Ann. Inst. Statist. Math.
Vol. 57, No. 1, 183–199 (2005)

EMPIRICAL CHARACTERISTIC FUNCTION APPROACH TO GOODNESS-OF-FIT TESTS FOR THE CAUCHY DISTRIBUTION WITH PARAMETERS ESTIMATED BY MLE OR EISE

MUNEYA MATSUI[1] AND AKIMICHI TAKEMURA[2]

[1] *The Graduate School of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

[2] *The Graduate School of Information Science and Technology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

(Received June 5, 2003; revised March 22, 2004)

Abstract. We consider goodness-of-fit tests of the Cauchy distribution based on weighted integrals of the squared distance between the empirical characteristic function of the standardized data and the characteristic function of the standard Cauchy distribution. For standardization of data Gürtler and Henze (2000, *Annals of the Institute of Statistical Mathematics*, **52**, 267–286) used the median and the interquartile range. In this paper we use the maximum likelihood estimator (MLE) and an equivariant integrated squared error estimator (EISE), which minimizes the weighted integral. We derive an explicit form of the asymptotic covariance function of the characteristic function process with parameters estimated by the MLE or the EISE. The eigenvalues of the covariance function are numerically evaluated and the asymptotic distributions of the test statistics are obtained by the residue theorem. A simulation study shows that the proposed tests compare well to tests proposed by Gürtler and Henze and more traditional tests based on the empirical distribution function.

Key words and phrases: Asymptotic distributions, Fredholm determinant, integrated squared error estimator, integral equations, maximum likelihood estimator, residue theorem.

1. Introduction

Let $C(\alpha, \beta)$ denote the Cauchy distribution with the location parameter α and the scale parameter β, with the density

$$f(x;\theta) = f(x;\alpha,\beta) = \frac{\beta}{\pi(\beta^2 + (x-\alpha)^2)}, \qquad \theta = (\alpha, \beta).$$

Given a random sample $x_1, \ldots, x_n$ from an unknown distribution F, we want to test the null hypothesis H_0 that F belongs to the family of Cauchy distributions. Since the Cauchy distributions form a location scale family, we consider affine invariant tests. The proposed tests are based on the empirical characteristic function

$$\Phi_n(t) = \Phi_n(t;\hat{\alpha},\hat{\beta}) = \frac{1}{n}\sum_{j=1}^{n} \exp(ity_j), \qquad y_j = \frac{x_j - \hat{\alpha}}{\hat{\beta}}, \tag{1.1}$$

of the standardized data y_j. Here $\hat{\alpha} = \hat{\alpha}_n = \hat{\alpha}_n(x_1, \ldots, x_n)$ and $\hat{\beta} = \hat{\beta}_n = \hat{\beta}_n(x_1, \ldots, x_n)$ are affine equivariant estimators of α and β satisfying

$$\hat{\alpha}_n(a + bx_1, \ldots, a + bx_n) = a + b\hat{\alpha}_n(x_1, \ldots, x_n),$$
$$\hat{\beta}_n(a + bx_1, \ldots, a + bx_n) = b\hat{\beta}_n(x_1, \ldots, x_n).$$

For $\hat{\alpha}_n$ and $\hat{\beta}_n$, we use the maximum likelihood estimator (MLE) and an equivariant integrated squared error estimator (EISE) defined in (2.6) below. The reason for considering the MLE is its asymptotic efficiency. Although optimality as estimators does not imply optimality for the goodness-of-fit tests, it seems natural to consider the MLE. The reason for considering the EISE is a possible extension to stable distributions other than the Cauchy distribution studied in this paper. Although the median and the interquartile range used by Gürtler and Henze (2000) are attractive estimators because of their simplicity, it seems theoretically more natural to consider the MLE and the EISE.

Following Gürtler and Henze (2000) we consider the test statistic

$$D_{n,\kappa} := n \int_{-\infty}^{\infty} |\Phi_n(t) - e^{-|t|}|^2 w(t)dt, \qquad w(t) = e^{-\kappa|t|}, \ \kappa > 0, \tag{1.2}$$

which is the weighted L^2-distance between $\Phi_n(t)$ and the characteristic function $e^{-|t|}$ of $C(0,1)$ with respect to the weight function $w(t) = e^{-\kappa|t|}$, $\kappa > 0$. This weight function is chosen for convenience, so that we can explicitly evaluate the asymptotic covariance function of the empirical characteristic function process under H_0. Using the relation

$$\int_{-\infty}^{\infty} \cos(ct)e^{-\kappa|t|}dt = \frac{2\kappa}{\kappa^2 + c^2},$$

the integral in (1.2) can be explicitly evaluated and an alternative convenient expression of $D_{n,\kappa}$ is given by

$$D_{n,\kappa} = \frac{2}{n}\sum_{j,k=1}^{n} \frac{\kappa}{\kappa^2 + (y_j - y_k)^2} - 4\sum_{j=1}^{n} \frac{1+\kappa}{(1+\kappa)^2 + y_j^2} + \frac{2n}{2+\kappa}. \tag{1.3}$$

Our test statistic $D_{n,\kappa}$ is a quadratic form of the empirical characteristic function process. Although we derive an explicit form of the asymptotic covariance function of the empirical characteristic function process, it is not trivial to derive the asymptotic distribution of $D_{n,\kappa}$ under H_0 from the covariance function, especially when the parameters are estimated (e.g. Chapter 7 of Durbin (1973*a*) and Durbin (1973*b*)). Therefore finite sample critical values of goodness-of-fit tests are often evaluated by Monte Carlo simulation, as was done in Gürtler and Henze (2000). Note that if we evaluate the critical values by Monte Carlo simulation only, there is no need to derive the explicit form of the asymptotic covariance function. Furthermore it is impossible to perform usual Monte Carlo simulation for the asymptotic case $n = \infty$. Therefore numerical evaluation of the asymptotic distribution is important in order to check the convergence of the finite sample distributions, which are evaluated by Monte Carlo simulations.

In this paper, we make use of the explicit form of the asymptotic covariance function for numerically evaluating the asymptotic critical values of the test statistics. We introduce a homogeneous integral equation of the second kind and consider the associated

Fredholm determinant, which can be approximated by numerically evaluating eigenvalues of the asymptotic covariance function. Then we apply the residue theorem in Lévy's inversion formula and evaluate the asymptotic distribution function of $D_{n,\kappa}$.

This paper is organized as follows. In Subsection 2.1 we first define and summarize properties of the MLE and the EISE. Then we state theoretical results on asymptotic distribution of $D_{n,\kappa}$ under H_0 in Theorems 2.1 and 2.2. Our method for numerically evaluating the asymptotic critical values of $D_{n,\kappa}$ is discussed in Subsection 2.2. In Section 3 we present computational studies of the proposed tests in comparison to other testing procedures. Appendix A gives proofs of the theoretical results of Subsection 2.1. We utilize theorems of Csörgő (1983) and Gürtler and Henze (2000). Appendix B gives calculation of residues for the results of Subsection 2.2.

