Ann. Inst. Statist. Math.
Vol. 52, No. 2, 267–286 (2000)

GOODNESS-OF-FIT TESTS FOR THE CAUCHY DISTRIBUTION BASED ON THE EMPIRICAL CHARACTERISTIC FUNCTION*

NORA GÜRTLER AND NORBERT HENZE

Institut für Mathematische Stochastik, Universität Karlsruhe, Englerstr. 2, 76128 Karlsruhe, Germany

(Received July 31, 1998; revised December 25, 1998)

Abstract. Let $X_1, \dots, X_n$ be independent observations on a random variable X. This paper considers a class of omnibus procedures for testing the hypothesis that the unknown distribution of X belongs to the family of Cauchy laws. The test statistics are weighted integrals of the squared modulus of the difference between the empirical characteristic function of the suitably standardized data and the characteristic function of the standard Cauchy distribution. A large-scale simulation study shows that the new tests compare favorably with the classical goodness-of-fit tests for the Cauchy distribution, based on the empirical distribution function. For small sample sizes and short–tailed alternatives, the uniformly most powerful invariant test of Cauchy versus normal beats all other tests under discussion.

Key words and phrases: Goodness-of-fit test, Cauchy distribution, empirical characteristic function, kernel transformed empirical process, stable distribution, uniformly most powerful invariant test.

1. Introduction

The Cauchy distribution has a long and rich history (see Stigler (1974)), and there are numerous characterizations of this probability model and methods of inference for its parameters (see Johnson *et al.* (1994)). However, there is still a paucity of genuine goodness-of-fit tests for the Cauchy family. In the spirit of approaches for assessing univariate and multivariate normality and exponentiality (see Epps and Pulley (1983), Epps and Pulley (1986), Baringhaus and Henze (1991), Henze (1993), Henze and Zirkler (1990)), this paper studies a new class of tests of fit for the Cauchy distribution, which are based on the empirical characteristic function. To be specific, suppose $X_1, \dots, X_n, \dots$ are independent random variables with unknown distribution function (df) F. The problem is to test the hypothesis

$$H_0 : F \in \mathcal{F} := \{F(\cdot, \vartheta) : \vartheta \in \Theta\}$$

against general alternatives, on the basis of the observations $X_1, \dots, X_n$. Here, $:=$ denotes definition, $\Theta := \{\vartheta = (\alpha, \beta)' : \alpha \in \mathbb{R}, \beta > 0\}$ is the two-dimensional parameter space (the prime denoting transpose), and

$$F(x; \vartheta) := 1/2 + \pi^{-1} \arctan((x - \alpha)/\beta)$$

is the df of a Cauchy law $C(\alpha, \beta)$ with median α and interquartile range 2β. Put in other words, the problem is whether data from an unknown distribution belong to the

* Work supported by the Deutsche Forschungsgemeinschaft.

location-scale family $\mathcal{F}$ generated by the standard Cauchy df $F_0(x) := F(x;(0,1)') = 1/2 + \pi^{-1}\arctan(x)$. Since $\mathcal{F}$ is closed with respect to affine transformations and the alternatives to H_0 are rarely known in practice, one is interested in affine invariant and consistent tests. The proposed tests are based on the empirical characteristic function

$$\Psi_n(t) := \frac{1}{n}\sum_{j=1}^{n}\exp(itY_j)$$

of the 'standardized' data $Y_j = (X_j - \hat{\alpha}_n)/\hat{\beta}_n$, $1 \le j \le n$. Here, $\hat{\alpha}_n = \hat{\alpha}_n(X_1,\ldots,X_n)$ and $\hat{\beta}_n = \hat{\beta}_n(X_1,\ldots,X_n)$ are estimators for α and β, respectively, such that

$$\hat{\alpha}_n(aX_1+b,\ldots,aX_n+b) = a\hat{\alpha}_n(X_1,\ldots,X_n)+b, \tag{1.1}$$

$$\hat{\beta}_n(aX_1+b,\ldots,aX_n+b) = a\hat{\beta}_n(X_1,\ldots,X_n) \tag{1.2}$$

for each $a > 0$ and $b \in \mathbb{R}$. Since by (1.1) and (1.2), $Y_1,\ldots,Y_n$ and hence $\Psi_n(t)$ do not depend on the median or on the interquartile range of the underlying distribution, we assume $\alpha = 0$ and $\beta = 1$ in what follows. The test statistic

$$D_{n,\lambda} := n\int_{-\infty}^{\infty}|\Psi_n(t) - e^{-|t|}|^2 e^{-\lambda|t|}dt \tag{1.3}$$

is the weighted L^2-distance between $\Psi_n(t)$ and the characteristic function $\exp(-|t|)$ of $C(0,1)$, λ denoting a fixed positive weighting parameter. Rejection of H_0 is for large values of $D_{n,\lambda}$. The rationale behind (1.3) is that, under H_0 and for a suitable choice of $\{(\hat{\alpha}_n,\hat{\beta}_n)\}_{n\ge1}$, $\Psi_n(t)$ converges in probability to $\exp(-|t|)$. Since

$$D_{n,\lambda} = \frac{2}{n}\sum_{j,k=1}^{n}\frac{\lambda}{\lambda^2+(Y_j-Y_k)^2} - 4\sum_{j=1}^{n}\frac{1+\lambda}{(1+\lambda)^2+Y_j^2} + \frac{2n}{2+\lambda}, \tag{1.4}$$

an efficient and numerically stable computer routine implementing the test is easily available. Straightforward algebra yields an alternative representation, which will be needed for determining the limit distribution of $D_{n,\lambda}$:

$$D_{n,\lambda} = \int_{-\infty}^{\infty}\hat{Z}_n(t)^2\hat{\beta}_n e^{-\hat{\beta}_n\lambda|t|}dt, \qquad \text{where}$$

$$\hat{Z}_n(t) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\{\cos(tX_j)+\sin(tX_j) - e^{-\hat{\beta}_n|t|}(\cos(t\hat{\alpha}_n)+\sin(t\hat{\alpha}_n))\}. \tag{1.5}$$

The paper is organized as follows. Section 2 presents theoretical results concerning the weak convergence of $D_{n,\lambda}$ under H_0, its limit distribution, and the consistency of the corresponding test. The proofs, which utilize the theory of weak convergence in the Fréchet space $C(\mathbb{R})$ of continuous functions on $\mathbb{R}$ and rely heavily on the work of Csörgő (1983), are deferred to Section 3. Due to the non-existing first moment of the Cauchy law, the derivations are more involved compared with the approach taken, for example, in Henze and Wagner (1997). Section 4 presents the results of a large-scale simulation study on the power of the new tests in comparison with several classical tests for the Cauchy distribution as well as the uniformly most powerful invariant test of Cauchy versus normal. An important message is that, for the goodness-of-fit problem under discussion, the tests of Kolmogorov-Smirnov, Cramér-von Mises, Anderson-Darling, and Watson, each based on the empirical df, should not be used as described in D'Agostino and Stephens (1986).

2. Theoretical results

Throughout the rest of the paper, $\xrightarrow{\mathcal{D}}$ denotes weak convergence of random variables or stochastic processes, $\xrightarrow{P}$ is convergence in probability, $o_P(1)$ stands for convergence in probability to 0, $\mathbf{1}\{A\}$ is the indicator function of a set A, and i.i.d. means 'independent and identically distributed'. An unspecified integral is over the real line.

To estimate $\vartheta = (\alpha, \beta)'$, we choose particular location and scale estimators $\hat{\alpha}_n$, $\hat{\beta}_n$ satisfying (1.1) and (1.2), respectively. For this purpose, denote ξ_p $(0 < p < 1)$ the p-quantile of the underlying distribution F and $\hat{\xi}_{pn}$ the sample p-quantile of $X_1, \ldots, X_n$. Writing $X_{(1)}, \ldots, X_{(n)}$ for the order statistics of the observations and $\lfloor x \rfloor$ for the largest integer not greater than x, let

$$\hat{\alpha}_n := \begin{cases} \frac{1}{2}(X_{(n/2)} + X_{(n/2+1)}), & \text{if } n \text{ is even} \\ X_{(\lfloor n/2 \rfloor + 1)}, & \text{otherwise} \end{cases} \tag{2.1}$$

be the unbiased empirical median and

$$\hat{\beta}_n := \frac{1}{2}(\hat{\xi}_{3/4n} - \hat{\xi}_{1/4n}) \tag{2.2}$$

the half-interquartile range of the sample. Under mild regularity conditions, these are consistent estimators for the median $\alpha(= 0)$ and half-interquartile range $\beta(= 1)$ of F.

The process $\hat{Z}_n$ defined in (1.5) can be considered as a random element in the Fréchet space $C(\mathbb{R})$ of continuous functions on $\mathbb{R}$, endowed with the metric

$$\rho(x, y) = \sum_{j=1}^{\infty} 2^{-j} \frac{\rho_j(x, y)}{1 + \rho_j(x, y)},$$

where $\rho_j(x, y) = \max_{|t| \le j} |x(t) - y(t)|$. The first result is about weak convergence of $\hat{Z}_n$, the second about weak convergence of the corresponding integral statistic $D_{n,\lambda}$.

