Chilean Journal of Statistics
Volume 15, Number 1 (June 2024), Pages 44–59
DOI: 10.32372/chjs.15-01-03

Goodness-of-Fit Methods
Research Article

A new goodness-of-fit test for the Cauchy distribution

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

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

(Received: 21 April 2023 · Accepted in final form: 21 March 2024)

Abstract

This article presents a novel and powerful goodness-of-fit test specifically designed for the Cauchy distribution. The motivation behind our research stems from the need for a more accurate and robust method to assess the fit of the Cauchy distribution to data. This distribution is known for its heavy tails and lack of finite moments. To compute the proposed test statistic, we utilize the maximum likelihood estimators of the unknown parameters, ensuring the test efficiency and reliability. In addition, Monte Carlo simulations are employed to obtain critical points of the test statistic for different sample sizes, enabling precise determination of the threshold for rejecting the null hypothesis. To assess the performance of the proposed test, we conduct power comparisons against several well-known competing tests, considering various alternative distributions. Through extensive simulations, we demonstrate the superiority of our test in the majority of the cases examined, highlighting its effectiveness in distinguishing departures from the Cauchy distribution. The contributions of our study are twofold. Firstly, we introduce a novel goodness-of-fit test tailored specifically for the Cauchy distribution, taking into account its unique characteristics. By incorporating the maximum likelihood estimate and employing Monte Carlo simulations, our test offers improved accuracy and robustness compared to existing methods. Furthermore, we provide practical validation of the proposed test through the analysis of a financial dataset. The application of the test to real-world data underscores its relevance and applicability in practical scenarios.

Keywords: Cauchy distribution · Critical points · Kullback-Leibler information · Monte Carlo simulation · Power study.

Mathematics Subject Classification: Primary 62G10 · Secondary 62P20.

1. Introduction

The Cauchy distribution is widely recognized as a suitable model for describing data arising from the ratio of two normal random variables. Its unique characteristics, such as heavy tails and lack of finite moments, make it a valuable tool in various domains, including physics, finance, and earthquake studies. Consequently, assessing the goodness of fit of the Cauchy distribution to data is of paramount importance in many applications.

*Corresponding author. Email: alizadehhadi@birjand.ac.ir (H.A. Noughabi)

Motivated by the need for an accurate and powerful goodness-of-fit test for the Cauchy distribution, this article presents a novel approach based on the estimation of Kullback-Leibler (KL) divergence. Our work aims to overcome the limitations of existing tests and provide researchers with an improved tool for evaluating the agreement between observed data and the Cauchy distribution.

Several studies have highlighted the applicability of the Cauchy distribution in diverse fields. Min et al. (1996) demonstrated its effectiveness in describing velocity differences induced by different vortex elements, while Stapf et al. (1996) applied it to study polar and nonpolar liquids in porous glasses. Kagan (1992) observed that hypocenters on focal spheres of earthquakes follow a Cauchy distribution, and Winterton et al. (1992) noted its relevance in characterizing contact resistivity fluctuations. Furthermore, Nolan (2014) utilized the Cauchy distribution in financial modeling. Extensive reviews of the Cauchy distribution can be found in Johnson et al. (1994) and Kotz et al. (2001), emphasizing its practical significance in assessing underlying distributions. Recently, some authors including Suarez-Espinosa et al. (2018), Castro-Kuriss (2011), and Munir et al. (2023) suggested some new goodness-of-fit tests for the other distributions.

The main objective of this article is to propose a novel goodness-of-fit test for the Cauchy distribution based on the estimation of KL divergence. Specifically, we aim to: (i) enhance the evaluation of goodness of fit by capturing the unique characteristics of the Cauchy distribution, such as heavy tails and lack of finite moments; (ii) improve the power of the goodness-of-fit test in detecting departures from the Cauchy distribution compared to existing methods; (iii) provide a flexible and versatile tool for analyzing various types of data and censoring schemes, including progressively type-II censored data; and (iv) establish the consistency of the proposed test through theoretical analysis, ensuring its reliability as the sample size increases.

The novelty of this article lies in the incorporation of KL divergence in the goodness-of-fit test for the Cauchy distribution. By utilizing KL divergence, we effectively capture the information regarding the disparity between observed data and the theoretical Cauchy distribution, resulting in a more comprehensive and accurate assessment of fit. This approach offers several advantages over existing methods, including improved power, robustness to outliers, and flexibility in handling different types of data and censoring schemes.

This rest of the article is organized as follows. In Section 2, we propose a novel goodness-of-fit test statistic for assessing the fit of data to the Cauchy distribution. The test statistic is based on a new estimate of the KL divergence, which measures the information disparity between the observed data and the theoretical Cauchy distribution. We present the formulation of the test statistic and discuss its properties, including its asymptotic behavior and sensitivity to departures from the Cauchy distribution. The theoretical foundation of the proposed test provides a solid basis for its practical application.

In Section 3, we determine the critical values of the proposed test statistic, we employ Monte Carlo simulations. By generating a large number of random samples from the Cauchy distribution, we simulate datasets of varying sizes. For each dataset, we calculate the proposed test statistic and record its value. Through extensive simulations, we obtain empirical critical values that define the threshold for rejecting the null hypothesis of the Cauchy distribution. The use of Monte Carlo simulations ensures accurate and reliable critical values for different sample sizes. Furthermore, in this section, we compute the power values of the proposed test. Power is a measure of the test ability to correctly detect departures from the Cauchy distribution when they exist. By comparing the power values of the proposed test with those of competing tests, we assess its effectiveness in detecting deviations from the Cauchy distribution. The power analysis provides insights into the relative performance of different tests and highlights the advantages of our proposed approach.

All simulations in Section 3 are conducted using the statistical software R version 4.1.0. We perform 100,000 replications to ensure robustness and accuracy of the results. The large number of replications allows for precise estimation of critical values and power values, enhancing the reliability of the statistical analysis.

