THE ENERGY GOODNESS-OF-FIT TEST FOR UNIVARIATE STABLE DISTRIBUTIONS

Guangyuan Yang

A Dissertation

Submitted to the Graduate College of Bowling Green
State University in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2012

Committee:

Maria L. Rizzo, Advisor

Kenneth J. Ryan,
Graduate Faculty Representative

James H. Albert

Craig L. Zirbel

ABSTRACT

Maria L. Rizzo, Advisor

The family of stable distributions is an important class of four-parameter continuous distributions. It has appealing properties such as being the only possible limit distribution of a suitably normalized sum of independent and identically distributed random variables. Therefore, it has wide applications in modeling distributions with heavy tails, such as the return of financial assets. However, there are also critics against using stable distribution in modeling financial assets such as stock and futures. It is very important to check the validity of the model assumption before making inferences based on the model.

Previous work has been done in the goodness-of-fit test for several special cases including normal distributions, Cauchy distributions and more generally, symmetric stable distributions. Classical goodness-of-fit methods such as the Kolmogorov-Smirnov test and the Anderson-Darling test are not able to handle the stable distributions directly because of the lack of closed-form probability density functions (PDF) and cumulative distribution functions (CDF). Since stable distributions can be fully characterized by their characteristic functions, goodness-of-fit tests based on the empirical characteristic function (ECF) have also been studied in recent years.

In this dissertation, a new goodness-of-fit test is proposed for general stable distributions based on the energy statistic, which is invariant under rigid motions. The test statistic is essentially a weighted L^2-norm of the distance between the empirical characteristic function and the hypothetical characteristic function of the null distribution, and it can also be expressed as a V-statistic with degenerate kernel. By asymptotic theory of degenerate kernel V-statistics, the test statistic converges in distribution to an infinite sum of weighted χ^2 random variables if the null hypothesis of stability is true. It can be proved that the test is consistent against a large class of alternatives. A relatively simple computation formula is derived for the test statistic, which involves numerical integration in general. Bootstrap

method and critical values based on the asymptotic distribution of the test statistic can be applied to implement the test.

The dissertation is organized as follows. In Chapter 1, the class of stable distributions and its properties are reviewed. In Chapter 2, existing methods of goodness-of-fit test for stable distributions will be discussed. In Chapter 3, theoretical properties of the test statistic, including the definition, computation issues and asymptotic results, are developed. In Chapter 4, simulation studies are presented to illustrate the empirical type I error and power of testing stable distributions against alternative distributions including stable distributions with different parameters and other interesting light-tailed and heavy-tailed distributions. Simulation results show that our test is sensitive in detecting the difference either in the center or extreme values in the tail. In Chapter 5, some basic work has been finished to study the asymptotic distribution of the energy statistic for testing Cauchy when parameters are estimated by maximum likelihood estimators (MLE).

ACKNOWLEDGEMENTS

I would like to thank my advisor, Professor Maria Rizzo, for her generous help and valuable advice throughout this research. I am also grateful to have her constant encouragement and support during my early career in a non-academic field. I also want to express my appreciation to other members of my committee, Professor Jim Albert, Professor Craig Zirbel and Professor Ken Ryan, for their valued time and advice.

I want to extend my gratitude to all my professors in the Department of Mathematics and Statistics, and the Department of Applied Statistics and Operations Research, for their guidance and help during my four-year study at Bowling Green.

Finally, I wish to express my deepest gratitude to my family for always being by my side. I wish to thank my parents for their never-ending love, support and care. My special thanks go to my parents-in-law for their selfless love and encouragement all the time. I thank my wife Xiao for her love and sacrifice.

Guangyuan Yang

Glen Allen, Virginia

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The stable distribution family is a parametric distribution family with four parameters, which has wide applications in modeling data from economics, physics, biology and other areas. An inevitable question after a statistical model is proposed is how well it fits the observed data set, and the question is generally addressed by a goodness-of-fit test. Although the term "goodness-of-fit" is also used in regression analysis, in this dissertation, we will focus on the fit of a distribution to a random sample, and to be more specific, the goodness-of-fit of stable distributions. The aim of a goodness-of-fit test is to measure the conformity of a sample of observations to a hypothesized distribution, or in other words, the discrepancy between them. For various types of goodness-of-fit tests proposed in the past, see [8, 9, 25, 52, 10]. A collection of goodness-of-fit tests applicable to stable distributions will be reviewed in Chapter 2.

The name "energy" was coined by G. J. Székely [53], by analogy with Newton's gravitational potential energy, who introduced the notion of energy statistics in the 1980s in several colloquium lectures given in Budapest, Hungary, in the Soviet Union, Germany (U. Dortmund and Technische U. of Munich), France (U. Pierre et Marie Curie), and in USA at MIT, Yale, and Columbia. A class of goodness-of-fit tests was proposed by Rizzo [44], based on the energy statistic, or E-statistic. Applications of energy statistics include testing multivariate normality [44, 55], testing for Pareto distribution [45], testing multivariate independence [58, 57], testing for equality of distributions [54], nonparametric extension of ANOVA [46], and cluster analysis [56]. The energy distance is essentially a weighted L^2-norm of the difference between the two characteristic functions, with a suitable weight function. Proof of the following fundamental inequality was published in [53, 44] for case $s = 1$ and in [56] for $0 < s < 1$.

Theorem 1.0.1 ***Energy Inequality***. *Suppose that X and Y are independent random vectors of the same dimension, X' denotes an independent copy of X, and Y' denotes an independent copy of Y. Then for all $0 < s < 2$,*

$$2E|X - Y|^s - E|X - X'|^s - E|Y - Y'|^s \geq 0, \tag{1.1}$$

with equality if and only if $X \stackrel{d}{=} Y$.

The result shows that it is possible to represent the distance between two distributions in terms of expected values of powers of Euclidean distances.

In this dissertation, a goodness-of-fit test is developed based on a generalized version of an energy statistic for stable distributions. The test statistic is a weighted L^2-norm of the distance between the empirical characteristic function and the characteristic function of the null distribution. It can also be expressed as a V-statistic with a kernel of degree two and with degeneracy of order one. The asymptotic distribution of the test statistic is developed, and consistency of the test is proved for the simple hypothesis testing problem. Power performance of the energy test against various alternatives is evaluated by simulation studies and compared to existing tests. An asymptotic result is proved for a special case, the Cauchy distribution, for the composite hypothesis testing problem.

1.1 Characterizations and parameterizations of stable distributions

Mathematicians and probabilists have a very long history of interest in the limiting distribution of sums of independent and identically distributed (iid) random variables. The classical central limit theorem (CLT) [59, p.32][9, p.7] states that such a distribution can only be normal if the variances of those variables are finite. If the condition of the theorem is relaxed by allowing the variances to be infinite, the limiting distribution then becomes a stable distribution [59, p.33]. The stable distribution was first studied by French mathematician Paul Lévy, who studied a type of stable distribution (Lévy distribution) in his 1925 book *Calcul des probabilités* [63, p.vii]. However, the fact that most members in the stable family do not have a distribution function that can be expressed in elementary functions, except for a few special cases, brings not only difficulties to the development of theories, but more importantly, more computational troubles to the application. Nevertheless, a series of papers by Mandelbrot [27, 29] and his successors [14, 47] suggested both theoretical and empirical evidence supporting non-normal stable distributions in certain economic models, such as financial asset returns, portfolio management and risk management. Stable distributions are also widely applied in other areas such as physics and biology. See [63, p.48, p.54] for more examples.