THEOREM 2.1. *Let $X_1, \ldots, X_n$ be i.i.d. $C(0,1)$-random variables, $\hat{Z}_n$ as defined in* (1.5), *and put*

$$J_1(s) := \int_0^{\infty} \frac{\sin(sx)}{1 + x^2} dx, \qquad J_2(s) := \int_0^1 \frac{\cos(sx)}{1 + x^2} dx. \tag{2.3}$$

There exists a zero mean Gaussian process Z in $C(\mathbb{R})$ having covariance kernel

$$\begin{aligned} c(s, t) = {} & e^{-|s-t|} - e^{-|s|-|t|} + \frac{\pi}{2} e^{-|s|-|t|} \left[\frac{\pi}{2} st + \frac{\pi}{2} |st| + |s| + |t| \right] \\ & - e^{-|t|}[tJ_1(s) + 2|t|J_2(s)] - e^{-|s|}[sJ_1(t) + 2|s|J_2(t)] \end{aligned} \tag{2.4}$$

$(s, t \in \mathbb{R})$, *such that $\hat{Z}_n \xrightarrow{\mathcal{D}} Z$ in $C(\mathbb{R})$.*

THEOREM 2.2. *Under the conditions of Theorem* 2.1, *we have for every positive λ*

$$D_{n,\lambda} = \int \hat{Z}_n(t)^2 \hat{\beta}_n e^{-\hat{\beta}_n \lambda |t|} dt \xrightarrow{\mathcal{D}} D_\lambda := \int Z(t)^2 e^{-\lambda |t|} dt.$$

Remark 2.1. Theorem 2.2 is not a trivial consequence of Theorem 2.1, since the functional

$$f \mapsto ||f||_\lambda^2 := \int f(t)^2 e^{-\lambda|t|} dt$$

is not continuous on $C(\mathbb{R})$. It is not even defined on $C(\mathbb{R})$, but only on the subset of functions that are square integrable with respect to $e^{-\lambda|t|}dt$. Things are even more complicated, since $D_{n,\lambda} = \hat{\beta}_n ||\hat{Z}_n||^2_{\hat{\beta}_n\lambda}$, where

$$||f||^2_{\hat{\beta}_n\lambda} := \int f(t)^2 e^{-\hat{\beta}_n\lambda|t|} dt,$$

i.e. $D_{n,\lambda}$ depends on the non-deterministic weight function $\exp(-\hat{\beta}_n\lambda|t|)$.

Remark 2.2. By Mercer's theorem (see e.g. Jörgens (1970), p. 152), the distribution of D_λ is that of $\sum_{j\geq 1} \eta_j(\lambda) N_j^2$, where $N_1, N_2, \ldots$ are independent unit normal random variables and $(\eta_j(\lambda))_{j\geq 1}$ are the nonzero eigenvalues of the integral operator A defined by

$$Ag(s) = \int c(s,t) g(t) e^{-\lambda|t|} dt.$$

Although not being able to solve the equation $Ag(s) = \eta g(s)$ and determining $\eta_j(\lambda)$ explicitly, we obtained the expectation of the limiting distribution via the relation

$$ED_\lambda = \int c(t,t) e^{-\lambda|t|} dt$$

by straightforward, but tedious manipulations of integrals:

$$\begin{aligned} ED_\lambda = & \frac{4(\lambda+1)}{\lambda^2(\lambda+2)^2} \left[4 \arctan\left(\frac{1}{\lambda+1}\right) - 2\ln(\lambda+1) - \pi \right] \\ & + \frac{2\pi^2}{(\lambda+2)^3} + \frac{4}{\lambda(\lambda+1)} + \frac{8}{\lambda(\lambda+2)(\lambda^2+2\lambda+2)}. \end{aligned}$$

Likewise, the variance of D_λ can be obtained via the formula

$$\operatorname{Var} D_\lambda = 2 \iint c(s,t)^2 e^{-\lambda|s|} e^{-\lambda|t|} ds dt,$$

but deriving an explicit expression requires immense calculations that seem to be disproportionate in view of the availability of efficient routines for numerical integration.

Besides affine invariance, the proposed test has the appealing feature of consistency against a large class of alternatives.

THEOREM 2.3. *For a given $\gamma \in (0,1)$ let $d_{n,\lambda}(\gamma)$ be the $(1-\gamma)$-quantile of $D_{n,\lambda}$ under the hypothesis. The test that rejects H_0 if $D_{n,\lambda} > d_{n,\lambda}(\gamma)$ is consistent against each alternative distribution having a unique median and unique upper and lower quartiles.*

Remark 2.3. Interestingly, the class of tests based on the family $\{D_{n,\lambda} : 0 < \lambda < \infty\}$ is 'closed at the boundaries' $\lambda \to 0$ and $\lambda \to \infty$. For a related class of statistics proposed in the context of testing for multivariate normality, this has been observed in

Henze (1997). As in Baringhaus *et al.* (2000), an Abelian theorem for Laplace transforms (see Widder (1959), p. 182) yields

$$\lim_{\lambda\to\infty} \lambda^3 D_{n,\lambda} = 4n\left(1+\overline{Y}_n^2\right) =: D_{n,\infty}$$

for fixed n and pointwise on the underlying probability space, with $\overline{Y}_n := n^{-1}\sum_{k=1}^n Y_k$. As for the behavior of $D_{n,\lambda}$ for $\lambda \to 0$, note that n of the summands of the double sum figuring in (1.4) are equal to $1/\lambda$. It is then obvious that

$$D_{n,\lambda} - \frac{2}{\lambda} \to n - 4\sum_{j=1}^n \frac{1}{1+Y_j^2} =: D_{n,0} \qquad \text{as} \quad \lambda \to 0.$$

3. Proofs

For short, put $\hat{\vartheta}_n := (\hat{\alpha}_n, \hat{\beta}_n)'$, $\vartheta := (\alpha,\beta)'$, and $\vartheta_0 := (\alpha_0,\beta_0)' := (0,1)'$.

Proof of Theorem 2.1. Weak convergence of the process $\hat{Z}_n$ is shown by fitting the present situation into the framework of Csörgő (1983). In the tradition of Durbin (1973), Komlós *et al.* (1975), Burke *et al.* (1978), Csörgő (1981) and many others, that paper summarizes results and gives handy conditions under which kernel transformations of the empirical parameter-estimated process converge in distribution. Thus, we first work in the space $C(S)$ of continuous functions on a compact subset S of $\mathbb{R}$, endowed with the supremum norm $\|f\|_\infty = \sup_{t\in S}|f(t)|$. Putting $k(x,t) := \cos(tx)+\sin(tx)$, we have the following representation of $\hat{Z}_n$:

$$\hat{Z}_n(t) = \frac{1}{\sqrt{n}}\sum_{j=1}^n\{\cos(tX_j)+\sin(tX_j) - e^{-\hat{\beta}_n|t|}(\cos(t\hat{\alpha}_n)+\sin(t\hat{\alpha}_n))\} \tag{3.1}$$

$$= \int k(x,t)d\{\sqrt{n}(F_n(x)-F(x,\hat{\vartheta}_n))\}. \tag{3.2}$$

Hence, $\hat{Z}_n$ can be regarded as a random element of $C(S)$, more precisely as the kernel transform corresponding to $k(x,t)$ of the parameter-estimated empirical process $\sqrt{n}(F_n(x) - F(x,\hat{\vartheta}_n))$. Section 3 of Csörgő (1983) deals with the weak convergence of $\hat{Z}_n$ to a centered Gaussian process in the space $(C(S), \|\cdot\|_\infty)$ under certain conditions (i)*, (ii)*, (iv), (v) and (vi), which will be checked in what follows.

Condition (i)* of Csörgő (1983) holds trivially, since for arbitrary $\delta > 0$,

$$\int \sup_{t\in S}|k(x,t)|^{2+\delta}dF_0(x) = \int\sup_{t\in S}|\cos(tx)+\sin(tx)|^{2+\delta}\frac{1}{\pi(1+x^2)}dx < \infty.$$

To verify condition (ii)*, we have to find a number $\varepsilon\in(0,1]$ and functions $v: S\times S\to S$ and $M:\mathbb{R}\times S\to\mathbb{R}$ with $\int\sup_{t\in S}M^2(x,t)dF_0(x)<\infty$, such that for all $s,t\in S$ and for every real x: $|k(x,s)-k(x,t)|\le|s-t|^\varepsilon M(x;v(s,t))$. This is true with $\varepsilon := 1/4$, $M(x,u) := 4|x|^{1/4}$ and arbitrary $v: S\times S\to S$, since

$$\begin{aligned}|k(x,s)-k(x,t)| &\le |\cos(sx)-\cos(tx)| + |\sin(sx)-\sin(tx)|\\ &= 2|\sin((s-t)x/2)|\cdot[|\sin((s+t)x/2)|+|\cos((s+t)x/2)|]\\ &\le 4\cdot|\sin((s-t)x/2)|\\ &\le 4\cdot|s-t|^{1/4}\cdot|x|^{1/4}.\end{aligned}$$

Conditions (i)* and (ii)* suffice to show weak convergence of the kernel transformed empirical process without estimated parameters. To account for regularity conditions concerning the parameter estimates, the following notations are needed. Put

$$\nabla_\vartheta F(x,\vartheta_*) := \left(\frac{\partial}{\partial\alpha}F(x,\vartheta), \frac{\partial}{\partial\beta}F(x,\vartheta)\right)'\Bigg|_{\vartheta=\vartheta_*=\binom{\alpha_*}{\beta_*}} = -\frac{(\beta_*, x-\alpha_*)'}{\pi(\beta_*^2+(x-\alpha_*)^2)}$$

and

$$H(t,\vartheta) := (H_1(t,\vartheta), H_2(t,\vartheta))' := \int k(x,t)d\nabla_\vartheta F(x,\vartheta).$$

In the present case, this leads to

$$H_1(t,\vartheta) = \frac{2}{\pi\beta}\int[\cos(t(\alpha+\beta y))+\sin(t(\alpha+\beta y))]\frac{y}{(1+y^2)^2}dy,$$
$$H_2(t,\vartheta) = -\frac{1}{\pi\beta}\int[\cos(t(\alpha+\beta y))+\sin(t(\alpha+\beta y))]\frac{1-y^2}{(1+y^2)^2}dy.$$