2. Main results

2.1 *Asymptotic theory of the proposed test statistics*

We first review maximum likelihood estimation of the Cauchy distributions and define an equivariant integrated squared error estimator. Except for differences in estimators, we follow the line of arguments in Gürtler and Henze (2000).

1. MLE. The log-likelihood function is given by

$$L = n\log\beta - \sum_{j=1}^{n}\log\{\beta^2 + (x_j-\alpha)^2\} - n\log\pi.$$

Differentiation of L with respect to (α,β) gives the likelihood equation

$$\frac{\partial L}{\partial\alpha} = 0 \Leftrightarrow \sum_{j=1}^{n}\frac{x_j-\alpha}{\beta^2+(x_j-\alpha)^2} = 0, \tag{2.1}$$

$$\frac{\partial L}{\partial\beta} = 0 \Leftrightarrow \sum_{j=1}^{n}\frac{\beta^2}{\beta^2+(x_j-\alpha)^2} = \frac{1}{2}n. \tag{2.2}$$

Equivariance of the MLE is easily checked. According to Copas (1975), except for pathological cases such that more than half of the observations are the same, the likelihood function L is unimodal. Therefore with probability one, a local maximum of the likelihood function is actually the global maximum and it is relatively easy to obtain the MLE by solving the likelihood equation.

2. EISE. Here we define an affine equivariant version of the ISE (integrated squared error) estimator proposed by Besbeas and Morgan (2001). The original ISE estimator of Besbeas and Morgan (2001) is not equivariant. The EISE is based on the empirical characteristic function of the standardized data. Let

$$\Phi_n(t;\alpha,\beta) = \frac{1}{n}\sum_{j=1}^{n}\exp(it(x_j-\alpha)/\beta),$$

which is the same as (1.1) with $\hat\alpha_n$ and $\hat\beta_n$ replaced by α and β. Write

$$I(\alpha,\beta) = \int_{-\infty}^{\infty}\left|\Phi_n(t;\alpha,\beta) - e^{-|t|}\right|^2 w(t)dt, \tag{2.3}$$

where we use the following weight function

$$w(t) = \exp(-\nu|t|), \qquad \nu > 0. \tag{2.4}$$

As in (1.3) the integral $I(\alpha, \beta)$ can be calculated as

$$I(\alpha,\beta) = \frac{2}{n^2}\sum_{j,k=1}^{n} \frac{\nu\beta^2}{\nu^2\beta^2 + (x_j - x_k)^2} - \frac{4}{n}\sum_{j=1}^{n} \frac{(1+\nu)\beta^2}{(1+\nu)^2\beta^2 + (x_j-\alpha)^2} + \frac{2}{2+\nu}. \tag{2.5}$$

Note that except for the factor n, $D_{n,\kappa}$ in (1.3) and $I(\alpha, \beta)$ are the same by the correspondence $\nu \leftrightarrow \kappa$, $x_j \leftrightarrow \alpha + \beta y_j$. The EISE $(\hat{\alpha}_n, \hat{\beta}_n)$ is defined to be the minimizer of $I(\alpha, \beta)$:

$$I(\hat{\alpha}_n, \hat{\beta}_n) = \min_{\alpha,\beta} I(\alpha, \beta). \tag{2.6}$$

It is easy to see that the EISE is affine equivariant by definition.

Note that the weighting constant κ in the test statistic (1.2) and the weighting constant ν in (2.4) for the EISE may be different. In our theoretical results on the EISE and $D_{n,\kappa}$ we treat ν and κ separately. However for performing goodness-of-fit test, it seems natural to set $\nu = \kappa$. In our simulation studies in Section 3 we set $\nu = \kappa$.

Setting $\partial I/\partial\alpha = \partial I/\partial\beta = 0$ in (2.5), we obtain the following estimating equations for the EISE.

$$\frac{\partial I}{\partial \alpha} = 0 \Leftrightarrow \sum_{j=1}^{n} \frac{x_j - \alpha}{((\nu+1)^2\beta^2 + (x_j-\alpha)^2)^2} = 0, \tag{2.7}$$

$$\frac{\partial I}{\partial \beta} = 0 \Leftrightarrow \frac{1}{n}\sum_{j,k=1}^{n} \frac{\nu(x_j - x_k)^2}{(\nu^2\beta^2 + (x_j-x_k)^2)^2} - \sum_{j=1}^{n} \frac{2(1+\nu)(x_j-\alpha)^2}{((1+\nu)^2\beta^2 + (x_j-\alpha)^2)^2} = 0. \tag{2.8}$$

Although these estimating equations are somewhat more complicated than the likelihood equations in (2.1) and (2.2), we can employ standard theory of U-statistics to study the asymptotic behavior of the estimating equations. We could not establish unimodality of $I(\alpha, \beta)$, but in our experiences the estimating equations can be solved numerically if an appropriate initial value is chosen and apparently produced a unique solution.

The test statistics $D_{n,\kappa}$ has yet another alternative representation, which is useful for obtaining its asymptotic distribution. We have

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

where

$$\hat{Z}_n(t) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\{\cos(tx_j) + \sin(tx_j) - e^{-\hat{\beta}_n|t|}(\cos(t\hat{\alpha}_n) + \sin(t\hat{\alpha}_n))\}. \tag{2.9}$$

$\hat{Z}_n(t)$ corresponds to the empirical characteristic process. We use the Fréchet space $C(\boldsymbol{R})$ of continuous functions on the real line $\boldsymbol{R}$ for considering the random processes. The metric of $C(\boldsymbol{R})$ is given by

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

where $\rho_j(x,y) = \max_{|t|\le j} |x(t) - y(t)|$.

In the rest of this paper we use the following notations. $\xrightarrow{D}$ means weak convergence of random variables or stochastic processes, $\xrightarrow{P}$ means convergence in probability and i.i.d. means "independently and identically distributed" as usual.

Now we state results on weak convergence of $\hat{Z}_n(t)$ and weak convergence of the test statistics $D_{n,\kappa}$ in the following two theorems. Note that because our tests are affine invariant, we can assume without loss of generality that $X_1,\dots,X_n$ is a random sample from $C(0,1)$.