There are many different ways of defining the family of stable distributions, and any one of them can be used as the original definition; that is, by starting from any form of definition, other forms of definition can be obtained as a result. Actually, different forms of definition can be very helpful for us in understanding the motivation of stable laws. Six equivalent definitions are given below.

Generalized central limit theorem The central limit theorems (CLT), as described in many textbooks [59, p.32][9, p.7], state that the sum of independent and identically distributed random variables converges weakly to a normal distribution, when the sum is centered and scaled appropriately and the number of terms in the sum goes to infinity. This theorem is widely used, but it is restricted to random variables with finite variance. For random variables with infinite variance, the iid sum may also converge, but to a stable distribution, not a normal distribution in general.

Definition 1.1.1 *[63, p.6] A distribution function G is said to be* stable *if it is the weak limit of the distribution functions F_n as $n \to \infty$, where $F_n(x) = P(Z_n < x)$ is the cumulative distribution function (CDF) of a linear normalized sum of independent and identically distributed random variables*

$$Z_n = (X_1 + \cdots + X_n)B_n^{-1} - A_n, \quad n = 1, 2, \ldots, \tag{1.2}$$

where $B_n > 0$ and $B_n \to \infty$ as $n \to \infty$.

The set of all such functions G is called the family of stable laws.

Closure under summation An interesting property of the class of stable distributions is that it is closed under summations.

Definition 1.1.2 *[63, p.6] A distribution function G is said to be stable if and only if for all positive numbers b_1 and b_2, there exist a positive number b and a real number a such that*

$$G\left(\frac{x}{b_1}\right) * G\left(\frac{x}{b_2}\right) = G\left(\frac{x-a}{b}\right). \tag{1.3}$$

If $a = 0$, G is said to be strictly stable.

The above definition can also be formulated as follows.

Definition 1.1.3 *[63, p.13] The distribution of a random variable X_1 belongs to the family of stable laws if and only if for all positive number b_1 and b_2, there exist a positive number b*

and a real number a, such that

$$b_1 X_1 + b_2 X_2 \stackrel{d}{=} b X_1 + a, \tag{1.4}$$

where X_2 is an independent copy of X_1, and $\stackrel{d}{=}$ denotes equality in distribution. If $a = 0$, then X_1 is said to be strictly stable.

Definition 1.1.4 *[63, p.14] The distribution of independent non-degenerate random variables $X_1 \stackrel{d}{=} X_2 \stackrel{d}{=} \cdots \stackrel{d}{=} X_n \stackrel{d}{=} \cdots$ belongs to the stable family if and only if for all $n \geq 2$ there exits a positive number b_n and a real number a_n such that*

$$X_1 + X_2 + \cdots + X_n \stackrel{d}{=} b_n X_1 + a_n. \tag{1.5}$$

Compared with Definition 1.1.3, Definition 1.1.4 requires closure under finite sum with possible scaling and shifting, instead of all linear combinations. Both conditions are sufficient and necessary conditions for X_1 to be stable, hence either one can be used as a characterization of stability.

Definition 1.1.5 *[63, p.6] The distribution of the random variables $X_1 \stackrel{d}{=} X_2 \stackrel{d}{=} X_3 \stackrel{d}{\neq} \text{constant}$ belongs to the stable family if and only if there exist positive numbers b_2 and b_3, and real numbers a_2 and a_3, such that*

$$X_1 + X_2 \stackrel{d}{=} b_2 X_1 + a_2, \tag{1.6}$$

$$X_1 + X_2 + X_3 \stackrel{d}{=} b_3 X_1 + a_3. \tag{1.7}$$

This definition reduces the sufficient condition in (1.1.4) to $n = 2$ and $n = 3$.

Characteristic function definition In general, a stable distribution does not have a probability density function or cumulative distribution function which can be expressed in terms of elementary functions, but it can be completely characterized by a closed-form characteristic function of four parameters.

Definition 1.1.6 *A random variable X is* stable *if and only if $X \stackrel{d}{=} aZ+b$, where $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $a > 0$, $b \in \mathbb{R}$ and Z is a random variable with characteristic function*

$$\varphi_X(t) = E(e^{itZ}) = \begin{cases} \exp(-|t|^\alpha[1 - i\beta \tan \frac{\pi\alpha}{2} \text{sign}(t)]) & \text{if } \alpha \neq 1, \\ \exp(-|t|[1 + i\beta \frac{2}{\pi} \text{sign}(t) \log |t|]) & \text{if } \alpha = 1. \end{cases} \tag{1.8}$$

In this dissertation, unified notation will be used for these parameters as suggested by Nolan [39]. Stable distributions are denoted as $S(\alpha, \beta, \gamma, \delta)$, where the parameters are interpreted as follows:

α, the tail index or characteristic exponent, describes the power rate at which the tail(s) of the density function decay;

β, the skewness index, describes how skewed the distribution is;

γ, the “scale” parameter;

δ, the “location” parameter.

The parameters a and b in the above definition are associated with parameters β, γ and δ in different ways in different parameterizations. One of the most important references for studying stable distributions, Zolotarev’s book [63], uses different notation, denoting the scale parameter as λ and the location parameter as γ. To avoid confusion, in this dissertation, results and theorems cited from Zolotarev’s book will be rewritten by using the unified notation as stated above by denoting the scale parameter as γ and the location parameter as δ.

Different parameterizations are used in the literature to accommodate different needs of solving specific problems. While one parameterization may be better for studying the analytic properties of the distribution, another may simplify the numerical computations or parameter estimation. There is no single parameterization that is best for all different purposes, and it is important to notice what type of parameterization is used. Several common parameterizations are described below and the pairwise conversions between them are also provided. Nolan [39] suggested two parameterizations: S_0 and S_1, both of which are defined as location and scale transformation of a random variable Z with characteristic function [39, p.8].

In the remaining part of the dissertation, sgn denotes the sign function, also known as signum function, which is defined as

$$\operatorname{sgn}(x) = \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases} \tag{1.9}$$

Definition 1.1.7 ***Nolan's*** S_0 ***parameterization*** *[39, p.8]. A random variable* X *is* $S(\alpha, \beta, \gamma, \delta; 0)$ *if*

$$X \overset{d}{=} \begin{cases} \gamma(Z - \beta \tan \frac{\pi\alpha}{2}) + \delta & \alpha \neq 1, \\ \gamma Z + \delta & \alpha = 1, \end{cases}$$

where $Z = Z(\alpha, \beta)$ *has characteristic function (1.8). In this case,* X *has characteristic function*

$$Ee^{itX} = \begin{cases} \exp\left(-\gamma^\alpha |t|^\alpha [1 + i\beta \tan(\frac{\pi\alpha}{2}) \operatorname{sgn}(t) \left(|\gamma t|^{1-\alpha} - 1\right)] + i\delta t\right) & \alpha \neq 1, \\ \exp\left(-\gamma |t| \left[1 + i\beta \frac{2}{\pi} \operatorname{sgn}(t) \log(\gamma |t|)\right] + i\delta t\right) & \alpha = 1. \end{cases} \tag{1.10}$$

Nolan [39] suggested using S_0 parametrization for numerical purposes and statistical inference. The S_0 parameterization admits a location-scale family in the sense that if $Z \sim S(\alpha, \beta, \gamma, \delta; 0)$, then for any $a \neq 0$, $b \in \mathbb{R}$, $aZ + b \sim S(\alpha, \operatorname{sgn}(a)\beta, |a|\gamma, a\delta + b; 0)$.

