+m)} - X_{(j-m)}\right)} , \qquad j = 1, ..., n \,,$$

where $X_{(j)} = X_{(1)}$ if $j \leq 1$ and $X_{(j)} = X_{(n)}$ if $j \geq n$.

Consequently, the density-based likelihood ratio test statistic to test the goodness-of-fit for the Cauchy distribution is

$$T_{mn} = \frac{\prod_{j=1}^{n} \frac{2m}{n\left(X_{(j+m)}-X_{(j-m)}\right)}}{\prod_{j=1}^{n} f_{H_0}(X_j;\hat{\theta})},$$

Clearly, the test statistic T_mn strongly depends on the value of m and for a given n, the value of m must be determined. It is not possible to have one value of m , for a given n, that would result in a test attaining the maximum power for all alternatives. Therefore, similar to Vexler and Gurevich [20-21], we propose the following test statistic.

$$T_n = \frac{\min\limits_{1\le m<n^{\delta}} \prod_{j=1}^{n} \frac{2m}{n\left(X_{(j+m)}-X_{(j-m)}\right)}}{\prod_{j=1}^{n} f_{H_0}(X_j;\hat{\theta})},$$

where $\delta \in (0,1)$. Here, we choose $\delta = 0.5$ for the power study of our test. The following theorems give some asymptotic properties of the test statistic. First, we denote

$$h(x,\theta) = \frac{\partial \log f_{H_0}(x;\theta)}{\partial \theta},$$

and $\theta = (\mu, \sigma)$. Assume the following conditions are hold.
(C1) $E(\log f(X_1))^2 < \infty$;
(C2) under the null hypothesis, $\left|\theta - \hat{\theta}\right| \to 0$ in probability as $n \to \infty$;
(C3) under the alternative hypothesis, $\hat{\theta} \to \theta_0$ as $n \to \infty$, where θ_0 is a constant vector with finite components;
(C4) There are open intervals $\Theta_0 \subseteq R^2$ and $\Theta_1 \subseteq R^2$ containing θ and θ_0 , respectively. There also exists a function $s(x)$ such that $|h(x,\xi)| \le s(x)$ for all $x \in R$ and $\xi \in \Theta_0 \cup \Theta_1$.

Theorem 2.2. *Assume that the conditions C1-C4 hold. Then, under H_0,*

$$\frac{1}{n}\log(T_n) \to 0,$$

in probability as $n \to \infty$.

Theorem 2.3. *Assume that the conditions C1-C4 hold. Then, under H_1,*

$$\frac{1}{n}\log(T_n) \to E\log\left(\frac{f_{H_1}(X_1)}{f_{H_0}(X_1;\theta_0)}\right),$$

in probability as $n \to \infty$. Hence, the test is consistent.

Vexler and Gurevich [20] proved that the above theorems are satisfied for any null family of distributions and hence for the null hypothesis of the Cauchy distribution Theorems 1 and 2 are hold.

We note that the proposed test statistic is invariant with respect to the location and scale transformations because $T_n(cx+d) = T_n(x)$, where $c > 0$ and $d \in R$ are constant values. Moreover, since the test statistic T_n is invariant and the parameter space (Ω) is transitive, the distribution of the proposed test statistic T_n does not depend on the unknown parameters μ and σ. Therefore, it is concluded that the critical values of the test statistic do not depend on the unknown parameters (μ, σ) and hence they can be obtained from a standard Cauchy distribution.

3. Simulation study

Since deriving the exact distribution of the proposed test statistic is complicated, we study the null distribution of the test statistic T_n via Monte Carlo simulations using 50,000 runs for each sample size. Upper tail percentiles are obtained for values 0.99, 0.95, and 0.90. These values are given in Table 1.

Table 1. Critical values

	α		
n	0.01	0.05	0.10
5	84.958	0.9583	0.9161
10	0.8954	0.6717	0.5835
15	0.6126	0.4688	0.4007
20	0.4661	0.3556	0.2930
25	0.3806	0.2765	0.2202
30	0.3183	0.2222	0.1686
40	0.2361	0.1469	0.0942
50	0.1715	0.0891	0.0392
100	0.0408	-0.0174	-0.0544

We also evaluate in Table 2 the estimated type I error control using the 0.05 percentiles of the proposed test ($\alpha = 0.05$). We generated random samples from a spectrum of Cauchy populations and then obtained the actual sizes of the tests. The results are presented in Table 2. The well-known EDF-tests are considered and the estimated type I error of these tests are reported. These tests are Cramer von Mises test W^2 , Watson test U^2, Kolmogorov-Smirnov test D, Anderson-Darling test A^2, and Kuiper test V.