THEOREM 2.1. *Let $X_1,\dots,X_n$ be i.i.d. $C(0,1)$ random variables and let $\hat{Z}_n$ be defined in* (2.9). *Then $\hat{Z}_n \xrightarrow{D} Z$ in $C(\boldsymbol{R})$, where Z is a zero mean Gaussian process with covariance functions given below for the MLE and the EISE, respectively.*

$$\text{(2.10)} \qquad MLE: \qquad \Gamma(s,t) = e^{-|t-s|} - \{1 + 2(st + |st|)\}e^{-|s|-|t|},$$

$$\begin{aligned}\text{(2.11)} \qquad EISE: \qquad \Gamma(s,t) = {} & e^{-|t-s|} - e^{-|s|-|t|} + M_1(st + |st|)e^{-|s|-|t|} \\ & - M_2\{(t\cdot \operatorname{sgn} s + |t|)(1 - e^{-\nu|s|}) \\ & \qquad + (s\cdot \operatorname{sgn} t + |s|)(1 - e^{-\nu|t|})\}e^{-|s|-|t|} \\ & + M_3(e^{-\nu|s|} + e^{-\nu|t|})(st + |st|)e^{-|s|-|t|},\end{aligned}$$

where

$$M_1 = \frac{(\nu+2)^2(5\nu^2+14\nu+10)}{16(\nu+1)^3}, \qquad M_2 = \frac{(\nu+1)(\nu+2)}{\nu^2}, \qquad M_3 = \frac{(\nu+2)^2}{2\nu},$$

and ν is the weighting constant in (2.4).

These asymptotic covariance functions do not involve definite integrals as was the case of the median and the interquartile range in Gürtler and Henze (2000). In particular the case of the MLE is very simple.

Note that for both cases $\Gamma(s,t)$ is symmetric with respect to the origin and $\Gamma(s,t) = 0$ for s,t such that $st < 0$. This implies that $\{Z(t) \mid t > 0\}$ and $\{Z(-t) \mid t > 0\}$ are independently and identically distributed for both cases.

THEOREM 2.2. *Under the conditions of Theorem* 2.1

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

By Fubini, the exact expectation of D_κ can be evaluated as

$$\mathrm{E}(D_\kappa) = \int_{-\infty}^{\infty} \mathrm{E}(Z(t)^2)e^{-\kappa|t|}dt = \int_{-\infty}^{\infty} \Gamma(t,t)e^{-\kappa|t|}dt.$$

Substituting (2.10) and (2.11) we obtain $\mathrm{E}(D_\kappa)$ for the case of the MLE and the EISE as

$$\text{(2.12)} \qquad \text{MLE}: \qquad \mathrm{E}(D_\kappa) = \frac{4}{\kappa(\kappa+2)} - \frac{16}{(\kappa+2)^3},$$

$$\text{(2.13)} \qquad \text{EISE}: \qquad \mathrm{E}(D_\kappa) = \frac{4}{\kappa(\kappa+2)} + \frac{8M_1}{(\kappa+2)^3} - \frac{8M_2}{(\kappa+2)^2} + \frac{8M_2}{(\kappa+\nu+2)^2} + \frac{16M_3}{(\kappa+\nu+2)^3}.$$

These exact expectations of D_κ for both cases are used as numerical checks in approximating the eigenvalues of the covariance function $\Gamma(s,t)e^{-\kappa(|s|+|t|)/2}$ in Subsection 3.3.

We now consider the consistency of our testing procedure corresponding to Theorem 2.3 of Gürtler and Henze (2000). Let $g(x)$ denote the true density of the observations $x_1, \dots, x_n$, not belonging to the Cauchy family. Concerning the MLE, define $(\alpha^*, \beta^*) = (\alpha^*_{MLE}, \beta^*_{MLE})$ by

$$\int_{-\infty}^{\infty} \log f(x;\alpha^*,\beta^*) g(x) dx = \max_{\alpha,\beta} \int_{-\infty}^{\infty} \log f(x;\alpha,\beta) g(x) dx,$$

where $f(x;\alpha,\beta)$ denotes the Cauchy density. Concerning the EISE, let $\Phi_g(t) = \int_{-\infty}^{\infty} e^{itx} g(x) dx$ and $\Phi_g(t;\alpha,\beta) = \int_{-\infty}^{\infty} e^{it(x-\alpha)/\beta} g(x) dx$. Define $(\alpha^*, \beta^*) = (\alpha^*_{EISE}, \beta^*_{EISE})$ to be the minimizer of the squared distance:

$$\int_{-\infty}^{\infty} |\Phi_g(t;\alpha^*,\beta^*) - e^{-|t|}|^2 w(t) dt = \min_{\alpha,\beta} \int_{-\infty}^{\infty} |\Phi_g(t;\alpha,\beta) - e^{-|t|}|^2 w(t) dt.$$

Then we have the following theorem.

THEOREM 2.3. *Assume that for the MLE and the EISE, (α^*, β^*) is uniquely determined and our estimator $(\hat{\alpha}_n, \hat{\beta}_n)$ converges to (α^*, β^*) almost surely for respective cases. Then the proposed tests that reject H_0 if $D_{n,\kappa}$ exceeds the upper ξ percentage points of the null distribution is consistent against the alternative $g(x)$.*

2.2 *Approximation of the asymptotic critical values of the proposed test statistics*

In this section we investigate the distribution of D_κ. We omit some technical details which are found in the preprint version of this paper (Matsui and Takemura (2003)) available from the authors upon request. Let $\Gamma(s,t)$ denote the asymptotic covariance functions of $Z(t)$ in Theorem 2.1 for the MLE (2.10) or the EISE (2.11). In order to evaluate the distribution of D_κ we want to evaluate the eigenvalues of the covariance kernel $\Gamma(s,t)e^{-\kappa(|s|+|t|)/2}$. We transform this kernel on $\boldsymbol{R}^2$ to $K(s,t)$ on $[-1,1]^2$ by a change of variables

$$(s,t) \mapsto (\operatorname{sgn} s \cdot (1-e^{-|s|}), \operatorname{sgn} t \cdot (1-e^{-|t|})),$$

because integral equations on $\boldsymbol{R}^2$ are theoretically difficult to treat and the eigenvalues of the covariance kernel are invariant for this transformation.