Since the function H is continuous and bounded on $S\times\Theta_0$, where Θ_0 is the closure of some neighbourhood of $\vartheta_0 = (0,1)'$, condition (vi) of Csörgő (1983) holds. Note that

$$H(t,\vartheta_0) = (te^{-|t|}, -|t|e^{-|t|})'.$$

Theorem 2.5.1 of Serfling (1980) gives the so-called *Bahadur representations* for the sample median $\hat\alpha_n$ and half-interquartile range $\hat\beta_n$:

$$\sqrt{n}\hat\alpha_n = \frac{1}{\sqrt{n}}\sum_{j=1}^n l_1(X_j)+r_n, \qquad \sqrt{n}(\hat\beta_n-1) = \frac{1}{\sqrt{n}}\sum_{j=1}^n l_2(X_j)+r_n, \tag{3.3}$$

where $r_n = O(n^{-1/4}(\log n)^{3/4})$ almost surely as $n\to\infty$, and l_1, l_2 are defined by

$$l_1(x) := \pi(1/2-\mathbf{1}\{x\le 0\}), \qquad l_2(x) := \pi(1/2-\mathbf{1}\{-1\le x\le 1\})$$

$(x\in\mathbb{R})$. In view of $El_1(X_1) = El_2(X_1) = E[l_1(X_1)l_2(X_1)] = 0$, $E[l_1(X_1)^2] = E[l_2(X_1)^2] = \pi^2/4$ and the fact that, with $l(x) := (l_1(x), l_2(x))'$, the matrix

$$E[l(X_1)l(X_1)'] = \begin{pmatrix}\pi^2/4 & 0\\ 0 & \pi^2/4\end{pmatrix}$$

is finite and positive definite, condition (iv) of Csörgő (1983) holds. Finally, also condition (v) of that paper is valid, since the functions l_1 and l_2 are bounded on $\mathbb{R}$ and since their derivatives exist on $\mathbb{R}\backslash\{-1,0,1\}$, i.e. almost surely on $\mathbb{R}$, and equal zero.

Conditions (i)*, (ii)*, (iv), (v) and (vi) imply the weak convergence of $\hat Z_n$ in the space $(C(S), \|\cdot\|_\infty)$ to a zero mean Gaussian process Z. We reproduce the idea of the proof, since the decomposition of $\hat Z_n$ will be needed for the proof of Theorem 2.2. Let $\langle\cdot,\cdot\rangle$ be the standard inner product and $\|\cdot\|$ the maximum norm on $\mathbb{R}^2$. Inserting $-F_0(x)+F(x,\vartheta_0)$ into $\sqrt{n}(F_n(x)-F(x,\hat\vartheta_n))$ of (3.2), using a bivariate Taylor expansion $F(x,\hat\vartheta_n)-F(x,\vartheta_0) = \langle\hat\vartheta_n-\vartheta_0, \nabla_\vartheta F(x,\vartheta_n^*)\rangle$ with $\|\vartheta_n^*-\vartheta_0\|\le\|\hat\vartheta_n-\vartheta_0\|\to 0$ a.s., and finally replacing $\sqrt{n}(\hat\vartheta_n-\vartheta_0)$ by its Bahadur representation (3.3), it follows that

$$\begin{aligned}\hat Z_n(t) &= \int k(x,t)d\{\sqrt{n}(F_n(x)-F_0(x))\} - \langle\sqrt{n}(\hat\vartheta_n-\vartheta_0), H(t,\vartheta_n^*)\rangle\\ &= Z_n^*(t)+\Delta_n^{(2)}(t)+\Delta_n^{(3)}(t).\end{aligned} \tag{3.4}$$

Here, the process

$$(3.5)\quad Z_n^*(t) := \int k(x,t)d\{\sqrt{n}(F_n(x)-F_0(x))\} - \left\langle \frac{1}{\sqrt{n}}\sum_{j=1}^n l(X_j), H(t,\vartheta_0)\right\rangle$$
$$= \frac{1}{\sqrt{n}}\sum_{j=1}^n [\cos(tX_j)+\sin(tX_j)-e^{-|t|}-te^{-|t|}l_1(X_j)+|t|e^{-|t|}l_2(X_j)]$$

is a sum of centered i.i.d. random variables, which also converges to Z. The remainder terms $\Delta_n^{(2)}$ and $\Delta_n^{(3)}$ are

$$(3.6)\qquad \Delta_n^{(2)}(t) := \langle \sqrt{n}(\hat{\vartheta}_n-\vartheta_0), H(t,\vartheta_0)-H(t,\vartheta_n^*)\rangle,$$
$$(3.7)\qquad \Delta_n^{(3)}(t) := -\langle \varepsilon_n, H(t,\vartheta_0)\rangle, \qquad \varepsilon_n = (r_n, r_n)',$$

and satisfy $\sup_{t\in S}|\Delta_n^{(2)}| \xrightarrow{P} 0$ and $\sup_{t\in S}|\Delta_n^{(3)}| \xrightarrow{P} 0$ by conditions (vi) and (iv). The limit process Z can be written as a stochastic integral

$$(3.8)\qquad Z(t) = \int k(x,t)dB_{F_0}(x) - \left\langle \int l(x)dB_{F_0}(x), H(t,\vartheta_0)\right\rangle,$$

where B_{F_0} is the Brownian bridge associated with the df F_0, i.e. a centered Gaussian process having covariance kernel $EB_F(s)B_F(t) = F_0(s\wedge t)-F_0(s)F_0(t)$. The first integral in (3.8) is the limit of the 'unestimated' part $\int k(x,t)d\{\sqrt{n}(F_n(x)-F_0(x))\}$ of Z_n^*, and its 'estimated' part $\langle n^{-1/2}\sum_{j=1}^n l(X_j), H(t,\vartheta_0)\rangle$ converges to the second term in (3.8). Both Z_n^* and Z have the covariance kernel

$$(3.9)\quad c(s,t) = \tilde{K}_0(s,t) - K_0(s)K_0(t) + H(s,\vartheta_0)'E[l(X_1)l(X_1)']H(t,\vartheta_0)$$
$$-\left\langle H(t,\vartheta_0), \int k(x,s)l(x)dF_0(x)\right\rangle - \left\langle H(s,\vartheta_0), \int k(x,t)l(x)dF_0(x)\right\rangle,$$

where $K_0(t) := \int k(x,t)dF_0(x) = \exp(-|t|)$ and $\tilde{K}_0(s,t) := \int k(x,s)k(x,t)dF_0(x) = \exp(-|s-t|)$. It should be remarked that the mixed terms figuring in the second line of (3.9) are missing in Csörgő (1983), p. 526. Since

$$\int k(x,s)l_1(x)dF_0(x) = \int_0^\infty \frac{\sin(sx)}{1+x^2}dx = J_1(s),$$
$$\int k(x,s)l_2(x)dF_0(x) = \frac{\pi}{2}e^{-|s|} - 2\int_0^1 \frac{\cos(sx)}{1+x^2}dx = \frac{\pi}{2}e^{-|s|} - 2J_2(s)$$

(cf. (2.3)), the covariance kernel is as stated in Theorem 2.1. The compact set S being arbitrary, $\hat{Z}_n$ converges weakly to Z in the Fréchet space $C(\mathbb{R})$, endowed with the metric ρ (adapt e.g. the reasoning in Karatzas and Shreve (1988), p. 62 f.). □

Remark 3.1. The process Z_n^* emerges from $\hat{Z}_n$ naturally by expanding the terms in definition (3.1) that contain the estimators $\hat{\alpha}_n$ and $\hat{\beta}_n$. More precisely, a Taylor expansion of the exponential function at $|t|$ and of the sine and cosine functions at 0 and an approximation of $\hat{\alpha}_n$ and $\hat{\beta}_n$ by their Bahadur representations (3.3) yield Z_n^*.

Remark 3.2. On principle, $\hat{\alpha}_n$ and $\hat{\beta}_n$ of (2.1) and (2.2) may be substituted by other estimators for α and β. Provided that these admit a representation of the form

(3.3) with suitable functions l_1, l_2 and a remainder term $r_n = o_P(1)$, Theorem 2.1 remains valid with $c(s,t)$ replaced by the more general covariance kernel given in (3.9).

PROOF OF THEOREM 2.2. By Tonelli's theorem, $ED_\lambda = E[\int Z(t)^2 \exp(-\lambda|t|)\, dt] = \int c(t,t)\exp(-\lambda|t|)dt$ with c defined in (2.4). The last integral being finite, D_λ is defined almost surely. In what follows, we assume the conditions of Theorem 2.2 and make use of the definitions and notations in the proof of Theorem 2.1. The rest of the reasoning is divided into several lemmas. Since the proof of the first of these is the same as for (2.17) of Henze and Wagner (1997), pp. 10–12, it will not be repeated.

LEMMA 3.1. $\int Z_n^*(t)^2 e^{-\lambda|t|}dt \xrightarrow{\mathcal{D}} \int Z(t)^2 e^{-\lambda|t|}dt.$

LEMMA 3.2. $\int(\hat{Z}_n(t) - Z_n^*(t))^2 e^{-\lambda|t|}dt \xrightarrow{P} 0.$

PROOF. Using decomposition (3.4) together with definitions (3.6) and (3.7), and writing $\tau_{n1} := \sqrt{n}\hat{\alpha}_n$, $\tau_{n2} := \sqrt{n}(\hat{\beta}_n - 1)$, it follows that

$$\begin{aligned}
&\int(\hat{Z}_n(t) - Z_n^*(t))^2 e^{-\lambda|t|}dt \\
&\quad = \int(\Delta_n^{(2)}(t) + \Delta_n^{(3)}(t))^2 e^{-\lambda|t|}dt \\
&\quad = \sum_{i,j=1}^{2} \tau_{ni}\tau_{nj} \int (H_i(t,\vartheta_0) - H_i(t,\vartheta_n^*))(H_j(t,\vartheta_0) - H_j(t,\vartheta_n^*))e^{-\lambda|t|}dt \\
&\qquad + 2\sum_{i,j=1}^{2} \tau_{ni} r_n \int (H_i(t,\vartheta_0) - H_i(t,\vartheta_n^*))H_j(t,\vartheta_0)e^{-\lambda|t|}dt \\
&\qquad + \sum_{i,j=1}^{2} r_n^2 \int H_i(t,\vartheta_0)H_j(t,\vartheta_0)e^{-\lambda|t|}dt.
\end{aligned}$$

Since for $i \in \{1,2\}$ the sequences τ_{ni} are tight, the functions H_i are bounded and continuous on $S \times \Theta_0$, $H_i(t,\vartheta_0) - H_i(t,\vartheta_n^*) \to 0$ a.s. and $r_n \to 0$ a.s., we are done. □

LEMMA 3.3. $\int(\hat{Z}_n(t) - Z_n^*(t))^2[e^{-\hat{\beta}_n\lambda|t|} - e^{-\lambda|t|}]dt \xrightarrow{P} 0.$

The proof being the same as that of Lemma 3.2, it will be omitted.