It is evident, from Table 2, that the actual sizes of considered tests are approximately equal to the nominal size 0.05. Therefore, we can conclude that the empirical percentiles presented in Table 1 provides an excellent type I error control.

Table 2. Type I error control of the tests for the nominal significance level $\alpha = 0.05$.

	n	W^2	D	V	U^2	A^2	Z_A	Z_C	Z_K	KL	T_n
$C(0, 0.5)$	10	0.0474	0.0454	0.0491	0.0479	0.0475	0.0485	0.0505	0.0498	0.0517	0.0511
	20	0.0522	0.0507	0.0509	0.0522	0.0496	0.0493	0.0491	0.0506	0.0494	0.0504
	30	0.0508	0.0492	0.0488	0.0508	0.0508	0.0503	0.0498	0.0494	0.0497	0.0518
	50	0.0509	0.0491	0.0515	0.0510	0.0520	0.0510	0.0507	0.0499	0.0503	0.0495
$C(0, 2)$	10	0.0489	0.0492	0.0514	0.0495	0.0474	0.0484	0.0490	0.0505	0.0513	0.0506
	20	0.0516	0.0500	0.0501	0.0516	0.0500	0.0503	0.0495	0.0504	0.0503	0.0483
	30	0.0514	0.0480	0.0503	0.0520	0.0513	0.0505	0.0502	0.0498	0.0520	0.0504
	50	0.0467	0.0492	0.0486	0.0468	0.0478	0.0498	0.0501	0.0502	0.0508	0.0483
$C(0, 4)$	10	0.0510	0.0512	0.0514	0.0512	0.0490	0.0495	0.0491	0.0489	0.0514	0.0495
	20	0.0503	0.0490	0.0498	0.0506	0.0492	0.0493	0.0502	0.0510	0.0493	0.0489
	30	0.0481	0.0480	0.0476	0.0486	0.0503	0.0501	0.0498	0.0493	0.0494	0.0507
	50	0.0510	0.0514	0.0507	0.0510	0.0503	0.0502	0.0499	0.0503	0.0508	0.0498
$C(0, 8)$	10	0.0514	0.0504	0.0528	0.0522	0.0501	0.0491	0.0495	0.0513	0.0516	0.0517
	20	0.0520	0.0503	0.0498	0.0522	0.0512	0.0505	0.0498	0.0507	0.0510	0.0488
	30	0.0499	0.0510	0.0512	0.0502	0.0511	0.0499	0.0503	0.0497	0.0503	0.0503
	50	0.0494	0.0499	0.0515	0.0493	0.0512	0.0504	0.0502	0.0508	0.0517	0.0497

Through Monte Carlo simulations, the power values of the proposed test against various alternatives are computed. Since the tests of fit based on the empirical distribution function (EDF) are commonly used in practice, we compare the performance of the EDF-tests and the proposed ELR based goodness of fit test under various alternative distributions. The well-known EDF-tests are Cramer von Mises test W^2, Watson test U^2, Kolmogorov-Smirnov test D, Anderson-Darling test A^2, and Kuiper test V.
Also, we consider the tests proposed by Zhang [34]. Briefly, these test statistics for the Cauchy distribution are as

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

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

$$Z_K = \max_{1\le i\le n}\left((i-0.5)\log\left\{\frac{i-0.5}{nF_0(X_{(i)};\hat{\mu},\hat{\sigma})}\right\} + \right.$$

$$\left.(n-i+0.5)\log\left\{\frac{n-i+0.5}{n(1-F_0(X_{(i)};\hat{\mu},\hat{\sigma}))}\right\}\right).$$

For large values of the above test statistics the null hypothesis H_0 will be rejected. The test statistics are invariant under any affine transformation on the sample data. Therefore, they are distribution-free within the Cauchy distribution family. Mahdizadeh and Zamanzade [11] investigated these statistics for testing the validity of Cauchy distribution.

Moreover, we consider the entropy-based test suggested by Mahdizadeh and Zamanzade [11] as a competitor test. The entropy-based test statistic for the

Cauchy distribution is as

$$KL = \exp\left\{-HV_{mn} - \frac{1}{n}\sum_{i=1}^{n}\log\left(f_0(X_i;\hat{\mu},\hat{\sigma})\right)\right\},$$

where HV_{mn} is Vasicek entropy estimator. Here, $\hat{\mu}$ and $\hat{\sigma}$ are estimated by $Median\,(X_i)$ and $(\xi_{0.75} - \xi_{0.25})/2$, respectively.

The following alternatives are considered in power comparison. These alternatives can divide into two groups, symmetric alternatives and asymmetric alternatives.