Definition 1.1.8 ***Nolan's*** S_1 ***parameterization*** *[39, p.8]. A random variable* X *is*

$S(\alpha, \beta, \gamma, \delta; 1)$ if

$$X \stackrel{d}{=} \begin{cases} \gamma Z + \delta & \alpha \neq 1, \\ \gamma Z + (\delta + \beta \frac{2}{\pi} \gamma \log \gamma) & \alpha = 1, \end{cases}$$

where $Z = Z(\alpha, \beta)$ has characteristic function (1.8). In this case, X has characteristic function

$$Ee^{itX} = \begin{cases} \exp\left(-\gamma^{\alpha}|t|^{\alpha}\left[1 - i\beta \mathrm{sgn}(t) \tan(\frac{\pi\alpha}{2})\right] + i\delta t\right) & \alpha \neq 1, \\ \exp\left(-\gamma|t|\left[1 + i\beta \frac{2}{\pi} \mathrm{sgn}(t) \log|t|\right] + i\delta t\right) & \alpha = 1. \end{cases} \tag{1.11}$$

Definition 1.1.9 ***Zolotarev's (A) parameterization*** *[63, p.9]. A random variable X is $S(\alpha, \beta, \gamma, \delta; A)$ if its characteristic function can be represented in the form*

$$Ee^{itX} = \begin{cases} \exp(\gamma[it\delta - |t|^{\alpha} + it|t|^{\alpha-1}\beta \tan\frac{\pi\alpha}{2}]) & \text{if } \alpha \neq 1, \\ \exp(\gamma[it\delta - |t|^{\alpha} - i\beta\frac{2}{\pi} t \log|t|]) & \text{if } \alpha = 1. \end{cases} \tag{A}$$

The characteristic functions of stable laws in form (A) are not continuous in the parameters determining them. They have discontinuities at all points of the form $\alpha = 1$, $\beta \neq 0$. Taking the limits $\alpha^{\star} \to 1$ ($\alpha^{\star} \neq 1$), $\beta^{\star} \to \beta \neq 0$, $\gamma^{\star} \to \gamma$, and $\delta^{\star} \to \delta$ not only does not yield the stable law with the parameters $\alpha = 1$, β, γ and δ, but does not even yield a proper distribution in the limit. The whole measure goes to infinity [63, p.11]. The discontinuity can be removed by adding shift $-\beta \tan(\pi\alpha/2)$ to the location parameter, which yields the following parameterization.

Definition 1.1.10 ***Zolotarev's (M) parameterization*** *[63, p.11]. A random variable X is $S(\alpha, \beta, \gamma, \delta; M)$ if its characteristic function can be represented in the form*

$$Ee^{itX} = \begin{cases} \exp(\gamma[it\delta - |t|^{\alpha} + it(|t|^{\alpha-1} - 1)\beta \tan\frac{\pi\alpha}{2}]) & \text{if } \alpha \neq 1, \\ \exp(\gamma[it\delta - |t|^{\alpha} - i\beta\frac{2}{\pi} t \log|t|]) & \text{if } \alpha = 1. \end{cases} \tag{M}$$

As one may notice, Nolan's S_0 parameterization is very similar to Zolotarev's M parameterization, with only modifications to make γ and δ more compliant with the classical sense of scale and location parameter. The same relationship applies to Nolan's S_1 and Zolotarev's A parameterization. For some parameterizations, γ is the scale parameter in the classical definition; that is, the CDF satisfies $F(x;\gamma) = F(x/\gamma;1)$. The scale parameter γ in both Nolan's S_0 and S_1 belong to this category. However, some parameterizations, such as Zolotarev's form (A), only mimic the scale parameter in the sense that it is a mixture of scale parameters and some other parameter(s). Besides the above parameterizations, Zolotarev also discussed another parameterization justified by its analytical nature.

Definition 1.1.11 ***Zolotarev's (B) parameterization*** *[63, p.12]. A random variable X is $S(\alpha, \beta, \gamma, \delta; B)$ if its characteristic function can be represented in the form*

$$Ee^{itX} = \begin{cases} \exp(\gamma[it\delta - |t|^\alpha \exp(-i\frac{\pi}{2}\beta K(\alpha)\mathrm{sgn}(t))]) & \text{if } \alpha \neq 1, \\ \exp(\gamma[it\delta - |t|^\alpha \left(\frac{\pi}{2} + i\beta \log|t| \,\mathrm{sgn}(t)\right)]) & \text{if } \alpha = 1, \end{cases} \tag{B}$$

where $K(\alpha) = \alpha - 1 + \mathrm{sgn}(1-\alpha)$, and the parameters have the same domain of variation as in the form (A).

In the form (B), as in (A), stable laws are not continuous at points of the form $\alpha = 1$. However, with (B) parameterization, the limit distribution exists and is a stable distribution, as $\alpha^\star \to 1_+$, $\beta^\star \to \beta$, $\gamma^\star \to \gamma$, and $\delta^\star \to \delta$. Here $\to 1_+$ denotes converging to 1 from above.

It is worthwhile to notice that γ and δ do not necessarily coincide with the scale or location parameter in the distribution functions of normal, Cauchy and Lévy distributions which are special cases of stable distributions explained on Page 13. It depends on the type of parameterizations used. It is important to determine the parameterization before the parameter estimation, random variable generation and hypothesis testing for the stable distributions. The conversions between different parameterizations are listed as follows. Note that the characteristic exponent α remains the same among all parameterizations mentioned in this dissertation.