LEMMA 3.4. $\int Z_n^*(t)^2 e^{-\hat{\beta}_n\lambda|t|}dt - \int Z_n^*(t)^2 e^{-\lambda|t|}dt \xrightarrow{P} 0.$

PROOF. Use the Taylor expansion

$$e^{-\hat{\beta}_n\lambda|t|} = e^{-\lambda|t|} - \lambda|t|e^{-\lambda|t|\Delta_n}(\hat{\beta}_n - 1)$$

with $\Delta_n \in (\min\{\hat{\beta}_n, 1\}, \max\{\hat{\beta}_n, 1\})$, uniformly in t, as well as Hölder's inequality to obtain

$$\begin{aligned}
&\left|\int Z_n^*(t)^2 e^{-\hat{\beta}_n\lambda|t|}dt - \int Z_n^*(t)^2 e^{-\lambda|t|}dt\right| \\
&\quad \le \lambda|\hat{\beta}_n - 1| \int Z_n^*(t)^2|t|e^{-\lambda|t|\Delta_n}dt \\
&\quad \le \lambda|\hat{\beta}_n - 1| \left(\int Z_n^*(t)^4 e^{-\lambda|t|}dt\right)^{1/2}\left(\int t^2 e^{-\lambda|t|(2\Delta_n - 1)}dt\right)^{1/2}.
\end{aligned}$$

Since $\hat{\beta}_n \to 1$ a.s. and since the last integral converges to $4/\lambda^3$ a.s., it remains to prove the tightness of the sequence $(V_n)_{n\geq 1}$, where

$$V_n := \left(\int Z_n^*(t)^4 e^{-\lambda|t|} dt \right)^{1/2}.$$

Now note that by (3.5), $Z_n^* = n^{-1/2} \sum_{j=1}^n h(X_j, t)$ with centered terms $h(X_j,t)$ that are bounded by some constant M for each ω and t, whence $E[Z_n^*(t)^4] \leq 4M^4$. By Jensen's inequality and Tonelli's theorem,

$$EV_n \leq \left(\int E[Z_n^*(t)^4] e^{-\lambda|t|} dt \right)^{1/2} \leq \left(\frac{8M^4}{\lambda} \right)^{1/2},$$

which, together with Markov's inequality concludes the proof of Lemma 3.4. □

PROOF OF THEOREM 2.2. Lemma 3.2 and Lemma 3.3 imply

$$\|[\hat{Z}_n - Z_n^*]\|^2_{\hat{\beta}_n \lambda} = \int (\hat{Z}_n(t) - Z_n^*(t))^2 e^{-\hat{\beta}_n \lambda |t|} dt \xrightarrow{P} 0,$$

hence $\|[\hat{Z}_n - Z_n^*]\|_{\hat{\beta}_n\lambda} \xrightarrow{P} 0$. The triangle inequality $|\|[\hat{Z}_n]\|_{\hat{\beta}_n\lambda} - \|[Z_n^*]\|_{\hat{\beta}_n\lambda}| \leq \|[\hat{Z}_n - Z_n^*]\|_{\hat{\beta}_n\lambda}$ yields $\|[\hat{Z}_n]\|_{\hat{\beta}_n\lambda} = \|[Z_n^*]\|_{\hat{\beta}_n\lambda} + o_P(1)$ and thus $\|[\hat{Z}_n]\|^2_{\hat{\beta}_n\lambda} = \|[Z_n^*]\|^2_{\hat{\beta}_n\lambda} + o_P(1)$. Then, by Lemma 3.1, Lemma 3.4 and Sluzky's lemma,

$$\|[\hat{Z}_n]\|^2_{\hat{\beta}_n\lambda} = (\|[\hat{Z}_n]\|^2_{\hat{\beta}_n\lambda} - \|[Z_n^*]\|^2_{\hat{\beta}_n\lambda}) + (\|[Z_n^*]\|^2_{\hat{\beta}_n\lambda} - \|[Z_n^*]\|^2_\lambda) + \|[Z_n^*]\|^2_\lambda \xrightarrow{\mathcal{D}} \|[Z]\|^2_\lambda$$

and thus $D_{n,\lambda} = \hat{\beta}_n \|[\hat{Z}_n]\|^2_{\hat{\beta}_n\lambda} \xrightarrow{\mathcal{D}} \|[Z]\|^2_\lambda$. □

PROOF OF THEOREM 2.3. Let $X_1, \ldots, X_n$ be i.i.d. random variables with df F having a unique median and unique upper and lower quartiles. Hence the empirical median $\hat{\alpha}_n$ and interquartile range $2\hat{\beta}_n$ converge almost surely to the median and interquartile range $\alpha := \xi_{1/2}$, $2\beta := \xi_{3/4} - \xi_{1/4}$ of F, respectively. Since $D_{n,\lambda}$ is affine invariant, assume $\alpha = 0$ and $\beta = 1$. The aim is to investigate the asymptotic behavior of $n^{-1}D_{n,\lambda}$, the first step in this direction being Lemma 3.5.

LEMMA 3.5. *For all* $\lambda > 0 : \int |e^{-\hat{\beta}_n\lambda|t|} - e^{-\lambda|t|}| dt \to 0$ a.s. *as* $n \to \infty$.

PROOF. Note that $\int_{-\infty}^{\infty} |e^{-\hat{\beta}_n\lambda|t|} - e^{-\lambda|t|}| dt = 2\int_0^\infty |e^{-\hat{\beta}_n\lambda t} - e^{-\lambda t}| dt$ and

$$\int_0^\infty |e^{-\hat{\beta}_n\lambda t} - e^{-\lambda t}| dt \leq \int_0^K |e^{-\hat{\beta}_n\lambda t} - e^{-\lambda t}| dt + \frac{1}{\hat{\beta}_n\lambda} e^{-\hat{\beta}_n\lambda K} + \frac{1}{\lambda} e^{-\lambda K} \tag{3.10}$$

for each $K > 0$. Pick $\varepsilon > 0$ and $K = K(\lambda, \varepsilon)$ satisfying $e^{-\lambda K}/\lambda < \varepsilon$. Use dominated convergence in the integral on the right-hand side of (3.10) and $\hat{\beta}_n \to 1$ a.s. to conclude that, almost surely, $\int_0^\infty |\exp(-\hat{\beta}_n\lambda t) - \exp(-\lambda t)| dt \leq 4\varepsilon$ for sufficiently large n. □

Write $n^{-1}D_{n,\lambda} = \hat{\beta}_n \int |A_n(t)|^2 \exp(-\hat{\beta}_n\lambda|t|) dt$, where $A_n(t) := n^{-1}\sum_{j=1}^n \exp(itX_j) - \exp(-\hat{\beta}_n|t| + i\hat{\alpha}_n t)$, and note that

$$|A_n(t)| \leq 2 \tag{3.11}$$

uniformly in n, t and on the underlying probability space. By the strong law of large numbers, $A_n(t) \to E[\exp(itX_1)] - \exp(|t|) =: A(t)$ a.s. as $n \to \infty$, and by Lemma 3.5,

$$\left| \int |A_n(t)|^2 (e^{-\hat{\beta}_n \lambda |t|} - e^{-\lambda |t|}) dt \right| \leq 4 \int |e^{-\hat{\beta}_n \lambda |t|} - e^{-\lambda |t|}| dt \to 0 \quad \text{a.s.} \tag{3.12}$$

By (3.11), Fubini's theorem and dominated convergence, we get $E[\int |A_n(t)|^2 e^{-\lambda|t|} dt] = \int E|A_n(t)|^2 e^{-\lambda|t|} dt \to \int |A(t)|^2 e^{-\lambda|t|} dt$. Likewise, $\mathrm{Var}(\int |A_n(t)|^2 e^{-\lambda|t|} dt) \to 0$, so that

$$\int |A_n(t)|^2 e^{-\lambda|t|} dt \xrightarrow{P} \int |A(t)|^2 e^{-\lambda|t|} dt \quad \text{as} \quad n \to \infty.$$

Combining this result with (3.12) yields

$$\frac{1}{n} D_{n,\lambda} \xrightarrow{P} \int |E[e^{itX_1}] - e^{-|t|}|^2 e^{-\lambda|t|} dt \quad \text{as} \quad n \to \infty.$$

Since this stochastic limit is zero if the underlying distribution is Cauchy and strictly positive for the alternatives considered in Theorem 2.3, an upper rejection region asymptotic test at some given level of significance is consistent against each such alternative distribution. □

4. Simulations

This section presents the results of a large-scale simulation study conducted to assess the power of the new tests in comparison with other tests of fit for the Cauchy distribution. The following procedures are compared:

1. *The tests based on $D_{n,\lambda}$ for $\lambda \in \{0.025, 0.1, 0.5, 1.0, 2.5, 5.0, 10.0\}$.* Their implementation is based on the computational form (1.4) with $Y_j = (X_j - \hat{\alpha}_n)/\hat{\beta}_n$ and $\hat{\alpha}_n$, $\hat{\beta}_n$ defined in (2.1), (2.2), respectively. Since the limit statistics $D_{n,0}$ and $D_{n,\infty}$ fail to be consistent and exhibit extremely poor power (cf. Fig. 1), they have been excluded from the study.