Group I: Symmetric alternatives:

- the standard normal distribution, denoted by $N(0,1)$;
- the student's distribution with 10 degrees of freedom, denoted by $t(10)$;
- the student's distribution with 3 degrees of freedom, denoted by $t(3)$;
- the standard Laplace distribution, denoted by $La(0,1)$;
- the standard logistic distribution, denoted by $Lo(0,1)$;
- the uniform distribution, denoted by $U(0,1)$;
- the Beta distribution, denoted by $Beta(2,2)$.

Group II: asymmetric alternatives:

- the exponential, $Exp(1)$;
- the Gamma, $\Gamma(0.5,1)$ and $\Gamma(2,1)$;
- the lognormal, $LN(0,0.5)$, $LN(0,1)$, $LN(0,2)$;
- the Weibull, $W(0.5,1)$ and $W(2,1)$;
- the extreme value distribution (Gumbel), $EV(0,1)$;
- the inverse Gaussian, $IG(1,0.5)$, $IG(1,1)$ and $IG(1,2)$;
- the skew normal distribution, $SN(0,1,0.5)$, $SN(0,1,2)$ and $SN(0,1,3)$;
- the skew Laplace distribution, $SL(0,1,0.5)$, $SL(0,1,2)$ and $SL(0,1,3)$.

We compute the power values of the tests under the above alternatives by Monte Carlo simulations as follows. Under each alternative 50,000 samples of size 10, 20, 30 and 50 are generated and 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 3 and 4 display and compare the power values of the tests 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 3. Empirical powers of the tests against symmetric distribution at significance level 5%.

	n	W^2	D	V	U^2	A^2	Z_A	Z_C	Z_K	KL	T_n
$N(0,1)$	10	0.0305	0.0315	0.0627	0.0645	0.0152	0.0209	0.0125	0.0125	0.2062	**0.2555**
	20	0.0689	0.0633	0.2064	0.1953	0.0592	0.2542	0.1795	0.0545	**0.7261**	0.7208
	30	0.1147	0.1048	0.3706	0.3457	0.1530	0.6741	0.5572	0.1848	**0.9757**	0.9628
	50	0.2722	0.2505	0.7030	0.6582	0.4968	0.9876	0.9677	0.6941	**1.0000**	0.9998
$t(10)$	10	0.0280	0.0298	0.0533	0.0552	0.0130	0.0176	0.0101	0.0110	0.1755	**0.2106**
	20	0.0571	0.0549	0.1603	0.1548	0.0442	0.1895	0.1295	0.0432	**0.6061**	0.6005
	30	0.0921	0.0834	0.2858	0.2711	0.1108	0.5415	0.4262	0.1352	**0.9121**	0.8896
	50	0.1997	0.1799	0.5583	0.5410	0.3673	0.9475	0.9009	0.5289	**0.9970**	0.9930
$t(3)$	10	0.0250	0.0276	0.0418	0.0417	0.0122	0.0148	0.0076	0.0115	0.1135	**0.1447**
	20	0.0416	0.0429	0.0920	0.0865	0.0282	0.0913	0.0565	0.0287	0.3356	**0.3407**
	30	0.0555	0.0563	0.1383	0.1330	0.0510	0.2488	0.1685	0.0663	**0.5807**	0.5483
	50	0.0941	0.0933	0.2597	0.2655	0.1383	0.6123	0.5006	0.2239	**0.8122**	0.7616
$La(0,1)$	10	0.0281	0.0289	0.0514	0.0527	0.0126	0.0168	0.0890	0.0114	0.1040	**0.1227**
	20	0.0538	0.0521	0.1438	0.1375	0.0417	0.1619	0.1093	0.0384	0.3219	**0.3258**
	30	0.0792	0.0735	0.2418	0.2328	0.0934	0.4767	0.3636	0.1105	**0.6359**	0.5955
	50	0.0661	0.0611	0.1758	0.1746	0.0954	0.5920	0.4691	0.1666	**0.9379**	0.8987
$Lo(0,1)$	10	0.0267	0.0292	0.0406	0.0397	0.0131	0.0149	0.0078	0.0112	0.1633	**0.1958**
	20	0.0379	0.0381	0.0717	0.0693	0.0241	0.0712	0.0439	0.0226	**0.5631**	0.5527
	30	0.0452	0.0445	0.0989	0.0941	0.0389	0.2003	0.1343	0.0460	**0.8813**	0.8538
	50	0.1709	0.1513	0.4893	0.4746	0.3074	0.9198	0.8565	0.4572	**0.9958**	0.9887
$U(0,1)$	10	0.0784	0.0856	0.0201	0.0187	0.0470	0.0799	0.0549	0.0453	0.5066	**0.5503**
	20	0.2347	0.2789	0.6888	0.5853	0.2658	0.7641	0.6649	0.3454	**0.9909**	0.9906
	30	0.4781	0.5731	0.9263	0.8379	0.6480	0.9873	0.9688	0.8313	**1.0000**	1.0000
	50	0.8628	0.9525	0.9985	0.9879	0.9753	1.0000	1.0000	0.9995	**1.0000**	1.0000
$Beta(2,2)$	10	0.0441	0.0437	0.1034	0.1032	0.0236	0.0348	0.0231	0.0184	0.3256	**0.3840**
	20	0.1185	0.1143	0.4083	0.3573	0.1203	0.5029	0.3940	0.1257	**0.9338**	0.9287
	30	0.2368	0.2306	0.6917	0.6014	0.3543	0.9203	0.8535	0.4648	**0.9997**	0.9994
	50	0.5675	0.6100	0.9646	0.9018	0.8299	0.9998	0.9992	0.9747	**1.0000**	1.0000