By a version of Mercer's theorem (Anderson and Darling (1952), Hammerstein (1927), Theorem 2.4 of Matsui and Takemura (2003)) D_κ has a series representation $\sum_{j=1}^{\infty} \frac{1}{\lambda_j} N_j^2$, where $N_1, N_2, \dots$ are i.i.d. standard normal variables and the eigenvalues $(\lambda_j)_{j\ge 1}$ satisfy the following integral equations

$$\text{(2.14)} \qquad \lambda \int_{-1}^{1} K(s,t) f(t) dt = f(s).$$

Although we could not solve the above integral equation and evaluate the eigenvalues λ explicitly, we numerically approximate the eigenvalues $(\lambda_j)_{j\geq 1}$ in Subsection 3.3. The characteristic function of D_κ is given as

$$\mathrm{E}(e^{itD_\kappa}) = \mathrm{E}\left[\exp\left(it\sum_{j=1}^{\infty} N_j^2/\lambda_j\right)\right] = \prod_{j=1}^{\infty}\mathrm{E}[\exp(itN_j^2/\lambda_j)] = \prod_{j=1}^{\infty}(1-2it/\lambda_j)^{-1/2}.$$

Note that $K(s,t)$ is symmetric with respect to the origin and $K(s,t)=0$ for $st<0$ as discussed in Subsection 2.1. This implies that in our problem the eigenvalues appear in pairs and we do not need the square root in the characteristic function. More precisely the characteristic function of D_κ is given by

$$\phi(t) = \mathrm{E}(e^{itD_\kappa}) = \frac{1}{D(2it)},$$

where $D(\lambda) = \prod_{j=1}^{\infty}(1-\lambda/\lambda_j)$ is the associated Fredholm determinant of our kernel $K(s,t)$ restricted to $[0,1]^2$.

Inverting the characteristic function gives the distribution function and the probability density function of the proposed statistics theoretically. However, practically there are two problems in treating the characteristic function in the form of Fredholm determinant. One is in the approximation of $D(\lambda)$ itself and the other is in the Lévy's inversion formula. We treat the former problem in Subsection 3.3.

Concerning the latter problem we use the residue theorem. The residue calculation is given in Appendix B. The reason we prefer the residue theorem over the direct numerical inversion of approximated characteristic function is that the characteristic function oscillates between negative and positive values and the convergence of the integral is very slow (see Section 6.1 of Tanaka (1996)). Assuming that the kernel $K(s,t)$ restricted to $[0,1]^2$ has only single eigenvalues, the corresponding density function and distribution function are calculated as

$$f_{D_\kappa}(y) = \sum_{j=1}^{\infty}\frac{\lambda_j}{2}\frac{\exp\left(-\dfrac{\lambda_j}{2}y\right)}{\prod_{k\neq j}^{\infty}\left(1-\dfrac{\lambda_j}{\lambda_k}\right)}, \tag{2.15}$$

$$F_{D_\kappa}(y) = 1-\sum_{j=1}^{\infty}\frac{\exp\left(-\dfrac{\lambda_j}{2}y\right)}{\prod_{k\neq j}^{\infty}\left(1-\dfrac{\lambda_j}{\lambda_k}\right)}. \tag{2.16}$$

Note that the series on the right hand side is alternating and we can bound $f_{D_\kappa}(y)$ and $F_{D_\kappa}(y)$ relatively easily.

3. Computational studies

In this section, some simulation results are given. Since the exact finite sample distributions are difficult to obtain, first we approximate the percentage points of $D_{n,\kappa}$ by Monte Carlo simulation. Then the power of both tests for the finite sample is evaluated. In the end of this section percentage points of D_κ are computed by the residue theorem.

Table 1. MLE: Upper 10 percentage points of $D_{n,\kappa}$.

$n \setminus \kappa$	0.1	0.5	1.0	2.5	5.0	10.0
10	21.23	3.054	1.111	0.307	0.127	0.0498
20	21.60	3.093	1.103	0.292	0.118	0.0451
50	21.83	3.119	1.105	0.287	0.115	0.0432
100	21.90	3.131	1.108	0.285	0.115	0.0430
200	21.89	3.136	1.111	0.286	0.114	0.0426

Table 2. MLE: Upper 5 percentage points of $D_{n,\kappa}$.

$n \setminus \kappa$	0.1	0.5	1.0	2.5	5.0	10.0
10	22.95	3.406	1.271	0.373	0.160	0.0675
20	23.31	3.481	1.263	0.349	0.147	0.0592
50	23.59	3.514	1.268	0.338	0.141	0.0543
100	23.66	3.546	1.271	0.337	0.138	0.0533
200	23.65	3.555	1.275	0.335	0.137	0.0524

Table 3. EISE: Upper 10 percentage points of $D_{n,\kappa}$.

$n \setminus \kappa$	0.1	0.5	1.0	2.5	5.0	10.0
10	19.88	2.806	0.992	0.209	0.0541	0.0109
20	20.74	2.916	1.040	0.228	0.0645	0.0154
50	21.22	3.006	1.078	0.241	0.0709	0.0192
100	21.39	3.026	1.085	0.244	0.0727	0.0202
200	21.45	3.040	1.089	0.245	0.0736	0.0209

Table 4. EISE: Upper 5 percentage points of $D_{n,\kappa}$.

$n \setminus \kappa$	0.1	0.5	1.0	2.5	5.0	10.0
10	20.95	3.103	1.118	0.234	0.0585	0.0113
20	22.20	3.277	1.182	0.264	0.0754	0.0180
50	22.85	3.394	1.231	0.281	0.0836	0.0226
100	23.07	3.416	1.241	0.285	0.0858	0.0240
200	23.14	3.438	1.251	0.286	0.0870	0.0247

For the MLE, the estimates are easily found by the Newton method. Convenient initial values are suggested by Copas (1975). But for the EISE the simple Newton method does notwork well and we often need grid search of initial values. In the case of the EISE we use the same initial value as the MLE for the Newton method. If it fails to converge we do grid search of initial values and we obtain the parameter value which minimizes (2.3) among, say, 20000 points, and use it as the initial value of the Newton method. When the values of the estimators have converged, we can compute $D_{n,\kappa}$ by (1.3). Based on 100,000 Monte Carlo replications, the upper 10 and 5 percentage points of the statistics $D_{n,\kappa}$, $\kappa \in \{0.1, 0.5, 1.0, 2.5, 5.0, 10.0\}$ are tabulated in Table 1, Table 2 for the MLE and Table 3, Table 4 for the EISE. In the preprint version of this paper (Matsui and Takemura (2003)) we also gave critical values for some classical procedures (i.e., Anderson-Darling, Cramér-von Mises, Kolmogorov-Smirnov and Watson statistics) based on the empirical distribution function of the Cauchy distribution when parameters are estimated by the MLE.

3.1 *Alternative hypotheses*

For studying the power functions of various tests considered, we use the following family of distributions containing the Cauchy distribution as a special case.

1. $t(j)$. Student's t distribution with j degrees of freedom for $j = 1, 2, 3, 4, 5, 10, \infty$. Note $t(1) = C(1, 0)$ and $t(\infty) = N(0, 1)$.
2. $st(a, b)$. Stable distributions with the characteristic function

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

Here we only consider symmetric stable distributions ($b = 0$). The characteristic exponent $a \in (0, 2]$ concerns the tail behavior of the distribution. Note that $st(1, 0) = C(0, 1)$ and $st(2, 0) = N(0, 1)$.