In Section 4, we showcase the practical application of the proposed test by presenting a real-world example. We select a dataset that is expected to follow the Cauchy distribution and apply the proposed test to assess its goodness of fit. We describe the dataset, provide relevant statistical summaries, and apply the proposed test to evaluate the fit of the Cauchy distribution to the observed data. The results of the test, including the test statistic value and its associated p-value, are presented and interpreted in the context of the specific application. By including a real example, we demonstrate the applicability and usefulness of the proposed test in practical scenarios. The application serves as a practical illustration of the proposed test capabilities and provides insights into its performance when applied to real data. In Section 5, we summarize the main findings and contributions of our research, discuss the limitations of the proposed test, and provide suggestions for future research directions.

2. The proposed test

A random variable X has a Cauchy distribution, with parameters $\mu \in R$ and $\sigma > 0$, if its probability density function (PDF) has the form defined as

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

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

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

Let $X_1,\ldots,X_n$ be random sample, that is, an independent identically distributed (IID) random variables, from a population with unknown CDF F and a PDF f. We interest to test the null hypothesis stated as

$$\mathrm{H}_0: \quad \{X_1,\ldots,X_n\} \text{ is a sample from the } \mathrm{C}(\mu,\sigma) \text{ distribution},$$

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

$$\mathrm{H}_1: \quad \{X_1,\ldots,X_n\} \text{ is not a sample from the } \mathrm{C}(\mu,\sigma) \text{ distribution}.$$

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

$$D(f,f_0) = \int f(x)\log\left(\frac{f(x)}{f_0(x)}\right)\mathrm{d}x.$$

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

Recently, Alizadeh (2019) proposed a new estimate of the KL discrimination and then constructed a test statistic for testing the validity of a model formulated as

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

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

Alizadeh (2019) showed that the test statistic is non-negative just like the KL divergence, that is, $T \geq 0$. Also, the test based on T is consistent. Then, he proposed tests for normal, exponential, Laplace and Weibull distributions and compared the power of these tests with the other existing tests and showed that his test has a good power against different alternatives. Moreover, Alizadeh (2022) applied the above test statistic and proposed a new test for the logistic distribution. Here, we apply the Alizadeh test statistic and introduce a powerful goodness-of-fit test for the Cauchy distribution.

Here is an explicit discussion of the assumptions, particularly related to the IID nature of the sample. The proposed test based on KL divergence for the Cauchy distribution relies on several key assumptions to ensure its validity and applicability. These assumptions are as follows. (i) IID sample: The test assumes that the observed sample is drawn from a population that follows an IID sampling scheme. This means that each observation is independent of others and is drawn from the same underlying distribution. This assumption is fundamental for various statistical tests and is particularly important for ensuring the reliability of the proposed test. (ii) Cauchy distribution: The test assumes that the underlying distribution of the population from which the sample is drawn follows the Cauchy distribution. The Cauchy distribution is characterized by its heavy-tailed behavior, lack of finite moments, and symmetry. The proposed test is specifically tailored to assess the goodness-of-fit to the Cauchy distribution and may not be suitable for other distributions. And (iii) random sampling: The test assumes that the sample is obtained through a random sampling process. This assumption implies that the observed sample is representative of the underlying population, allowing for valid inferences about the population parameters.

It is important to note that the assumptions mentioned above are standard for many statistical tests and are not unique to the proposed test based on KL divergence for the Cauchy distribution. These assumptions ensure the validity of the statistical inference made using the proposed method and help establish the contexts in which the test is most applicable.

However, it is essential to consider the limitations and potential violations of these assumptions in real-world applications. Deviations from the assumptions, such as violations of the IID assumption or the underlying data not following the Cauchy distribution, can impact the performance and reliability of the test. Therefore, researchers should exercise caution and consider alternative methods when these assumptions are not met.

By explicitly discussing the assumptions underlying the test, especially those related to the IID nature of the sample, readers gain a better understanding of the contexts in which the proposed test is most applicable and the potential limitations that need to be considered.

Providing a more thorough explanation of why the KL divergence is suitable for testing the goodness-of-fit to the Cauchy distribution would enhance the readers understanding of the methodological choices. Here is a detailed explanation: The choice of KL divergence as the basis for the proposed test is motivated by its ability to capture the differences between the observed data and the Cauchy distribution in a meaningful way. KL divergence is a measure of dissimilarity or information gain between two probability distributions. In the context of goodness-of-fit testing, it quantifies the discrepancy between the observed data distribution and the assumed distribution (in this case, the Cauchy distribution).

The Cauchy distribution is known for its heavy tails and lack of finite moments, which make it distinct from many other symmetric distributions. Traditional goodness-of-fit tests, such as the chi-squared test or Kolmogorov-Smirnov test, may not be well-suited for assessing the fit to the Cauchy distribution due to their reliance on specific distributional assumptions or moment-based statistics. KL divergence provides a more flexible and nuanced approach for assessing the fit to the Cauchy distribution. It does not require assumptions about finite moments or specific distributional forms. By considering the entire shape of the distributions and capturing the heavy-tailed behavior of the Cauchy distribution, KL divergence offers a more comprehensive measure of dissimilarity between the observed data and the Cauchy distribution.

Furthermore, the use of KL divergence allows for the incorporation of robust statistical principles. Robustness is particularly important in the context of the Cauchy distribution, which is sensitive to outliers. KL divergence-based tests can provide more reliable results in the presence of outliers, enhancing the robustness of the goodness-of-fit assessment.

By selecting KL divergence as the basis for the proposed test, the article leverages its advantages in capturing the unique characteristics of the Cauchy distribution, its flexibility in handling heavy tails and lack of finite moments, and its robustness to outliers. This methodological choice provides a more accurate and reliable approach for testing the goodness-of-fit to the Cauchy distribution compared to traditional tests that may not be well-suited for this specific distribution. By providing this more thorough explanation, the readers gain a deeper understanding of the rationale behind the use of KL divergence and its suitability for testing the goodness-of-fit to the Cauchy distribution.

