



Cauchy or not Cauchy? New goodness-of-fit tests for the Cauchy distribution

Bruno Ebner¹ · Lena Eid¹ · Bernhard Klar¹ 

Received: 13 October 2021 / Revised: 8 November 2022 / Accepted: 1 December 2022
© The Author(s) 2022

Abstract

We introduce a new characterization of the Cauchy distribution and propose a class of goodness-of-fit tests for the Cauchy family. The limit distribution is derived in a Hilbert space framework under the null hypothesis. The new tests are consistent against a large class of alternatives. A comparative Monte Carlo simulation study shows that the test is a good competitor for the state of the art procedures, and we apply the tests to log-returns of cryptocurrencies.

Keywords Goodness-of-fit · Cauchy distribution · Hilbert-space valued random elements

Mathematics Subject Classification Primary 62G10 · Secondary 62E10

1 Introduction

In this article we dedicate our studies to answer the question of whether a data set of univariate real numbers belongs to the famous family of Cauchy distributions. The Cauchy distribution is undoubtedly the standard example for a distribution without existing mean value and was studied in the mathematical world for more than 300 years, having wide applicability in diverse fields ranging from modeling resonances in physics (then often called Lorentz distribution) to cryptocurrencies in finance, see Szczygielski et al. (2020). It is also known as the Breit–Wigner distribution, for an extensive historical overview, see Stigler (1974). To be precise, we write shorthand $C(\alpha, \beta)$, $\alpha \in \mathbb{R}$, $\beta > 0$, for the Cauchy distribution with location parameter α and scale parameter β , having density

✉ Bernhard Klar
Bernhard.Klar@kit.edu

Bruno Ebner
Bruno.Ebner@kit.edu

¹ Institute of Stochastics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

$$f(x, \alpha, \beta) = \frac{1}{\pi} \frac{\beta}{\beta^2 + (x - \alpha)^2}, \quad x \in \mathbb{R}.$$

For a detailed discussion on this family, see Johnson et al. (1994), Chapter 16. The Cauchy distribution is a so called heavy tailed distribution and is a member of the stable distributions, see Nolan (2020). Note that $X \sim C(\alpha, \beta)$ if, and only if, $(X - \alpha)/\beta \sim C(0, 1)$ and hence the Cauchy distribution belongs to the location-scale family of distributions. In the following we denote the family of Cauchy distributions by $\mathcal{C} := \{C(\alpha, \beta) : \alpha \in \mathbb{R}, \beta > 0\}$, a family of distributions which is closed under translation and rescaling. We test the composite hypothesis

$$H_0 : \mathbb{P}^X \in \mathcal{C} \quad (1)$$

against general alternatives on the basis of independent identical copies $X_1, \dots, X_n$ of X defined on an underlying probability space $(\Omega, \mathcal{A}, \mathbb{P})$. This testing problem has been considered in the literature: Gürtler and Henze (2000) propose a test procedure based on the empirical characteristic function and Matsui and Takemura (2005) extend this test by considering alternative estimation methods. More recently, Mahdizadeh and Zamanzade (2017) propose to use the likelihood ratio as in Zhang (2002) as well as the Kullback–Leibler (KL) distance, an idea that is extended in Mahdizadeh and Zamanzade (2019). A quantile based method is proposed in Rublik (2001) and compared to the classical omnibus procedures. An empirical power study of goodness-of-fit tests for the Cauchy model based on the empirical distribution function as the Kolmogorov–Smirnov test, the Cramér–von Mises test, the Kuiper test, the Anderson Darling and the Watson test is found in Onen et al. (2001). In Litvinova (2005), two tests for the standard Cauchy distribution based on characterizations are given; however, they are not designed for the composite hypothesis (1).

The novel procedure is based on the following new characterization of the standard Cauchy distribution.

Theorem 1.1 *Let X be a random variable with absolutely continuous density p and $\mathbb{E} \left[\frac{|X|}{1+X^2} \right] < \infty$. Then X has a Cauchy distribution $C(0, 1)$ if and only if*

$$\mathbb{E} \left[\left(it - \frac{2X}{1+X^2} \right) \exp(itX) \right] = 0 \quad (2)$$

holds for all $t \in \mathbb{R}$, where i denotes the imaginary unit.

Proof For $X \sim C(0, 1)$ a direct calculation shows the assertion. Let X be a random variable with absolutely continuous density function $p(x)$ such that

$$\mathbb{E} \left[\left(it - \frac{2X}{1+X^2} \right) \exp(itX) \right] = 0$$

holds for all $t \in \mathbb{R}$. Note that since $-it\mathbb{E}[\exp(itX)]$ is the Fourier-Stieltjes transform of the derivative $p'(x)$ of $p(x)$ we have

$$0 = \mathbb{E} \left[\left(it - \frac{2X}{1+X^2} \right) \exp(itX) \right] = \int_{-\infty}^{\infty} \left(-p'(x) - \frac{2x}{1+x^2} p(x) \right) \exp(itx) dx$$

for all $t \in \mathbb{R}$. By properties of the Fourier-Stieltjes transform, we hence note that $p(x)$ must satisfy the ordinary differential equation

$$p'(x) + \frac{2x}{1+x^2} p(x) = 0$$

for almost all $x \in \mathbb{R}$. The only solution satisfying $\int_{-\infty}^{\infty} p(x) dx = 1$ is $p(x) = f(x, 0, 1)$, $x \in \mathbb{R}$, and $X \sim C(0, 1)$ follows. $\square$

Remark 1.2 The characterization in Theorem 1.1 is related to the spectral representation of the Stein operator of the so called density approach pioneered in Stein et al. (2004). Note that the quotient in (2) is $f'(x, 0, 1)/f(x, 0, 1) = -2x/(1+x^2)$, $x \in \mathbb{R}$, and the set of so called test functions is given by $\{\exp(itx) : t \in \mathbb{R}\}$. For details on this approach see Anastasiou et al. (2022), Sect. 5.4.2.

This paper is organized as follows. In Sect. 2, we introduce a family of test statistics, denoted by $T_{n,a}$, which is based on the characterization in Theorem 1.1. In Sect. 3, the limit distribution of $T_{n,a}$ is derived in a Hilbert space framework under the null hypothesis. Furthermore, we derive the limit distribution of the statistic $\tilde{T}_{n,0} = \lim_{a \rightarrow 0} aT_{n,a}$. Consistency of the new tests against a large class of alternatives is shown in Sect. 4. An extensive Monte Carlo simulation study in Sect. 5 shows that the test is a good competitor for the state of the art procedures. In Sect. 6, the tests are applied to log-returns of cryptocurrencies. Finally, conclusions are given in Sect. 7.

2 A new class of goodness of fit tests for the Cauchy distribution

The testing problem under discussion is invariant with respect to transformations of the kind $x \rightarrow ax + b$, $x \in \mathbb{R}$, where $a \in \mathbb{R}$ and $b > 0$. Consequently, a decision in favor or against H_0 should be the same for $X_1, \dots, X_n$ and $aX_1 + b, \dots, aX_n + b$. This goal is achieved if the test statistic T_n , say, is based on the standardized data $Y_{n,1}, \dots, Y_{n,n}$, given by

$$Y_{n,j} = \frac{X_j - \hat{\alpha}_n}{\hat{\beta}_n}, \quad j = 1, \dots, n. \quad (3)$$

Here, $\hat{\alpha}_n = \hat{\alpha}_n(X_1, \dots, X_n)$ and $\hat{\beta}_n = \hat{\beta}_n(X_1, \dots, X_n)$ denote consistent estimators of $\alpha \in \mathbb{R}$ and $\beta > 0$ such that

$$\hat{\alpha}_n(aX_1 + b, \dots, aX_n + b) = a\hat{\alpha}_n(X_1, \dots, X_n) + b, \quad (4)$$

$$\hat{\beta}_n(aX_1 + b, \dots, aX_n + b) = a\hat{\beta}_n(X_1, \dots, X_n), \quad (5)$$

holds for each $a > 0$ and $b \in \mathbb{R}$. By (4) and (5) it is easy to see that $Y_{n,j}$, $j = 1, \dots, n$, do not depend on the location nor on the scale parameter. Hence, T_n has the property $T_n(aX_1 + b, \dots, aX_n + b) = T_n(X_1, \dots, X_n)$, and we may and do assume $\alpha = 0$ and $\beta = 1$ in the following. Motivated by Theorem 1.1, we choose the test statistic

$$T_n = n \int_{-\infty}^{\infty} \left| \frac{1}{n} \sum_{j=1}^n \left(it - \frac{2Y_{n,j}}{1 + Y_{n,j}^2} \right) e^{itY_{n,j}} \right|^2 \omega(t) dt, \quad (6)$$

which is the weighted L^2 -distance from (2) to 0. Here, $|\cdot|$ is the complex absolute value and $\omega : \mathbb{R} \rightarrow (0, \infty)$ denotes a positive weight function that is given by $\omega(t) = \omega_a(t) = \exp(-a|t|)$, $t \in \mathbb{R}$. Note that ω is symmetric around the origin, i.e. $\omega(t) = \omega(-t)$ holds for all $t \in \mathbb{R}$, and that

$$\int_{-\infty}^{\infty} t^6 \omega(t) dt < \infty \quad (7)$$

holds. We have the integration-free, numerical stable formula

$$\begin{aligned} T_n = T_{n,a} = & \frac{1}{n} \sum_{j,k=1}^n \left(\frac{8aY_{n,j}Y_{n,k}}{(1 + Y_{n,j}^2)(1 + Y_{n,k}^2)((Y_{n,j} - Y_{n,k})^2 + a^2)} \right. \\ & \left. - \frac{16aY_{n,j}(Y_{n,j} - Y_{n,k})}{(1 + Y_{n,j}^2)((Y_{n,j} - Y_{n,k})^2 + a^2)^2} + \frac{4a^3 - 12a(Y_{n,j} - Y_{n,k})^2}{((Y_{n,j} - Y_{n,k})^2 + a^2)^3} \right), \end{aligned} \quad (8)$$

and hence a whole family of tests depending on the so called tuning parameter $a > 0$. The next result deals with the limit behavior of $T_{n,a}$ for $a \rightarrow 0$ and $a \rightarrow \infty$.

Theorem 2.1 *For fixed n , we have*

$$\lim_{a \rightarrow 0} a \left(T_{n,a} - \frac{4}{a^3} \right) = \frac{8}{n} \sum_{j=1}^n \frac{Y_{n,j}^2}{(1 + Y_{n,j}^2)^2} = \tilde{T}_{n,0}, \quad (9)$$

and

$$\lim_{a \rightarrow \infty} a T_{n,a} = \frac{8}{n} \left(\sum_{j=1}^n \frac{Y_{n,j}}{1 + Y_{n,j}^2} \right)^2. \quad (10)$$

Proof Splitting the sum in (8) in a diagonal and a non-diagonal part results in

$$T_{n,a} = \frac{1}{n} \sum_{j,k=1}^n R_{j,k,a} = \frac{1}{n} \sum_{j=1}^n R_{j,j,a} + \frac{1}{n} \sum_{j \neq k} R_{j,k,a} = T_{n,a}^d + T_{n,a}^{nd},$$

say. Since

$$T_{n,a}^d = \frac{8}{na} \sum_{j=1}^n Y_{n,j}^2 / (1 + Y_{n,j}^2)^2 + \frac{4}{a^3}$$

and $\lim_{a \rightarrow 0} T_{n,a}^{nd} = 0$, we obtain

$$\lim_{a \rightarrow 0} a \left(T_{n,a} - \frac{4}{a^3} \right) = \lim_{a \rightarrow 0} a \left(T_{n,a}^d - \frac{4}{a^3} \right) = \frac{8}{n} \sum_{j=1}^n \frac{Y_{n,j}^2}{(1 + Y_{n,j}^2)^2}.$$

By the symmetry of the weight function $\omega_a(\cdot)$, straightforward calculations show

$$T_{n,a} = \int_{-\infty}^{\infty} Z_n^2(t) \omega_a(t) dt,$$

where

$$Z_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \left(\frac{2Y_{n,j}}{1 + Y_{n,j}^2} + t \right) \cos(tY_{n,j}) + \left(t - \frac{2Y_{n,j}}{1 + Y_{n,j}^2} \right) \sin(tY_{n,j}), \quad t \in \mathbb{R}. \quad (11)$$

Then, $aT_{n,a} = \int_{-\infty}^{\infty} Z_n^2(s/a) \exp(-|s|) ds$, where Z_n satisfies $Z_n^2(s/a) \leq 4n(1 + |s|)^2$, for $a > 1$, and

$$\lim_{t \rightarrow 0} Z_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \frac{2Y_{n,j}}{1 + Y_{n,j}^2}.$$

Now, the second assertion follows from an application of the dominated convergence theorem. $\square$

3 Limit distribution under the null hypothesis

The asymptotic theory is derived in the Hilbert space $\mathbb{H}$ of measurable, square integrable functions $\mathbb{H} = L^2(\mathbb{R}, \mathcal{B}, \omega(t)dt)$, where $\mathcal{B}$ is the Borel- σ -field of $\mathbb{R}$ and $\omega(t)$ is the weight function defined in the introduction. Notice that $T_n = \int_{-\infty}^{\infty} Z_n^2(t) \omega(t) dt$, where Z_n is defined in (11), is a real-valued $(\mathcal{A} \otimes \mathcal{B}, \mathcal{B})$ -measurable random element of $\mathbb{H}$. We denote by

$$\|f\|_{\mathbb{H}} = \left(\int_{-\infty}^{\infty} |f(t)|^2 \omega(t) dt \right)^{1/2}, \quad \langle f, g \rangle_{\mathbb{H}} = \int_{-\infty}^{\infty} f(t)g(t) \omega(t) dt$$

the usual norm and inner product in $\mathbb{H}$. In the following, we assume that the estimators $\widehat{\alpha}_n$ and $\widehat{\beta}_n$ allow linear representations

$$\sqrt{n}\widehat{\alpha}_n = \frac{1}{\sqrt{n}} \sum_{j=1}^n \psi_1(X_j) + o_{\mathbb{P}}(1), \quad (12)$$

$$\sqrt{n}(\widehat{\beta}_n - 1) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \psi_2(X_j) + o_{\mathbb{P}}(1), \quad (13)$$

where $X_1, \dots, X_n \sim C(0, 1)$ are independent random variables, $o_{\mathbb{P}}(1)$ denotes a term that converges to 0 in probability, and ψ_1 and ψ_2 are measurable functions with

$$\begin{aligned} \mathbb{E}[\psi_1(X_1)] &= \mathbb{E}[\psi_2(X_1)] = 0, \text{ and } \mathbb{E}[(\psi_1(X_1), \psi_2(X_1))^{\top} (\psi_1(X_1), \psi_2(X_1))] \\ &= C \cdot I_2; \end{aligned}$$

here, C is a positive constant, $^{\top}$ stands for the transpose of a vector and I_2 is the 2×2 -identity matrix. See Remark 3.2 for examples of estimation procedures satisfying these assumptions.