3.2 *Analysis of finite sample power*

For the significance levels $\xi = 0.1$, 0.05, finite sample power of the tests were studied using sample sizes $n = 50$, 100, 200 for estimations of the MLE and using sample sizes $n = 50$, 100 for the EISE based on 10,000 Monte Carlo replications. The case of the EISE with the sample size $n = 200$ was computationally very heavy due to frequent need of grid search of initial values under various alternative hypotheses. In this paper we only give results of the case $n = 50$ in Tables 5 and 6, since the results for other sample sizes were qualitatively similar to the case of $n = 50$ (see Matsui and Takemura (2003) for detail). In these tables '*' stands for 100, i.e. the power of 100%.

In comparison to the test proposed by Gürtler and Henze (2000) for the nominal level of 10%, we derived the following main conclusions. In the case of the Cauchy distribution, the total power of a goodness-of-fit test based on the empirical characteristic function is not affected very much by the choice of the estimators of (α, β).

Some tendency concerning the weight κ can be observed for each estimator. The test by the MLE tends to have more power than the other two tests when the weight κ is small (e.g. $\kappa = 0.1$, 0.5, 1.0) for most alternatives and is most powerful for $st(0.5, 0)$ regardless of the weight. Although the test by the EISE has more power than the other two tests when the weight is around $\kappa = 2.5$ for most cases, it has the least power when the weight is large (e.g. $\kappa = 10.0$), except for the case of $N(0, 1)$. The test proposed by Gürtler and Henze has more power than the other two tests when the weight is large (e.g. $\kappa = 10.0$) for most alternatives and tends to have less power for other weights.

We also see that our tests compare well to more traditional tests based on the empirical distribution function tabulated in Gürtler and Henze (2000).

For future research, it is worth investigating the problem of choosing weights κ and ν depending on alternatives in order to maximize power.

3.3 *Approximation of D_κ*

Because we can not obtain the exact distribution of D_κ as a result of difficulty in solving integral equation, we have to approximate the distribution of D_κ. This can be

Table 5. MLE: Power of $D_{n,\kappa}$ (significance levels $\xi = 0.1$, 0.05, $n = 50$).

	0.1						0.05					
ξ / κ	0.1	0.5	1.0	2.5	5.0	10.0	0.1	0.5	1.0	2.5	5.0	10.0
$C(0,1)$	10	10	10	11	10	10	5	5	5	6	5	5
$N(0,1)$	34	72	87	96	98	98	22	60	78	90	92	86
$t(2)$	15	22	25	26	27	24	8	13	16	15	14	9
$t(3)$	19	34	43	50	55	52	10	24	30	34	36	25
$t(4)$	22	43	54	64	71	70	13	30	42	49	52	40
$t(5)$	24	48	62	73	80	80	14	35	49	58	62	50
$t(10)$	29	60	75	87	92	92	17	47	63	77	80	71
$st(0.5,0)$	75	90	94	97	98	98	64	84	89	94	96	97
$st(0.8,0)$	14	21	27	35	40	42	7	13	17	24	29	31
$st(0.9,0)$	10	12	15	18	19	21	5	7	8	10	12	13
$st(1.1,0)$	6	8	9	10	11	10	2	3	4	5	6	6
$st(1.2,0)$	8	11	12	14	15	12	2	4	6	8	8	6
$st(1.5,0)$	20	34	40	42	44	34	11	23	28	28	26	15

Table 6. EISE: Power of $D_{n,\kappa}$ (significance levels $\xi = 0.1$, 0.05, $n = 50$).

	0.1						0.05					
ξ / κ	0.1	0.5	1.0	2.5	5.0	10.0	0.1	0.5	1.0	2.5	5.0	10.0
$C(0,1)$	10	10	10	10	10	10	5	5	5	5	5	5
$N(0,1)$	30	66	86	98	*	*	20	53	76	95	97	92
$t(2)$	13	19	24	30	29	19	8	11	15	19	16	8
$t(3)$	17	30	41	56	58	43	10	20	30	41	40	22
$t(4)$	19	37	52	71	74	62	11	25	40	57	58	36
$t(5)$	21	42	60	80	84	74	13	30	47	67	69	48
$t(10)$	24	53	73	91	95	91	15	40	62	84	88	73
$st(0.5,0)$	50	88	94	93	87	72	32	80	89	88	79	60
$st(0.8,0)$	12	21	27	31	28	25	6	12	17	20	18	15
$st(0.9,0)$	10	12	14	16	16	15	5	6	8	9	9	8
$st(1.1,0)$	7	8	9	10	10	10	3	3	4	5	5	6
$st(1.2,0)$	7	10	13	13	13	12	3	5	7	7	7	6
$st(1.5,0)$	18	29	39	47	41	24	10	19	28	33	26	11

done by numerically approximating the eigenvalues of the integral equation. Furthermore we have to approximate the infinite sum and infinite products in the series (2.16) by a finite sum and finite products.

First we discuss numerical approximation of the eigenvalues of kernel $K(s,t)$. $N \times N$ regularly placed points are taken from $(0,1) \times (0,1)$ to approximate $K(s,t)$ by an $N \times N$ matrix $\tilde{K}$. By standard argument of the theory of integral equations, we evaluate N

eigenvalues of $\tilde{K}$ and define

$$\tilde{D}_N(\lambda) = \left| I - \frac{\lambda}{N}\tilde{K} \right| = \prod_{j=1}^{N} \left(1 - \frac{\lambda}{\tilde{\lambda}_j}\right), \qquad 0 < \tilde{\lambda}_1 \le \cdots \le \tilde{\lambda}_N,$$

where $1/\tilde{\lambda}_j = 1/\tilde{\lambda}_j(N)$ are the corresponding eigenvalues of $\tilde{K}/N$. Note that as $N \to \infty$, $\tilde{D}_N(\lambda)$ converges to the associated Fredholm determinant $D(\lambda)$ which is the characteristic function of D_κ in our case. We found that we can easily evaluate $N = 500$ eigenvalues by this simple method except for the case $\kappa = 0.1$. For the case of $\kappa = 0.1$, we had some numerical difficulty and the approximated sum of $2\sum_{j=1}^{\infty} 1/\lambda_j$ did not converge to $\mathrm{E}[D_\kappa]$ quickly. Therefore we omit the case $\kappa = 0.1$ and present results in Tables 7 and 8 for $\kappa \in \{0.5, 1.0, 2.5, 5.0, 10.0\}$. Note that the powers of both tests are the lowest for the case of $\kappa = 0.1$ and we do not recommend using $\kappa = 0.1$. We remark that in the literature on numerical treatment of integral equations (e.g. Baker (1977)), many other approximations of the eigenvalues are considered.