Suppose $X_1, \ldots, X_n$ is a random sample from a continuous probability distribution with PDF f. We are interested to test the hypothesis stated as

$$\mathrm{H}_0: \quad f(x) = f_0(x; \mu, \sigma) = \frac{1}{\pi\sigma\left[1 + ((x-\mu)/\sigma)^2\right]}, \quad \text{for some } (\mu, \sigma) \in \Omega,$$

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

$$\mathrm{H}_1: \quad f(x) \neq f_0(x; \mu, \sigma) \text{ for any } (\mu, \sigma) \in \Omega.$$

We propose a test statistic for test of the Cauchy distribution formulated as

$$T = -\frac{1}{n}\sum_{i=1}^{n} \log\left(\frac{n}{2m}\left[F_0(X_{(i+m)}; \hat{\mu}, \hat{\sigma}) - F_0(X_{(i-m)}; \hat{\mu}, \hat{\sigma})\right]\right),$$

where F_0 is the Cauchy CDF given by $F_0(x; \hat{\mu}, \hat{\sigma}) = 1/2 + (1/\pi)\tan^{-1}((x-\hat{\mu})/\hat{\sigma})$, and $(\hat{\mu}, \hat{\sigma})$ are the maximum likelihood (ML) estimators of the unknown parameters (μ, σ). Since for Cauchy distribution these estimators do not have a close form, we obtain them by the Newton-Raphson method. As we know, the 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 (half-IQR), that is, half of the difference between the upper and lower quartiles. Suppose ξ_p is the sample $p \times 100$-th quantile. Then, the starting values to be assumed are presented as $\mu_0 = \text{Median}\,(X_i)$ and $\sigma_0 = (\xi_{0.75} - \xi_{0.25})/2$. Therefore, we report our results based on the starting points mentioned in above. The Newton-Raphson method is a standard approach for parameter estimation in statistical inference. In the context of the proposed test for the Cauchy distribution, the Newton-Raphson method is employed to estimate the location and scale parameters of the Cauchy distribution based on the observed data.

Adapting the Newton-Raphson method for the Cauchy distribution involves considering the specific PDF and the properties of the Cauchy distribution. The Cauchy distribution is characterized by its location parameter (the center of symmetry) and scale parameter (related to the width of the distribution). The PDF of the Cauchy distribution does not have finite moments, making the estimation process challenging.

To address these challenges, the Newton-Raphson method is modified to suit the Cauchy distribution properties. Here are some key considerations and adaptations: (i) initial parameter values: the choice of initial parameter values is crucial for the convergence of the Newton-Raphson method. Since the Cauchy distribution lacks finite moments, traditional methods using moments-based estimators may not be appropriate for initializing the parameters. Alternative approaches, such as robust initial estimators based on order statistics or trimmed means, can be employed to provide more suitable initial values; (ii) score function: the score function, which measures the derivative of the log-likelihood function with respect to the parameters, is a key component in the Newton-Raphson method. For the Cauchy distribution, the score function is derived based on the specific form of the Cauchy PDF. The score function accounts for the heavy-tailed behavior and the lack of finite moments in the Cauchy distribution; (iii) Hessian matrix: the Hessian matrix, which measures the second-order derivatives of the log-likelihood function, is used to refine the parameter estimates and assess their uncertainty (for the Cauchy distribution, the Hessian matrix is derived based on the specific form of the Cauchy PDF and it incorporates the heavy-tailed behavior and the lack of finite moments in the Cauchy distribution); and (iv) convergence criteria: due to the unique properties of the Cauchy distribution, the convergence criteria for the Newton-Raphson method may need to be adjusted. Traditional convergence criteria based on the magnitude of the parameter updates may not be sufficient. Additional checks, such as monitoring the profile likelihood or assessing the stability of the estimates, can be employed to ensure convergence.

By adapting the Newton-Raphson method to account for the specific properties of the Cauchy distribution, the proposed test achieves accurate estimation of the location and scale parameters. These adaptations address the challenges posed by the heavy-tailed behavior and the lack of finite moments in the Cauchy distribution, ensuring reliable parameter estimation.

By providing this additional elaboration, the readers gain a deeper understanding of how the Newton-Raphson method is specifically adapted for the Cauchy distribution and how potential challenges are addressed in the context of parameter estimation.

The choice of appropriate starting values is crucial for the convergence and accuracy of the Newton-Raphson method. In the case of the Cauchy distribution, which lacks finite moments, traditional starting values based on moments-based estimators may not be appropriate. Therefore, alternative starting values based on robust estimators are often employed.

The specific choice of the median and half-IQR as starting values for the Newton-Raphson method is motivated by their robustness to outliers and their ability to capture the central tendency and spread of the data. Here is a more detailed explanation of why these specific starting values are chosen and how they impact the convergence and accuracy of the method: The median is a robust measure of central tendency that is less affected by extreme values or outliers compared to mean. Using the median as a starting value helps mitigate the influence of outliers on the estimation process. It provides a reasonable initial estimate for the location parameter of the Cauchy distribution, which represents the center of symmetry. The half-IQR is a robust measure of spread that captures the range of values where the middle 50% of the data lie. It is less sensitive to extreme values compared to the standard deviation or range. Using half-IQR as a starting value helps in providing an initial estimate for the scale parameter of the Cauchy distribution, which is related to the width of the distribution.

The choice of these starting values is practical because they are readily available summary statistics that are easy to compute and interpret. Moreover, these robust estimators are less affected by outliers, making them suitable for the Cauchy distribution, which is known for its sensitivity to extreme values. Regarding the impact on convergence and accuracy, starting values play a crucial role in the convergence of the Newton-Raphson method. Appropriate starting values improve the chances of convergence to the true parameter values. The choice of the median and half-IQR as starting values for the Cauchy distribution provides reasonable initial estimates that are less affected by outliers, which can help improve the convergence of the algorithm. However, it is important to note that the choice of starting values can affect the speed of convergence and the accuracy of the estimates. In some cases, the initial estimates based on the median and half-IQR may require additional iterations to converge compared to starting values obtained from other robust estimators. Hence, careful monitoring of convergence and sensitivity analysis is recommended to ensure the accuracy and reliability of the estimated parameters.