Remark 4.1. Instead of $\hat{\alpha}_n$ and $\hat{\beta}_n$, we also considered the asymptotically efficient estimators proposed in Chernoff *et al.* (1967), i.e.,

$$\tilde{\alpha}_n := \sum_{j=1}^{n} c_j X_{(j)} \quad \text{and} \quad \tilde{\beta}_n := \sum_{j=1}^{n} d_j X_{(j)}, \tag{4.1}$$

where

$$c_j := \frac{\sin[4\pi(j/(n+1) - 1/2)]}{n \tan[\pi(j/(n+1) - 1/2)]}, \qquad d_j := \frac{8 \tan[\pi(j/(n+1) - 1/2)]}{n \sec^4[\pi(j/(n+1) - 1/2)]},$$

as well as their affine-invariant counterparts, obtained by replacing the coefficients c_j by $nc_j/(n+1)$. However, since using these estimators resulted in a substantial loss of power of the test based on $D_{n,\lambda}$ for all the values of λ considered and for most of the alternatives chosen, the empirical power results regarding the new class of tests only refer to the standardization of the data using $\hat{\alpha}_n$ and $\hat{\beta}_n$.

2. *The tests proposed in D'Agostino and Stephens* (1986), pp. 160–164. Each of these classical procedures is based on a measure of discrepancy between the empirical df of $Z_{(j)} := 1/2 + \pi^{-1}\arctan[(X_{(j)} - \tilde{\alpha}_n)/\tilde{\beta}_n]$, $j = 1, \dots, n$, where $\tilde{\alpha}_n$ and $\tilde{\beta}_n$ are given in (4.1), and the df of the uniform distribution on the unit interval. In particular, we considered the Kolmogorov-Smirnov statistic KS, the Cramér-von Mises statistic CM, the Anderson-Darling statistic AD and the Watson statistic W, given by

$$KS = \max\left\{\max_{1\le j\le n}(j/n - Z_{(j)}), \max_{1\le j\le n}(Z_{(j)} - (j-1)/n)\right\},$$

$$CM = \sum_{j=1}^{n}[Z_{(j)} - (2j-1)/(2n)]^2 + 1/(12n),$$

$$AD = -n - n^{-1}\sum_{j=1}^{n}\{(2j-1)\log Z_{(j)} + (2n+1-2j)\log(1 - Z_{(j)})\},$$

$$W = CM - n\left(n^{-1}\sum_{j=1}^{n} Z_{(j)} - 1/2\right)^2 .$$

Alternatively, we implemented KS, CM, AD and W using the estimators $\hat{\alpha}_n$ and $\hat{\beta}_n$, i.e., putting $Z_{(j)} := 1/2 + \pi^{-1}\arctan[(X_{(j)} - \hat{\alpha}_n)/\hat{\beta}_n]$ in the definitions above. It will be seen that this choice has a striking influence on the power of the tests.

3. *The uniformly most powerful invariant test against normality.* Franck (1981) developed the uniformly most powerful scale and location invariant test of normality versus the family of Cauchy distributions. We use his results to derive the corresponding invariant test of 'Cauchy versus normal'. To this end, let $h(x)$ be the density of X_1, let $f_0(x) := 1/(\pi(1+x^2))$, $\varphi(x) := \exp(-x^2/2)/\sqrt{2\pi}$, $x \in \mathbb{R}$, and let '$\sim$' mean 'is proportional to'. According to Hájek and Šidák (1967), p. 49, the most powerful scale and location invariant test of

$$\widetilde{H}_0 : h(x) = \lambda f_0(\lambda x + u) \quad \text{for some} \quad \lambda > 0 \text{ and } u \in \mathbb{R} \text{ against}$$
$$\widetilde{H}_1 : h(x) = \lambda \varphi(\lambda x + u) \quad \text{for some} \quad \lambda > 0 \text{ and } u \in \mathbb{R}$$

rejects $\widetilde{H}_0$ for large values of $UMP := I_1/I_2$, where (cf. Franck (1981))

$$I_1 := \left(\frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X}_n)^2\right)^{-(n-1)/2} \sim \int_0^\infty \int_{-\infty}^\infty \varphi(\lambda x_1 - u, \dots, \lambda x_n - u)\lambda^{n-2} du d\lambda,$$

$$I_2 := \begin{cases} (-1)^{(n-3)/2}\displaystyle\sum_{k=1}^{n}\sum_{j<k}\frac{(X_j - X_k)^{n-1}\log|X_j - X_k|}{\prod_{l\ne j}(X_l - X_j)\prod_{l\ne k}(X_l - X_k)}, & \text{if } n \text{ is odd} \\ (-1)^{n/2}\displaystyle\sum_{k=1}^{n}\sum_{j<k}\frac{|X_j - X_k|^{n-1}}{\prod_{l\ne j}(X_l - X_j)\prod_{l\ne k}(X_l - X_k)}, & \text{if } n \text{ is even} \end{cases}$$

$$\sim \int_0^\infty \int_{-\infty}^\infty f_0(\lambda x_1 - u, \dots, \lambda x_n - u)\lambda^{n-2} du d\lambda.$$

Table 1. Percentage points for $D_{n,\lambda}$, $\lambda \in \{0.025, 0.1, 0.5, 1.0, 2.5, 5.0, 10.0\}$, $\gamma = 0.05$.

n	0.025	0.1	0.5	1.0	2.5	5.0	10.0
10	93.4	24.8	3.99	1.37	0.35	0.14	0.053
20	93.5	24.8	4.14	1.52	0.40	0.15	0.052
30	92.8	24.6	4.03	1.51	0.41	0.15	0.052
40	92.8	24.5	4.04	1.54	0.42	0.16	0.053
50	92.8	24.5	4.03	1.53	0.42	0.16	0.054
100	92.5	24.5	4.01	1.54	0.43	0.16	0.055
200	92.5	24.5	4.00	1.54	0.43	0.16	0.055

Table 2. Percentage points for $D_{n,\lambda}$, $\lambda \in \{0.025, 0.1, 0.5, 1.0, 2.5, 5.0, 10.0\}$, $\gamma = 0.1$.

n	0.025	0.1	0.5	1.0	2.5	5.0	10.0
10	88.2	22.5	3.43	1.20	0.30	0.11	0.036
20	88.7	22.6	3.55	1.31	0.34	0.12	0.040
30	88.6	22.6	3.51	1.32	0.35	0.13	0.042
40	88.7	22.6	3.53	1.33	0.35	0.13	0.042
50	88.6	22.6	3.52	1.33	0.36	0.13	0.043
100	88.6	22.6	3.52	1.34	0.36	0.13	0.044
200	88.7	22.6	3.52	1.34	0.37	0.13	0.045

Remark 4.2. Meintanis (1997) suggested two goodness-of-fit tests for the Cauchy family which are based on the empirical characteristic function of $X_1, \ldots, X_n$, evaluated at two points, and he presented simulation results on the power of these tests for samples of size 200, 500, and 1000. While his statistics are free of standardization and have simple asymptotic null distributions, the tests are consistent only against certain subclasses of alternatives. In view of the results in Table 1 of Meintanis (1997) and the empirical power given in Table 9 below, we conclude that only the second test of Meintanis can nearly compete with the procedures based on $D_{n,\lambda}$ for each of the values $\lambda = 1$, $\lambda = 2.5$, and $\lambda = 5$.

All calculations were done on an IBM RS/6000 SP parallel computer at the Rechenzentrum of the University of Karlsruhe, using at least double precision arithmetic in FORTRAN 90 and routines from the NAG and the IMSL libraries, whenever available.

The empirical critical values for the statistics $D_{n,\lambda}$, $\lambda \in \{0.025, 0.1, 0.5, 1.0, 2.5, 5.0, 10.0\}$, based on 100000 Monte Carlo replications, are given in Table 1 and Table 2 for the significance levels $\gamma = 0.05$ and $\gamma = 0.1$, respectively. Likewise, Table 3 and Table 4 exhibit critical values for KS, CM, AD and W. In these tables, an entry $a \mid b$ means that a and b refer to two different implementations of the test statistics, a using the estimators given in (4.1), and b the estimators $\hat{\alpha}_n$ and $\hat{\beta}_n$. The entries for a are in complete accordance with the values given in D'Agostino and Stephens (1986), p. 163, and in Stephens (1991).