Theorem 3.1 *Let $X_1, \dots, X_n$ be i.i.d. $C(\alpha, \beta)$ distributed random variables. Then there exists a centred Gaussian random process $\mathcal{Z}$ in $\mathbb{H}$ with covariance kernel*

$$\begin{aligned} K(s, t) &= \frac{1}{2}(s^2 + t^2 + |s - t| + 1)e^{-|s-t|} \\ &\quad - \frac{1}{2}(t^2 + |t| + 1)e^{-|t|}\mathbb{E}\left[\left(\left(\frac{2X_1}{1+X_1^2} + s\right)\cos(sX_1) + \left(s - \frac{2X_1}{1+X_1^2}\right)\sin(sX_1)\right)\psi_1(X_1)\right] \\ &\quad + \frac{1}{2}t(|t| + 1)e^{-|t|}\mathbb{E}\left[\left(\left(\frac{2X_1}{1+X_1^2} + s\right)\cos(sX_1) + \left(s - \frac{2X_1}{1+X_1^2}\right)\sin(sX_1)\right)\psi_2(X_1)\right] \\ &\quad - \frac{1}{2}(s^2 + |s| + 1)e^{-|s|}\mathbb{E}\left[\left(\left(\frac{2X_1}{1+X_1^2} + t\right)\cos(tX_1) + \left(t - \frac{2X_1}{1+X_1^2}\right)\sin(tX_1)\right)\psi_1(X_1)\right] \\ &\quad + \frac{1}{2}s(|s| + 1)e^{-|s|}\mathbb{E}\left[\left(\left(\frac{2X_1}{1+X_1^2} + t\right)\cos(tX_1) + \left(t - \frac{2X_1}{1+X_1^2}\right)\sin(tX_1)\right)\psi_2(X_1)\right] \\ &\quad + \frac{1}{4}(s^2 + |s| + 1)(t^2 + |t| + 1)e^{-|s|-|t|}\mathbb{E}[\psi_1^2(X_1)] + \frac{1}{4}s(|s| + 1)t(|t| + 1)e^{-|s|-|t|}\mathbb{E}[\psi_2^2(X_1)] \end{aligned}$$

for $s, t \in \mathbb{R}$, such that $T_n \xrightarrow{D} \|\mathcal{Z}\|_{\mathbb{H}}^2$ as $n \rightarrow \infty$.

A proof of Theorem 3.1 is found in Appendix A.1. It is well known that the distribution of $\|\mathcal{Z}\|_{\mathbb{H}}^2$ is that of $\sum_{j=1}^{\infty} \lambda_j N_j^2$, where $N_1, N_2, \dots$ are i.i.d. standard normal random variables and (λ_j) is a decreasing sequence of positive eigenvalues of the integral operator

$$\mathcal{K}g(s) = \int_{-\infty}^{\infty} K(s, t)g(t)\omega(t) dt.$$

Due to the complexity of the covariance kernel K as given in Theorem 3.1 it seems hopeless to solve the integral equation $\mathcal{K}g(s) = \lambda g(s)$ and find explicit values of λ_j ,

$j \geq 1$. For a numerical approximation method we refer to Subsect. 3.3 in Matsui and Takemura (2005). Note that

$$\mathbb{E}\|\mathcal{Z}\|_{\mathbb{H}}^2 = \int_{-\infty}^{\infty} K(t, t)\omega(t) dt \quad (14)$$

and

$$\text{Var}\|\mathcal{Z}\|_{\mathbb{H}}^2 = 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K^2(s, t)\omega(t)\omega(s) dt ds \quad (15)$$

can be derived for specific estimation procedures. Such results in the theory of goodness-of-fit tests for the Cauchy family are sparse, for some explicit formulae for mean values, see Gürtler and Henze (2000) and Matsui and Takemura (2005).

Remark 3.2 The generality of Theorem 3.1 in view of the linear representations of the estimators and hence dependence on the functions ψ_1 and ψ_2 leads to explicit covariance kernels for different parameter estimation procedures. To estimate α and β we choose the following location and scale estimators $\hat{\alpha}_n$ and $\hat{\beta}_n$, which all satisfy (4) and (5) respectively. For a compact notation, we write $\psi(x) = (\psi_1(x), \psi_2(x))^T$. Some derivations were partially provided by the computer algebra system Maple, see Maplesoft (2019).

- 1. Median and interquartile-distance estimators:** Let ξ_p , $p \in (0, 1)$, denote the p -quantile of the underlying distribution F , $\hat{\xi}_{p,n}$ the sample p -quantile, and $X_{(1)} \leq \dots \leq X_{(n)}$ the order statistics of $X_1, \dots, X_n$. With $\lfloor \cdot \rfloor$ denoting the floor function, let

$$\hat{\alpha}_n = \begin{cases} \frac{1}{2}(X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}), & \text{if } n \text{ even,} \\ X_{(\lfloor \frac{n}{2} \rfloor + 1)}, & \text{otherwise,} \end{cases} \quad (16)$$

be the unbiased empirical median and

$$\hat{\beta}_n = \frac{1}{2}(\hat{\xi}_{\frac{3}{4},n} - \hat{\xi}_{\frac{1}{4},n}) \quad (17)$$

the half-interquartile range (iqr) of the sample. Under mild regularity conditions $\hat{\alpha}_n$ and $\hat{\beta}_n$ are consistent estimators of α and β . Display (3.3) of Gürtler and Henze (2000) then gives the so-called Bahadur representations (see Theorem 2.5.1 in Serfling (1980)) with

$$\psi_1(x) = \pi \left(\frac{1}{2} - \mathbf{1}\{x \leq 0\} \right), \quad \text{and} \quad \psi_2(x) = \pi \left(\frac{1}{2} - \mathbf{1}\{-1 \leq x \leq 1\} \right), \quad x \in \mathbb{R}.$$

It is easy to see that $\mathbb{E}[\psi_1(X_1)] = \mathbb{E}[\psi_2(X_1)] = 0$ and $\mathbb{E}[\psi(X_1)\psi(X_1)^\top] = \frac{\pi^2}{4}I_2$ holds. With these representations, we get the covariance kernel of $\mathcal{Z}$ in Theorem 3.1

$$\begin{aligned} K_{MIQ}(s, t) &= \frac{1}{2}(s^2 + t^2 + |s - t| + 1)e^{-|s-t|} \\ &\quad + \left[t(|t| + 1)(2J_1(s) - sJ_2(s)) - (t^2 + |t| + 1)\left(\frac{s}{2}J_3(s) + J_4(s)\right) \right] e^{-|t|} \\ &\quad + \left[s(|s| + 1)(2J_1(t) - tJ_2(t)) - (s^2 + |s| + 1)\left(\frac{t}{2}J_3(t) + J_4(t)\right) \right] e^{-|s|} \\ &\quad + \frac{\pi^2}{16} \left[(s^2 + |s| + 1)(t^2 + |t| + 1) + st(|s| + 1)(|t| + 1) \right] e^{-|s| - |t|}, \end{aligned}$$

for $s, t \in \mathbb{R}$, where

$$\begin{aligned} J_1(t) &= \int_0^1 \frac{x \sin(tx)}{(1+x^2)^2} dx, \quad J_2(t) = \int_0^1 \frac{\cos(tx)}{1+x^2} dx, \\ J_3(t) &= \int_0^\infty \frac{\sin(tx)}{1+x^2} dx, \quad J_4(t) = \int_0^\infty \frac{x \cos(tx)}{(1+x^2)^2} dx. \end{aligned}$$

Direct calculations of integrals lead to

$$\begin{aligned} \mathbb{E}\|\mathcal{Z}\|_{\mathbb{H}}^2 &= \int_{-\infty}^{\infty} K_{MIQ}(t, t)\omega_a(t) dt \\ &= (8(a+2)^5(1+a)^3a^3(a^2+2a+2)^3)^{-1} \left[(\pi^2-8)a^{16} + (19\pi^2-152)a^{15} \right. \\ &\quad + (173\pi^2-1368)a^{14} + (1003\pi^2-7720)a^{13} + (4126\pi^2-30192)a^{12} \\ &\quad + (12594\pi^2-84304)a^{11} + (29128\pi^2-163520)a^{10} \\ &\quad + (51460\pi^2-188832)a^9 + (69320\pi^2-8256)a^8 \\ &\quad + (70296\pi^2+457664)a^7 + (52176\pi^2+1025920)a^6 \\ &\quad + (26848\pi^2+1323264)a^5 + (8576\pi^2+1151488)a^4 \\ &\quad \left. + (1280\pi^2+693248)a^3 + 280576a^2 + 69632a + 8192 \right]. \end{aligned}$$

2. **Maximum likelihood estimators:** Matsui and Takemura (2005) show in Lemma A.1 that for the maximum-likelihood estimator $\hat{\alpha}_n$ and $\hat{\beta}_n$ in the Cauchy family the linear representations are given by

$$\psi_1(x) = \frac{4x}{1+x^2}, \quad \text{and} \quad \psi_2(x) = \frac{2(x^2-1)}{1+x^2}, \quad x \in \mathbb{R}.$$

Again, straightforward calculations show $\mathbb{E}[\psi_1(X_1)] = \mathbb{E}[\psi_2(X_1)] = 0$ as well as $\mathbb{E}[\psi(X_1)\psi(X_1)^\top] = 2I_2$. Note that there are no closed form expressions for the estimators, such that the log-likelihood equations have to be solved numerically.

This leads to the covariance kernel of $\mathcal{Z}$ in Theorem 3.1 given by

$$\begin{aligned} K_{ML}(s, t) = & \frac{1}{2}(s^2 + t^2 + |s - t| + 1)e^{-|s-t|} \\ & - \frac{1}{2}(t^2 + |t| + 1)(s^2 + |s| + 1)e^{-|t|-|s|} \\ & - \frac{1}{2}t(|t| + 1)s(|s| + 1)e^{-|t|-|s|}, \quad s, t \in \mathbb{R}. \end{aligned}$$

With this explicit formula for the covariance kernel, we compute

$$\mathbb{E}\|\mathcal{Z}\|_{\mathbb{H}}^2 = \int_{-\infty}^{\infty} K_{ML}(t, t)\omega_a(t) dt = \frac{8a^4 + 80a^3 + 352a^2 + 320a + 128}{a^3(a+2)^5}$$

and

$$\begin{aligned} \text{Var}\|\mathcal{Z}\|_{\mathbb{H}}^2 &= 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K_{ML}^2(s, t)\omega_a(s)\omega_a(t) dt ds \\ &= \left(\frac{1}{2}a^5(a+1)^7(a+2)^{10}\right)^{-1} \left[10a^{14} + 270a^{13} + 3521a^{12} \right. \\ &\quad + 27987a^{11} + 146819a^{10} + 510582a^9 + 1194078a^8 + 1914216a^7 \\ &\quad + 2134432a^6 + 1671456a^5 + 928192a^4 + 369792a^3 + 104192a^2 \\ &\quad \left. + 18432a + 1536\right]. \end{aligned}$$

- Equivariant integrated squared error estimator:** In Matsui and Takemura (2005) the authors propose an equivariant version of the minimal integrated squared error estimator introduced by Besbeas and Morgan (2001) for the Cauchy distribution. The estimators are derived by minimization of the weighted L^2 -distance

$$I(\alpha, \beta) = \int_{-\infty}^{\infty} \left| \varphi_n(t; \alpha, \beta) - e^{-|t|} \right|^2 \omega(t) dt,$$

where $\varphi_n(t; \alpha, \beta) = \frac{1}{n} \sum_{j=1}^n \exp(it(X_j - \alpha)/\beta)$, $t \in \mathbb{R}$, is the empirical characteristic function of a distribution from the location scale family in dependence of the parameters. As weight function the authors chose the same weight function $\omega_\nu(t) = \exp(-\nu|t|)$, $t \in \mathbb{R}$, $\nu > 0$, and hence get families of estimators $\hat{\alpha}_{n,\nu}$ and $\hat{\beta}_{n,\nu}$ in dependence of ν . Note that the optimization problem has to be solved numerically and no closed-form formula is known for the estimators. We denote this class of estimators hereinafter EISE. Lemma A.1 in Matsui and Takemura (2005) provides the linear asymptotic representations

$$\begin{aligned} \psi_1(x, \nu) &= (\nu + 1)(\nu + 2)^3 \frac{x}{((\nu + 1)^2 + x^2)^2}, \\ \psi_2(x, \nu) &= \frac{1}{2}(\nu + 2) - \frac{1}{2}(\nu + 2)^3 \frac{(\nu + 1)^2 - x^2}{((\nu + 1)^2 + x^2)^2}, \end{aligned}$$

for $x \in \mathbb{R}$, which lead for every $\nu > 0$ to an explicit expression of the covariance kernel of $\mathcal{Z}$ in Theorem 3.1, which can be found in Appendix A.3.