TABLE 4. Empirical powers of the tests against asymmetric distribution at significance level 5%.

	n	W^2	D	V	U^2	A^2	Z_A	Z_C	Z_K	KL	T_n
$Exp(1)$	10	0.1531	0.2078	0.1043	0.1965	0.2116	0.1404	0.0772	0.1279	**0.4027**	0.3852
	20	0.3870	0.5782	0.3673	0.5890	0.6412	0.7159	0.4979	0.6514	**0.9185**	0.9129
	30	0.6179	0.8174	0.6685	0.8788	0.8813	0.9789	0.8902	0.9662	0.9931	**0.9937**
	50	0.9268	0.9858	0.9754	0.9982	0.9963	1.0000	0.9997	1.0000	1.0000	**1.0000**
$\Gamma(0.5, 1)$	10	0.3513	0.4163	0.2835	0.4524	0.4498	0.3616	0.2189	0.3615	**0.6334**	0.4706
	20	0.7263	0.8669	0.7138	0.9096	0.9143	0.9368	0.7730	0.9345	**0.9769**	0.8783
	30	0.9308	0.9810	0.9473	0.9947	0.9931	**0.9995**	0.9861	0.9992	0.9976	0.9238
	50	0.9988	0.9998	0.9999	1.0000	1.0000	**1.0000**	1.0000	1.0000	0.9996	0.9536
$\Gamma(2, 1)$	10	0.0754	0.1191	0.0442	0.0882	0.1170	0.0596	0.0325	0.0460	0.2854	**0.3189**
	20	0.1976	0.3660	0.1810	0.2777	0.3984	0.4894	0.3303	0.3186	0.8458	**0.8463**
	30	0.3485	0.6006	0.4034	0.5361	0.6551	0.8967	0.7537	0.7673	0.9861	**0.9892**
	50	0.6823	0.9016	0.8410	0.9264	0.9426	0.9999	0.9963	0.9985	1.0000	**1.0000**
$LN(0, 0.5)$	10	0.0651	0.0977	0.0363	0.0721	0.0935	0.0470	0.0240	0.0366	0.2426	**0.2734**
	20	0.1643	0.2967	0.1483	0.2108	0.3146	0.4061	0.2595	0.2414	0.7527	**0.7581**
	30	0.2826	0.5021	0.3202	0.4101	0.5282	0.8203	0.6549	0.6402	0.9591	**0.9641**
	50	0.5829	0.8316	0.7481	0.8217	0.8619	0.9986	0.9858	0.9923	0.9985	**0.9997**
$LN(0, 1)$	10	0.1840	0.2140	0.1348	0.2259	0.2164	0.1688	0.0868	0.1550	**0.3401**	0.3007
	20	0.4457	0.5631	0.4293	0.6153	0.6030	0.7152	0.4651	0.6740	**0.8112**	0.7882
	30	0.6798	0.8062	0.7174	0.8812	0.8463	**0.9770**	0.8494	0.9649	0.9220	0.9491
	50	0.9463	0.9825	0.9804	0.9976	0.9898	**1.0000**	0.9994	1.0000	0.9626	0.9968
$LN(0, 2)$	10	0.5288	0.5366	0.5194	**0.6124**	0.5778	0.5997	0.4332	0.5606	0.6026	0.2944
	20	0.8864	0.9364	0.9119	0.9647	0.9596	**0.9874**	0.9116	0.9793	0.8867	0.6136
	30	0.9856	0.9949	0.9940	0.9989	0.9975	**1.0000**	0.9980	0.9999	0.8433	0.7047
	50	1.0000	1.0000	1.0000	1.0000	1.0000	**1.0000**	1.0000	1.0000	0.6499	0.7229
$W(0.5, 1)$	10	0.5208	0.5592	0.4752	0.6271	0.6058	0.5612	0.3795	0.5497	**0.7122**	0.3951
	20	0.8875	0.9504	0.8946	0.9730	0.9721	**0.9846**	0.8984	0.9832	0.9739	0.7221
	30	0.9883	0.9967	0.9932	0.9993	0.9987	**1.0000**	0.9978	0.9999	0.9817	0.7426
	50	1.0000	1.0000	1.0000	1.0000	1.0000	**1.0000**	1.0000	1.0000	0.9571	0.7472
$W(2, 1)$	10	0.0432	0.0859	0.0224	0.0459	0.0827	0.0309	0.0182	0.0204	0.2569	**0.3037**
	20	0.1041	0.2757	0.0955	0.1178	0.3045	0.3812	0.2736	0.1304	0.8362	**0.8399**
	30	0.1978	0.4862	0.2611	0.2435	0.5421	0.8273	0.7101	0.4569	**0.9937**	0.9916
	50	0.4556	0.8138	0.7001	0.6225	0.8780	0.9989	0.9932	0.9692	**1.0000**	1.0000
$EV(0, 1)$	10	0.0457	0.0784	0.0239	0.0493	0.0745	0.0322	0.0176	0.0219	0.2150	**0.2492**
	20	0.1101	0.2302	0.0961	0.1271	0.2426	0.3068	0.2029	0.1313	**0.7197**	0.7190
	30	0.1906	0.4011	0.2242	0.2414	0.4261	0.7303	0.5778	0.4119	**0.9607**	0.9540
	50	0.4171	0.7333	0.6122	0.5811	0.7645	0.9939	0.9718	0.9318	**0.9996**	0.9994