$S_1 \to S_0$

$$\beta_0 = \beta_1, \gamma_0 = \gamma_1, \delta_0 = \begin{cases} \delta_1 + \beta\gamma \tan\frac{\pi\alpha}{2} & \text{if } \alpha \neq 1, \\ \delta_1 + \beta\frac{2}{\pi}\gamma \ln\gamma & \text{if } \alpha = 1. \end{cases}$$

$S_0 \to S_1$

$$\beta_1 = \beta_0, \gamma_1 = \gamma_0, \delta_1 = \begin{cases} \delta_0 - \beta\gamma \tan\frac{\pi\alpha}{2} & \text{if } \alpha \neq 1, \\ \delta_0 - \beta\frac{2}{\pi}\gamma \ln\gamma & \text{if } \alpha = 1. \end{cases}$$

$(B) \to (A)$

$$\beta_A = \beta_B, \delta_A = \frac{2}{\pi}\delta_B, \gamma_A = \frac{\pi}{2}\gamma_B, \qquad \text{If } \alpha = 1,$$

$$\begin{cases} \beta_A = \cot(\frac{\pi}{2}\alpha)\tan(\frac{\pi}{2}\beta_B K(\alpha)), \\ \delta_A = \frac{\delta_B}{\cos(\frac{\pi}{2}\beta_B K(\alpha))}, \\ \gamma_A = \gamma_B \cos(\frac{\pi}{2}\beta_B K(\alpha)), \end{cases} \qquad \text{If } \alpha \neq 1.$$

$(A) \to (B)$

$$\beta_B = \beta_A, \delta_B = \frac{\pi}{2}\delta_A, \gamma_B = \frac{2}{\pi}\gamma_A, \qquad \text{If } \alpha = 1,$$

$$\begin{cases} \beta_B = \frac{2}{\pi K(\alpha)} \arctan\left(\frac{\beta_A}{\cos\frac{\pi}{2}\alpha}\right), \\ \delta_B = \delta_A \left(\frac{\cos^2\frac{\pi}{2}\alpha}{\beta_A^2 + \cos^2\frac{\pi}{2}\alpha}\right)^{1/2}, \\ \gamma_B = \gamma_A \left(\frac{\cos^2\frac{\pi}{2}\alpha}{\beta_A^2 + \cos^2\frac{\pi}{2}\alpha}\right)^{-1/2}, \end{cases} \qquad \text{If } \alpha \neq 1.$$

$(A) \to (M)$

$$\beta_M = \beta_A, \delta_M = \delta_A, \gamma_M = \gamma_A, \qquad \text{If } \alpha = 1,$$

$$\beta_M = \beta_A, \delta_M = \delta_A + \beta_A \tan\frac{\pi\alpha}{2}, \gamma_M = \gamma_A, \qquad \text{If } \alpha \neq 1.$$

$(M) \rightarrow (A)$

$$\begin{aligned} &\beta_A = \beta_M, \delta_A = \delta_M, \gamma_A = \gamma_M, && \text{If } \alpha = 1, \\ &\beta_A = \beta_M, \delta_A = \delta_M - \beta_M \tan\frac{\pi\alpha}{2}, \gamma_A = \gamma_M, && \text{If } \alpha \neq 1. \end{aligned}$$

$(A) \rightarrow S_1$

$$\beta_1 = \beta_A, \delta_1 = \gamma_A \delta_A, \gamma_1^\alpha = \gamma_A.$$

$S_1 \rightarrow (A)$

$$\beta_A = \beta_1, \gamma_A = \gamma_1^\alpha, \delta_A = \frac{\delta_1}{\gamma_1^\alpha}.$$

For all parameterizations, the notation $S(\alpha, \beta) = S(\alpha, \beta, 1, 0)$ will be used. For example, it is assumed that $S(\alpha, \beta; 1) = S(\alpha, \beta, 1, 0; 1)$ for parameterization S_1.

1.2 Analytical properties of stable distributions

Except for several special cases, stable distributions, in general, do not have closed form probability density functions (PDF) or cumulative distribution functions (CDF). Zolotarev [63] stated in detail the integral form of the density functions of stable distributions, and Nolan [36] discussed the numerical method of computing the densities. Although there are other expressions of the density functions, few of them seem practical in computing the densities. The integral formula is quite complicated, which is one of the obstacles for the applications of stable distributions. Thanks to the location-scale property of distribution functions, only the density functions of standard stable distributions need to be considered.

To state Zolotarev's integral formula in the (M) parameterization [36], define

$$\zeta = \zeta(\alpha,\beta) = \begin{cases} -\beta \tan \frac{\pi\alpha}{2} & \alpha \neq 1 \\ 0 & \alpha = 1 \end{cases}$$

$$\theta_0 = \theta_0(\alpha,\beta) = \begin{cases} \frac{1}{\alpha}\arctan(\beta \tan \frac{\pi\alpha}{2}) & \alpha \neq 1 \\ \frac{\pi}{2} & \alpha = 1 \end{cases}$$

$$c_1(\alpha,\beta) = \begin{cases} \frac{1}{\pi}(\frac{\pi}{2} - \theta_0) & \alpha < 1 \\ 0 & \alpha = 1 \\ 1 & \alpha > 1 \end{cases}$$

$$V(\theta;\alpha,\beta) = \begin{cases} (\cos \alpha\theta_0)^{\frac{1}{\alpha-1}} \left(\frac{\cos\theta}{\sin\alpha(\theta_0+\theta)}\right)^{\frac{\alpha}{\alpha-1}} \frac{\cos(\alpha\theta_0+(\alpha-1)\theta)}{\cos\theta} & \alpha \neq 1 \\ \frac{2}{\pi}\left(\frac{\pi/2+\beta\theta}{\cos\theta}\right)\exp\left(\frac{1}{\beta}\left(\frac{\pi}{2}+\beta\theta\right)\tan\theta\right) & \alpha = 1, \beta \neq 0. \end{cases}$$

Theorem 1.2.1 *Let $X \sim S(\alpha,\beta;0)$. The density f and distribution function F of X are given by*

1. *When $\alpha \neq 1$ and $x > \zeta$,*

$$f(x;\alpha,\beta) = \frac{\alpha(x-\zeta)^{\frac{1}{\alpha-1}}}{\pi\,|\alpha-1|}\int_{-\theta_0}^{\frac{\pi}{2}} V(\theta;\alpha,\beta)\exp(-(x-\zeta)^{\frac{\alpha}{\alpha-1}}V(\theta;\alpha,\beta))d\theta, \tag{1.12}$$

and

$$F(x;\alpha,\beta) = c_1(\alpha,\beta) + \frac{\operatorname{sgn}(1-\alpha)}{\pi}\int_{-\theta_0}^{\frac{\pi}{2}} \exp(-(x-\zeta)^{\frac{\alpha}{\alpha-1}}V(\theta;\alpha,\beta))d\theta. \tag{1.13}$$

2. When $\alpha \neq 1$ and $x = \zeta$,

$$f(\zeta;\alpha,\beta) = \frac{\Gamma(1+\frac{1}{\alpha})\cos\theta_0}{\pi(1+\zeta^2)^{\frac{1}{2\alpha}}}, \tag{1.14}$$

and

$$F(\zeta;\alpha,\beta) = \frac{1}{\pi}\left(\frac{\pi}{2} - \theta_0\right). \tag{1.15}$$

3. When $\alpha \neq 1$ and $x < \zeta$,

$$f(x;\alpha,\beta) = f(-x;\alpha,-\beta), \tag{1.16}$$

and

$$F(x;\alpha,\beta) = 1 - F(-x;\alpha,-\beta). \tag{1.17}$$

4. When $\alpha = 1$,

$$f(x;1,\beta) = \begin{cases} \frac{1}{2|\beta|}e^{-\frac{\pi x}{2\beta}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} V(\theta;1,\beta)\exp(-e^{-\frac{\pi x}{2\beta}}V(\theta;1,\beta))d\theta & \beta \neq 0, \\ \frac{1}{\pi(1+x^2)} & \beta = 0, \end{cases} \tag{1.18}$$

and

$$F(x;1\beta) = \begin{cases} \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \exp(-e^{-\frac{\pi x}{2\beta}}V(\theta;1,\beta))d\theta & \beta > 0, \\ \frac{1}{2} + \frac{1}{\pi}\arctan x & \beta = 0, \\ 1 - F(x;\alpha,-\beta) & \beta < 0. \end{cases} \tag{1.19}$$