As demonstrated in Remark 3.2, the estimation procedure has some influence on the limit null distribution of T_n . We refer to Chen (2011); Fegyverneki (2013) for alternative estimators of the parameters of the Cauchy distribution and to Cohen Freue (2007) for a Pitman estimator of the location parameter α when β is known.

3.1 Asymptotic normality of the limit statistic under the null hypothesis

The next result gives the limiting distribution of the scaled limiting statistic $aT_{n,a}$ for $a \rightarrow 0$ given in (9).

Theorem 3.3 (a) Let $\text{arcsine}(a, b)$ denote the arcsine distribution on $[a, b]$ with distribution function

$$F(x) = \frac{2}{\pi} \arcsin \sqrt{\frac{x-a}{b-a}}, \quad a \leq x \leq b.$$

If $X \sim C(0, 1)$, then

$$\frac{4X^2}{(1+X^2)^2} \sim \text{arcsine}(0, 1).$$

(b) Let $X_1, \dots, X_n$ be i.i.d. $C(0, 1)$ distributed random variables, and let $Y_{n,j}$ be the corresponding scaled residuals, using any estimator $(\hat{\alpha}, \hat{\beta})$ with linear representation. Then,

$$\tilde{T}_{n,0} = \sqrt{2n} \left(\frac{8}{n} \sum_{j=1}^n \frac{Y_{n,j}^2}{(1+Y_{n,j}^2)^2} - 1 \right) \xrightarrow{\mathcal{D}} N(0, 1), \quad (18)$$

where $N(0, 1)$ denotes the standard normal distribution.

Proof (a) Let U have a uniform distribution on $(-\pi/2, \pi/2)$, and put $X = \tan U$. Then, $X \sim C(0, 1)$. Further, $\sin U$, $X^2/(1+X^2) = \sin^2 U$, $1/(1+X^2) = \cos^2 U$ as well as $2 \sin U \cos U$ have an $\text{arcsine}(-1, 1)$ distribution (see, e.g., Norton (1983)). If $Z \sim \text{arcsine}(-1, 1)$, then $Z^2 \sim \text{arcsine}(0, 1)$. Hence,

$$\frac{4X^2}{(1+X^2)^2} \sim 4 \sin^2 U \cos^2 U \sim \text{arcsine}(0, 1).$$

(b) A Taylor expansion around $(\alpha_0, \beta_0) = (0, 1)$ yields

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{j=1}^n \frac{(X_j - \alpha)^2 / \beta^2}{(1 + (X_j - \alpha)^2 / \beta^2)^2} \\ &= \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ \frac{X_j^2}{(1 + X_j^2)^2} + \frac{2X_j(X_j^2 - 1)\alpha}{(1 + X_j^2)^3} + \frac{2X_j^2(X_j^2 - 1)(\beta - 1)}{(1 + X_j^2)^3} \right. \\ & \quad \left. + f_1(X_j, \tilde{\alpha}, \tilde{\beta})\alpha^2 + f_2(X_j, \tilde{\alpha}, \tilde{\beta})\alpha(\beta - 1) + f_3(X_j, \tilde{\alpha}, \tilde{\beta})(\beta - 1)^2 \right\}, \end{aligned}$$

where $(\tilde{\alpha}, \tilde{\beta})$ lies between (α, β) and $(0, 1)$, and

$$\begin{aligned} f_1(x, \alpha, \beta) &= \frac{2}{\beta^2} \left(1 + \frac{(x - \alpha)^2}{\beta^2} \right)^{-2} - \frac{20(x - \alpha)^2}{\beta^4} \left(1 + \frac{(x - \alpha)^2}{\beta^2} \right)^{-3} \\ & \quad + \frac{24(x - \alpha)^4}{\beta^6} \left(1 + \frac{(x - \alpha)^2}{\beta^2} \right)^{-4}, \\ f_2(x, \alpha, \beta) &= \frac{4(x - \alpha)}{\beta^3} \left(1 + \frac{(x - \alpha)^2}{\beta^2} \right)^{-2} - \frac{24(x - \alpha)^3}{\beta^5} \left(1 + \frac{(x - \alpha)^2}{\beta^2} \right)^{-3} \\ & \quad + \frac{24(x - \alpha)^5}{\beta^7} \left(1 + \frac{(x - \alpha)^2}{\beta^2} \right)^{-4}, \\ f_3(x, \alpha, \beta) &= \frac{6(x - \alpha)^2}{\beta^4} \left(1 + \frac{(x - \alpha)^2}{\beta^2} \right)^{-2} - \frac{28(x - \alpha)^4}{\beta^6} \left(1 + \frac{(x - \alpha)^2}{\beta^2} \right)^{-3} \\ & \quad + \frac{24(x - \alpha)^6}{\beta^8} \left(1 + \frac{(x - \alpha)^2}{\beta^2} \right)^{-4}. \end{aligned}$$

If we replace (α, β) by any estimator $(\hat{\alpha}_n, \hat{\beta}_n)$ with linear representation, the second and third term converge to zero in probability, since, by the results given in the proof of part a),

$$\frac{1}{n} \sum_{j=1}^n \frac{X_j(X_j^2 - 1)}{(1 + X_j^2)^3} \xrightarrow{a.s.} \mathbb{E} \frac{X_1(X_1^2 - 1)}{(1 + X_1^2)^3} = 0, \quad (19)$$

$$\frac{1}{n} \sum_{j=1}^n \frac{X_j^2(X_j^2 - 1)}{(1 + X_j^2)^3} \xrightarrow{a.s.} \mathbb{E} \frac{X_1^2(X_1^2 - 1)}{(1 + X_1^2)^3} = 0, \quad (20)$$

and $\sqrt{n}\hat{\alpha}_n$ and $\sqrt{n}(\hat{\beta}_n - 1)$ have limiting normal distributions. Putting $z = (x - \alpha)/\beta$, the functions f_j , $j = 1, 2, 3$, involve terms of the form $z^k(1 + z^2)^{-l}/\beta^2$,

where $k < 2l$. Hence,

$$\frac{1}{n} \sum_{j=1}^n |f_j(X_j, \tilde{\alpha}_n, \tilde{\beta}_n)| \leq \frac{c_j}{\tilde{\beta}_n^2}, \quad j = 1, 2, 3,$$

where $c_j > 0$, and $(\tilde{\alpha}_n, \tilde{\beta}_n)$ lies between $(\hat{\alpha}_n, \hat{\beta}_n)$ and $(0, 1)$. Since $\sqrt{n}\hat{\alpha}_n^2$, $\sqrt{n}\hat{\alpha}_n(\hat{\beta}_n - 1)$ and $\sqrt{n}(\hat{\beta}_n - 1)^2$ converge to zero in probability, and $\tilde{\beta}_n \xrightarrow{\mathbb{P}} 1$, we obtain

$$\frac{8}{\sqrt{n}} \sum_{j=1}^n \frac{Y_{n,j}^2}{(1 + Y_{n,j}^2)^2} = \frac{8}{\sqrt{n}} \sum_{j=1}^n \frac{X_j^2}{(1 + X_j^2)^2} + R_n,$$

where $R_n \xrightarrow{\mathbb{P}} 0$ for $n \rightarrow \infty$. By a), $8X_1^2/(1 + X_1^2)^2$ has expected value 1 and variance 1/2, and the assertion follows by the central limit theorem. $\square$

Remark 3.4 (i) Let $X \sim F$, and assume that $8X^2/(1 + X^2)^2$ has expected value 1 and finite variance. Then, the test based on $\tilde{T}_{n,0}$ is not consistent against F , i.e. the test is not an omnibus test.

(ii) Typically, the estimation method enters the asymptotic distribution via the functions ψ_1 and ψ_2 in (12) and (13) and the first order terms in the Taylor expansion. Here, this is not the case, because the corresponding expected values in (19) and (20) vanish.

Remark 3.5 Note that the sum appearing on the right hand side of (10) involves essentially the first component of the score function of the Cauchy distribution (see the second part of Remark 3.2). Hence, when using the maximum-likelihood estimator for (α, β) , $\lim_{a \rightarrow \infty} aT_{n,a} \equiv 0$. For other estimators, the limit does not vanish, but a goodness-of-fit test based on it has very low power. Hence, we don't give further results for this statistic.

4 Consistency

In this section we show that the new tests are consistent against all alternatives satisfying a weak moment condition. Let $X, X_1, X_2, \dots$ be i.i.d. random variables with cumulative distribution function F having a unique median as well as unique upper and lower quartiles and $\mathbb{E}|X|^4/(1 + X^2)^2 < \infty$. Since the tests T_n are affine invariant, we assume w.l.o.g. that the true median $\alpha = 0$ and the half interquartile range $\beta = 1$.

Under these assumptions we clearly have that $(\hat{\alpha}_n, \hat{\beta}_n) \xrightarrow{\mathbb{P}} (0, 1)$ as $n \rightarrow \infty$ for the MIQ estimation method in Remark 3.2, for details see Gürtler and Henze (2000), and assume the convergence in all other cases.

Theorem 4.1 *Under the standing assumptions, we have*

$$\frac{T_n}{n} \xrightarrow{\mathbb{P}} \int_{-\infty}^{\infty} \left| \mathbb{E} \left[\left(it - \frac{2X}{1 + X^2} \right) e^{itX} \right] \right|^2 \omega(t) dt = \Delta_F, \quad \text{as } n \rightarrow \infty.$$

Table 1 Empirical 0.95-quantiles, based on 10 000 MC samples, for $T_{n,1}$ and $T_{n,5}$ using median/half-IQR and ML estimator

	Median and half-IQR		Maximum Likelihood	
	$T_{n,1}$	$T_{n,5}$	$T_{n,1}$	$T_{n,5}$
$n = 10$	7.72	0.56	6.26	0.139
$n = 20$	7.62	0.80	6.45	0.130
$n = 50$	7.37	0.98	6.54	0.128
$n = 100$	7.32	1.08	6.59	0.127
$n = 200$	7.34	1.12	6.58	0.127
$n = 500$	7.32	1.15	6.60	0.127

A proof of Theorem 4.1 is deferred to Appendix A.2. In view of the characterization in Theorem 1.1 $\Delta_F = 0$ if and only if $F = C(0, 1)$. Hence, we conclude that the tests T_n are able to detect all alternatives satisfying the assumptions of this section.

Remark 4.2 In this remark we fix the weight function $\omega_1(t) = \exp(-|t|)$, $t \in \mathbb{R}$. By numerical integration we calculate that if $X \sim \mathcal{U}(-\sqrt{3}, \sqrt{3})$, then $\Delta_{\mathcal{U}} \approx 2.332019$, if $X \sim N(0, 1)$ then $\Delta_N \approx 0.021839$ and if $X \sim \text{Log}(0, 1)$, then $\Delta_L \approx 0.041495$.

5 Simulation results

In this section, we describe an extensive simulation study to compare the power of the new tests with established ones. For all computations in this and the following section, we used the statistical software R (R Core Team, 2021). All simulations have been done at nominal level $\alpha = 0.05$ and for sample sizes $n = 20$ and $n = 50$. The unknown parameters are estimated either by median and half-IQR estimators or by the method of maximum likelihood, see Remark 3.2. The convergence of the critical values of the new test statistics to its asymptotic values is generally quite slow. Table 1 shows a few selected cases. For $T_{n,a}$ in (8) with weight $a = 1$, asymptotic values could be used for sample sizes larger than 100; however, for $a = 5$ and the median and half-IQR estimators, sample sizes of 500 or more are necessary. Hence, in a first step, critical values are determined for each test statistic by a Monte Carlo simulation with 10^5 replications. In a second step, empirical power of the different tests is calculated based on 10^4 replications.