Continued.

	n	W^2	D	V	U^2	A^2	Z_A	Z_C	Z_K	KL	T_n
$IG(1,0.5)$	10	0.2722	0.2990	0.2165	0.3341	0.3098	0.2645	0.1447	0.2516	**0.4187**	0.3235
	20	0.6069	0.7100	0.6034	0.7800	0.7516	0.8475	0.6169	0.8253	**0.8661**	0.7970
	30	0.8395	0.9141	0.8750	0.9608	0.9347	**0.9942**	0.9361	0.9907	0.9441	0.9418
	50	0.9903	0.9964	0.9975	0.9999	0.9984	**1.0000**	1.0000	1.0000	0.9781	0.9822
$IG(1,1)$	10	0.1725	0.2077	0.1218	0.2116	0.2068	0.1544	0.0804	0.1412	**0.3513**	0.3197
	20	0.4269	0.5540	0.4060	0.5894	0.5912	0.6986	0.4662	0.6485	**0.8402**	0.8299
	30	0.6618	0.7991	0.7012	0.8672	0.8354	**0.9733**	0.8554	0.9576	0.9641	0.9731
	50	0.9395	0.9800	0.9763	0.9966	0.9879	**1.0000**	0.9994	1.0000	0.9958	0.9988
$IG(1,2)$	10	0.1033	0.1391	0.0638	0.1208	0.1370	0.0829	0.0438	0.0699	0.2887	**0.3017**
	20	0.2670	0.4067	0.2465	0.3757	0.4314	0.5362	0.3475	0.4243	0.8110	**0.8111**
	30	0.4563	0.6524	0.4990	0.6616	0.6882	0.9187	0.7614	0.8487	0.9688	**0.9789**
	50	0.7865	0.9271	0.8938	0.9633	0.9488	**0.9999**	0.9961	0.9994	0.9990	0.9998
$SN(0,1,0.5)$	10	0.0319	0.0632	0.0148	0.0313	0.0587	0.0198	0.0110	0.0128	0.2057	**0.2482**
	20	0.0661	0.1911	0.0571	0.0607	0.2094	0.2585	0.1811	0.0555	**0.7249**	0.7174
	30	0.1108	0.3423	0.1463	0.1024	0.3680	0.6690	0.5601	0.1893	**0.9740**	0.9633
	50	0.2722	0.6593	0.4984	0.2475	0.7036	0.9858	0.9649	0.6945	**1.0000**	0.9999
$SN(0,1,2)$	10	0.0353	0.0671	0.0173	0.0371	0.0639	0.0227	0.0133	0.0142	0.2026	**0.2496**
	20	0.0762	0.2041	0.0671	0.0768	0.2128	0.2731	0.1934	0.0762	**0.7231**	0.7187
	30	0.1370	0.3577	0.1733	0.1370	0.3756	0.6838	0.5591	0.2448	**0.9747**	0.9658
	50	0.3118	0.6755	0.5270	0.3397	0.7163	0.9892	0.9679	0.7777	**0.9999**	0.9998
$SN(0,1,3)$	10	0.0402	0.0762	0.0213	0.0427	0.0714	0.0277	0.0160	0.0180	0.2275	**0.2705**
	20	0.1044	0.2416	0.0945	0.1102	0.2513	0.3255	0.2333	0.1146	**0.7638**	0.7580
	30	0.1785	0.4214	0.2216	0.2045	0.4428	0.7481	0.6211	0.3538	**0.9797**	0.9720
	50	0.4019	0.7491	0.6237	0.5024	0.7851	0.9954	0.9814	0.8906	**1.0000**	1.0000
$SL(0,1,0.5)$	10	0.0503	0.0671	0.0266	0.0525	0.0634	0.0316	0.0159	0.0252	0.1446	**0.1629**
	20	0.1125	0.1591	0.0912	0.1130	0.1520	0.1815	0.1156	0.0895	0.4346	**0.4366**
	30	0.1839	0.2540	0.1794	0.1859	0.2340	0.4264	0.3219	0.2088	**0.7418**	0.7147
	50	0.3557	0.4717	0.4239	0.3586	0.4266	0.8470	0.7576	0.5339	**0.9644**	0.9512
$SL(0,1,2)$	10	0.0518	0.0695	0.0274	0.0542	0.0643	0.0326	0.0163	0.0250	0.1478	**0.1633**
	20	0.1129	0.1593	0.0913	0.1148	0.1512	0.1797	0.1158	0.0894	0.4382	**0.4406**
	30	0.1849	0.2548	0.1795	0.1841	0.2327	0.4280	0.3211	0.2093	**0.7409**	0.7184
	50	0.3533	0.4675	0.4219	0.3528	0.4228	0.8409	0.7513	0.5291	**0.9662**	0.9524
$SL(0,1,3)$	10	0.0797	0.1049	0.0474	0.0855	0.0956	0.0552	0.0286	0.0445	0.1955	**0.2129**
	20	0.1987	0.2716	0.1721	0.2165	0.2563	0.3235	0.2144	0.2004	**0.5765**	0.5740
	30	0.3302	0.4436	0.3389	0.3704	0.4100	0.6545	0.5227	0.4386	**0.8463**	0.8329
	50	0.6035	0.7449	0.6989	0.6718	0.6950	0.9658	0.9240	0.8390	**0.9840**	0.9821