Although the statistic *UMP* of the most powerful invariant test of 'Cauchy versus normal' may be calculated on a computer fairly easily, round-off errors cause severe problems even for moderate values of n. Double precision arithmetic suffices to give accurate results if $n \leq 20$. For larger values of n, extended precision arithmetic is needed, but even then the results are wrong if $n > 50$. We strongly recommend to adapt the numerically stable algorithm given in Franck (1981) to the present situation. The critical values for *UMP* were calculated with extended precision arithmetic and represent

Table 3. Percentage points for KS, CM, AD, W, $\gamma = 0.05$; location and scale estimators are $\tilde{\alpha}_n$, $\tilde{\beta}_n$ of (4.1) (left) and $\hat{\alpha}_n$, $\hat{\beta}_n$ (right).

n	KS	CM	AD	W
10	1.76|0.289	0.81|0.14	3.72|0.94	0.321|0.093
20	1.75|0.209	0.82|0.15	3.89|1.04	0.288|0.096
30	1.53|0.168	0.60|0.15	3.03|1.03	0.198|0.089
40	1.38|0.147	0.46|0.15	2.46|1.06	0.148|0.089
50	1.27|0.130	0.38|0.15	2.11|1.05	0.119|0.087
100	1.07|0.092	0.26|0.15	1.57|1.06	0.080|0.086
200	0.98|0.065	0.20|0.15	1.36|1.06	0.069|0.085

Table 4. Percentage points for KS, CM, AD, W, $\gamma = 0.10$; location and scale estimators are $\tilde{\alpha}_n$, $\tilde{\beta}_n$ of (4.1) (left) and $\hat{\alpha}_n$, $\hat{\beta}_n$ (right).

n	KS	CM	AD	W
10	1.41|0.255	0.47|0.11	2.24|0.77	0.201|0.075
20	1.33|0.187	0.41|0.12	2.11|0.84	0.156|0.078
30	1.20|0.152	0.32|0.12	1.75|0.84	0.116|0.074
40	1.11|0.132	0.27|0.12	1.54|0.86	0.096|0.074
50	1.06|0.118	0.23|0.12	1.39|0.86	0.084|0.073
100	0.93|0.084	0.17|0.12	1.14|0.87	0.065|0.073
200	0.86|0.060	0.15|0.12	1.03|0.87	0.057|0.072

Table 5. Empirical percentage points for UMP ($\gamma = 0.05$ and $\gamma = 0.1$).

γ	10	20	25	30	35	40	45	50
0.05	3.177	0.0680	0.0_2259	0.0_3163	0.0_5267	0.0_7845	0.0_9725	$0.0_{10}133$
0.10	2.057	0.0155	0.0_3319	0.0_4102	0.0_7896	0.0_8138	$0.0_{11}623$	$0.0_{13}516$

the 20%–trimmed mean of 10 Monte Carlo estimates, each based on 100000 replications. They are given in Table 5 for sample sizes $n \in \{10, 20, 25, 30, 35, 40, 45, 50\}$, where an entry like 0.0_3319 means 0.000319.

As alternatives to the Cauchy distribution, we considered several transitions from $\mathcal{C}(0,1)$ to the standard normal distribution $\mathcal{N}(0,1)$:

- Mixtures $p\mathcal{N}(0,1) + (1-p)\mathcal{C}(0,1)$ of $\mathcal{N}(0,1)$ and $\mathcal{C}(0,1)$, denoted by $\mathcal{NC}(p, 1-p)$, for mixing probabilities $p \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$.
- Student's distribution with k degrees of freedom, denoted by Student(k), for $k \in \{2, 3, 4, 5, 7, 10\}$. Remember that Student(1) is the Cauchy law and that for $k \to \infty$, Student(k) approaches the normal distribution.
- Stable distributions, denoted by Stable (a, b), with characteristic function

$$\phi(t) = \begin{cases} \exp(-|t|^a[1 - ib\,\mathrm{sgn}(t)\tan(a\pi/2)]), & \text{if } \ a \neq 1 \\ \exp(-|t|[1 + ib(2/\pi)\,\mathrm{sgn}(t)\log|t|]), & \text{if } \ a = 1. \end{cases}$$

We considered various combinations of the characteristic index a, which determines the basic properties of the law, and the parameter b, which describes the skewness. For $b = 0$, the so-called symmetric stable distributions perform a transition from the Cauchy law (Stable $(1, 0)$) to the normal law (Stable $(2, 0)$).

Table 6. Percentage of 10000 Monte Carlo samples declared significant by various tests for the Cauchy distribution ($\gamma = 0.1, n = 20$).

alternative	0.025	0.1	0.5	1.0	2.5	5.0	10.0	KS	CM	AD	W	UMP
$\mathcal{C}$(0,1)	10	10	10	10	10	10	10	9|10	9|10	9|10	9|10	10
$\mathcal{N}$(0,1)	12	14	19	24	37	43	17	0|13	0|15	0|17	1|30	97
$\mathcal{NC}$(0.1,0.9)	9	10	9	9	10	9	9	9|10	9|10	9|9	9|10	12
$\mathcal{NC}$(0.3,0.7)	10	10	9	9	9	8	7	7|8	7|9	7|8	7|9	15
$\mathcal{NC}$(0.5,0.5)	10	10	10	10	11	10	5	5|9	6|8	6|8	6|11	24
$\mathcal{NC}$(0.7,0.3)	10	12	12	13	16	16	6	3|9	3|10	3|9	3|15	41
$\mathcal{NC}$(0.9,0.1)	11	13	16	20	29	30	11	1|12	1|13	1|13	2|24	72
Student(2)	10	10	9	9	11	11	4	0|8	0|9	0|7	0|12	44
Student(3)	11	10	10	11	16	18	6	0|9	0|10	0|9	0|15	64
Student(4)	11	11	13	14	20	23	8	0|10	0|10	0|10	0|18	75
Student(5)	11	11	13	15	23	27	9	0|10	0|11	0|10	0|20	82
Student(7)	11	12	14	17	26	31	11	0|10	0|12	0|12	0|22	88
Student(10)	11	12	16	20	30	35	13	0|11	0|13	0|14	1|25	92
Stable(0.5,0)	36	46	50	55	58	65	74	70|38	67|43	66|72	77|50	0
Stable(1.2,0)	10	9	9	8	9	7	4	3|8	3|8	3|7	3|9	23
Stable(1.5,0)	11	10	11	12	14	15	5	0|9	1|9	1|8	1|15	50
Stable(1.7,0)	12	12	14	16	22	25	8	0|10	0|11	0|10	0|20	69
Stable(1.9,0)	11	13	17	21	32	37	14	0|12	0|14	0|15	1|27	89
Stable(0.5,-1)	88	95	98	99	95	76	68	98|99	95|98	95|99	97|98	0
Stable(1,-1)	27	39	59	66	56	35	25	34|71	26|62	29|64	31|65	14
Stable(1.5,-1)	13	16	22	28	32	30	15	1|27	1|24	1|24	2|33	53
Stable(2,-1)	12	13	18	24	37	44	17	0|13	0|15	0|17	1|30	97
Stable(0.5,1)	77	90	97	98	84	68	79	98|98	95|96	96|98	97|95	0
Stable(1,1)	21	31	48	54	34	24	24	33|58	26|50	28|56	31|48	14
Stable(1.5,1)	12	15	20	23	22	22	11	1|19	1|18	1|18	2|23	52
Stable(2,1)	12	13	19	25	38	43	17	0|13	0|15	0|17	1|29	97
Tukey(1.0)	11	11	12	12	13	12	15	13|11	13|12	13|13	13|11	7
Tukey(0.2)	11	11	11	12	16	18	6	0|9	0|10	0|9	0|16	67
Tukey(0.1)	11	12	14	17	25	29	10	0|10	0|11	0|11	0|21	86
Tukey(0.05)	12	13	17	20	31	36	13	0|11	0|13	0|14	0|25	92
$\mathcal{U}$(0,1)	25	38	61	73	81	83	56	0|44	0|44	0|53	9|70	*
Logistic	11	12	14	17	26	31	12	0|11	0|12	0|12	0|23	89
Laplace	10	10	10	10	14	16	6	0|8	0|9	0|8	0|13	69
Gumbel	13	16	24	29	34	37	18	0|19	0|19	0|20	1|30	83

- Tukey distributions, denoted by Tukey(h), for $h \in \{0.05, 0.1, 0.2, 1.0\}$. Tukey($h$) is the distribution of $Z\exp(hZ^2/2)$ with $Z \sim \mathcal{N}(0,1)$. Hence, Tukey(0) is the normal law, while the tails of Tukey (1.0) are Cauchy-like.

Moreover, several classical short-tailed alternatives were taken into account, namely the uniform distribution on the unit interval, denoted by $\mathcal{U}(0,1)$, the logistic distribution having density function $f(x) = e^x/(1+e^x)^2$, the Laplace distribution with density function $f(x) = \exp(-|x|)/2$, and the Gumbel extreme value distribution with df $F(x) =$

Table 7. Percentage of 10000 Monte Carlo samples declared significant by various tests for the Cauchy distribution ($\gamma = 0.1, n = 50$).