Besides the new tests based on $T_{n,a}$ with weights $a = 1, 2, 3, 4, 5, 6$, we use the limiting test $\tilde{T}_{n,0}$ in (18). Power decreased for larger weights for all distributions used in the study; hence, we omit the results for weights larger than 6 as well as for the test based on the limiting statistic $T_{n,\infty}$ in (10), see also Remark 3.5. Moreover, we use the test based on the KL distance described in Mahdizadeh and Zamanzade (2017). For the nonparametric estimation of the entropy, the authors recommend the window size $m = 4$ for $n = 20$ and $m = 20$ for $n = 50$ (Mahdizadeh and Zamanzade 2017, p. 1109), and we followed this suggestion. For computing the Vasicek estimate of differential Shannon entropy, we used the function `entropy.estimate` in the

R package `vsgoftest` (Lequesne and Regnault 2020). From the classical tests that utilize the empirical distribution function (so-called edf tests), we choose the following: Kolmogorov-Smirnov test (KS), Cramér-von Mises test (CM), Anderson-Darling test (AD) and Watson test (W). Finally, we apply the tests based on the statistics $D_{n,\lambda}$ considered in Gürtler and Henze (2000) and Matsui and Takemura (2005), with weights $\lambda = 1, 2, 3, 4, 5, 6$.

Of course, there are further goodness of fit tests for the Cauchy distribution. For example, one can use the procedure proposed by Chen and Balakrishnan (2012) which consists in transforming the observations to approximately normally distributed random variables, and then applying any edf test for normality with estimated parameters. For the latter, there exist tables and software, so there is no need for, say, a parametric bootstrap procedure. This is a nice feature for distributions with shape parameters or in more complicated situations (Klar and Meintanis 2012; Goldmann et al. 2015). However, for location-scale families and parameter-free test statistics, where the null distribution can be obtained by Monte Carlo simulation, this is only a minor advantage. On the other hand, these tests don't maintain exactly the theoretical level. Further, besides the results derived in Goldmann et al. (2015), not much is known about asymptotic properties of the tests.

Similarly as in Chen and Balakrishnan (2012), Villaseñor and González-Estrada (2021) suggest to transform the observations to approximately exponentially distributed random variables. Then, the authors apply the Anderson-Darling test for exponentiality with estimated parameters. They use this procedure to test for the Cauchy distribution, but recommend to obtain critical values by Monte Carlo simulation. Their results show that neither the classical Anderson-Darling test nor the Anderson-Darling tests based on the two transformations outperforms the other two in all cases. Clearly, one could construct a multitude of further tests by (approximate) transformations to an arbitrary continuous distribution, and using any of the well-known edf test statistics. For the above reasons, and since we don't aim at a complete overview over gof tests for the Cauchy distribution, we decided not to include such transformation based tests in our power study.

Besides the standard Cauchy and normal distribution $C(0, 1)$ and $N(0, 1)$, for short), we use Cauchy-normal mixtures $CN(p) = (1 - p)C(0, 1) + pN(0, 1)$ with $p = 0.5$ and $p = 0.8$. Further, we employ Student's t -distribution $(Student(k))$ with $k = 2, 3, 5, 10$ degrees of freedom, the (symmetric) Tukey g-h distribution with $g = 0$ and $h = 0.2, 0.1, 0.05$, denoted by $Tukey(h)$, and the Tukey lambda distribution $(Tukey-L(\lambda))$ with $\lambda = -3, -2, -0.5, 0.5$. Note that the Tukey lambda distribution approximates the Cauchy distribution for $\lambda = -1$, it coincides with the logistic distribution for $\lambda = 0$, it is close to the normal for $\lambda = 0.14$, and coincides with the uniform distribution for $\lambda = 1$. In the family of stable distributions $S(\alpha, \beta)$, we choose symmetric distributions $S(\alpha, 0)$ with $\alpha = 0.4, 0.7, 1.2, 1.5, 1.8$, and skewed distributions $S(\alpha, 1)$ with $\alpha = 0.5, 1, 1.5, 2$. Moreover, we use the uniform, logistic, Laplace, Gumbel and exponential distribution, and the Mittag-Leffler distribution $ML(\alpha)$ with $\alpha = 0.25, 0.5, 0.75$. The $ML(\alpha)$ distribution has Laplace transform $(1 + t^\alpha)^{-1}$, and coincides with the exponential distribution for $\alpha = 1$.

Power estimates of the tests under discussion are given in Tables 2, 3, 4 and 5. All entries are the percentage of rejection of H_0 , rounded to the nearest integer. In

Table 2 Percentage of 10,000 MC samples declared significant by various tests for the Cauchy distribution using median and half-IQR estimator ($\alpha = 0.05$, $n = 20$)

	$T_{n,1}$	$T_{n,2}$	$T_{n,3}$	$T_{n,4}$	$T_{n,5}$	$T_{n,6}$	$\tilde{T}_{n,0}$	KL	KS
C(0,1)	5	5	5	5	5	5	5	5	5
N(0,1)	8	4	1	0	0	0	29	73	5
CN(0.5)	4	3	3	3	3	3	9	14	4
CN(0.8)	6	3	1	1	1	1	19	36	4
Student(2)	4	2	2	2	2	2	9	22	3
Student(3)	5	2	1	1	1	1	14	34	3
Student(5)	6	3	1	1	1	1	20	49	4
Student(10)	7	3	1	1	1	1	24	61	4
Stable(0.4,0)	55	52	39	26	18	13	79	2	39
Stable(0.7,0)	16	15	12	9	8	7	21	2	12
Stable(1.2,0)	4	3	3	3	3	3	6	10	4
Stable(1.5,0)	5	2	1	1	1	2	14	28	4
Stable(1.8,0)	7	3	1	1	1	1	22	52	5
Stable(0.5,1)	91	97	90	76	61	50	40	65	94
Stable(1.5,1)	10	10	7	6	6	6	19	49	9
Stable(2,1)	8	4	0	0	0	0	28	73	5
Tukey(0.2)	5	2	1	1	1	1	14	36	3
Tukey(0.1)	5	3	1	1	1	1	21	52	4
Tukey(0.05)	7	3	1	0	0	0	24	62	4
Tukey-L(-3)	43	42	33	24	17	12	64	2	31
Tukey-L(-2)	21	22	18	14	11	10	33	1	16
Tukey-L(-0.5)	4	3	2	2	2	2	7	15	3
Tukey-L(0.5)	17	11	1	0	0	0	45	94	10
Uniform	32	21	1	1	0	0	56	99	24
Logistic	6	3	1	1	1	1	21	56	4
Laplace	4	2	1	1	1	1	10	32	3
Gumbel	10	9	3	3	2	2	26	70	9
ML(0.25)	100	100	100	97	92	83	93	89	100
ML(0.5)	98	99	93	79	63	49	55	81	99
ML(0.75)	78	87	69	52	43	39	20	75	85
Exponential	39	44	23	15	13	11	27	92	48
	CM	AD	W	$D_{n,1}$	$D_{n,2}$	$D_{n,3}$	$D_{n,4}$	$D_{n,5}$	$D_{n,6}$
C(0,1)	5	5	5	5	5	5	5	5	5
N(0,1)	6	6	15	15	21	26	26	23	19
CN(0.5)	3	3	5	5	5	5	5	4	3
CN(0.8)	4	4	8	8	11	13	12	10	8
Student(2)	4	3	5	4	5	5	5	5	3
Student(3)	4	3	7	6	7	9	9	8	6

Table 2 continued

	CM	AD	W	$D_{n,1}$	$D_{n,2}$	$D_{n,3}$	$D_{n,4}$	$D_{n,5}$	$D_{n,6}$
Student(5)	4	4	9	8	11	14	14	12	9
Student(10)	5	5	11	11	15	18	19	16	13
Stable(0.4,0)	45	83	58	62	64	66	69	71	74
Stable(0.7,0)	13	23	14	16	18	20	22	24	25
Stable(1.2,0)	4	3	4	4	4	4	4	3	3
Stable(1.5,0)	4	3	6	6	7	8	8	6	5
Stable(1.8,0)	5	5	11	10	14	16	16	14	11
Stable(0.5,1)	90	97	88	95	76	57	46	44	46
Stable(1.5,1)	9	9	11	13	11	12	11	10	8
Stable(2,1)	6	6	14	14	20	24	24	22	17
Tukey(0.2)	4	3	7	6	8	10	9	8	7
Tukey(0.1)	4	4	9	8	12	15	15	13	10
Tukey(0.05)	5	5	12	11	15	19	19	17	13
Tukey-L(-3)	35	79	44	52	58	62	65	68	71
Tukey-L(-2)	17	41	19	24	28	32	35	38	41
Tukey-L(-0.5)	4	3	5	4	4	4	4	3	3
Tukey-L(0.5)	12	15	31	36	45	50	50	47	40
Uniform	22	28	51	60	67	69	68	65	59
Logistic	5	4	11	9	13	16	17	15	11
Laplace	3	3	6	5	6	8	8	7	5
Gumbel	8	8	15	17	18	20	20	18	15
ML(0.25)	100	100	100	100	99	98	96	96	96
ML(0.5)	96	99	97	99	86	69	58	55	57
ML(0.75)	71	79	72	83	52	33	25	23	23
Exponential	31	32	41	49	34	30	28	27	25

Tables 2 and 3, the results using the median and interquartile-distance estimators are given, with sample size $n = 20$ and $n = 50$, respectively. Tables 4 and 5 show the corresponding results for the maximum likelihood estimator.

The main conclusions that can be drawn from the simulation results are the following:

- As always in similar situations, there exists no uniformly most powerful test, which is in accordance to the results in Janssen (2000).
- As expected, the power of nearly all test statistics increases for increasing sample size for all alternative distributions. An exception is the KL distance based test. Its power decreases for certain alternatives, in particular for the Mittag-Leffler distribution. This behavior has been reported previously, see Tables 3 to 6 in Mahdizadeh and Zamanzade (2019), and may be due to the choice of the window size.

Table 3 Percentage of 10,000 MC samples declared significant by various tests for the Cauchy distribution using median and half-IQR estimator ($\alpha = 0.05$, $n = 50$)

	$T_{n,1}$	$T_{n,2}$	$T_{n,3}$	$T_{n,4}$	$T_{n,5}$	$T_{n,6}$	$\tilde{T}_{n,0}$	KL	KS
C(0,1)	5	5	4	5	5	5	5	5	5
N(0,1)	24	24	4	1	0	0	72	100	26
CN(0.5)	8	5	3	3	3	3	23	14	7
CN(0.8)	16	13	2	1	1	1	51	46	14
Student(2)	6	4	2	2	2	2	23	52	6
Student(3)	10	7	2	2	1	1	38	82	9
Student(5)	14	12	2	1	1	1	52	96	13
Student(10)	18	18	3	1	1	0	64	100	18
Stable(0.4,0)	91	90	76	47	25	15	99	0	80
Stable(0.7,0)	23	23	14	10	8	7	46	0	16
Stable(1.2,0)	6	4	3	3	3	3	12	19	6
Stable(1.5,0)	11	8	3	2	2	2	38	53	9
Stable(1.8,0)	18	16	3	1	1	1	62	88	18
Stable(0.5,1)	100	100	100	100	97	87	71	13	100
Stable(1.5,1)	28	41	26	18	15	13	52	81	56
Stable(2,1)	23	23	4	1	1	0	72	100	26
Tukey(0.2)	10	7	2	1	1	1	40	87	9
Tukey(0.1)	15	13	2	1	1	1	56	98	15
Tukey(0.05)	19	17	3	1	1	1	64	100	18
Tukey-L(-3)	70	74	57	35	20	13	96	0	64
Tukey-L(-2)	33	35	24	16	12	10	64	0	26
Tukey-L(-0.5)	5	4	3	3	3	3	13	37	5
Tukey-L(0.5)	56	61	11	1	0	0	90	100	68
Uniform	87	87	18	1	0	0	97	100	95
Logistic	15	14	3	1	1	1	59	100	15
Laplace	6	4	2	1	1	1	21	94	6
Gumbel	33	43	18	8	6	5	68	100	58
ML(0.25)	100	100	100	100	100	100	100	4	100
ML(0.5)	100	100	100	100	98	91	86	15	100
ML(0.75)	100	100	100	93	78	62	30	45	100
Exponential	94	96	75	44	32	26	64	100	100
	CM	AD	W	$D_{n,1}$	$D_{n,2}$	$D_{n,3}$	$D_{n,4}$	$D_{n,5}$	$D_{n,6}$
C(0,1)	5	5	5	5	5	5	5	5	5
N(0,1)	29	55	67	65	83	89	92	93	94
CN(0.5)	6	7	13	13	15	16	16	15	13
CN(0.8)	14	23	38	37	49	54	55	55	54
Student(2)	6	8	14	10	15	19	21	23	23