The power values in Table 3 shows a uniform superiority of the density based empirical likelihood ratio test to all other tests for sample size $n = 10$ against symmetric alternatives. For large sample sizes, the KL test has the most power. Also, a uniform superiority of the proposed test to the EDF-based tests is evident. Moreover, the power differences between the proposed test T_n and the EDF-based tests are substantial.
From Table 4, against asymmetric alternatives, it is seen that the test based on T_n statistic has the most power against some alternatives such as $\Gamma(2, 1)$, $LN(0, 0.5)$, $W(2, 1)$, and $IG(1, 2)$. It is evident that the power values of the proposed test in compared with the EDF-based tests (with the exception of a few alternatives) has the most power and the power differences between the test T_n and these tests are substantial. Also, we can see the KL test has a good performance against asymmetric alternatives and the power difference between KL test and the proposed test is small.
In general, Tables 3 and 4 reveal a uniform superiority of the density based empirical likelihood ratio procedure to the EDF-based tests as it outperforms all competing EDF-based tests. When we compare the proposed test with the entropy-based test (KL), we can conclude that sometimes the proposed test has a higher power, and sometimes entropy test does. We can generally conclude that the proposed test T_n and the KL test have a good performance and therefore can be used in practice.

4. An illustrative example

In this section, we illustrate how the proposed test can be applied to test the goodness-of-fit for the Cauchy distribution when the observations are available. The stock market price is usually modeled by lognormal distribution, that is to say stock market returns follow the Gaussian law. The feature of stock market return distribution is a sharp peak and heavy tails. Therefore, the Cauchy distribution may be a potential model. We apply the proposed test to 30 returns of closing prices of the German Stock Index (DAX). The data are observed daily from 1 January 1991, excluding weekends and public holidays.

Table 5. Scores for 30 returns of closing prices of DAX.

		Observations $n = 30$			
0.0011848	-0.0057591	-0.0051393	-0.0051781	0.0020043	0.0017787
0.0026787	-0.0066238	-0.0047866	-0.0052497	0.0004985	0.0068006
0.0016206	0.0007411	-0.0005060	0.0020992	-0.0056005	0.0110844
-0.0009192	0.0019014	-0.0042364	0.0146814	-0.0002242	0.0024545
-0.0003083	-0.0917876	0.0149552	0.0520705	0.0117482	0.0087458

The data (rounded up to seven decimal places) are presented in Table 5. In Figure 1, the histogram, superimposed by a Cauchy density function, is