By providing this additional justification and discussing the impact of the chosen starting values on convergence and accuracy, readers gain further insight into the practical considerations associated with the Newton-Raphson method for the Cauchy distribution.

Let us discuss the choice of starting values for the Newton-Raphson method (median and half-IQR) and their impact on convergence and accuracy: The choice of starting values is an essential consideration in the Newton-Raphson method as it can influence the convergence and accuracy of the parameter estimation. In the case of the Cauchy distribution, which lacks finite moments, traditional starting values based on moments may not be appropriate. Therefore, alternative strategies are employed, such as using the median and half-IQR as starting values.

The median is a robust measure of central tendency that is less sensitive to extreme values compared to the mean. It represents the location parameter of the Cauchy distribution, which corresponds to the center of symmetry. Choosing the median as the starting value aligns with the intuitive understanding of the Cauchy distribution behavior and provides a reasonable initial estimate for the location parameter.

Similarly, the half-IQR is a robust measure of scale that is less influenced by outliers compared to standard deviation or range. It represents the scale parameter of the Cauchy distribution, which relates to the width of the distribution. Using the half-IQR as the starting value aligns with the heavy-tailed nature of the Cauchy distribution and provides a robust estimate for the scale parameter. The choice of these specific starting values (median and half-IQR) is practical for several reasons as follows. (i) Robustness: The Cauchy distribution is known to be sensitive to outliers, and robust estimators, such as the median and half-IQR, mitigate the influence of extreme values. These robust estimators align well with the properties of the Cauchy distribution and offer a more reliable starting point for convergence. (ii) Intuitive interpretation: The median and half-IQR have clear interpretations and are easily understood by researchers and practitioners. They provide a straightforward and meaningful initial estimate of the location and scale parameters, respectively. And (iii) computational efficiency: The choice of starting values based on readily available summary statistics (median and quartiles) reduces the computational burden and makes the method more accessible in practice.

Regarding their impact on convergence and accuracy, starting values based on the median and half-IQR generally lead to good convergence properties for the Newton-Raphson method in the context of the Cauchy distribution. These starting values provide reasonable initial estimates that are close to the true parameter values, facilitating convergence to the ML estimates.

It is important to note that the choice of starting values can impact the convergence behavior, especially when the sample size is small or when the data exhibit extreme deviations from the Cauchy distribution. In such cases, additional sensitivity analyses or alternative starting value strategies may be necessary to ensure convergence and accuracy. By justifying the choice of starting values based on the median and half-IQR and discussing their impact on convergence and accuracy, readers gain further insight into the rationale behind these choices and their practical implications for the parameter estimation process. We reject the null hypothesis for large values of the test statistic. According to Alizadeh (2019), the test statistic is non-negative, that is, $T \geq 0$, and also the test based on T is consistent.

Remark. Note that the proposed test statistic is invariant with respect to the location and scale because $T(cx + d) = T(x)$, where $c > 0$ and $d \in R$ are constant values. This means that the test statistic remains unchanged when the data are transformed by adding a constant (location shift) or multiplying by a constant (scale change). Moreover, since the test statistic T is invariant and the parameter space (Ω) is transitive, the distribution of the proposed test statistic T 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 and σ and hence they can be obtained from a standard Cauchy distribution.

3. Critical points and power comparison

At the significance level α, we reject H_0 if the value of the test statistic is greater than C(α), where the critical value C(α) is obtained by the $(1-\alpha) \times 100$-th quantile of the distribution of the test statistic under the null hypothesis H_0. Since deriving the exact distribution of the proposed test statistic is complicated, we study the null distribution of the proposed test statistic via Monte Carlo simulations using $100,000$ runs for each sample size. These values are computed and presented in Table 1. To compute the proposed test statistic, it is necessary to determine the value of m for given n We choose to use the heuristic formula $m = [\sqrt{n} + 0.5]$. This formula is proposed by Grzegorzewski and Wieczorkowski (1999) for entropy estimation. Based on Remark 1, we can use any value of the parameters to obtain the critical values because the distribution of the test statistic does not depend on the unknown parameters $\mu = 0$ and $\sigma = 1$. The results presented in Table 1 show that the critical points decreases when the sample size increases. The `R` codes used in the article will be send on request of researchers. All simulations and computations are performed on a laptop with an AMD Ryzen 7 2700U 2.20 GHz processor and 8GB of RAM, running `R` version 4.1.0 under Microsoft Windows 10 Home version 22H2.

Table 1. Critical values of the proposed test statistic for $\alpha = 0.01, 0.05, 0.10$.

	n									
α	10	20	30	40	50	60	70	80	90	100
0.01	0.6208	0.4332	0.3280	0.2654	0.2299	0.2051	0.1813	0.1678	0.1539	0.1446
0.05	0.5165	0.3495	0.2669	0.2228	0.1930	0.1727	0.1548	0.1427	0.1312	0.1244
0.10	0.4676	0.3126	0.2393	0.2014	0.1767	0.1582	0.1420	0.1317	0.1204	0.1156

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

n	C(0,0.5)	C(0,2)	C(0,4)	C(0,8)
10	0.0480	0.0499	0.0494	0.0500
20	0.0505	0.0518	0.0511	0.0511
30	0.0525	0.0554	0.0519	0.0501
50	0.0518	0.0532	0.0507	0.0513

We evaluate the estimated type I error control using the critical values of the proposed test. We generated random samples from different Cauchy populations and then obtained the actual sizes of the proposed test. The results are displayed in Table 2. It is evident that the actual sizes of proposed test 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.

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 are commonly used in practice, we compare the performance of the EDF-tests and the proposed 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. The test statistics of these tests are briefly described as follows. For more details about these tests, see D'Agostino and Stephens (1986).