Table 3 continued

	CM	AD	W	$D_{n,1}$	$D_{n,2}$	$D_{n,3}$	$D_{n,4}$	$D_{n,5}$	$D_{n,6}$
Student(3)	10	16	27	21	32	40	44	47	49
Student(5)	15	28	41	36	52	62	67	70	72
Student(10)	20	41	54	49	70	78	83	85	86
Stable(0.4,0)	91	99	98	97	96	97	97	98	98
Stable(0.7,0)	17	39	26	29	34	37	41	45	48
Stable(1.2,0)	5	5	8	6	7	8	8	8	8
Stable(1.5,0)	9	13	24	20	29	34	36	37	37
Stable(1.8,0)	19	35	50	46	63	71	74	77	77
Stable(0.5,1)	100	100	100	100	100	100	100	99	99
Stable(1.5,1)	38	53	60	61	71	72	70	69	68
Stable(2,1)	28	55	67	65	84	89	92	94	94
Tukey(0.2)	9	16	27	22	34	42	47	50	52
Tukey(0.1)	16	32	44	39	57	67	72	75	77
Tukey(0.05)	21	41	54	50	70	78	83	85	87
Tukey-L(-3)	73	99	91	90	94	95	96	97	98
Tukey-L(-2)	29	72	47	48	58	64	68	72	75
Tukey-L(-0.5)	5	5	8	7	8	10	11	12	12
Tukey-L(0.5)	63	89	92	96	99	100	100	100	100
Uniform	86	98	99	100	100	100	100	100	100
Logistic	17	35	48	43	63	72	77	80	82
Laplace	7	12	18	13	24	32	37	40	42
Gumbel	42	64	73	74	86	89	90	91	91
ML(0.25)	100	100	100	100	100	100	100	100	100
ML(0.5)	100	100	100	100	100	100	100	100	100
ML(0.75)	100	100	100	100	100	100	98	94	91
Exponential	92	98	99	100	100	99	98	97	97

- The new tests $T_{n,a}$, $a = 1, \dots, 6$, perform better using the maximum likelihood estimator than with median and half-IQR. Hence, the latter estimators should not be used. For all other test statistics, including $\tilde{T}_{n,0}$, performance is comparable, or even better when using median and half-IQR. In any case, the choice of the estimation method can have a pronounced influence on the performance of the tests.
- The power of $\tilde{T}_{n,0}$ is comparable under both estimation methods.
- For alternatives with finite first and second moments, the KL distance based test has the highest power among all competitors. On the other hand, its power breaks down completely for some alternatives without existing first moment, and it is low for some alternatives with infinite second moment. Hence, the test can not really be seen as an omnibus test.

Table 4 Percentage of 10,000 MC samples declared significant by various tests for the Cauchy distribution using ML estimation ($\alpha = 0.05$, $n = 20$)

	$T_{n,1}$	$T_{n,2}$	$T_{n,3}$	$T_{n,4}$	$T_{n,5}$	$T_{n,6}$	$\tilde{T}_{n,0}$	KL	KS
C(0,1)	5	5	5	5	5	5	5	5	5
N(0,1)	16	30	34	28	10	3	26	76	5
CN(0.5)	8	10	8	5	2	2	9	14	4
CN(0.8)	13	19	18	12	5	2	18	37	5
Student(2)	8	9	8	6	3	2	10	22	4
Student(3)	10	14	14	10	4	2	15	37	4
Student(5)	12	19	20	14	6	2	19	51	5
Student(10)	15	25	27	20	8	3	22	64	5
Stable(0.4,0)	53	64	72	77	80	81	76	0	42
Stable(0.7,0)	10	14	19	22	25	26	19	1	12
Stable(1.2,0)	7	6	5	4	2	2	7	11	4
Stable(1.5,0)	10	13	12	8	3	2	14	29	4
Stable(1.8,0)	14	23	25	18	7	3	21	55	5
Stable(0.5,1)	55	70	78	81	83	84	26	54	99
Stable(1.5,1)	14	21	22	17	11	7	16	50	14
Stable(2,1)	17	29	34	28	10	4	26	76	5
Tukey(0.2)	10	14	14	10	4	2	14	38	4
Tukey(0.1)	13	20	21	16	6	2	19	55	5
Tukey(0.05)	15	25	28	21	7	3	23	65	5
Tukey-L(-3)	34	49	60	67	72	75	60	0	35
Tukey-L(-2)	13	21	29	35	40	42	29	1	17
Tukey-L(-0.5)	6	7	6	4	3	2	8	16	4
Tukey-L(0.5)	31	55	64	59	24	8	37	96	11
Uniform	48	74	81	78	43	14	45	99	22
Logistic	14	22	24	18	7	3	21	60	5
Laplace	8	10	10	8	4	2	10	35	3
Gumbel	18	31	35	30	16	9	24	74	13
ML(0.25)	99	99	100	100	100	100	92	77	100
ML(0.5)	80	86	90	91	92	92	43	72	100
ML(0.75)	52	63	66	66	65	64	12	70	96
Exponential	38	53	57	54	45	36	20	92	64
	CM	AD	W	$D_{n,1}$	$D_{n,2}$	$D_{n,3}$	$D_{n,4}$	$D_{n,5}$	$D_{n,6}$
C(0,1)	5	5	5	5	5	5	5	5	5
N(0,1)	3	2	29	23	17	10	7	4	2
CN(0.5)	4	3	9	7	3	2	2	2	1
CN(0.8)	3	2	18	14	8	5	3	2	1

Table 4 continued

	CM	AD	W	$D_{n,1}$	$D_{n,2}$	$D_{n,3}$	$D_{n,4}$	$D_{n,5}$	$D_{n,6}$
Student(2)	3	2	10	7	4	2	2	1	1
Student(3)	2	1	14	10	6	3	2	2	1
Student(5)	3	2	19	14	9	5	4	3	2
Student(10)	3	2	24	19	13	8	5	3	2
Stable(0.4,0)	33	77	77	71	79	83	86	87	88
Stable(0.7,0)	11	18	15	18	24	26	28	29	30
Stable(1.2,0)	4	3	7	5	3	2	2	2	2
Stable(1.5,0)	3	2	14	9	5	3	2	2	1
Stable(1.8,0)	3	2	23	17	11	7	5	3	2
Stable(0.5,1)	99	99	82	79	82	85	87	88	89
Stable(1.5,1)	12	9	22	17	11	9	7	6	5
Stable(2,1)	3	2	28	23	17	10	7	4	3
Tukey(0.2)	3	2	14	10	6	4	3	2	1
Tukey(0.1)	3	2	20	15	10	5	4	2	2
Tukey(0.05)	3	2	24	19	13	8	5	3	2
Tukey-L(-3)	27	72	60	58	71	76	79	81	82
Tukey-L(-2)	15	34	26	27	38	42	45	48	49
Tukey-L(-0.5)	3	2	7	5	3	2	2	1	1
Tukey-L(-0.5)	6	5	50	50	43	28	19	11	7
Uniform	13	10	66	72	65	48	35	24	14
Logistic	3	2	22	17	11	6	4	3	2
Laplace	2	1	10	8	5	3	2	2	1
Gumbel	10	7	31	25	19	14	12	9	7
ML(0.25)	100	100	100	100	100	100	100	100	100
ML(0.5)	100	100	96	93	92	94	95	95	96
ML(0.75)	93	91	72	66	62	64	66	67	68
Exponential	50	43	55	50	41	39	38	37	35

- Among the new tests, values of the tuning parameter around $a = 3$ result in a quite homogeneous power against all alternatives. For the tests based on $D_{n,\lambda}$, $\lambda = 5$ seems to be a good choice. Both classes of tests perform similarly for the ML estimator; for the median and half-IQR estimator, the latter is preferable.
- Among the group of edf tests, W outperform the other tests in most cases.
- For symmetric alternatives without existing first moment as Stable(0.4,0), Stable(0.7,0), Tukey-L(-3) and Tukey-L(-2), the tests based on $\tilde{T}_{n,0}$, AD and $D_{n,6}$ perform best.

Table 5 Percentage of 10,000 MC samples declared significant by various tests for the Cauchy distribution using ML estimation ($\alpha = 0.05$, $n = 50$)

	$T_{n,1}$	$T_{n,2}$	$T_{n,3}$	$T_{n,4}$	$T_{n,5}$	$T_{n,6}$	$\tilde{T}_{n,0}$	KL	KS
C(0,1)	5	5	5	5	5	5	5	5	5
N(0,1)	40	76	90	95	96	96	64	100	16
CN(0.5)	14	20	20	15	10	6	22	14	6
CN(0.8)	28	53	61	60	52	39	48	46	11
Student(2)	12	20	24	23	20	15	23	52	5
Student(3)	19	36	45	48	46	38	38	82	7
Student(5)	26	51	66	71	71	66	48	96	9
Student(10)	32	65	81	86	88	86	57	100	13
Stable(0.4,0)	95	98	99	99	99	99	99	0	82
Stable(0.7,0)	23	34	40	45	47	49	44	0	18
Stable(1.2,0)	8	10	10	8	6	4	12	18	5
Stable(1.5,0)	19	34	40	39	33	25	35	53	7
Stable(1.8,0)	34	62	76	80	78	71	56	88	12
Stable(0.5,1)	98	100	100	100	100	100	46	6	100
Stable(1.5,1)	32	61	74	78	76	68	44	80	57
Stable(2,1)	41	77	90	95	97	96	65	100	17
Tukey(0.2)	20	36	46	49	48	42	37	87	7
Tukey(0.1)	27	55	70	76	77	73	51	98	9
Tukey(0.05)	33	66	81	88	89	87	57	100	13
Tukey-L(-3)	79	92	96	97	98	98	96	0	68
Tukey-L(-2)	33	51	63	69	73	76	63	0	29
Tukey-L(-0.5)	8	11	13	12	11	8	13	37	4
Tukey-L(0.5)	75	98	100	100	100	100	83	100	54
Uniform	95	100	100	100	100	100	91	100	91
Logistic	28	58	74	81	82	80	52	100	10
Laplace	10	21	30	36	38	36	19	94	4
Gumbel	44	79	91	94	96	94	59	100	57
ML(0.25)	100	100	100	100	100	100	100	0	100
ML(0.5)	100	100	100	100	100	100	76	6	100
ML(0.75)	97	99	100	100	100	99	14	38	100
Exponential	87	98	99	100	100	99	48	100	100
	CM	AD	W	$D_{n,1}$	$D_{n,2}$	$D_{n,3}$	$D_{n,4}$	$D_{n,5}$	$D_{n,6}$
C(0,1)	5	5	5	5	5	5	5	5	5
N(0,1)	9	15	77	77	87	90	91	91	91
CN(0.5)	4	3	19	16	13	10	9	8	7
CN(0.8)	6	6	50	48	50	48	46	43	39

Table 5 continued

	CM	AD	W	$D_{n,1}$	$D_{n,2}$	$D_{n,3}$	$D_{n,4}$	$D_{n,5}$	$D_{n,6}$
Student(2)	3	3	21	16	15	14	15	14	13
Student(3)	4	4	38	30	33	34	36	35	35
Student(5)	5	6	54	47	55	58	60	61	60
Student(10)	6	10	67	63	74	78	80	80	80
Stable(0.4,0)	81	99	100	99	99	100	100	100	100
Stable(0.7,0)	15	31	38	36	44	49	52	53	55
Stable(1.2,0)	4	3	10	8	6	5	5	4	4
Stable(1.5,0)	4	4	34	28	28	27	27	25	24
Stable(1.8,0)	6	9	64	60	67	68	69	68	67
Stable(0.5,1)	100	100	100	100	100	100	100	100	100
Stable(1.5,1)	43	46	62	60	66	66	67	67	66
Stable(2,1)	9	16	78	77	88	90	92	92	91
Tukey(0.2)	3	4	39	31	35	36	38	38	38
Tukey(0.1)	5	7	57	52	61	65	67	67	67
Tukey(0.05)	6	10	68	65	75	79	81	82	82
Tukey-L(-3)	60	98	96	94	97	98	99	99	99
Tukey-L(-2)	23	62	60	57	69	75	78	80	81
Tukey-L(-0.5)	3	3	12	9	8	8	7	7	7
Tukey-L(0.5)	27	50	97	99	100	100	100	100	100
Uniform	53	80	100	100	100	100	100	100	100
Logistic	5	8	60	56	66	70	72	73	72
Laplace	2	2	23	20	25	25	26	26	26
Gumbel	36	43	80	80	88	90	91	91	91
ML(0.25)	100	100	100	100	100	100	100	100	100
ML(0.5)	100	100	100	100	100	100	100	100	100
ML(0.75)	100	100	100	100	100	100	100	100	100
Exponential	96	97	98	99	99	99	99	99	99

6 Real data example: log-returns of cryptocurrencies

In this section, we apply the tests for the Cauchy distribution to log-returns of various cryptocurrencies, namely Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), BitcoinCash (BCH), EOS (EOS), BinanceCoin (BNB) and Tron(TRX). The Cauchy distribution is found to be a comparably good model for such data sets in Szczygielski et al. (2020). There, 58 hypothetical distributions have been fitted to 15 major cryptocurrencies, and the Cauchy model turned out to be the best fitting distribution for 10 cryptocurrencies (including all currencies considered here). In Szczygielski et al. (2020), the number of observations of the various data sets varied widely from 638 to 2255. Further, with very large data sets, each model will be rejected in the end. Hence, we decided to consider shorter time series: a series with daily observations