Secondly, the infinite sum and the infinite products in (2.16) have to be approximated by a finite sum and finite products. Let l and m ($l < m$) denote the number of terms in the sum and the products, respectively. Then $F_{D_\kappa}(y)$ in (2.16) is approximated as

$$F_{D_\kappa}(y) \approx 1 - \sum_{j=1}^{l} \frac{\exp\left(-\frac{\lambda_j}{2}y\right)}{\prod_{k\neq j}^{m}\left(1 - \frac{\lambda_j}{\lambda_k}\right)}.$$

For the remaining part of the product $1/\prod_{m+1}^{\infty}(1 - \frac{\lambda_j}{\lambda_k})$, we can give bounds by using the relation $1/(1-x) = \exp(\sum_{n=1}^{\infty} x^n/n)$. A lower bound of the j-th term of the series is given by

$$\frac{\exp\left(-\frac{\lambda_j}{2}y\right)}{\prod_{k\neq j}^{m}\left(1 - \frac{\lambda_j}{\lambda_k}\right)} \exp\left(\lambda_j \sum_{k=m+1}^{\infty} \frac{1}{\lambda_k}\right),$$

and an upper bound is given by

$$\frac{\exp\left(-\frac{\lambda_j}{2}y\right)}{\prod_{k\neq j}^{m}\left(1 - \frac{\lambda_j}{\lambda_k}\right)} \exp\left(\lambda_j \left\{1 + \frac{\lambda_j}{2(\lambda_{m+1} - \lambda_j)}\right\} \sum_{k=m+1}^{\infty} \frac{1}{\lambda_k}\right).$$

For evaluating $\sum_{k=m+1}^{\infty} 1/\lambda_k$, we can utilize the expected value $\mathrm{E}(D_\kappa) = 2\sum_{k=1}^{\infty} 1/\lambda_k$. The sum $\sum_{k=1}^{m} 1/\lambda_k$ is evaluated numerically by approximating the first m eigenvalues as $\sum_{k=1}^{m} 1/\lambda_k \approx \sum_{k=1}^{m} 1/\tilde{\lambda}_k$ and then $\mathrm{E}(D_\kappa)/2 - \sum_{k=1}^{m} 1/\tilde{\lambda}_k$ approximates $\sum_{k=m+1}^{\infty} 1/\lambda_k$.

Note that the series is alternating. Therefore the range of the critical values can be obtained by substituting the above bounds for positive terms and negative terms separately, i.e., by substituting the upper bound for positive terms and the lower bound for negative terms, or vice versa. The ranges for the percentage points in Tables 7 and 8 are less than 0.01% of each value for any κ if l and m are $l > 50$ and $m > 100$. Therefore accuracy of approximation of the infinite sum and infinite product is very good.

Table 7. MLE: Upper ξ percentage points of D_κ.

$\xi\backslash\kappa$	0.5	1.0	2.5	5.0	10.0
0.1	3.153	1.111	0.286	0.114	0.0431
0.05	3.571	1.276	0.336	0.137	0.0527

Table 8. EISE: Upper ξ percentage points of D_κ.

$\xi\backslash\kappa$	0.5	1.0	2.5	5.0	10.0
0.1	3.057	1.093	0.248	0.0750	0.0213
0.05	3.458	1.256	0.290	0.0886	0.0254

The difficulty lies more in the approximation of the eigenvalues. In Matsui and Takemura (2003) the values of $\mathrm{E}(D_\kappa)/2 - \sum_{k=1}^{m} 1/\tilde{\lambda}_k$ are given. This error is less than 1% of true value $\mathrm{E}(D_\kappa)/2$ if l and m are $l > 50$ and $m > 100$. Therefore the true sum of eigenvalues $\mathrm{E}(D_\kappa)/2$ is well approximated by the sum of numerically obtained eigenvalues.

Note that more accurate approximations may be obtained if we evaluate higher order moments of D_κ.

Acknowledgements

The authors wish to thank the referees for constructive comments and detailed stylistic corrections.

Appendix

A. *Proofs of the results in Subsection* 2.1

Proofs of the theorems are essentially the same as those of Gürtler and Henze (2000) based on Csörgő (1983). Therefore we state only the differences in our case from Gürtler and Henze (2000) along their framework.

Before considering Fréchet space $C(\boldsymbol{R})$, we first consider the restricted space $C(S)$ of continuous functions on a compact subset S with the supremum norm $\|f\|_\infty = \sup_{t\in S}|f(t)|$ so that we can utilize arguments of Csörgő (1983). Define $k(x,t) = \cos(tx) + \sin(tx)$. Then an alternative representation of $\hat{Z}_n(t)$ is given by

$$\begin{aligned}\hat{Z}_n(t) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\{\cos(tX_j) + \sin(tX_j) - e^{-\hat{\beta}_n|t|}(\cos(t\hat{\alpha}_n) + \sin(t\hat{\alpha}_n))\} \\ &= \int k(x,t)d\{\sqrt{n}(F_n(x) - F(x,\hat{\theta}_n))\}.\end{aligned}$$

By checking the conditions of Csörgő (1983), the weak convergence of $\hat{Z}_n(t)$ to a zero mean Gaussian process Z is proved in the space $(C(S), \|\cdot\|_\infty)$. Since the compact set S is arbitrary, the space $(C(s), \|\cdot\|_\infty)$ can be extended to Fréchet space $C(\boldsymbol{R})$ easily.

Let $\langle\cdot,\cdot\rangle$ denote the standard inner product and let Z have the general covariance function obtained by Gürtler and Henze (2000)

$$\Gamma(s,t) = \tilde{K}_0(s,t) - K_0(s)K_0(t) + H(s,\theta_0)'\mathrm{E}[l(X_1)l(X_1)']H(t,\theta_0) \tag{A.1}$$

$$-\left\langle H(t,\theta_0), \int k(x,s)l(x)dF_0(x)\right\rangle$$
$$-\left\langle H(s,\theta_0), \int k(x,t)l(x)dF_0(x)\right\rangle.$$

We only need to evaluate terms of (A.1) i.e., $\mathrm{E}[l(X_1)l(X_1)']$ and $\int k(x,s)l(x)dF_0(x)$.

The first step is to obtain an asymptotically linear representation of an estimation error

$$\sqrt{n}\hat{\alpha}_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} l_1(X_j) + r_{1n}, \qquad \sqrt{n}(\hat{\beta}_n - 1) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} l_2(X_j) + r_{2n}, \qquad r_{1n}, r_{2n} \xrightarrow{P} 0,$$

for the standard Cauchy case $C(0,1)$.