In certain special cases, the density functions or distribution functions of stable distributions can be expressed explicitly by simple functions. When $\alpha = 2$, the stable distribution $S(2,\beta,\gamma,\delta;0)$ is a normal (Gaussian) distribution, with mean $\mu = \delta$ and variance $\sigma^2 = 2\gamma^2$;

that is

$$S(2, \beta, \gamma, \delta; 0) = S(2, \beta, \gamma, \delta; 1) = N(\delta, 2\gamma^2),$$

where the normal distribution $N(\mu, \sigma^2)$ has density function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

When $\alpha = 1$ and $\beta = 0$, the stable distribution $S(1, 0, \gamma, \delta; 0)$ is a Cauchy distribution, with location parameter δ and scale parameter γ; that is

$$S(1, 0, \gamma, \delta; 0) = S(1, 0, \gamma, \delta; 1) = \text{Cauchy}(\gamma, \delta),$$

where the Cauchy(γ, δ) distribution has density function

$$f(x) = \frac{1}{\pi\gamma \left[1 + (\frac{x-\delta}{\gamma})^2\right]}.$$

When $\alpha = \frac{1}{2}$ and $\beta = 1$, the stable distribution $S(\frac{1}{2}, 1, \gamma, \delta; 0)$ is a Lévy distribution with location parameter δ and scale parameter γ; that is

$$S\left(\frac{1}{2}, 1, \gamma, \delta; 1\right) = S\left(\frac{1}{2}, 1, \gamma, \delta + \gamma; 0\right) = \text{Lévy}(\gamma, \delta),$$

where the Lévy distribution Lévy(γ, δ), has density function

$$f(x; \delta, \gamma) = \sqrt{\frac{\gamma}{2\pi}} \frac{e^{-\frac{\gamma}{2(x-\delta)}}}{(x-\delta)^{3/2}}, \text{ if } x > \delta.$$

The normal distribution is widely applied in every field of statistics partly because of its nice analytical properties, many of which are also shared by other members of the stable distribution family.

Property 1.2.2 *Let $X \sim S(\alpha, \beta, \gamma, \delta)$, and $f(x)$ and $F(x)$ be its probability density function and cumulative distribution function, respectively.*

1. ***Paretian tail density*** *[39, p.14]. Both tail probabilities and densities of non-normal stable distributions are asymptotically power laws. If $0 < \alpha < 2$ and $-1 < \beta \leq 1$, then as $x \to \infty$,*

$$\frac{1 - F(x)}{\gamma^{\alpha} c_{\alpha}(1 + \beta)x^{-\alpha}} \to 1, \quad \frac{f(x)}{\alpha\gamma^{\alpha} c_{\alpha}(1 + \beta)x^{-(\alpha+1)}} \to 1, \tag{1.20}$$

where $c_{\alpha} = \sin(\frac{\pi\alpha}{2})\Gamma(\alpha)/\pi$. Using Property 1.2.6 below (the reflection property), the lower tail properties are similar: for $-1 \leq \beta < 1$, as $x \to \infty$

$$\frac{F(-x)}{\gamma^{\alpha} c_{\alpha}(1 - \beta)x^{-\alpha}} \to 1, \quad \frac{f(-x)}{\alpha\gamma^{\alpha} c_{\alpha}(1 - \beta)x^{-(\alpha+1)}} \to 1. \tag{1.21}$$

2. ***Unimodality*** *[63, p.134]. Each stable distribution is unimodal.*

3. *Stable laws have densities with uniformly bounded derivatives of every order.*

Property 1.2.3 *[63, p.61] Every admissible parameter quadruples $(\alpha, \beta_k, \gamma_k, \delta_k)$ and every real numbers h and c_k, $k = 1, \ldots, n$, uniquely determine a parameter quadruple $(\alpha, \beta, \gamma, \delta)$ such that*

$$S(\alpha, \beta, \gamma, \delta) \stackrel{d}{=} \sum_{k} c_k S(\alpha, \beta_k, \gamma_k, \delta_k) + h.$$

With parameterization form (A), the dependence of the quadruple $(\alpha, \beta, \gamma, \delta)$ on the chosen *parameters and numbers is as follows:*

$$\delta = \sum_k \delta_k |c_k|^\alpha,$$

$$\delta\beta = \sum_k \delta_k \beta_k |c_k|^\alpha \mathrm{sgnc_k},$$

$$\delta\gamma = \sum_k \delta_k \gamma_k c_k + h_0,$$

where $h_0 = h$ *if* $\alpha \neq 1$*, and* $h_0 = h - \frac{2}{\pi} \sum_k \delta_k \beta_k c_k \log |c_k|$ *if* $\alpha = 1$*.*

Property 1.2.4 *[63, p.60] Any two admissible parameter quadruples* $(\alpha, \beta, \gamma, \delta)$ *and* $(\alpha, \beta, \gamma', \delta')$ *uniquely determine real numbers* $a > 0$ *and* b *such that*

$$S(\alpha, \beta, \gamma, \delta) \stackrel{d}{=} aS(\alpha, \beta, \gamma', \delta') + \lambda b.$$

In the form (A) the dependence of a *and* b *on the parameters is expressed as follows:*

$$a = (\gamma/\gamma')^{1/\alpha}, \tag{1.22}$$

$$b = \begin{cases} \delta - \delta'(\gamma/\gamma')^{1/\alpha - 1} & \text{if } \alpha \neq 1, \\ \delta - \delta' + \frac{2}{\pi}\beta \log(\gamma/\gamma') & \text{if } \alpha = 1. \end{cases} \tag{1.23}$$

This property can be used to standardize any stable distribution by letting $\delta = 0$ and $\gamma = 1$.

Property 1.2.5 *An arbitrary admissible parameter quadruple* $(\alpha, \beta, \gamma, \delta)$ *and any* β' *and* β'' *with* $-1 \leq \beta' \leq \beta \leq \beta'' \leq 1$ *determine unique positive numbers* c' *and* c'' *and a real number* l *such that*

$$Y(\alpha, \beta, \gamma, \delta) \stackrel{d}{=} c'Y(\alpha, \beta') + c''Y(\alpha, \beta'') + l.$$

In the form (A) the dependence of the parameters and the numbers is expressed as follows:

$$c' = \left(\delta \frac{\beta'' - \beta}{\beta'' - \beta'}\right)^{1/\alpha}, c'' = \left(\delta \frac{\beta - \beta'}{\beta'' - \beta'}\right)^{1/\alpha},$$

$$l = \begin{cases} \delta\gamma & \text{if } \alpha \neq 1, \\ \delta\gamma + \frac{2}{\pi}(\beta' c' \log c' + \beta'' c'' \log c'') & \text{if } \alpha = 1. \end{cases}$$

Property 1.2.6 *[63, p.60] If X and X' are independent and identically distributed with $S(\alpha, \beta, \gamma, \delta)$, then $-X \sim S(\alpha, -\beta, -\gamma, \delta)$, and subsequently, $X - X' \sim S(\alpha, 0, 0, 2\delta)$.*

1.3 Simulating stable distributions

In general, because of lack of closed-form expressions of the cumulative distribution functions, it is not possible to use common random variable generation methods such as inverse transformation method. The Chambers-Mallows-Stuck method [7, 61, 39, 59] provides a general solution to generating independent and identically distributed random variables from stable distribution $S(\alpha, \beta, \gamma, \delta)$ for a given parameter set. With Property 1.2.4, one only needs to consider generating sequence of random variables with distribution $S(\alpha, \beta) = S(\alpha, \beta, 1, 0)$.