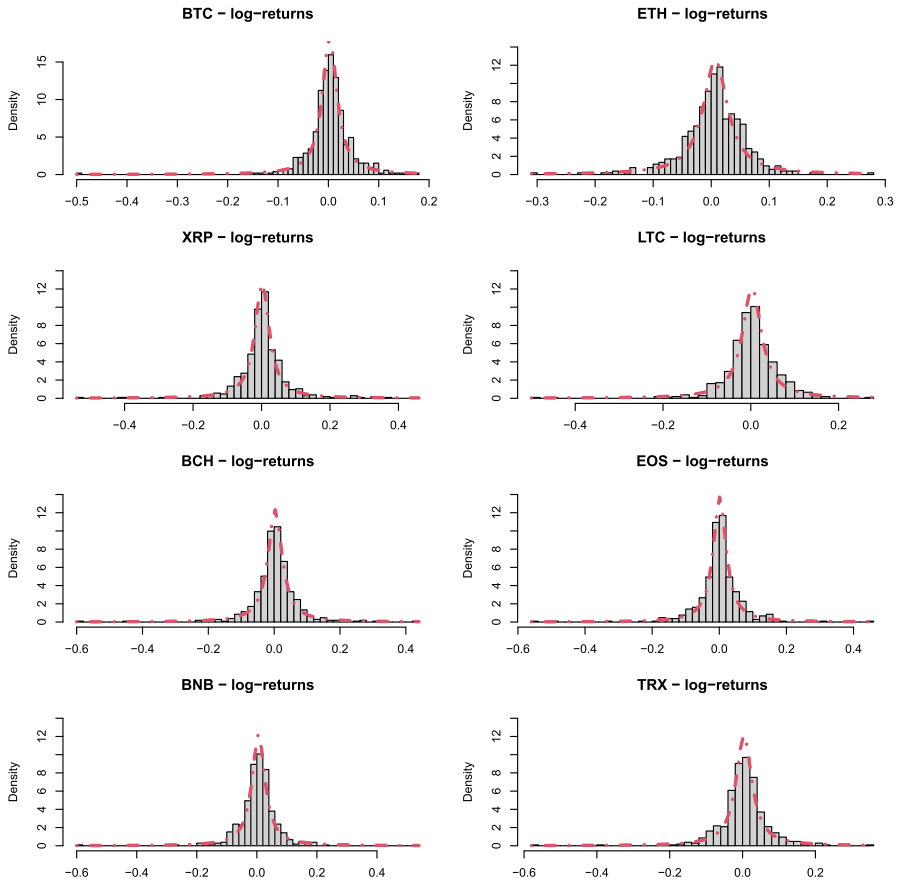


Fig. 1 Histograms of cryptocurrency log-returns from January 01, 2020 to June 10, 2021, together with fitted Cauchy densities

from January 01, 2020 to June 10, 2021, with sample size 527 (526 for ETH), and an even shorter series from January 01, 2021 to June 10, 2021, with sample size 161 (160 for ETH). All prices are closing values in U.S. dollars, freely accessible from CryptoDataDownload via the link www.cryptodatadownload.com/data. Returns are estimated by taking logarithmic differences. Days with zero trading volume are omitted. Figure 1 shows histograms of the datasets with larger time span, together with the densities of fitted Cauchy distributions. Visually, the Cauchy model seems to be a reasonable approximation.

As in the above cited literature, we assume in the following the independence of daily log return data. This is justified in view of the very weak serial correlations of all datasets. Figure 2 shows the autocorrelations of the first four datasets with larger time span, which are all quite small and not significant in most cases. The findings are similar for the other four cryptocurrencies and the shorter time span.

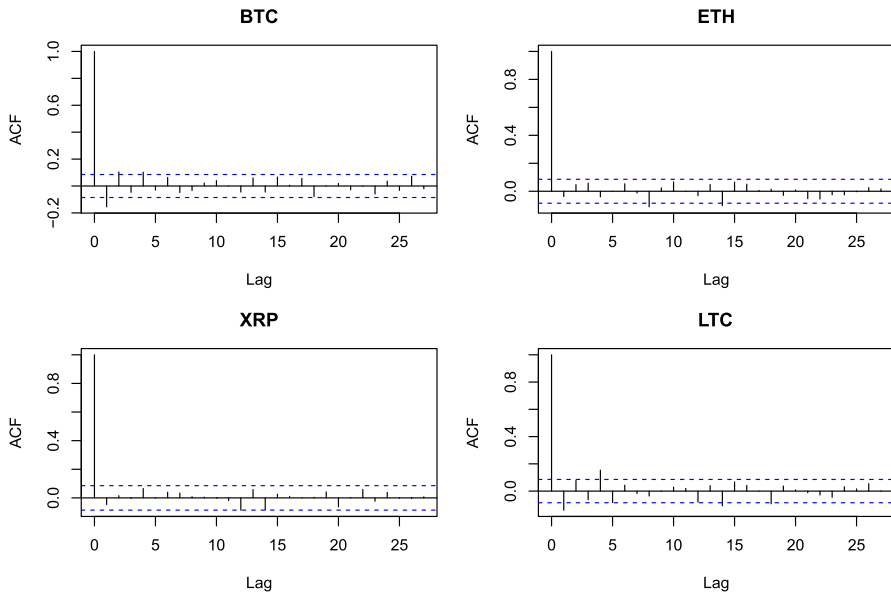


Fig. 2 Autocorrelations of cryptocurrency log-returns from January 01, 2020 to June 10, 2021

Tables 6 and 7 report the results of the different tests for the Cauchy model for the two time series. The p -values are based on Monte Carlo simulation under the Cauchy model with 10^4 replications. For the test based on the KL distance, we choose the window length $m = 100$ for the longer time series, and $m = 50$ for the shorter one. The results show that even if the Cauchy distribution fits better than many other distributional models, it is still not an acceptable fit in any of the cases, possibly with the exception of EOS data. The tests based on $T_{n,5}$, KL and $D_{n,5}$ result in p -values of 0.000 for all currencies for the longer time series, and p -values smaller than 0.01 for all currencies for the shorter one. The edf based tests have generally larger p -values, with the Watson test having smallest, and Cramér–von Mises test having largest p -values.

Overall, the EOS data seem to be most compatible with the hypothesis of a Cauchy distribution. For the shorter series, the p -values of all edf based tests are larger than 0.1, and the tests based on $T_{n,1}$, $\tilde{T}_{n,0}$, KS and CM don't reject H_0 on the 5%-level even for the larger data set. These tests also show relatively large values for Ripple (XRP) and BitcoinCash (BCH). On the whole, the results confirm the findings of the simulation study concerning the comparative power of the test.

7 Conclusion

We have proposed a family of tests for the Cauchy family of distributions with desirable theoretical and computational features as consistency and competitiveness to existing procedures. The authors suggest in view of the results of the simulation study to use the tuning parameter $a = 4$ in applications for good power against most alternatives.

Table 6 *p*-values of cryptocurrency log-returns (Jan. 01, 2020–June 10, 2021) for various tests for the Cauchy distribution

	$T_{n,1}$	$T_{n,3}$	$T_{n,5}$	$\tilde{T}_{n,0}$	KL	KS	CM	AD	W	$D_{n,1}$	$D_{n,3}$	$D_{n,5}$
BTC	0.122	0.000	0.000	0.001	0.000	0.005	0.066	0.003	0.000	0.000	0.000	0.000
ETH	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.000	0.000	0.000	0.000	0.000
XRP	0.357	0.000	0.000	0.016	0.000	0.064	0.191	0.019	0.004	0.003	0.000	0.000
LTC	0.008	0.000	0.000	0.000	0.000	0.000	0.023	0.001	0.000	0.000	0.000	0.000
BCH	0.036	0.000	0.000	0.000	0.000	0.035	0.063	0.004	0.000	0.000	0.000	0.000
EOS	0.308	0.009	0.000	0.559	0.000	0.119	0.277	0.044	0.024	0.006	0.000	0.000
BNB	0.001	0.000	0.000	0.000	0.000	0.002	0.029	0.002	0.000	0.000	0.000	0.000
TRX	0.008	0.000	0.000	0.001	0.000	0.015	0.042	0.002	0.000	0.000	0.000	0.000

Table 7 p -values of cryptocurrency log-returns (Jan. 01, 2021–June 10, 2021), for various tests for the Cauchy distribution

	$T_{n,1}$	$T_{n,3}$	$T_{n,5}$	$\tilde{T}_{n,0}$	KL	KS	CM	AD	W	$D_{n,1}$	$D_{n,3}$	$D_{n,5}$
BTC	0.134	0.000	0.000	0.019	0.000	0.062	0.151	0.041	0.005	0.002	0.000	0.000
ETH	0.000	0.000	0.000	0.000	0.000	0.004	0.022	0.006	0.000	0.000	0.000	0.000
XRP	0.216	0.003	0.001	0.014	0.000	0.147	0.247	0.090	0.010	0.016	0.002	0.002
LTC	0.132	0.000	0.000	0.015	0.000	0.073	0.163	0.043	0.002	0.001	0.000	0.000
BCH	0.049	0.001	0.001	0.001	0.001	0.316	0.156	0.068	0.004	0.011	0.002	0.001
EOS	0.387	0.011	0.001	0.339	0.001	0.222	0.458	0.153	0.108	0.021	0.005	0.006
BNB	0.023	0.000	0.000	0.032	0.001	0.020	0.074	0.026	0.000	0.000	0.000	0.000
TRX	0.010	0.000	0.000	0.000	0.000	0.070	0.074	0.027	0.000	0.001	0.000	0.000

We conclude the paper by pointing out some open problems for further research. The explicit formula of the covariance kernels in Remark 3.2 open ground to numerical approximation of the eigenvalues of the integral operator $\mathcal{K}$. Especially an approximation of the largest eigenvalue would offer more theoretical insights, since, as is well known, it is useful for efficiency statements in the sense of Bahadur, see Bahadur (1960) and Nikitin (1995). A profound knowledge of the eigenvalues leads to explicit values of cumulants of the null distribution, such that a fit of the null distribution to the Pearson distribution system is possible. This usually gives a very accurate approximation of the quantiles of the distribution and hence of the critical values of the test, see i.e. in the context of normality testing (Ebner 2020; Henze 1990). In view of the results of Baringhaus et al. (2017), we conjecture that for a large class of fixed alternatives $\sqrt{n}(T_n/n - \Delta_F)$ converges weakly to a centred normal distribution, where the variance is positive and may be consistently estimated. Since the derivation involves heavy algebra, the authors postpone results of such kind to future work. Another desirable extension for the family of tests is a data dependent choice of the tuning parameter a , see Tenreiro (2019) for a procedure which might be applicable. Finally, it is clear that the choice of the weight function $\omega_a(t)$ has an impact on the power of the tests, so maybe other choices lead to better performances.

Acknowledgements The authors are grateful for the suggestions and comments of three anonymous reviewers which helped to improve the paper.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

A Proofs

A.1 Proof of Theorem 3.1

Recall that $T_n = \|Z_n\|_{\mathbb{H}}^2$ with Z_n defined in (11). In view of the affine invariance of T_n we assume w.l.o.g. $\alpha = 0$ and $\beta = 1$. Denoting

$$h(t, x) = \frac{2x}{1+x^2} (\cos(tx) - \sin(tx)) + t (\cos(tx) + \sin(tx)), \quad t, x \in \mathbb{R}, \quad (21)$$

we have $Z_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^n h(t, Y_{n,j})$. A simple change of variable and a multivariate Taylor expansion around $(\alpha, \beta) = (0, 1)$ shows

$$\|Z_n\|_{\mathbb{H}}^2 = \widehat{\beta}_n \int_{-\infty}^{\infty} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n \frac{2Y_{n,j}}{1+Y_{n,j}^2} (\cos(uX_j - u\widehat{\alpha}_n) - \sin(uX_j - u\widehat{\alpha}_n)) \right|^2 du$$

$$\begin{aligned}
& + \widehat{\beta}_n u \left(\cos(u X_j - u \widehat{\alpha}_n) + \sin(u X_j - u \widehat{\alpha}_n) \right) \Big|^2 \omega(\widehat{\beta}_n u) \, du \\
& = \widehat{\beta}_n \int_{-\infty}^{\infty} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n h(u, X_j) + g_1(u, X_j) \widehat{\alpha}_n + g_2(u, X_j) (\widehat{\beta}_n - 1) + R_j(u) \right|^2 \omega(\widehat{\beta}_n u) \, du,
\end{aligned}$$

where

$$\begin{aligned}
g_1(u, x) = \frac{1}{(1+x^2)^2} & \left((u^2 x^4 - 2ux^3 + (2u^2 - 2)x^2 - 2ux + u^2 + 2) \cos(ux) \right. \\
& \left. - (u^2 x^4 + 2ux^3 + (2u^2 - 2)x^2 + 2ux + u^2 + 2) \sin(ux) \right), \quad (22)
\end{aligned}$$

and

$$\begin{aligned}
g_2(u, x) = \frac{1}{(1+x^2)^2} & \left((ux^4 + 2ux^2 + 2x^3 + u - 2x) \cos(ux) \right. \\
& \left. + (ux^4 + 2ux^2 - 2x^3 + u + 2x) \sin(ux) \right), \quad u, x \in \mathbb{R}. \quad (23)
\end{aligned}$$