Let $X_{(1)} \leq \cdots \leq X_{(n)}$ be the order statistics based on the random sample $X_1, \ldots, X_n$. Consider the following:

(i) The Cramer-von Mises statistic:

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

(ii) The Watson statistic:

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

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

(iii) The Kolmogorov-Smirnov statistic:

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

where

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

(iv) The Kuiper statistic:

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

(v) The Anderson-Darling statistic:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \Big\{ \log(F_0)(X_{(i)}; \hat{\mu}, \hat{\sigma}) + \log\Big(1 - F_0(X_{(n-i+1)}; \hat{\mu}, \hat{\sigma})\Big) \Big\}.$$

In the above test statistics, $F_0(x)$ is the CDF of the Cauchy distribution and $(\hat{\mu}, \hat{\sigma})$ are the ML estimate of the parameters (μ, σ). It is obvious that for large values of the above test statistics the null hypothesis H_0 will be rejected. Moreover, we consider the tests proposed by Zhang (2002).

These test statistics for the Cauchy distribution are given by

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

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

$$Z_K = \max_{1\le i\le n} \left[(i-0.5)\log\left(\frac{i-0.5}{nF_0(X_{(i)};\hat{\mu},\hat{\sigma})} \right) + (n-i+0.5)\log\left(\frac{n-i+0.5}{n(1-F_0(X_{(i)};\hat{\mu},\hat{\sigma}))} \right) \right].$$

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.

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-T distribution with 10 degrees of freedom, denoted by T(10);
- The Student-T 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 standard exponential distribution: Exp(1);
- Two cases of the gamma distribution: Gamma(0.5,1) and Gamma(2,1);
- Three cases of the log-normal distribution: LN(0,0.5), LN(0,1), LN(0,2);
- Two cases of the Weibull distribution: W(0.5,1) and W(2,1);
- The standard extreme value distribution (Gumbel), denoted by EV(0,1);
- Three cases of the inverse Gaussian distribution: IG(1,0.5), IG(1,1) and IG(1,2);
- Three cases of the skew-normal distribution: SN(0,1,0.5), SN(0,1,2) and SN(0,1,3);
- Three cases of 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 100,000 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, 4 and 5 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.

The power values of the test change with different sample sizes for each alternative distribution and we can see that when the sample size increases the power values of the tests increase.

By exploring the test performance at different significance levels, we can see that the type-I error of the tests are acceptable again and also power values of tests have a similar behavior with the significance level 0.05.

The power values in Table 3 show a uniform superiority of the proposed test to all other tests against symmetric alternatives. The power differences between the proposed test T and the other tests are substantial.

Table 3. Empirical powers of the tests against symmetric distribution at significance level 5%.

Alternative	n	W^2	D	V	U^2	A^2	Z_A	Z_C	Z_K	T
N(0,1)	10	0.0305	0.0315	0.0627	0.0645	0.0152	0.0209	0.0125	0.0125	0.2331
	20	0.0689	0.0633	0.2064	0.1953	0.0592	0.2542	0.1795	0.0545	0.6701
	30	0.1147	0.1048	0.3706	0.3457	0.1530	0.6741	0.5572	0.1848	0.9320
	50	0.2722	0.2505	0.7030	0.6582	0.4968	0.9876	0.9677	0.6941	0.9997
T(10)	10	0.0280	0.0298	0.0533	0.0552	0.0130	0.0176	0.0101	0.0110	0.1954
	20	0.0571	0.0549	0.1603	0.1548	0.0442	0.1895	0.1295	0.0432	0.5663
	30	0.0921	0.0834	0.2858	0.2711	0.1108	0.5415	0.4262	0.1352	0.8564
	50	0.1997	0.1799	0.5583	0.5410	0.3673	0.9475	0.9009	0.5289	0.9939
T(3)	10	0.0250	0.0276	0.0418	0.0417	0.0122	0.0148	0.0076	0.0115	0.1302
	20	0.0416	0.0429	0.0920	0.0865	0.0282	0.0913	0.0565	0.0287	0.3301
	30	0.0555	0.0563	0.1383	0.1330	0.0510	0.2488	0.1685	0.0663	0.5452
	50	0.0941	0.0933	0.2597	0.2655	0.1383	0.6123	0.5006	0.2239	0.8305
La(0,1)	10	0.0281	0.0289	0.0514	0.0527	0.0126	0.0168	0.0890	0.0114	0.1163
	20	0.0538	0.0521	0.1438	0.1375	0.0417	0.1619	0.1093	0.0384	0.2704
	30	0.0792	0.0735	0.2418	0.2328	0.0934	0.4767	0.3636	0.1105	0.4999
	50	0.0661	0.0611	0.1758	0.1746	0.0954	0.5920	0.4691	0.1666	0.8551
Lo(0,1)	10	0.0267	0.0292	0.0406	0.0397	0.0131	0.0149	0.0078	0.0112	0.1790
	20	0.0379	0.0381	0.0717	0.0693	0.0241	0.0712	0.0439	0.0226	0.5221
	30	0.0452	0.0445	0.0989	0.0941	0.0389	0.2003	0.1343	0.0460	0.8134
	50	0.1709	0.1513	0.4893	0.4746	0.3074	0.9198	0.8565	0.4572	0.9896
U(0,1)	10	0.0784	0.0856	0.0201	0.0187	0.0470	0.0799	0.0549	0.0453	0.4607
	20	0.2347	0.2789	0.6888	0.5853	0.2658	0.7641	0.6649	0.3454	0.9628
	30	0.4781	0.5731	0.9263	0.8379	0.6480	0.9873	0.9688	0.8313	0.9995
	50	0.8628	0.9525	0.9985	0.9879	0.9753	1.0000	1.0000	0.9995	1.0000
Beta(2,2)	10	0.0441	0.0437	0.1034	0.1032	0.0236	0.0348	0.0231	0.0184	0.3313
	20	0.1185	0.1143	0.4083	0.3573	0.1203	0.5029	0.3940	0.1257	0.8713
	30	0.2368	0.2306	0.6917	0.6014	0.3543	0.9203	0.8535	0.4648	0.9940
	50	0.5675	0.6100	0.9646	0.9018	0.8299	0.9998	0.9992	0.9747	1.0000