Let Θ and W be independent with Θ uniformly distributed on the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$ and W exponentially distributed with mean 1; that is, $\Theta \sim \text{Unif}(-\frac{\pi}{2}, \frac{\pi}{2})$ and $W \sim \text{Exp}(1)$. For $0 < \alpha \leq 2$ and $-1 \leq \beta \leq 1$, define $\theta_0 = \frac{1}{\alpha} \arctan(\beta \tan \frac{\pi\alpha}{2})$. Then

$$Z = \begin{cases} \frac{\sin \alpha(\theta_0 + \Theta)}{(\cos \alpha\theta_0 \cos \Theta)^{1/\alpha}} \left[\frac{\cos(\alpha\theta_0 + (\alpha - 1)\Theta)}{W}\right]^{\frac{1-\alpha}{\alpha}} & \text{if } \alpha \neq 1 \\ \frac{2}{\pi}\left[(\frac{\pi}{2} + \beta\Theta) \tan \Theta - \beta \log\left(\frac{\frac{\pi}{2} W \cos \Theta}{\frac{\pi}{2} + \beta\Theta}\right)\right] & \text{if } \alpha = 1 \end{cases} \tag{1.24}$$

has a $S(\alpha, \beta, 1, 0; 1)$ distribution [39].

1.4 Series representation of stable density functions

Let $g(x, \alpha, \beta)$ be the density function of stable law $S(\alpha, \beta)$. If $\alpha > 1$, then for all admissible β and all real x, [63, p.89]

$$g(x, \alpha, \beta) = \frac{1}{\pi} \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\Gamma(n\alpha' + 1)}{\Gamma(n+1)} \sin(\pi n \rho) x^{n-1}, \tag{1.25}$$

where $\alpha' = 1/\alpha$.

If $\alpha < 1$, then for any admissible β and any real x,

$$g(x, \alpha, \beta) = \frac{1}{\pi} \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\Gamma(n\alpha + 1)}{\Gamma(n+1)} \sin(\pi n \rho) x^{-n\alpha-1}. \tag{1.26}$$

If $\alpha = 1$ and $\beta > 0$, then for any real x,

$$g(x, 1, \beta) = \frac{1}{\pi} \sum_{n=1}^{\infty} (-1)^{n-1} n b_n x^{n-1}, \tag{1.27}$$

where

$$b_n = \frac{1}{\Gamma(n+1)} \int_0^{\infty} \exp(-\beta u \log u) u^{n-1} \sin\left[(1+\beta) u \frac{\pi}{2}\right] du. \tag{1.28}$$

1.5 Estimation of parameters

Unlike normal distributions, there are no simple estimators for stable distribution parameters which possess nice properties such as unbiasedness and minimum variance, due to the unavailability of an explicit form of the density function. Methods of estimating stable distribution parameters have been proposed to bypass this obstacle [16, 1, 41, 35, 3, 22, 42, 38, 26].

Fama and Roll [15, 16] proposed using sample fractiles to estimate α, γ and δ for symmetric stable distributions with $\alpha > 1$. This method was then improved by McCulloch [35] who generalized the method to estimate all four parameters when $\alpha \geq 0.5$. Koutrouvelis

[23, 22] proposed a regression-type estimator. A comparative study was done by Akgiray and Lamoureux [1], who compared the performance of the iterative regression method and the fractile method using both simulated and actual data, and concluded by recommending the regression-type estimator.

DuMouchel [12] discussed the approximate maximum-likelihood estimators for stable distributions, and pointed out that if both α and δ are unknown, then the likelihood function will have no maximum within $0 < \alpha \leq 2$, $-\infty < \delta < \infty$. Maximum likelihood estimators for general stable distribution were studied by Nolan [38] to estimate all four parameters.

Methods have also been proposed for estimating parameters of special interest, especially the characteristic exponent α. Fan [17] developed an unbiased estimator for α with the structure of a U-statistic. Fan [18] constructed a minimum distance estimator for α by minimizing the Kolmogorov distance or the Cramér von-Mises distance between the empirical distribution function and a class of distributions defined based on the sum-preserving property of stable random variables.

CHAPTER 2

EXISTING GOODNESS-OF-FIT TESTS FOR STABLE DISTRIBUTIONS

There has been a long-standing debate on whether it is proper and necessary to fit models using stable distributions, especially for financial asset returns. While normality is the core assumption of modern portfolio theory, it is often observed that the distribution of financial asset returns, such as stock returns, possesses characteristics such as heavy tail and skewness, which cannot be justified by the normality assumption. Stable distributions are proposed to be a better model for financial returns for both theoretical and empirical reasons; see [28, 14, 47]. However, there are also voices against the correctness of modeling the return of financial assets by a stable distribution. Blattberg and Gonedes [5] found evidence that weekly and monthly returns have significantly higher characteristic exponent estimates than daily returns do, which contradicts the "closure under summation" properties of stable distributions. They proposed to use Student's t distribution instead. Lau et al. [24] investigated the behavior of higher moments of stock data, and argued that stock returns, when taken as a group, should not be assumed from non-normal stable distributions.

DuMouchel [13] criticized that if the true distribution is not stable, then the measure of tail behavior (thickness) is biased when assuming a stable model. In this case, he suggested fitting the tail data separately using a generalized Pareto model. Nevertheless, Uchaikin and Zolotarev [59, p.482] responded to the empirical objection to stable distribution in financial areas, and tried to justify the inconformity of theoretical properties of stable distributions with empirical findings.

In a formal framework, the goodness-of-fit problem for stable distributions will be formulated as follows: let $x_1, \ldots, x_n$ be independent and identically distributed observations from distribution F. Two types of null hypotheses to be tested are

$$H_0 : F = S_\theta,$$

where the parameter vector $\theta = (\alpha, \beta, \gamma, \delta)$ is fully specified; and

$$H_0 : F = S_\theta \in \mathbb{S}_\Theta,$$

where $\mathbb{S}_\Theta = \{S_\theta : \theta \in \Theta\}$ is a family of stable distributions and Θ is the parameter space. The second null hypothesis corresponds to the case when the parameters are partly unknown, or all unknown. A goodness-of-fit test for the first null hypothesis is a test for a simple hypothesis, and that for the second null hypothesis is a test for a composite hypothesis.