The remainder term $R_j(u) = (\widehat{\alpha}_n, \widehat{\beta}_n - 1) A_j(u, \xi_n) (\widehat{\alpha}_n, \widehat{\beta}_n - 1)^\top$, $t \in \mathbb{R}$, and $A_j(u, \xi_n)$ is the 2×2 matrix given by the second order partial derivatives in the Taylor expansion. Here, each component of the matrix is evaluated at $\xi_n = (\xi_{1,n}, \xi_{2,n}) \in [0, \widehat{\alpha}_n] \times [1, \widehat{\beta}_n]$, where the order of the endpoints of the intervals is reversed whenever $\widehat{\alpha}_n < 0$ or $\widehat{\beta}_n < 1$. Note that by assumption, we have

$$\sqrt{n} \widehat{\alpha}_n = \frac{1}{\sqrt{n}} \sum_{j=1}^n \psi_1(X_j) + r_{1n}, \quad \sqrt{n}(\widehat{\beta}_n - 1) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \psi_2(X_j) + r_{2n}, \quad (24)$$

with $r_{1n}, r_{2n} = o_{\mathbb{P}}(1)$ and $\mathbb{E}[\psi_1(X)] = \mathbb{E}[\psi_2(X)] = 0$ and $\mathbb{E}[\psi(X)\psi(X)^\top] = C \cdot I_2$, where $X \sim C(0, 1)$, $\psi(X) = (\psi_1(X), \psi_2(X))^\top$, C is a positive constant and I_2 and is the 2×2 unit matrix. This shows that $(\widehat{\alpha}_n, \widehat{\beta}_n) \xrightarrow{\mathbb{P}} (0, 1)$ and consequently $\xi_n \xrightarrow{\mathbb{P}} (0, 1)$ when $n \rightarrow \infty$. Next, we denote by $\|\cdot\|_{\mathbb{H}_n}$ the norm in the Hilbert space $\mathbb{H}_n = L^2(\mathbb{R}, \mathcal{B}, \omega(\widehat{\beta}_n u) du)$, equipped with the usual scalar product. By the Cauchy-Schwarz inequality, we have

$$\|R_j\|_{\mathbb{H}_n}^2 \leq \|(\widehat{\alpha}_n, \widehat{\beta}_n - 1)\|^4 \int_{-\infty}^{\infty} \|A_j(u, \xi_n)\|_F^2 \omega(\widehat{\beta}_n u) \, du,$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Denote by $A_{j,1}(u, \xi_n)$ the upper left component of the matrix $A_j(u, \xi_n)$. Writing $\kappa_{m,\ell}(x) = x^m(1+x^2)^{-\ell}$ for all $x \in \mathbb{R}$ and

$m, \ell \in \mathbb{N}$, we have $|\kappa_{m,\ell}(x)| \leq 1$ for $m \leq \ell \in \mathbb{N}$, and hence

$$\begin{aligned} |A_{j,1}(u, \xi_n)| &\leq \frac{8}{\xi_{2,n}} |u| |\kappa_{0,1}((X_j - \xi_{1,n})/\xi_{2,n})| + 4u^2 |\kappa_{1,1}((X_j - \xi_{1,n})/\xi_{2,n})| \\ &\quad + \frac{24}{\xi_{2,n}^2} |\kappa_{1,2}((X_j - \xi_{1,n})/\xi_{2,n})| + \frac{16}{\xi_{2,n}^2} |\kappa_{2,2}((X_j - \xi_{1,n})/\xi_{2,n})| \\ &\quad + \frac{32}{\xi_{2,n}^2} |\kappa_{3,3}((X_j - \xi_{1,n})/\xi_{2,n})| + 2\xi_{2,n} |u|^3 \\ &\leq \frac{8}{\xi_{2,n}} |u| + 4u^2 + \frac{56}{\xi_{2,n}^2} + \frac{16}{\xi_{2,n}} + 2\xi_{2,n} |u|^3 \end{aligned}$$

follows. Similar expressions hold for all components of the matrix $A_j(u, \xi_n)$. Therefore by a change of variable we have

$$\widehat{\beta}_n \int_{-\infty}^{\infty} \|A_j(u, \xi_n)\|_F^2 \omega(\widehat{\beta}_n u) du \leq C \max\{1/\xi_{2,n}^4, 1/\xi_{2,n}^2, \xi_{2,n}^2\} \int_{-\infty}^{\infty} u^6 \omega(u) du, \quad (25)$$

where C is a positive constant.

Hence, defining

$$Z_n^*(u) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \left[h(u, X_j) + g_1(u, X_j) \widehat{\alpha}_n + g_2(u, X_j) (\widehat{\beta}_n - 1) \right], \quad u \in \mathbb{R}, \quad (26)$$

we have by the triangle inequality and (25)

$$\begin{aligned} \left| \|Z_n\|_{\mathbb{H}} - \sqrt{\widehat{\beta}_n} \|Z_n^*\|_{\mathbb{H}_n} \right| &\leq \sqrt{\widehat{\beta}_n} \frac{1}{\sqrt{n}} \sum_{j=1}^n \|R_j\|_{\mathbb{H}_n} \\ &\leq \sqrt{n} \|(\widehat{\alpha}_n, \widehat{\beta}_n - 1)\|^2 \frac{1}{n} \sum_{j=1}^n \left(\widehat{\beta}_n \int_{-\infty}^{\infty} \|A_j(u, \xi_n)\|_F^2 \omega(\widehat{\beta}_n u) du \right)^{1/2} \\ &\leq C \max\{1/\xi_{2,n}^2, 1/\xi_{2,n}, \xi_{2,n}\} \times \dots \\ &\quad \dots \times \sqrt{n} (\widehat{\alpha}_n, \widehat{\beta}_n - 1) (\widehat{\alpha}_n, \widehat{\beta}_n - 1)^\top \left(\int_{-\infty}^{\infty} u^6 \omega(u) du \right)^{1/2}. \end{aligned}$$

The last term is $o_{\mathbb{P}}(1)$ since $\sqrt{n} (\widehat{\alpha}_n, \widehat{\beta}_n - 1)$ is tight due to the linearity assumption (24) and the multivariate central limit theorem, $\xi_{2,n} \xrightarrow{\mathbb{P}} 1$, the continuous mapping theorem, as well as (7) and $(\widehat{\alpha}_n, \widehat{\beta}_n - 1) = o_{\mathbb{P}}(1)$. This concludes that $\|Z_n\|_{\mathbb{H}}^2 = \widehat{\beta}_n \|Z_n^*\|_{\mathbb{H}_n}^2 + o_{\mathbb{P}}(1)$.

We obtain using symmetry arguments

$$\mathbb{E}[g_1(t, X_1)] = -\frac{1}{2}(t^2 + |t| + 1)e^{-|t|}, \quad \text{and} \quad \mathbb{E}[g_2(t, X_1)] = \frac{1}{2}t(|t| + 1)e^{-|t|}.$$

By applying the Cauchy-Schwarz inequality, we directly see that both functions belong to $\mathbb{H}$, since

$$\begin{aligned} \int_{-\infty}^{\infty} |\mathbb{E}[g_1(t, X_1)]|^2 \omega(t) dt &\leq \frac{13}{8} \int_{-\infty}^{\infty} \omega(t) dt < \infty, \\ \int_{-\infty}^{\infty} |\mathbb{E}[g_2(t, X_1)]|^2 \omega(t) dt &\leq \frac{7}{8} \int_{-\infty}^{\infty} \omega(t) dt < \infty. \end{aligned}$$

By the law of large numbers in Hilbert spaces, we have for $l = 1, 2$

$$\left\| \frac{1}{n} \sum_{j=1}^n g_l(\cdot, X_j) - \mathbb{E}[g_l(\cdot, X_1)] \right\|_{\mathbb{H}} = o_{\mathbb{P}}(1). \quad (27)$$

Replacing $\sqrt{n}\hat{\alpha}_n$ and $\sqrt{n}(\hat{\beta}_n - 1)$ by the linear representations (24), and substituting the arithmetic means by their asymptotic counterparts leads to

$$\begin{aligned} \tilde{Z}_n(t) &= \frac{1}{\sqrt{n}} \sum_{j=1}^n h(t, X_j) - \frac{1}{2}(t^2 + |t| + 1)e^{-|t|}\psi_1(X_j) \\ &\quad + \frac{1}{2}t(|t| + 1)e^{-|t|}\psi_2(X_j), \quad t \in \mathbb{R}. \end{aligned} \quad (28)$$

Since for all $a_1, a_2 \in \mathbb{R}$, $(a_1 + a_2)^2 \leq 2a_1^2 + 2a_2^2$ holds, we have

$$\begin{aligned} \|Z_n^* - \tilde{Z}_n\|_{\mathbb{H}}^2 &= \int_{-\infty}^{\infty} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n g_1(t, X_j)\hat{\alpha}_n + g_2(t, X_j)(\hat{\beta}_n - 1) \right. \\ &\quad \left. + \frac{1}{2}(t^2 + |t| + 1)e^{-|t|}\psi_1(X_j) - \frac{1}{2}t(|t| + 1)e^{-|t|}\psi_2(X_j) \right|^2 \omega(t) dt \\ &\stackrel{(24)}{\leq} \int_{-\infty}^{\infty} \left| \left(\frac{1}{n} \sum_{j=1}^n g_1(t, X_j) + \frac{1}{2}(t^2 + |t| + 1)e^{-|t|} \right) \frac{1}{\sqrt{n}} \sum_{k=1}^n \psi_1(X_k) \right. \\ &\quad \left. + \left(\frac{1}{n} \sum_{j=1}^n g_2(t, X_j) - \frac{1}{2}t(|t| + 1)e^{-|t|} \right) \frac{1}{\sqrt{n}} \sum_{k=1}^n \psi_2(X_k) \right|^2 \omega(t) dt + o_{\mathbb{P}}(1) \\ &\leq 2 \left\| \frac{1}{n} \sum_{j=1}^n g_1(\cdot, X_j) - \mathbb{E}[g_1(\cdot, X_1)] \right\|_{\mathbb{H}}^2 \left(\frac{1}{\sqrt{n}} \sum_{k=1}^n \psi_1(X_k) \right)^2 \\ &\quad + 2 \left\| \frac{1}{n} \sum_{j=1}^n g_2(\cdot, X_j) - \mathbb{E}[g_2(\cdot, X_1)] \right\|_{\mathbb{H}}^2 \left(\frac{1}{\sqrt{n}} \sum_{k=1}^n \psi_2(X_k) \right)^2 + o_{\mathbb{P}}(1). \end{aligned}$$

By (27), the tightness of $\frac{1}{\sqrt{n}} \sum_{j=1}^n \psi_l(X_j)$, $l = 1, 2$ and Slutsky's Theorem, we have $\|Z_n^* - \tilde{Z}_n\|_{\mathbb{H}} \xrightarrow{\mathbb{P}} 0$, as $n \rightarrow \infty$. Hence Z_n^* and $\tilde{Z}_n$ share the same weak limit in $\mathbb{H}$. Writing $\tilde{Z}_n(t) = n^{-1/2} \sum_{j=1}^n \tilde{Z}_{n,j}(t)$, $t \in \mathbb{R}$, where $\tilde{Z}_{n,j}(t) = h(t, X_j) - \frac{1}{2}(t^2 + |t| + 1)e^{-|t|}\psi_1(X_j) + \frac{1}{2}t(|t| + 1)e^{-|t|}\psi_2(X_j)$, $t \in \mathbb{R}$, $j = 1, \dots, n$, we have $\mathbb{E}[\tilde{Z}_{n,1}(t)] = 0$ and $\tilde{Z}_{n,1}, \tilde{Z}_{n,2}, \dots$ are i.i.d. centred random elements in $\mathbb{H}$. Straightforward calculations show that $K(s, t) = \mathbb{E}[\tilde{Z}_{n,1}(s)\tilde{Z}_{n,1}(t)]$ has the stated formula and the weak limit of $\tilde{Z}_n$ follows by the central limit theorem in Hilbert spaces. To conclude the proof, we have by the Cauchy-Schwarz inequality

$$\left| \|Z_n^*\|_{\mathbb{H}_n} - \|Z_n^*\|_{\mathbb{H}} \right| \leq \left(\int_{-\infty}^{\infty} (Z_n^*(t))^4 \omega(t) dt \right)^{1/2} \left(\int_{-\infty}^{\infty} \left| \frac{\omega(\hat{\beta}_n t)}{\omega(t)} - 1 \right|^2 \omega(t) dt \right)^{1/2}.$$