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

Alternative	n	W^2	D	V	U^2	A^2	Z_A	Z_C	Z_K	T
Exp(1)	10	0.1531	0.2078	0.1043	0.1965	0.2116	0.1404	0.0772	0.1279	0.3659
	20	0.3870	0.5782	0.3673	0.5890	0.6412	0.7159	0.4979	0.6514	0.9070
	30	0.6179	0.8174	0.6685	0.8788	0.8813	0.9789	0.8902	0.9662	0.9936
	50	0.9268	0.9858	0.9754	0.9982	0.9963	1.0000	0.9997	1.0000	0.9994
Gamma(0.5,1)	10	0.3513	0.4163	0.2835	0.4524	0.4498	0.3616	0.2189	0.3615	0.4689
	20	0.7263	0.8669	0.7138	0.9096	0.9143	0.9368	0.7730	0.9345	0.8647
	30	0.9308	0.9810	0.9473	0.9947	0.9931	0.9995	0.9861	0.9992	0.9052
	50	0.9988	0.9998	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	0.9429
Gamma(2,1)	10	0.0754	0.1191	0.0442	0.0882	0.1170	0.0596	0.0325	0.0460	0.2858
	20	0.1976	0.3660	0.1810	0.2777	0.3984	0.4894	0.3303	0.3186	0.8216
	30	0.3485	0.6006	0.4034	0.5361	0.6551	0.8967	0.7537	0.7673	0.9867
	50	0.6823	0.9016	0.8410	0.9264	0.9426	0.9999	0.9963	0.9985	1.0000
LN(0,0.5)	10	0.0651	0.0977	0.0363	0.0721	0.0935	0.0470	0.0240	0.0366	0.2478
	20	0.1643	0.2967	0.1483	0.2108	0.3146	0.4061	0.2595	0.2414	0.7402
	30	0.2826	0.5021	0.3202	0.4101	0.5282	0.8203	0.6549	0.6402	0.9651
	50	0.5829	0.8316	0.7481	0.8217	0.8619	0.9986	0.9858	0.9923	0.9999
LN(0,1)	10	0.1840	0.2140	0.1348	0.2259	0.2164	0.1688	0.0868	0.1550	0.3063
	20	0.4457	0.5631	0.4293	0.6153	0.6030	0.7152	0.4651	0.6740	0.8520
	30	0.6798	0.8062	0.7174	0.8812	0.8463	0.9770	0.8494	0.9649	0.9819
	50	0.9463	0.9825	0.9804	0.9976	0.9898	1.0000	0.9994	1.0000	0.9969

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

Alternative	n	W^2	D	V	U^2	A^2	Z_A	Z_C	Z_K	T
LN(0,2)	10	0.5288	0.5366	0.5194	0.6124	0.5778	0.5997	0.4332	0.5606	0.3887
	20	0.8864	0.9364	0.9119	0.9647	0.9596	0.9874	0.9116	0.9793	0.6860
	30	0.9856	0.9949	0.9940	0.9989	0.9975	1.0000	0.9980	0.9999	0.6870
	50	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.6690
W(0.5,1)	10	0.5208	0.5592	0.4752	0.6271	0.6058	0.5612	0.3795	0.5497	0.4459
	20	0.8875	0.9504	0.8946	0.9730	0.9721	0.9846	0.8984	0.9832	0.7058
	30	0.9883	0.9967	0.9932	0.9993	0.9987	1.0000	0.9978	0.9999	0.6921
	50	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.6913
W(2,1)	10	0.0432	0.0859	0.0224	0.0459	0.0827	0.0309	0.0182	0.0204	0.2711
	20	0.1041	0.2757	0.0955	0.1178	0.3045	0.3812	0.2736	0.1304	0.7809
	30	0.1978	0.4862	0.2611	0.2435	0.5421	0.8273	0.7101	0.4569	0.9784
	50	0.4556	0.8138	0.7001	0.6225	0.8780	0.9989	0.9932	0.9692	0.9999
EV(0,1)	10	0.0457	0.0784	0.0239	0.0493	0.0745	0.0322	0.0176	0.0219	0.2334
	20	0.1101	0.2302	0.0961	0.1271	0.2426	0.3068	0.2029	0.1313	0.6854
	30	0.1906	0.4011	0.2242	0.2414	0.4261	0.7303	0.5778	0.4119	0.9427
	50	0.4171	0.7333	0.6122	0.5811	0.7645	0.9939	0.9718	0.9318	0.9997
IG(1,0.5)	10	0.2722	0.2990	0.2165	0.3341	0.3098	0.2645	0.1447	0.2516	0.3390
	20	0.6069	0.7100	0.6034	0.7800	0.7516	0.8475	0.6169	0.8253	0.8554
	30	0.8395	0.9141	0.8750	0.9608	0.9347	0.9942	0.9361	0.9907	0.9474
	50	0.9903	0.9964	0.9975	0.9999	0.9984	1.0000	1.0000	1.0000	0.9766
IG(1,1)	10	0.1725	0.2077	0.1218	0.2116	0.2068	0.1544	0.0804	0.1412	0.3156
	20	0.4269	0.5540	0.4060	0.5894	0.5912	0.6986	0.4662	0.6485	0.8594
	30	0.6618	0.7991	0.7012	0.8672	0.8354	0.9733	0.8554	0.9576	0.9838
	50	0.9395	0.9800	0.9763	0.9966	0.9879	1.0000	0.9994	1.0000	0.9981
IG(1,2)	10	0.1033	0.1391	0.0638	0.1208	0.1370	0.0829	0.0438	0.0699	0.2831
	20	0.2670	0.4067	0.2465	0.3757	0.4314	0.5362	0.3475	0.4243	0.8103
	30	0.4563	0.6524	0.4990	0.6616	0.6882	0.9187	0.7614	0.8487	0.9850
	50	0.7865	0.9271	0.8938	0.9633	0.9488	0.9999	0.9961	0.9994	1.0000
SN(0,1,0.5)	10	0.0319	0.0632	0.0148	0.0313	0.0587	0.0198	0.0110	0.0128	0.2296
	20	0.0661	0.1911	0.0571	0.0607	0.2094	0.2585	0.1811	0.0555	0.6669
	30	0.1108	0.3423	0.1463	0.1024	0.3680	0.6690	0.5601	0.1893	0.9346
	50	0.2722	0.6593	0.4984	0.2475	0.7036	0.9858	0.9649	0.6945	0.9995
SN(0,1,2)	10	0.0353	0.0671	0.0173	0.0371	0.0639	0.0227	0.0133	0.0142	0.2294
	20	0.0762	0.2041	0.0671	0.0768	0.2128	0.2731	0.1934	0.0762	0.6711
	30	0.1370	0.3577	0.1733	0.1370	0.3756	0.6838	0.5591	0.2448	0.9348
	50	0.3118	0.6755	0.5270	0.3397	0.7163	0.9892	0.9679	0.7777	0.9997
SN(0,1,3)	10	0.0402	0.0762	0.0213	0.0427	0.0714	0.0277	0.0160	0.0180	0.2429
	20	0.1044	0.2416	0.0945	0.1102	0.2513	0.3255	0.2333	0.1146	0.7103
	30	0.1785	0.4214	0.2216	0.2045	0.4428	0.7481	0.6211	0.3538	0.9541
	50	0.4019	0.7491	0.6237	0.5024	0.7851	0.9954	0.9814	0.8906	0.9997
SL(0,1,0.5)	10	0.0503	0.0671	0.0266	0.0525	0.0634	0.0316	0.0159	0.0252	0.1541
	20	0.1125	0.1591	0.0912	0.1130	0.1520	0.1815	0.1156	0.0895	0.4087
	30	0.1839	0.2540	0.1794	0.1859	0.2340	0.4264	0.3219	0.2088	0.6857
	50	0.3557	0.4717	0.4239	0.3586	0.4266	0.8470	0.7576	0.5339	0.9552
SN(0,1,2)	10	0.0518	0.0695	0.0274	0.0542	0.0643	0.0326	0.0163	0.0250	0.1565
	20	0.1129	0.1593	0.0913	0.1148	0.1512	0.1797	0.1158	0.0894	0.4143
	30	0.1849	0.2548	0.1795	0.1841	0.2327	0.4280	0.3211	0.2093	0.6811
	50	0.3533	0.4675	0.4219	0.3528	0.4228	0.8409	0.7513	0.5291	0.9557
SN(0,1,3)	10	0.0797	0.1049	0.0474	0.0855	0.0956	0.0552	0.0286	0.0445	0.2040
	20	0.1987	0.2716	0.1721	0.2165	0.2563	0.3235	0.2144	0.2004	0.5634
	30	0.3302	0.4436	0.3389	0.3704	0.4100	0.6545	0.5227	0.4386	0.8420
	50	0.6035	0.7449	0.6989	0.6718	0.6950	0.9658	0.9240	0.8390	0.9918