LEMMA A.1. *The asymptotically linear representation* $l(x) = (l_1(x), l_2(x))$ *and their covariance matrices for the MLE and for the EISE are given by*

1. *MLE*

$$l_1(x) = \frac{4x}{1+x^2}, \qquad l_2(x) = \frac{2(x^2-1)}{1+x^2},$$
$$\mathrm{E}[l(X)l(X)'] = 2 \times I_2.$$

2. *EISE*

$$l_1(x) = (\nu+1)(\nu+2)^3\frac{x}{((\nu+1)^2+x^2)^2},$$
$$l_2(x) = \frac{1}{2}(\nu+2) - \frac{1}{2}(\nu+2)^3\frac{(\nu+1)^2-x^2}{((\nu+1)^2+x^2)^2},$$
$$\mathrm{E}[l(X)l(X)'] = \frac{(\nu+2)^2(5\nu^2+14\nu+10)}{16(\nu+1)^3} \times I_2,$$

where ν *is the weighting constant in* (2.4) *and* I_2 *is the* 2×2 *identity matrix.*

PROOF. In the case of the MLE, l_1 and l_2 are easily obtained from the score functions and the Fisher information matrix.

For the EISE we apply the delta method to the estimating equations (2.7), (2.8) for the case of $C(\alpha,\beta)$ and obtain

$$\sqrt{n}(\hat{\alpha}_n - \alpha) = \frac{1}{\sqrt{n}}\frac{\sum_{j=1}^n g_1(X_j)}{\sum_{j=1}^n g_2(X_j)/n} + r_{1n}, \qquad r_{1n} \xrightarrow{P} 0,$$

$$\sqrt{n}(\hat{\beta}_n - \beta) = \frac{\sqrt{n}}{4\beta}\frac{\frac{1}{n(n-1)}\sum_{j,k=1}^n h_1(X_j,X_k) - \frac{1}{n-1}\sum_{j=1}^n 2h_2(X_j)}{\frac{1}{n(n-1)}\sum_{j,k=1}^n h_3(X_j,X_k) - \frac{1}{n-1}\sum_{j=1}^n 2h_4(X_j)} \tag{A.2}$$
$$+ r_{2n}, \qquad r_{2n} \xrightarrow{P} 0,$$

where

$$g_1(x) = -\frac{(x-\alpha)}{((x-\alpha)^2+(\nu+1)^2\beta^2)^2}, \qquad g_2(x) = \frac{3(x-\alpha)^2-(\nu+1)^2\beta^2}{((x-\alpha)^2+(\nu+1)^2\beta^2)^3},$$

$$h_1(x_1,x_2) = \frac{\nu(x_1-x_2)^2}{((x_1-x_2)^2+\nu^2\beta^2)^2}, \qquad h_3(x_1,x_2) = \frac{\nu^3(x_1-x_2)^2}{((x_1-x_2)^2+\nu^2\beta^2)^3},$$

$$h_2(x) = \frac{(\nu+1)(x-\alpha)^2}{((x-\alpha)^2+(\nu+1)^2\beta^2)^2}, \qquad h_4(x) = \frac{(\nu+1)^3(x-\alpha)^2}{((x-\alpha)^2+(\nu+1)^2\beta^2)^3}.$$

Returning to the standard case $(\alpha,\beta)=(0,1)$, the numerator of $\sqrt{n}(\hat{\beta}_n-1)$ in (A.2) can be expressed in the form of a U-statistic

$$\sqrt{n}\left\{U_n - \frac{1}{n(n-1)}\sum_{j=1}^n 2h_2(X_j)\right\} = \sqrt{n}U_n + r_{3n}, \qquad r_{3n} \xrightarrow{P} 0,$$

where

$$U_n = \binom{n}{2}^{-1}\sum_{1\le j<k\le n} h(X_j,X_k) = \frac{2}{n(n-1)}\sum_{1\le j<k\le n}\{h_1(X_j,X_k)-h_2(X_j)-h_2(X_k)\}.$$

By standard arguments on U-statistics (Chapter 3 of Maesono (2001), Chapter 5 of Serfling (1980)) we only need to evaluate

$$a(x_1) = \mathrm{E}[h(X_1,X_2) \mid X_1 = x_1],$$

since

$$\sqrt{n}U_n = \frac{1}{\sqrt{n}}\sum_{j=1}^n 2a(X_j) + r_{4n}, \qquad r_{4n} \xrightarrow{P} 0.$$

It can be shown that $a(x_1)$ is written as

$$a(x_1) = \frac{1}{2}\frac{(\nu+1)^2-x_1^2}{(x_1^2+(\nu+1)^2)^2} - \frac{1}{2(\nu+2)^2}.$$

The denominators of $\sqrt{n}\hat{\alpha}_n$ and $\sqrt{n}(\hat{\beta}_n-1)$ converge in probability to their expectations or limiting expectations

$$\mathrm{E}\left[\sum_{j=1}^n g_2(X_j)/n\right] = -\frac{1}{(\nu+1)(\nu+2)^3},$$

$$\lim_{n\to\infty}\mathrm{E}\left[\frac{4}{n(n-1)}\sum_{j,k=1}^n h_3(X_j,X_k) - \frac{8}{n-1}\sum_{j=1}^n h_4(X_j)\right]$$
$$= \lim_{n\to\infty}\left[\frac{3\nu+2}{(\nu+2)^3} - \frac{n}{n-1}\frac{3\nu+4}{(\nu+2)^3}\right] = -\frac{2}{(\nu+2)^3}.$$

Thus the asymptotically linear representation of the EISE is given by

$$\sqrt{n}\hat{\alpha}_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}(\nu+1)(\nu+2)^3\frac{X_j}{(X_j^2+(\nu+1)^2)^2} + r_{1n},$$

$$\sqrt{n}(\hat{\beta}_n - 1) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left\{\frac{(\nu+2)^3}{2}\frac{{X_j}^2-(\nu+1)^2}{\left({X_j}^2+(\nu+1)^2\right)^2} + \frac{(\nu+2)}{2}\right\} + r_{2n}. \qquad \square$$

Note that the covariance matrix $\mathrm{E}[l(X)l(X)']$ for both the MLE and the EISE is finite and positive definite. Therefore condition (iv) of Csörgő (1983) is satisfied. Since l_1 and l_2 are bounded and differentiable, condition (v) of Csörgő (1983) is satisfied.

Write $F_0(x) = F(x, \theta_0)$ for simplicity. The kernel transform of the asymptotically linear representation are given as follows.