alternative	0.025	0.1	0.5	1.0	2.5	5.0	10.0	KS	CM	AD	W	UMP
$\mathcal{C}$(0,1)	9	9	10	10	10	10	10	9|11	9|10	9|10	9|10	10
$\mathcal{N}$(0,1)	19	28	57	78	94	97	98	5|45	1|51	14|77	64|80	99
$\mathcal{NC}$(0.1,0.9)	10	10	10	10	10	10	9	9|10	9|10	9|10	10|11	11
$\mathcal{NC}$(0.3,0.7)	11	11	12	13	13	12	9	6|11	6|10	6|10	9|13	14
$\mathcal{NC}$(0.5,0.5)	12	14	18	22	26	25	17	5|15	4|14	5|15	13|24	23
$\mathcal{NC}$(0.7,0.3)	14	18	30	39	50	52	41	4|22	2|23	4|29	26|42	39
$\mathcal{NC}$(0.9,0.1)	17	24	47	65	82	87	83	4|36	1|40	8|58	49|68	72
Student(2)	11	13	16	19	28	35	36	1|14	0|15	1|18	11|26	70
Student(3)	13	15	24	34	51	63	67	2|19	0|21	2|32	23|42	92
Student(4)	14	18	31	43	64	76	81	2|24	0|28	3|44	31|52	98
Student(5)	15	19	35	50	72	82	88	2|27	0|32	4|50	38|58	99
Student(7)	16	21	41	59	80	90	93	3|32	0|37	6|58	45|65	*
Student(10	16	23	46	65	86	93	96	3|35	1|41	8|64	51|71	*
Stable(0.5,0)	58	73	86	89	91	93	97	92|69	91|79	92|96	99|91	0
Stable(1.2,0)	11	11	12	13	14	15	12	3|11	2|11	3|10	6|15	30
Stable(1.5,0)	13	16	24	32	45	51	49	2|19	1|21	2|28	21|38	66
Stable(1.7,0)	15	20	36	51	69	78	77	2|29	1|32	4|48	37|58	85
Stable(1.9,0)	17	26	50	70	87	93	95	4|39	1|45	10|68	56|74	96
Stable(0.5,-1)	*	*	*	*	*	*	*	*|*	*|*	*|*	*|*	0
Stable(1,-1)	46	73	96	99	98	89	81	98|99	90|96	96|99	96|97	13
Stable(1.5,-1)	19	29	59	75	85	83	80	33|73	13|59	29|75	57|76	68
Stable(2,-1)	19	28	57	78	94	98	99	5|45	1|52	13|77	64|81	99
Stable(0.5,1)	*	*	*	*	*	*	*	*|*	*|*	*|*	*|*	0
Stable(1,1)	45	72	96	98	98	89	81	98|99	89|97	96|99	96|97	13
Stable(1.5,1)	19	30	59	76	85	83	80	34|72	13|58	29|74	57|75	68
Stable(2,1)	19	28	57	78	94	98	99	5|44	1|51	14|76	64|79	99
Tukey(1.0)	11	11	12	13	14	15	17	15|11	15|11	14|14	15|12	6
Tukey(0.2)	12	15	24	34	53	66	71	1|19	0|23	2|34	24|43	96
Tukey(0.1)	15	20	37	54	76	86	91	2|28	1|33	5|54	41|61	*
Tukey(0.05	17	23	46	66	86	94	96	3|36	1|42	9|65	52|71	*
$\mathcal{U}$(0,1)	58	86	99	*	*	*	*	75|99	29|96	83|*	99|*	90
Logistic	15	21	40	58	80	90	94	2|30	0|36	6|58	45|64	*
Laplace	11	12	16	24	41	56	64	1|14	0|17	1|25	12|31	*
Gumbel	21	34	67	85	95	97	97	31|76	8|65	30|84	72|85	99

$\exp(-\exp(-x))$.

Standard routines of the IMSL and the NAG library were used to generate random numbers from the distributions $\mathcal{C}(0,1)$, $\mathcal{N}(0,1)$, $\mathcal{U}(0,1)$, Student(k), Stable(a, b), and from the logistic distribution, whereas the inversion method was adopted to generate random numbers from the Laplace or the Gumbel distribution.

For the nominal level 10%, power estimates of the tests under discussion are shown in Tables 6–9, the entries being the percentages of 10000 Monte Carlo samples that resulted

Table 8. Percentage of 10000 Monte Carlo samples declared significant by various tests for the Cauchy distribution ($\gamma = 0.1, n = 100$).

alternative	0.025	0.1	0.5	1.0	2.5	5.0	10.0	KS	CM	AD	W
$\mathcal{C}$(0,1)	10	10	10	10	10	10	10	10|10	10|10	10|10	10|10
$\mathcal{N}$(0,1)	30	53	94	99	*	*	*	71|91	45|92	96|*	99|99
$\mathcal{NC}$(0.1,0.9)	10	10	10	10	10	10	9	9|11	8|10	9|10	10|11
$\mathcal{NC}$(0.3,0.7)	11	13	17	18	20	19	13	8|15	6|13	8|14	18|20
$\mathcal{NC}$(0.5,0.5)	14	19	33	42	49	48	38	11|25	5|23	11|31	41|42
$\mathcal{NC}$(0.7,0.3)	20	29	58	75	84	86	79	23|48	9|48	31|66	75|75
$\mathcal{NC}$(0.9,0.1)	25	44	85	97	*	*	*	51|80	27|81	79|96	97|96
Student(2)	14	16	28	38	57	70	76	8|23	3|25	13|43	45|46
Student(3)	17	23	48	68	88	95	98	16|41	7|46	39|77	76|76
Student(4)	19	29	62	82	96	99	*	25|54	11|60	58|89	87|87
Student(5)	21	33	68	88	98	*	*	32|63	16|69	69|94	92|92
Student(7)	22	37	78	94	*	*	*	42|72	23|77	81|97	96|95
Student(10)	24	42	83	97	*	*	*	51|79	29|82	87|99	98|97
Stable(0.5,0)	83	94	99	*	99	*	*	99|95	99|99	*|*	*|*
Stable(1.2,0)	11	12	16	18	24	28	27	6|15	3|14	6|16	19|22
Stable(1.5,0)	17	24	48	65	81	87	88	16|40	6|43	29|66	71|70
Stable(1.7,0)	21	35	72	89	97	99	99	34|64	16|69	65|91	92|91
Stable(1.9,0)	27	46	89	98	*	*	*	58|85	34|87	90|99	99|98
Stable(0.5,-1)	*	*	*	*	*	*	*	*|*	*|*	*|*	*|*
Stable(1,-1)	78	97	*	*	*	*	*	*|*	*|*	*|*	*|*
Stable(1.5,-1)	30	55	93	99	*	*	*	98|99	74|95	97|*	99|98
Stable(2,-1)	29	52	94	99	*	*	*	71|91	45|91	96|*	*|99
Stable(0.5,1)	*	*	*	*	*	*	*	*|*	*|*	*|*	*|*
Stable(1,1)	77	97	*	*	*	*	*	*|*	*|*	*|*	*|*
Stable(1.5,1)	31	55	93	99	*	*	*	98|99	75|94	97|*	99|98
Stable(2,1)	29	52	93	99	*	*	*	70|91	44|92	96|*	*|99
Tukey(1.0)	11	12	14	14	15	18	19	16|12	16|13	16|16	18|13
Tukey(0.2)	17	24	50	71	90	96	98	16|42	7|47	43|79	79|77
Tukey(0.1)	21	34	73	92	99	*	*	36|67	19|73	75|96	94|93
Tukey(0.05)	25	41	84	97	*	*	*	51|80	29|83	88|99	98|97
$\mathcal{U}$(0,1)	93	*	*	*	*	*	*	*|*	*|*	*|*	*|*
Logistic	23	37	77	94	*	*	*	40|71	21|76	80|97	96|94
Laplace	13	15	29	49	81	93	98	8|26	3|31	22|64	54|57
Gumbel	37	65	98	*	*	*	*	99|*	77|97	99|*	*|*

in rejection of H_0, rounded to the nearest integer. An asterisk denotes power 100%. For the statistics KS, CM, AD and W, and entry like $u \mid v$ means that the estimated power is u if the test is implemented as recommended in D'Agostino and Stephens (1986), i.e., using the estimators of (4.1), and the estimated power is v if the estimators $\hat{\alpha}$ and $\hat{\beta}$ are used instead. Figure 1 displays the empirical power of the tests based on $D_{n,\lambda}$ as a function of the parameter λ, for some selected alternative distributions. The main conclusions that can be drawn from the simulation results are the following:

Table 9. Percentage of 10000 Monte Carlo samples declared significant by various tests for the Cauchy distribution ($\gamma = 0.1, n = 200$).

alternative	0.025	0.1	0.5	1.0	2.5	5.0	10.0	KS	CM	AD	W
$\mathcal{C}$(0,1)	10	10	10	10	10	10	10	10|10	10|10	10|10	10|10
$\mathcal{N}$(0,1)	55	90	*	*	*	*	*	*|*	99|*	*|*	*|*
$\mathcal{NC}$(0.1,0.9)	10	11	12	11	11	11	10	9|11	8|10	9|11	12|12
$\mathcal{NC}$(0.3,0.7)	14	17	27	32	33	32	24	14|21	8|18	12|22	35|32
$\mathcal{NC}$(0.5,0.5)	20	31	61	74	80	79	69	34|48	16|46	38|61	77|72
$\mathcal{NC}$(0.7,0.3)	31	55	92	98	99	99	98	76|86	53|86	87|96	99|97
$\mathcal{NC}$(0.9,0.1)	46	80	*	*	*	*	*	99|*	94|*	*|*	*|*
Student(2)	17	25	54	73	91	96	99	33|48	17|53	65|85	86|80
Student(3)	25	42	85	97	*	*	*	72|83	52|87	97|99	99|98
Student(4)	30	55	95	99	*	*	*	89|95	75|96	*|*	*|*
Student(5)	35	62	98	*	*	*	*	96|98	85|98	*|*	*|*
Student(7)	40	70	99	*	*	*	*	99|99	93|99	*|*	*|*
Student(10)	43	78	*	*	*	*	*	*|*	96|*	*|*	*|*
Stable(0.5,0)	98	*	*	*	*	*	*	*|*	*|*	*|*	*|*
Stable(1.2,0)	13	16	26	33	43	51	52	12|22	7|21	16|32	42|39
Stable(1.5,0)	26	44	84	95	99	*	*	65|78	46|82	90|97	98|96
Stable(1.7,0)	36	66	98	*	*	*	*	94|97	84|98	*|*	*|*
Stable(1.9,0)	49	83	*	*	*	*	*	*|*	98|*	*|*	*|*
Stable(0.5,-1)	*	*	*	*	*	*	*	*|*	*|*	*|*	*|*
Stable(1,-1)	99	*	*	*	*	*	*	*|*	*|*	*|*	*|*
Stable(1.5,-1)	59	91	*	*	*	*	*	*|*	*|*	*|*	*|*
Stable(2,-1)	55	89	*	*	*	*	*	*|*	99|*	*|*	*|*
Stable(0.5,1)	*	*	*	*	*	*	*	*|*	*|*	*|*	*|*
Stable(1,1)	99	*	*	*	*	*	*	*|*	*|*	*|*	*|*
Stable(1.5,1)	59	91	*	*	*	*	*	*|*	*|*	*|*	*|*
Stable(2,1)	56	89	*	*	*	*	*	*|*	99|*	*|*	*|*
Tukey(1.0)	11	13	16	18	20	23	25	17|13	17|13	17|18	21|17
Tukey(0.2)	25	43	86	98	*	*	*	74|86	55|88	98|*	99|99
Tukey(0.1)	37	66	98	*	*	*	*	97|99	88|99	*|*	*|*
Tukey(0.05)	45	78	*	*	*	*	*	*|*	96|*	*|*	*|*
$\mathcal{U}$(0,1)	*	*	*	*	*	*	*	*|*	*|*	*|*	*|*
Logistic	39	70	99	*	*	*	*	99|*	92|99	*|*	*|*
Laplace	17	24	62	90	*	*	*	47|66	21|65	91|98	95|91
Gumbel	69	97	*	*	*	*	*	*|*	*|*	*|*	*|*

1. For short-tailed alternatives and for small sample sizes like $n = 20$, the test based on *UMP* outperforms all other tests under discussion. However, it has no power against the three (long-tailed) stable alternatives with characteristic exponent 0.5. A disadvantage of this procedure is that round-off errors preclude its applicability for sample sizes larger than 50. We presume that the peculiar behavior of *UMP* under $\mathcal{U}(0, 1)$ (a power of 100% for $n = 20$ and of only 90% for $n = 50$) may already be due to this effect.