From Table 4 and 5, against asymmetric alternatives, it is seen that the test based on T statistic (with the exception of a few alternatives) has the most power. The power differences between the test T and the other tests are substantial.

In general, Tables 3, 4 and 5 reveal a uniform superiority of the proposed test to all other tests as it outperforms all other competing tests. We can also conclude that the proposed test T has 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. The data (rounded up to seven decimal places) are presented in Table 6. In Figure 1, the histogram, superimposed by a Cauchy PDF is displayed.

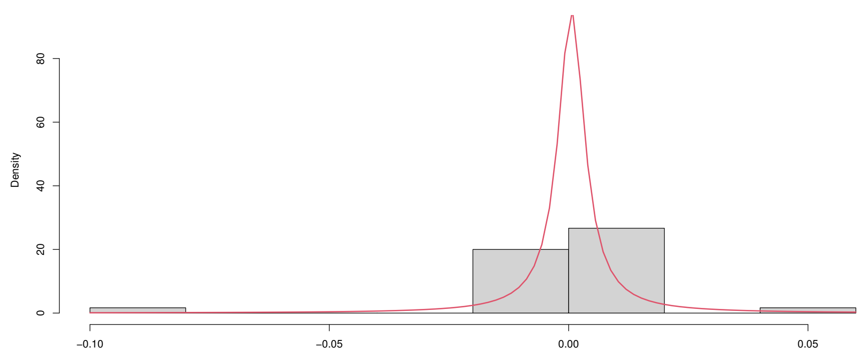

Figure 1. The histogram of the 30 returns along with fitted Cauchy PDF.

Table 6. 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 ML estimates are

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

The value of the test statistic is $T = 0.1789$ and the critical value at the 5% is obtained as 0.2669. Since the value 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. Conclusions, limitations, and future research

In this article, we have proposed and utilized the Kullback-Leibler information as a measure of fitness assessment for the Cauchy distribution. Our proposed approach includes the development of a novel goodness-of-fit test, leveraging the Kullback-Leibler information as its core metric for evaluation. Then, we have investigated some properties of the test statistic. To further validate the performance and applicability of our proposed test, we have conducted extensive Monte Carlo simulations. Through these simulations, we were able to determine the critical points and actual sizes of the test, thereby ensuring its accuracy in practical scenarios.

We have conducted a comparative study to assess the performance of our proposed test in comparison to other existing methods. Our results provide evidence of the superior performance of the Kullback-Leibler information-based test in detecting deviations from the Cauchy distribution under certain alternative hypotheses. This highlights the effectiveness of this information as a tool for assessing the goodness of fit of the Cauchy distribution.

To demonstrate the practical application and relevance of our proposed test, we have illustrated its usage using a real-world dataset. This practical example serves to showcase the test ability to accurately assess the fit of the Cauchy distribution to observed data and reinforces the utility of the Kullback-Leibler information-based approach in real data analysis scenarios. Overall, this comparative study and practical illustration enhance our understanding of the proposed Kullback-Leibler information-based test and its superiority in evaluating the goodness of fit for the Cauchy distribution. These findings contribute to the broader field of statistical analysis by providing a valuable alternative approach for assessing distributional fit.