1. MLE

$$\text{(A.3)} \qquad \begin{aligned} \int k(x,s)l_1(x)dF_0(x) &= 2se^{-|s|}. \\ \int k(x,s)l_2(x)dF_0(x) &= -2|s|e^{-|s|}. \end{aligned}$$

2. EISE

$$\text{(A.4)} \qquad \begin{aligned} \int k(x,s)l_1(x)dF_0(x) &= \frac{(\nu+1)(\nu+2)}{\nu^2}\operatorname{sgn} s(e^{-|s|} - e^{-(\nu+1)|s|}) \\ &\quad - \frac{(\nu+2)^2}{2\nu}se^{-(\nu+1)|s|}. \\ \int k(x,s)l_2(x)dF_0(x) &= \frac{(\nu+1)(\nu+2)}{\nu^2}(e^{-(\nu+1)|s|} - e^{-|s|}) \\ &\quad + \frac{(\nu+2)^2}{2\nu}|s|e^{-(\nu+1)|s|}. \end{aligned}$$

Evaluating (A.1) for the case of the MLE and the EISE using (A.3) and (A.4) proves Theorem 2.1.

We here remark a relation between the asymptotically linear representation l^M of the MLE and the asymptotically linear representation l^I of the EISE. From the asymptotic efficiency of the MLE it follows that l^M and $l^I - l^M$ are orthogonal, i.e.,

$$\mathrm{E}[l^M(l^I - l^M)] = 0.$$

The proofs of Theorems 2.2 and 2.3 are clear from detailed proofs of Gürtler and Henze (2000) and we omit them.

B. *Calculation of residues for the results in Subsection* 2.2

Here we derive the density function and the distribution function in (2.15) and (2.16). By the inversion formula

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-ixt}}{\prod_{j=1}^{\infty}\left(1-\dfrac{2it}{\lambda_j}\right)}dt.$$

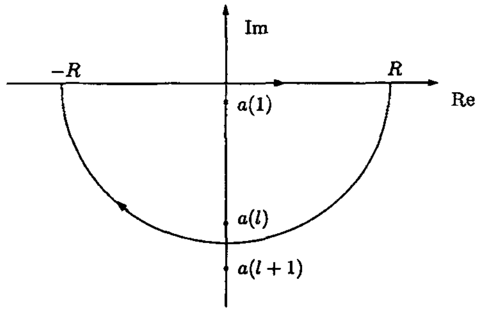

Fig. 1. Complex integration.

We consider complex integration

$$\frac{1}{2\pi}\oint_{-C}\frac{e^{-ixt}}{\prod_{j=1}^{\infty}\left(1-\dfrac{2it}{\lambda_j}\right)}dt,$$

where a closed curve C consists of a semicircle C_R with radius R and a line segment $[-R,R]$ and $-C$ or $-C_R$ means the clockwise direction. See Fig. 1.

In the region $D=\{re^{-i\theta}\mid 0<r<R, 0<\theta<\pi\}$, the integrand has singular points at $a_j=-i\lambda_j/2$. Except for these points, the integrand is regular and continuous. The residue theorem tells us

$$\begin{aligned}\frac{1}{2\pi}\oint_{-C}\frac{e^{-ixt}}{\prod_{j=1}^{\infty}\left(1-\dfrac{2it}{\lambda_j}\right)}dt &= -i\sum_{j=1}^{\infty}\mathrm{Res}_{t=a_j}\left[\frac{e^{-ixt}}{\prod_{j=1}^{\infty}\left(1-\dfrac{2it}{\lambda_j}\right)}\right]\\ &=\sum_{j=1}^{\infty}\frac{\lambda_j}{2}\frac{\exp\left(-\dfrac{\lambda_j}{2}y\right)}{\prod_{k\neq j}^{\infty}\left(1-\dfrac{\lambda_j}{\lambda_k}\right)}.\end{aligned}$$

In the integral on C_R we transform t by R and θ as $t=Re^{-i\theta}$. Then

$$\frac{1}{2\pi}\oint_{C_R}\frac{e^{-ixt}}{\prod_{j=1}^{\infty}\left(1-\dfrac{2it}{\lambda_j}\right)}dt=\frac{1}{2\pi}\int_0^{\pi}\frac{iR\exp\{-i(xRe^{-i\theta}+\theta)\}}{\prod_{j=1}^{\infty}\left(1-\dfrac{2iRe^{-i\theta}}{\lambda_j}\right)}d\theta\to 0,$$

as $R\to\infty$. Here we can take R to be the midpoint $(a_j+a_{j+1})/2$ of neighboring a_j's, so that the denominator of the integrand never vanishes. Although the integrand is a function of all the eigenvalues λ_j, the convergence to zero of the integral over C_R is easily justified if the entire function $\prod_{j=1}^{\infty}(1-2iu/\lambda_j)^{1/2}$ is of exponential order less than unity (see Slepian (1957)). It can be easily shown that $\prod_{j=1}^{\infty}(1-2iu/\lambda_j)^{1/2}$ is of exponential order less than unity based on the fact $\sum_j 1/\lambda_j<\infty$. In general $1/\lambda_j=O(1/j^2)$ as discussed in Section 4 of Anderson and Darling (1952).

References

Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes, *The Annals of Mathematical Statistics*, **23**, 193–212.

Baker, C. T. H. (1977). *The Numerical Treatment of Integral Equations*, Clarendon Press, Oxford.

Besbeas, P. and Morgan, B. (2001). Integrated squared error estimation of Cauchy parameters, *Statistics & Probability Letters*, **55**, 397–401.

Copas, J. B. (1975). On the unimodality of the likelihood for the Cauchy distribution, *Biometrika*, **62**, 701–704.

Csörgő, S. (1983). Kernel-transformed empirical process, *Journal of Multivariate Analysis*, **13**, 517–533.

Durbin, J. (1973*a*). *Distribution Theory for Tests Based on the Sample Distribution Function*, SIAM, Philadelphia.

Durbin, J. (1973*b*). Weak convergence of the sample distribution function when parameters are estimated, *The Annals of Statistics*, **1**, 279–290.

Gürtler, N. and Henze, N. (2000). Goodness-of-fit tests for the Cauchy distribution based on the empirical characteristic function, *Annals of the Institute of Statistical Mathematics*, **52**, 267–286.

Hammerstein, A. (1927). Über Entwicklungen gegebener Funktionen nach Eigenfunktionen von Randwertaufgaben, *Mathematische Zeitschrift*, **27**, 269–311.

Maesono, Y. (2001). *Toukeitekisuisoku no Zenkinriron*, Kyushu University Press, Fukuoka, Japan (in Japanese).

Matsui, M. and Takemura, A. (2003). Empirical characteristic function approach to goodness-of-fit tests for the Cauchy distribution with parameters estimated by MLE or EISE, Discussion Paper CIEJE 2003-CF-226, Center for International Research on the Japanese Economy, Faculty of Economics, University of Tokyo, available from `http://www.e.u-tokyo.ac.jp/cirje/research/dp/2003/list.htm`.

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

Slepian, D. (1957). Fluctuations of random noise power, *Bell System Technical Journal*, **37**, 163–184.

Tanaka, K. (1996). *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*, John Wiley, New York.