2. Except for the stable alternatives with characteristic exponent 0.5, there is a

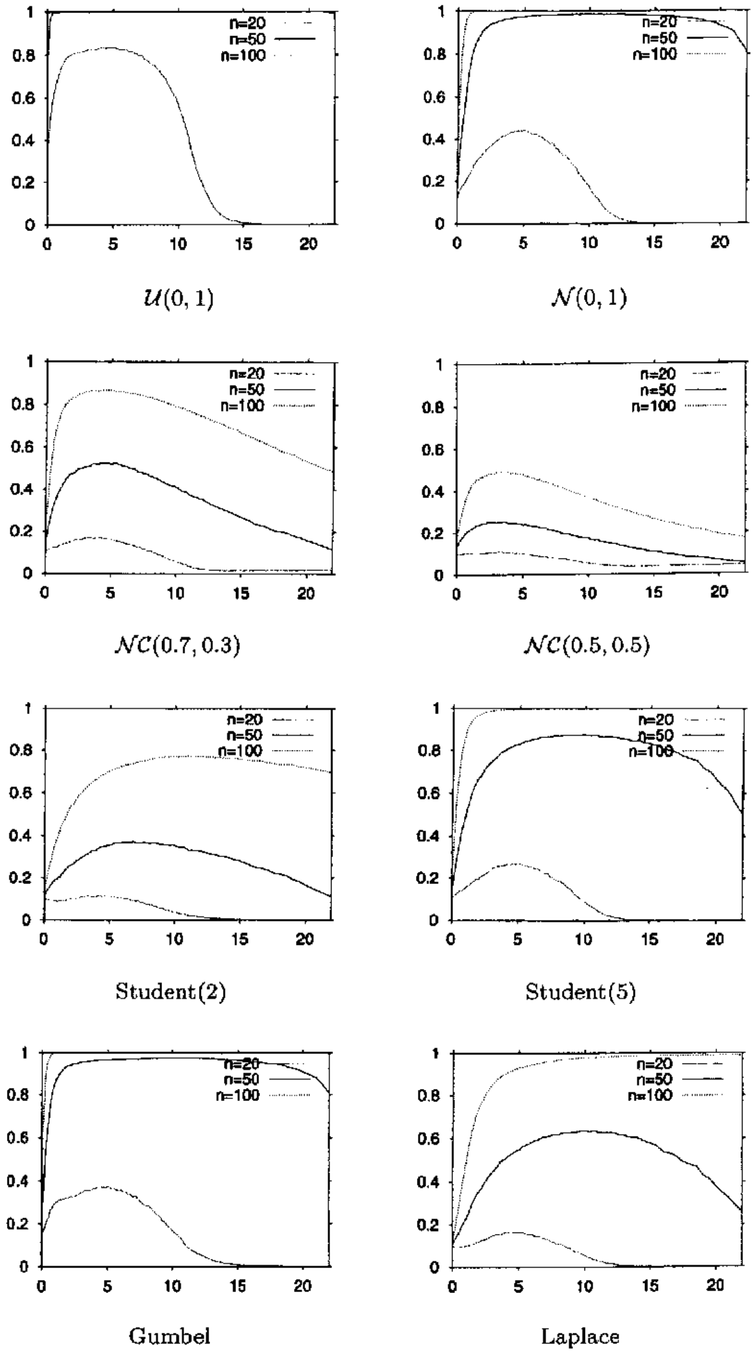

Fig. 1. Empirical power of the tests based on $D_{n,\lambda}$ as a function of the parameter λ for some selected alternatives ($\gamma = 0.10$).

striking increase in power of the empirical distribution function (EDF) tests based on KS, CM, AD and W if, instead of the estimators (4.1) recommended by D'Agostino and Stephens (1986), the estimators $\hat{\alpha}_n$ and $\hat{\beta}_n$ are used. In some respect, the power of the EDF tests is 'complementary' to that of *UMP*. Although being a little inferior with respect to some of the tests based on $D_{n,\lambda}$, the Watson test outperforms the other

EDF tests for the sample sizes 50, 100, and 200 and thus, when implemented with the estimators $\hat{\alpha}_n$ and $\hat{\beta}_n$, is a reasonable procedure for the testing problem under discussion. However the EDF tests, at least as described in D'Agostino and Stephens (1986), should not be used when testing for the Cauchy distribution.

3. For short-tailed alternatives in combination with very small sample sizes ($n \leq 20$), the new tests based on $D_{n,\lambda}$, as well as the EDF tests, are distinctly inferior to the test based on *UMP*. But especially for values of λ between 2.5 and 10.0, for long-tailed alternatives in combination with any sample size and for each alternative when $n \geq 50$, this class of tests is very competitive. Note that the W test with $\hat{\alpha}_n$ and $\hat{\beta}_n$, which is the best procedure among the EDF tests, is dominated by $D_{n,1}$ for $n = 20$ and by $D_{n,2.5}$ if $n \geq 50$ over the whole range of alternatives considered.

As for the new class of tests based on $D_{n,\lambda}$, $0 < \lambda < \infty$, there is a natural idea that suggests itself when looking at Fig. 1, namely letting λ depend on the data $X_1, \ldots, X_n$ in order to maximize power. This problem is an interesting topic of future research. A further basic problem is to give a theoretical explanation for the striking effect that different estimators for α and β may have on the power of the tests under discussion.

References

Baringhaus, L. and Henze, N. (1991). A class of consistent tests for exponentiality based on the empirical Laplace transform, *Ann. Inst. Statist. Math.*, **43**, 551–564.

Baringhaus, L., Gürtler, N. and Henze, N. (2000). Weighted and components of smooth tests of fit, *Australian and New Zealand Journal of Statistics*, **42** (to appear).

Burke, M., Csörgő, M., Csörgő, S. and Révész, P. (1978). Approximations of the empirical process when parameters are estimated, *Ann. Probab.*, **7**(5), 790–810.

Chernoff, H., Gastwirth, J. and Johns, M. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation, *Ann. Math. Statist.*, **38**, 52–72.

Csörgő, S. (1981). The empirical characteristic process when parameters are estimated, *Contributions to probability* (eds. J. Gani and V. Rohatgi), 215–230, Academic Press, New York.

Csörgő, S. (1983). Kernel-transformed empirical processes, *J. Multivariate. Anal.*, **13**, 517–533.

D'Agostino, R. and Stephens, M. (1986). *Goodness-of-Fit Techniques*, Marcel Dekker, New York.

Durbin, J. (1973). Weak convergence of the sample distribution function when parameters are estimated, *Ann. Statist.*, **1**(2), 279–290.

Epps, T. and Pulley, L. (1983). A test for normality based on the empirical characteristic function, *Biometrika*, **70**(3), 723–726.

Epps, T. and Pulley, L. (1986). A test of exponentiality vs. monotone-hazard alternatives derived from the empirical characteristic function, *J. Roy. Statist. Soc. Ser. B*, **48**(2), 206–213.

Franck, W. (1981). The most powerful invariant test of normal versus Cauchy with applications to stable alternatives, *J. Amer. Statist. Assoc.*, **76**(376), 1002–1005.

Hájek, J. and Šidák, Z. (1967). *Theory of rank tests*, Academic Press, New York.

Henze, N. (1993). A new flexible class of omnibus tests for exponentiality, *Comm. Statist. Theory Methods.*, **22**, 115–133.

Henze, N. (1997). Extreme smoothing and testing for multivariate normality, *Statist. Probab. Lett.*, **35**, 203–213.

Henze, N. and Wagner, T. (1997). A new approach to the BHEP tests for multivariate normality, *J. Multivariate Anal.*, **62**(1), 1–23.

Henze, N. and Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality, *Comm. Statist. Theory Methods.*, **19**(10), 3595–3617.

Johnson, N., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, Vol. 1, 2nd ed., Wiley, New York.

Jörgens, K. (1970). *Lineare Integraloperatoren*, Teubner, Stuttgart.

Karatzas, I. and Shreve, S. (1988). *Brownian Motion and Stochastic Calculus*, Springer, New York.

Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent rv's, and the sample df, *Z. W. theorie Verw. Geb.*, **32**, 111–131.

Meintanis, S. (1997). *Goodness-of-Fit tests for Cauchy Distributions Derived from the Empirical Characteristic Function* (Preprint). Greece: University of Patras.

Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

Stephens, M. (1991). Tests of fit for the Cauchy distribution based on the empirical distribution function, Teck. Report, No. 449, Department of Statistics, Stanford, California.

Stigler, S. (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**, 375–379.

Widder, D. (1959). *The Laplace Transform*, 5th ed, Princeton University Press, Princeton.