Discussing potential limitations of the test when applied to real-world data is crucial to provide a balanced view and guide practitioners on when the test is most appropriate. Here are some potential limitations that could be considered as follows. (i) Sensitivity to extreme outliers: The test performance may be affected by the presence of extreme outliers in the data. The Cauchy distribution is known for its heavy tails, but extreme outliers can still have a significant impact on the estimation of parameters and the goodness of fit. It is important to assess the robustness of the test in the presence of extreme outliers and consider alternative approaches or modifications to handle such situations. (ii) Small sample sizes: The test reliability may be compromised when applied to datasets with very small sample sizes. Limited data may lead to imprecise parameter estimation, resulting in less reliable conclusions. It is important to be cautious when interpreting the test results in such cases and consider alternative methods or approaches that are more suitable for small sample sizes. (iii) Assumption of independent and identically distributed observations: The test assumes that the data are independent and identically distributed, which may not always hold in real-world scenarios. Financial data, for example, often exhibit time dependencies or heteroscedasticity. Violations of these assumptions can affect the test performance and lead to incorrect conclusions. It is essential to carefully consider the data characteristics and whether the assumptions of the test are appropriate. (iv) Limited generalizability: While the test may perform well on certain datasets or in specific domains, its generalizability across different types of data and sectors may be limited. The Cauchy distribution may not be the most appropriate model for all financial data, and other distributional assumptions might be more suitable in certain cases. It is important to consider the context and characteristics of the data before applying the test. And (v) impact of parameter estimation methods: The choice of parameter estimation method can influence the test performance and outcomes. Different estimation techniques may lead to varying parameter estimates, which in turn affect the conclusions about the data distribution. It is important to carefully select and justify the chosen estimation method and assess its potential impact on the test results.

Future research directions can further enhance the understanding and applicability of the proposed test for the Cauchy distribution. Some potential avenues for future investigation include applying the test to other types of data beyond financial markets, such as environmental or biological datasets, to assess its validity and performance in different domains. Additionally, exploring modifications or extensions to the test to enhance its robustness to outliers or to relax certain assumptions, such as the independence and identical distribution of observations, could be fruitful. Furthermore, investigating the test performance in dynamic or time-varying environments, where the parameters of the Cauchy distribution may change over time, would provide valuable insights. These future research efforts can contribute to the refinement and broader applicability of the test in diverse fields.

The goodness-of-fit tests considered here are all based on complete samples. But data arising from reliability and life-testing experiments are often censored. There are a number of goodness-of-fit tests in this case in the literature. Therefore, it will be of interest to extend the goodness-of-fit test proposed here to the situation when the data are censored or progressively censored. Another problem of interest will be to develop more general goodness-of-fit tests based on phi-divergence measures. We are currently working on these problems and hope to report these finding in a future article.

In summary, our research contributes a novel Kullback-Leibler information-based test for assessing the goodness of fit for the Cauchy distribution. Its practical applications span various fields, including finance, physics, and engineering, among others. By addressing potential limitations and exploring future research directions, we aim to refine and enhance the test performance and applicability in diverse real-world scenarios.

Author contributions Conceptualization: H.A.N.; data curation: H.A.N., M.S.N.; formal analysis: H.A.N.; investigation: M.S.N.; methodology: H.A.N.; software: M.S.N.; supervision: H.A.N., M.S.N.; validation: H.A.N.; visualization: H.A.N.; writing —original draft: M.S.N.; writing —review and editing: H.A.N., M.S.N. All authors have read and agreed to the published version of the manuscript.

Acknowledgements The authors thank the editors, associate editor and the anonymous referees for their valuable comments to improve this manuscript.

Funding The authors received no financial support for the research, authorship, and/or publication of this article.

Conflicts of interest The authors declare no conflict of interest.

References

Alizadeh, H.N., 2019. A new estimator of Kullback–Leibler information and its application in goodness of fit tests. Journal of Statistical Computation and Simulation, 89, 1914–1934.

Alizadeh, H.N., 2022. A new goodness of fit test for the logistic distribution. Sankhya B, 84, 303–319.

Castro-Kuriss, C., 2011. On a goodness-of-fit test for censored data from a location-scale distribution with applications. Chilean Journal of Statistics, 2(1), 115–136.

D'Agostino, R.B. and Stephens, M.A., 1986. Goodness-of-fit Techniques. Marcel Dekker, New York, NY, US.

Grzegorzewski, P. and Wieczorkowski, R., 1999. Entropy-based goodness-of-fit test for exponentiality. Communications in Statistics: Theory and Methods, 28, 1183–1202.

Johnson, N.L., Kotz, S. and Balakrishnan, N., 1994. Continuous Univariate Distributions. Wiley, New York, NY, US.

Kagan, Y.Y., 1992. Correlations of earthquake focal mechanisms. Geophysical Journal International, 110, 305–320.

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, New York, NY, US.

Min, I.A., Mezic, I. and Leonard, A., 1996. Levy stable distributions for velocity and velocity difference in systems of vortex elements. Physics of Fluids, 8, 1169–1180.

Munir, S., Raza, S. and Abdul Haq, K., 2023. Critical value functions for likelihood-ratio tests under normality. Chilean Journal of Statistics, 14, 48–64.

Nolan, J.P., 2014. Financial modeling with heavy-tailed stable distributions. WIREs Computational Statistics, 6, 45–55.

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. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 115, 107–114.

Suarez-Espinosa, J., Villasenor-Alva, J.A., Hurtado-Jaramillo, A., and Perez-Rodríguez, P., 2018. A goodness of fit test for the Pareto distribution. Chilean Journal of Statistics, 9(1), 33-46.

Winterton, S.S., Smy, T.J., Tarr, N.G., 1992. On the source of scatter in contact resistance data. Journal of Electronic Materials, 21, 917–921.

Zhang, J., 2002. Powerful goodness-of-fit tests based on the likelihood ratio. Journal of Royal Statistical Society B, 64, 281–294.