Analogous to the distance between two points on the real line, the statistical "distance" between the sample and the hypothesized distribution can be defined. For a random sample $x_1, \ldots, x_n$, the empirical distribution function (EDF) is defined as

$$F_n(t) = \frac{1}{n} \sum_{j=1}^{n} I(x_j \leq t),$$

where I is the indicator function. The empirical characteristic function (ECF) is defined as

$$\varphi_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{itx_j},$$

where i is the imaginary unit. Since the empirical distribution function and the empirical characteristic function converge almost surely to the cumulative distribution function and characteristic function, respectively, it is natural to measure the goodness-of-fit by measuring the discrepancy between the empirical functions and their counterparts. Many powerful tests have been proposed for testing normality, but few can be extended to the general stable family. Therefore, the focus of the dissertation will be on testing the goodness-of-fit for stable distributions, or “stability”.

2.1 Tests based on empirical distribution function

Tests based on the empirical distribution function test the goodness-of-fit by measuring the difference between the empirical distribution function of the test sample and the null distribution. Some well-known tests based on empirical distribution function (EDF) include the Kolmogorov-Smirnov (K-S) test, the Cramer-von Mises (C-M) test and the Anderson-Darling (A-D) test [9, p.421]. Since most stable distributions do not have a closed-form distribution function, few results are found in the literature that use an EDF-based method to test goodness-of-fit for stable distributions, except for some special cases such as normal distributions and Cauchy distributions. These tests are usually implemented in a “distribution-free” way such that the random sample is first transformed by the CDF of the null distribution and then tested for uniformity. Let $x_1, \ldots, x_n$ be an independent and identically distributed sample, and $x_{(1)}, \ldots, x_{(n)}$ be the order statistics of the sample. Let $U_{(j)} = F_0(x_{(j)})$. The

Kolmogorov-Smirnov test statistic is

$$D_n = \max_{1 \le j \le n} \max \left\{ \frac{j}{n} - U_{(j)}, U_{(j)} - \frac{j-1}{n} \right\},$$

the Cramer-von-Mises statistic is

$$C_n = \frac{1}{12n} + \sum_{j=1}^{n} \left(U_{(j)} - \frac{2j-1}{n} \right)^2,$$

and the Anderson-Darling statistic is

$$A_n = -n - \frac{1}{n} \left[\sum_{j=1}^{n} (2j-1) \left(\log U_{(j)} + \log(1 - U_{(n-j+1)}) \right) \right].$$

The above goodness-of-fit test statistics measure how close the transformed order statistics of the sample are to uniform order statistics.

2.2 Tests based on empirical characteristic function

Stable distributions have characteristic functions that can be expressed in a relatively simple form, so the goodness-of-fit problem can be considered by measuring the discrepancy between the empirical characteristic function and the characteristic function of the null distribution. A goodness-of-fit test for Cauchy distributions based on the empirical characteristic function was proposed by Gürtler and Henze (G-H) [19] for the composite hypothesis, with the location parameter estimated by the sample median and the scale parameter estimated by the half-interquartile range. The same form of the test statistic was then adopted by Matsui and Takemura (M-T) [33, 34] in testing Cauchy distributions and symmetric stable distributions, with parameters estimated by maximum likelihood estimators, as well as equivariant integrated squared error estimator (EISE). The test statistic $D_{n,\lambda}$ computes the weighted L^2 distance between the empirical characteristic function of the "standardized" data and the hypothetical characteristic function of the standard distribution, with a weight function $w(t) = e^{-\kappa|t|}$. Let $x_1, \ldots, x_n$ be a random sample, and $\hat{\alpha}$, $\hat{\gamma}$ and $\hat{\delta}$ be affine equivariant estimators of α, γ and δ, and $\varphi_n(t) := \frac{1}{n}\sum_{j=1}^{n}\exp(ity_j)$ be the empirical characteristic function of the standardized data $y_j = (x_j - \hat{\delta})/\hat{\gamma}$. The statistic for testing symmetric stable distributions is

$$\begin{aligned} D_{n,\kappa} :=& n\int_{-\infty}^{\infty} \|\varphi_n(t) - \exp(-|t|^{\hat{\alpha}})\|^2 e^{-\kappa|t|} dt \\ =& \frac{1}{n}\sum_{j,k}\frac{2\kappa}{\kappa^2 + (y_j - y_k)^2} - 4\sum_j \int_0^{\infty} \cos(ty_j)\exp(-t^{\hat{\alpha}} - \kappa t) dt \\ &+ 2n\int_0^{\infty} \exp(-2t^{\hat{\alpha}} - \kappa t) dt. \end{aligned}$$

When the characteristic exponent α is known and equal to one; that is, when the null distributions for testing are Cauchy distributions, the test statistic can be simplified as

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

It is noted that no result in the literature is found to extend a G-H type of test to asymmetric stable distributions.

2.3 Other types of goodness-of-fit tests

Graphical methods such as a Quantile-Quantile (Q-Q) plot or a Probability-Probability (P-P) plot can be useful in comparing a sample of data to the hypothesized distribution in a visual, intuitive way. A Q-Q plot displays the sample quantiles plotted against the theoretical quantiles of the hypothesized distribution, while a P-P plot displays the empirical distribution function plotted against the theoretical cumulative distribution function. Nolan [37] recommended a "variance stabilized" P-P plot in diagnosing the fit of a stable model. Although these methods have great value in showing the goodness-of-fit of the model, especially useful in determining the difference of the tail weight between data and model, it still depends on subjective judgment and is not reliable against all alternatives.

Saniga and Hayya [48] discussed a simple goodness-of-fit test using sample kurtosis b_2 to distinguish among symmetric stable distributions with different tail index α. Saniga and Miles [49] then examined the power of several standard goodness-of-fit tests of normality against asymmetric stable alternatives and concluded that a test based on kurtosis statistic b_2 performs generally better for large sample size ($n \geq 50$) and less skewed cases ($|\beta| \leq .75$).

Some research has been done by taking advantage of the unique properties of stable distributions. Brcich et al. [6] developed a goodness-of-fit test based on one of the characterizations of a symmetric stable distribution, that it can be fully characterized by the condition: $X_1 + X_2 \stackrel{d}{=} C_2 X$ and $X_1 + X_2 + X_3 \stackrel{d}{=} C_3 X$, but this test is powerful only when the sample size is large.

CHAPTER 3

THE ENERGY GOODNESS-OF-FIT TEST FOR STABLE DISTRIBUTIONS

In this chapter, a new goodness-of-fit test for stable distributions is presented based on the energy distance between the empirical characteristic function (ECF) and the characteristic function of the hypothesized stable distribution.