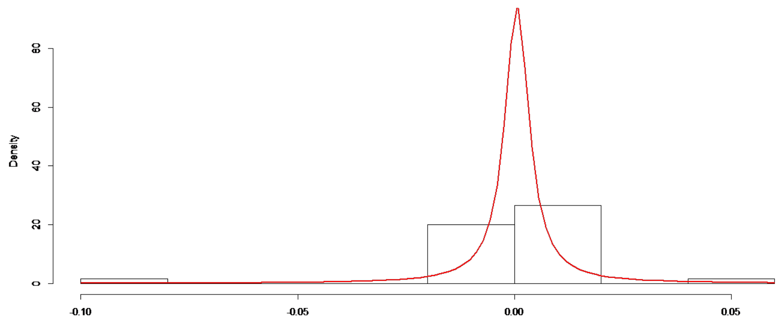

FIGURE 1. The histogram of the 30 returns along with fitted Cauchy density.

displayed. The estimated parameters by using Newton-Raphson method are

$$\hat{\mu} = 0.0005769257 \quad and \quad \hat{\sigma} = 0.003328893.$$

The value of the test statistic is $T_n = 0.004077402$ and the critical value at the 5% is obtained as 784.7944. Since the values of the test statistic is smaller than the critical value, the null hypothesis that the data follow the Cauchy distribution is not rejected at 0.05 significance level. This conclusion seems fairly reliable given the good performance of the proposed test in simulation studies.

5. Conclusion

In this paper, we have proposed a density based empirical likelihood ratio goodness-of-fit test for the Cauchy distribution. The properties and critical values of the test statistic have been derived. We have carried out an extensive power comparison using Monte Carlo simulations and observed that the proposed test outperforms the competing EDF-goodness-of-fit tests which are commonly used in practice. Also, we compared the proposed test with the entropy-based test, and concluded that sometimes the proposed test has a higher power, and sometimes entropy test does. Finally, we have presented a financial real data set and have illustrated how the proposed test can be applied to test the goodness-of-fit for the Cauchy distribution when a sample is available.

6. Aknowledgement

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

References

[1] Berger, J.O. and Wolpert, R.L. (1984), The Likelihood Principle. IMS Lecture Notes, Vol. 6. Institue of Mathematical Statistics, Haywood, CA.

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

[3] Ebner, B., Eid, L., Klar, B. (2021), Cauchy or not Cauchy? New goodness-of-fit tests for the Cauchy distribution. arXiv preprint arXiv:2106.13073.

[4] Glover, S. and Dixon, P. (2004), Likelihood ratios: A simple and flexible statistic for empirical psychologists, Psychonomic Bulletin and Review, 11, 791–806.

[5] Gurevich, G. and Vexler, A. (2011), A Two-sample empirical likelihood ratio test based on samples entropy, Statistics and Computing, 21, 657-670.

[6] Gurevich, G. and Vexler, A. (2018), A density based empirical likelihood approach for testing bivariate normality, Journal of Statistical Computation and Simulation, doi.org/10.1080/00949655.2018.1476516.

[7] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994), Continuous Univariate Distributions, Wiley. ISBN 0-471-58495-9.

[8] Kagan, Y.Y. (1992), Correlations of earthquake focal mechanisms, Geophys. J. Int., 110, 305–320.

[9] Kotz, S., Kozubowski, T.J. and Podgórski, K. (2001), The Laplace distribution and generalizations: a revisit with applications to Communications, Economics, Engineering and Finance. Birkhauser. ISBN 9780817641665.

[10] Lehmann, E.L. (1986), Testing Statistical Hypotheses, 2nd Edition. Springer-Verlag, New York.

[11] Mahdizadeh, M. and Zamanzade, E. (2017), New goodness of fit tests for the Cauchy distribution, Journal of Applied Statistics, 44, 1106-1121.

[12] Mahdizadeh, M. and Zamanzade, E. (2019), Goodness-of-fit testing for the Cauchy distribution with application to financial modeling, Journal of King Saud University – Science, 31, 1167-1174.

[13] Min, I.A., Mezic, I. and Leonard, A. (1996), Levy stable distributions for velocity and velocity difference in systems of vortex elements, Phys. Fluids, 8, 1169–1180.

[14] Nolan, J.P. (2014), Financial modeling with heavy-tailed stable distributions, WIREs Comput. Stat., 6, 45–55.

[15] Shan, G., Vexler, A., Wilding, G. E. and Hutson A. D. (2011), Simple and exact empirical likelihood ratio tests for normality based on moment relations, Communications in Statistics-Simulation and Computation, 40, 141–158.

[16] Shepherd, L.A., Tsai, W-M., Vexler, A. and Miecznikowski, J.C. (2013), dbEmpLikeNorm: test for joint assessment of normality, R package. Available at: http://cran.r-project.org/web/packages/dbEmpLikeNorm/index.html.