Since by assumption $\omega^{1/2}$ is a weight function, we directly see using the Cauchy-Schwarz inequality and the continuous mapping theorem that $\left(\int_{-\infty}^{\infty} (Z_n^*(t))^4 \omega(t) dt \right)^{1/2}$ is tight. Since by direct calculations $\int_{-\infty}^{\infty} \left| \frac{\omega(\hat{\beta}_n t)}{\omega(t)} - 1 \right|^2 \omega(t) dt = o_{\mathbb{P}}(1)$ holds, we see that $\|Z_n\|_{\mathbb{H}}^2 = \hat{\beta}_n \|Z_n^*\|_{\mathbb{H}}^2 + o_{\mathbb{P}}(1)$. Finally, by Slutsky's theorem, $\|Z_n\|_{\mathbb{H}}^2$ has the same weak limit as $\|\tilde{Z}_n\|_{\mathbb{H}}^2$ and the assertion follows. $\square$

A.2 Proof of Theorem 4.1

Let Z_n and Z_n^* be defined as in (11) and (26). Since by assumption we have $(\hat{\alpha}_n, \hat{\beta}_n) \xrightarrow{\mathbb{P}} (0, 1)$, we can show in complete analogy to the proof of Theorem 3.1, that $n^{-1} \left| \|Z_n\|_{\mathbb{H}}^2 - \|Z_n^*\|_{\mathbb{H}}^2 \right| \xrightarrow{\mathbb{P}} 0$ holds. Let $g_j(\cdot, \cdot)$, $j = 1, 2$, be defined as in (22) and (23), and put $Z_n^0(t) = n^{-1/2} \sum_{j=1}^n h(t, X_j)$, $t \in \mathbb{R}$. Then $n^{-1/2}(Z_n^*(t) - Z_n^0(t)) = \hat{\alpha}_n/n \sum_{j=1}^n g_1(t, X_j) + (\hat{\beta}_n - 1)/n \sum_{j=1}^n g_2(t, X_j)$ follows and by the triangle inequality we have

$$\begin{aligned} \|n^{-1/2}(Z_n^* - Z_n^0)\|_{\mathbb{H}}^2 &= \int_{-\infty}^{\infty} \left| \hat{\alpha}_n \frac{1}{n} \sum_{j=1}^n g_1(t, X_j) + (\hat{\beta}_n - 1) \frac{1}{n} \sum_{j=1}^n g_2(t, X_j) \right|^2 \omega(t) dt \\ &\leq \int_{-\infty}^{\infty} \left(\left| \hat{\alpha}_n \right| \left| \frac{1}{n} \sum_{j=1}^n g_1(t, X_j) \right| + |\hat{\beta}_n - 1| \left| \frac{1}{n} \sum_{j=1}^n g_2(t, X_j) \right| \right)^2 \omega(t) dt \\ &\leq 2|\hat{\alpha}_n|^2 \left\| \frac{1}{n} \sum_{j=1}^n g_1(t, X_j) \right\|_{\mathbb{H}}^2 + 2|\hat{\beta}_n - 1|^2 \left\| \frac{1}{n} \sum_{j=1}^n g_2(t, X_j) \right\|_{\mathbb{H}}^2. \end{aligned}$$

By the law of large numbers in Hilbert spaces and $(\hat{\alpha}_n, \hat{\beta}_n) \xrightarrow{\mathbb{P}} (0, 1)$, the right hand side is $o_{\mathbb{P}}(1)$. Note that the existence of the mean values is guaranteed by the assumptions. Again, the law of large numbers in $\mathbb{H}$ shows $n^{-1/2} Z_n^0(t) \xrightarrow{a.s.} \mathbb{E}h(t, X)$,

as $n \rightarrow \infty$, and by the symmetry of the weight function, we have

$$\begin{aligned} \frac{T_n}{n} &= \left\| \frac{1}{\sqrt{n}} Z_n \right\|_{\mathbb{H}}^2 \xrightarrow{\mathbb{P}} \int_{-\infty}^{\infty} |\mathbb{E}h(t, X)|^2 \omega(t) dt \\ &= \int_{-\infty}^{\infty} \left| \mathbb{E} \left[\left(it - \frac{2X}{1+X^2} \right) e^{itX} \right] \right|^2 \omega(t) dt, \end{aligned}$$

as $n \rightarrow \infty$. □

A.3 Covariance kernel under EISE estimators

We have for the covariance kernel in Theorem 3.1 in case of EISE estimators

$$\begin{aligned} K_{EISE}(s, t; v) &= \frac{1}{2} (s^2 + t^2 + |s - t| + 1) e^{-|s-t|} \\ &\quad - \frac{1}{2} (t^2 + |t| + 1)(v + 1) \left(\frac{2|s|(v + 2)v^3 + (v + 1)(1 - |s|) + |s| + 3}{v^3} \right) e^{-|s|-|t|} \\ &\quad + \frac{1}{2} (t^2 + |t| + 1) \left(\frac{(v + 1)(|s|(v + 1)^2 + 3(v + 1) - |s|) - 1}{v^3} \right. \\ &\quad \left. + (v + 1)(s^2(v + 2)(v + 1) + 2|s|(v + 2) + s^2) \right) e^{-(v+1)|s|-|t|} \\ &\quad - \frac{1}{2} t(|t| + 1)(v + 1) \left(\frac{s(v + 2)^3(v + 1)}{2v^2} + s(v + 1)^2 + s - 4 \operatorname{sgn}(s) \right) e^{-|s|-|t|} \\ &\quad + \frac{1}{2} t(|t| + 1) \left(\frac{s(v + 2)^3(|s|(v + 1)^3 - 3(v + 1)^2 + |s|(v + 1) + 1)}{4(v + 1)v^2} \right. \\ &\quad \left. - s(v + 1)^2 + s - 4 \operatorname{sgn}(s)(v + 1) \right) e^{-|s|-|t|} \\ &\quad + \frac{1}{2\pi} t(|t| + 1)(v + 2)^3 (s J_1(s) - 2 J_2(s)) e^{-|t|} \\ &\quad - \frac{1}{2} (s^2 + |s| + 1)(v + 1) \left(\frac{2|t|(v + 2)v^3 + (v + 1)(1 - |t|) + |t| + 3}{v^3} \right) e^{-|s|-|t|} \\ &\quad + \frac{1}{2} (s^2 + |s| + 1) \left(\frac{(v + 1)(|t|(v + 1)^2 + 3(v + 1) - |t|) - 1}{v^3} \right. \\ &\quad \left. + (v + 1)(t^2(v + 2)(v + 1) + 2|t|(v + 2) + t^2) \right) e^{-(v+1)|s|-|t|} \\ &\quad - \frac{1}{2} s(|s| + 1)(v + 1) \left(\frac{t(v + 2)^3(v + 1)}{2v^2} + t(v + 1)^2 + t - 4 \operatorname{sgn}(t) \right) e^{-|s|-|t|} \\ &\quad + \frac{1}{2} s(|s| + 1) \left(\frac{t(v + 2)^3(|t|(v + 1)^3 - 3(v + 1)^2 + |t|(v + 1) + 1)}{4(v + 1)v^2} \right. \\ &\quad \left. - t(v + 1)^2 + t - 4 \operatorname{sgn}(t)(v + 1) \right) e^{-|s|-|t|} \\ &\quad + \frac{1}{2\pi} s(|s| + 1)(v + 2)^3 (t J_1(t) - 2 J_2(t)) e^{-|s|} \\ &\quad + ((s^2 + |s| + 1)(t^2 + |t| + 1) + s(|s| + 1)t(|t| + 1)) \frac{(v + 2)^2(5v^2 + 14v + 10)}{64(v + 1)^3} \cdot e^{-|s|-|t|}, \end{aligned}$$

where $\text{sgn}(\cdot)$ is the sign function, and

$$J_1(t; \nu) = \int_0^\infty \frac{x^2 \cos(tx)}{(1+x^2)((\nu+1)^2+x^2)^2} dx,$$

$$J_2(t; \nu) = \int_0^\infty \frac{x^3 \sin(tx)}{(1+x^2)^2((\nu+1)^2+x^2)^2} dx.$$

References

- Anastasiou A, Barp A, Briol F-X, Ebner B, Gaunt RE, Ghaderinezhad F, Gorham J, Gretton A, Ley C, Liu Q, Mackey L, Oates CJ, Reinert G, Swan Y (2022) Stein's method meets computational statistics: a review of some recent developments. *Stat Sci*. <https://doi.org/10.1214/22-STS863>
- Bahadur RR (1960) Stochastic comparison of tests. *Ann Math Stat* 31:276–295
- Baringhaus L, Ebner B, Henze N (2017) The limit distribution of weighted L^2 -Goodness-of-Fit statistics under fixed alternatives, with applications. *Ann Inst Stat Math* 69(5):969–995
- Besbeas P, Morgan BJ (2001) Integrated squared error estimation of Cauchy parameters. *Stat Probab Lett* 55:397–401
- Chen G, Balakrishnan N (2012) A general purpose approximate goodness-of-fit test. *J Qual Technol* 27:154–161
- Chen Z (2011) A simple method for estimating parameters of the location-scale distribution family. *J Stat Comput Simul* 81(1):49–58
- Cohen Freue GV (2007) The Pitman estimator of the Cauchy location parameter. *J Stat Plan Inference* 137(6):1900–1913
- Ebner B (2020) On combining the zero bias transform and the empirical characteristic function to test normality. *ALEA* 18:1029–1045
- Fegyverneki S (2013) A simple robust estimation for parameters of Cauchy distribution. *Miskolc Math. Notes* 14(3):887–892
- Goldmann G, Klar B, Meintanis S (2015) Data transformations and goodness-of-fit tests for type-II right censored samples. *Metrika* 78:59–83
- Gürtler N, Henze N (2000) Goodness-of-fit tests for the Cauchy distribution based on the empirical characteristic function. *Ann Inst Stat Math* 52(2):267–286
- Henze N (1990) An approximation to the limit distribution of the Epps-Pulley test statistic for normality. *Metrika* 37(1):7–18
- Janssen A (2000) Global power functions of goodness of fit tests. *Ann Stat* 28(1):239–253
- Johnson NL, Kotz S, Balakrishnan N (1994) Continuous univariate distributions, vol 1, 2nd edn. Wiley, New York
- Klar B, Meintanis S (2012) Specification tests for the response distribution in generalized linear models. *Comput Stat* 27:251–267
- Lequesne J, Regnault P (2020) vsgoftest: An R package for goodness-of-fit testing based on Kullback–Leibler divergence. *J Stat Softw* 96(1):1–26
- Litvinova VV (2005) Two criteria of goodness-of-fit for Cauchy distributions based on characterizations. *J Math Sci* 127(1):1752–1756
- Mahdizadeh M, Zamanzade E (2017) New goodness of fit tests for the Cauchy distribution. *J Appl Stat* 44(6):1106–1121
- Mahdizadeh M, Zamanzade E (2019) Goodness-of-fit testing for the Cauchy distribution with application to financial modeling. *J King Saud Univ* 31(4):1167–1174
- Maplesoft.: Maple. Maplesoft, Waterloo (2019)
- Matsui M, Takemura A (2005) Empirical characteristic function approach to goodness-of-fit tests for the Cauchy distribution with parameters estimated by mle or eise. *Ann Inst Stat Math* 57(1):183–199
- Nikitin Y (1995) Asymptotic efficiency of nonparametric tests. Cambridge University Press, Cambridge
- Nolan JP (2020) Univariate stable distributions: models for heavy tailed data. Springer series in operations research and financial engineering. Springer, Cham
- Norton RM (1983) A characterization of the Cauchy distribution. *Sankhyā* 45:247–252

- Onen BH, Dietz DC, Yen VC, Moore AH (2001) Goodness-of-fit tests for the Cauchy distribution. *Comput Stat* 16(1):97–107
- R Core Team (2021) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Rublik F (2001) A quantile goodness-of-fit test for Cauchy distribution, based on extreme order statistics. *Appl Math* 46(5):339–351
- Serfling RJ (1980) Approximation theorems of mathematical statistics. Wiley, New York
- Stein C, Diaconis P, Holmes S, Reinert G (2004) Use of exchangeable pairs in the analysis of simulations. In: Diaconis P, Holmes S (eds) Stein's method, volume 46 of Lecture Notes—Monograph Series. Institute of Mathematical Statistics, Beachwood, pp 1–25
- Stigler SM (1974) Studies in the history of probability and statistics. XXXIII: Cauchy and the witch of Agnesi: An historical note on the Cauchy distribution. *Biometrika* 61:375–380
- Szczygielski J, Karathanasopoulos A, Zaremba A (2020) One shape fits all? A comprehensive examination of cryptocurrency return distributions. *Appl. Econ. Lett.* 27:1567–1573
- Tenreiro C (2019) On the automatic selection of the tuning parameter appearing in certain families of goodness-of-fit tests. *J Stat Comput Simul* 89(10):1780–1797
- Villaseñor J, González-Estrada E (2021) Goodness-of-fit tests for Cauchy distributions using data transformations. In: Ghosh I, Balakrishnan N, Ng H (eds) Advances in statistics: theory and applications. Emerging topics in statistics and biostatistics. Springer, Berlin
- Zhang J (2002) Powerful goodness-of-fit tests based on the likelihood ratio. *J R Stat Soc B* 64(2):281–294

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.