[17] Solomon, D.L. (1975), A note on the non-equivalence of the Neyman-Pearson and generalized likelihood ratio tests for testing a simple null versus a simple alternative hypothesis (PDF), The American Statistician, 29, 101–102.

[18] Stapf, S. and Kimmich, R. and Seitter, R.O., Maklakov, A.I. and Skirda, V.D. (1996), Proton and deuteron fieldcycling NMR relaxometry of liquids confined in porous glasses, Coll. Surf. Physicochem. Eng. Aspects, 115, 107–114.

[19] Tanajian, H., Vexler. A. and Hutson, A.D. (2013), Novel and efficient density based empirical likelihood procedures for symmetry and K-sample comparisons: STATA package. Available at http://sphhp.buffalo.edu/biostatistics/research-and-acilities/software/stata.html.

[20] Vexler, A. and Gurevich, G. (2010), Empirical likelihood ratios applied to goodness-of-fit tests based on sample entropy, Computational Statistics and Data Analysis, 54, 531–545.

[21] Vexler, A. and Gurevich, G. (2011), A note on optimality of hypothesis testing, Mathematics in Engineering, Science and Aerospace, 2, 243-250.

[22] Vexler, A. and Yu, J. (2011), Two-sample density-based empirical likelihood tests for incomplete data in application to a pneumonia study, Biometrical Journal, 53, 628–651.

[23] Vexler, A., Kim, Y.M., Yu, J., Lazar, N.A. and Hutson, A.D. (2014), Computing critical values of exact tests by incorporating Monte Carlo simulations combined with statistical tables, Scandinavian Journal of Statistics, DOI: 10.1111/sjos.12079.

[24] Vexler, A., Liu, S. and Schisterman, E.F. (2011c), Nonparametric likelihood inference based on cost-effectively-sampled-data, Journal of Applied Statistics, 38, 769–783.

[25] Vexler, A., Shan, G., Kim, S., Tsai, W. M., L. Tian and Hutson, A. D. (2011a), An empirical likelihood ratio based goodness-of-fit test for Inverse Gaussian distributions, Journal of Statistical Planning and Inference, 141, 6, 2128-2140.

[26] Vexler, A., Tsai, W-M. and Malinovsky, Y. (2012a), Estimation and testing based on data subject to measurement errors: from parametric to non-parametric likelihood methods, Statistics in Medicine, 31, 2498-2512.

[27] Vexler, A., Tsai, W-M., Gurevich, G. and Yu, J. (2012b), Two-sample density-based empirical likelihood ratio tests based on paired data, with application to a treatment study of Attention-Deficit/Hyperactivity Disorder and Severe Mood Dysregulation, Statistics in Medicine, 31, 1821–1837.

[28] Vexler, A., Yu, J. and Hutson, A.D. (2011b), Likelihood testing populations modeled by autoregressive process subject to the limit of detection in applications to longitudinal biomedical data, Journal of Applied Statistics. 38, 1333-1346.

[29] Villaseñor, J. A., González-Estrada, E. (2021), Goodness-of-Fit Tests for Cauchy Distributions Using Data Transformations. In Advances in Statistics-Theory and Applications (pp. 271-282). Springer, Cham.

[30] Winterton, S.S., Smy, T.J. and Tarr, N.G. (1992), On the source of scatter in contact resistance data, J. Electron. Mater., 21, 917–921.

[31] Yu, J., Vexler, A. and Tian, L. (2010), Analyzing incomplete data subject to a threshold using empirical likelihood methods: an application to a pneumonia risk study in an ICU setting, Biometrics, 66, 123-130.

[32] Yu, J., Vexler, A., Kim, S. and Hutson, A. D. (2011), Two-sample empirical likelihood ratio tests for medians in application to biomarker evaluations, The Canadian Journal of Statistics. 39, 671–689.

[33] Yu, J., Yang, L., Vexler, A., Hutson, A. D. (2016), A generalized empirical likelihood approach for two-group comparisons given a U-statistic constraint, Statistics: A Journal of Theoretical and Applied Statistics, 50, 435-453.

[34] Zhang, J. (2002), Powerful goodness-of-fit tests based on the likelihood ratio, Journal of Royal Statistical Society, Series B, 64, 281-294.

[35] Zhao, Y., Vexler, A., Hutson, A., Chen, X. (2017), A statistical software procedure for exact parametric and nonparametric likelihood-ratio tests for two-sample comparisons, Communications in Statistics - Simulation and Computation, 46, 2829-2841.

Hadi Alizadeh Noughabi
Orcid number: 0000-0002-7515-1896
Department of Statistics
University of Birjand
Birjand, Iran
Email address: alizadehhadi@birjand.ac.ir