3.1 Preliminaries

Let $\varphi_F(t)$ and $\varphi_G(t)$ be characteristic functions corresponding to distributions F and G, respectively. The difference between two distributions can be measured by

$$Q(F,G) = \int_{-\infty}^{\infty} |\varphi_F(t) - \varphi_G(t)|^2 w(t)dt, \tag{3.1}$$

where $w(t)$ is a proper weight function. By letting $F = F_n$ be the empirical distribution of independent and identically distributed $x_1, \ldots, x_n$, and $G = F_0$ be the distribution to be tested, one can get a family of goodness-of-fit tests based on empirical characteristic functions. For example, when $w(t) = e^{-\kappa|t|}$, one has the goodness-of-fit test statistic studied by Gürtler and Henze [19], and Matsui and Takemura [33, 34]. The original energy statistic discussed by Rizzo [44] corresponds to a weight function $w(t) = \frac{1}{\pi t^2}$ for the goodness-of-fit test, which has an alternative expression

$$\begin{aligned} Q_n =& n \int_{-\infty}^{\infty} |\varphi_{F_n}(t) - \varphi_{F_0}(t)|^2 \frac{1}{\pi t^2} dt \\ =& n \left\{ \frac{2}{n} \sum_{j=1}^{n} E|x_j - X| - E|X - X'| - \frac{1}{n^2} \sum_{j,k=1}^{n} |x_j - x_k| \right\}, \end{aligned} \tag{3.2}$$

where X is a random variable with distribution F_0, and X' is an independent copy of X. This expression is useful in both evaluating the test statistic and deriving its asymptotic distribution. It follows from (3.2) that the original energy statistic is applicable only if $E|x_j - X| < \infty$, and this condition is not necessarily satisfied for arbitrary F_0. For example, when X follows stable distribution with tail index $\alpha < 2$, $E|x_j - X|$ is not finite. The following theorem [56] generalizes the original energy inequality, based on which a modified energy statistic can be applied to test goodness-of-fit for stable distributions.

Theorem 3.1.1 *If the d-dimensional random variables X and Y are independent, and there exists $0 < s < 2$ such that $E\|X\|^s + E\|Y\|^s < \infty$, and φ_X and φ_Y denote their respective characteristic functions, then*

$$\frac{1}{C(d,s)} \int_{R^d} \frac{|\varphi_X(t) - \varphi_Y(t)|^2}{\|\boldsymbol{t}\|^{d+s}} d\boldsymbol{t} = 2E\|X - Y\|^s - E\|X - X'\|^s - E\|Y - Y'\|^s \geq 0, \tag{3.3}$$

with equality if and only if $X \overset{d}{=} Y$. Here

$$C(d,s) = \frac{2\pi^{d/2}\Gamma(1 - s/2)}{s 2^s \Gamma((d+s)/2)}. \tag{3.4}$$

For a proof, see [53, 56]. The above theorem applies to multivariate distributions. In this chapter, only the univariate goodness-of-fit test problem is considered. For simplicity, we denote

$$C(s) = C(1, s) = \frac{2\pi^{1/2}\Gamma(1 - s/2)}{s2^s\Gamma((1 + s)/2)}. \tag{3.5}$$

In the special case, when $s = 1$, $C(1, 1) = \pi$, we get the original energy statistic (3.2).

As pointed out in Chapter 1, stable distributions with tail index $\alpha < 2$ do not have finite variance (second moment), and those with $\alpha \leq 1$ do not have finite mean (first moment). However, some of the non-integer moments and absolute moments exist for stable distributions, depending on α.

Proposition 3.1.2 *If $X \sim S(\alpha, \beta, \gamma, \delta)$ and $s > 0$, then $E|X|^s < \infty$ if and only if $s < \alpha$.*

Proof Suppose $f(x) = f(x; \alpha, \beta, \gamma, \delta)$ is the density function of X. Then

$$E|X|^s = \int_0^\infty x^s f(x)dx + \int_{-\infty}^0 (-x)^s f(x)dx = \int_0^\infty x^s f(x)dx + \int_0^\infty x^s f(-x)dx.$$

By (1.20), there exists some constant C, such that $\frac{f(x)}{Cx^{-(\alpha+1)}} \to 1$, as $x \to \infty$. Therefore, for all $0 < \epsilon < 1$, there exists $M > 0$ such that for all $x > M$, $1 - \epsilon < \frac{f(x)}{Cx^{-(\alpha+1)}} < 1 + \epsilon$. When $0 < s < \alpha$,

$$\begin{aligned} 0 < \int_0^\infty x^s f(x)dx &\leq \int_0^M x^s f(x)dx + \int_M^\infty x^s(1 + \epsilon)Cx^{-(\alpha+1)}dx \\ &= \int_0^M x^s f(x)dx + (1 + \epsilon)C\frac{M^{-\alpha+s}}{\alpha - s} < \infty. \end{aligned}$$

Similarly, we can get $0 < \int_0^\infty x^s f(-x)dx < \infty$, and thus $E|X|^s < \infty$.

When $s \geq \alpha$,

$$\begin{aligned}
\int_0^\infty x^s f(x)dx &\geq \int_0^M x^s f(x)dx + \int_M^\infty x^s (1-\epsilon) C x^{-(\alpha+1)} dx \\
&= \int_0^M x^s f(x)dx + (1-\epsilon) C \int_M^\infty x^{s-\alpha-1} dx \\
&\geq \int_0^M x^s f(x)dx + (1-\epsilon) C \int_M^\infty x^{-1} dx
\end{aligned}$$

Notice that $\int_0^M x^s f(x)dx \geq 0$ and $\frac{1}{x}$ is not integrable over the interval (M, ∞), hence $x^s f(x)$ is not integrable on $(0, \infty)$. By the same argument, $x^s f(-x)$ is not integrable on $(0, \infty)$, either. Therefore, $E|X|^s$ is not finite. $\blacksquare$

Let X' be an independent copy of x. By Property 1.2.6, $X - X'$ is also stable with tail index α, so $E|X - X'|^s < \infty$, for all $0 < s < \alpha$.

3.2 Energy statistic for testing stable distributions

Let $x_1, \ldots, x_n$ be a random sample. Let X and X' be independent and identically distributed with $F = S(\alpha, \beta, \gamma, \delta; 1)$. Nolan's S_1 parameterization is adopted because the family of stable distributions is a location-scale family under such parameterization.

The proposed energy goodness-of-fit test statistic is defined as follows:

$$Q_{n,s} = n \left\{ \frac{2}{n} \sum_{j=1}^{n} E|x_j - X|^s - E|X - X'|^s - \frac{1}{n^2} \sum_{j,k=1}^{n} |x_j - x_k|^s \right\}, \tag{3.6}$$

where s is a test parameter chosen to be less than α. The statistic $Q_{n,s}/n$ is an estimate of the energy distance $\mathcal{E}_s(X, Y) \equiv 2E|X - Y|^s - E|X - X'|^s - E|Y - Y'|^s$. The test statistic $Q_{n,s}$ is affine invariant (location-scale invariant), and since stable distributions consist of a location-scale family, the test problem can be simplified by only considering the case when $\gamma = 1$ and $\delta = 0$. Therefore, in the following sections of this chapter, it will be assumed that

$\gamma = 1$ and $\delta = 0$, and the corresponding characteristic function of X is

$$\varphi_X(t) = \begin{cases} \exp\left(-|t|^\alpha \left[1 - i\beta(\mathrm{sgn}t)\tan\frac{\pi\alpha}{2}\right]\right) & \alpha \neq 1, \\ \exp\left(-|t| \left[1 + i\beta\frac{2}{\pi}(\mathrm{sgn}t)\log|t|\right]\right) & \alpha = 1. \end{cases} \tag{3.7}$$

The statistic $Q_{n,s}/n = \frac{1}{n^2}\sum_{i,j=1}^n h(X_i, X_j)$ is a V-statistic with kernel of degree two given by

$$h(x, y) = E|x - X|^s + E|y - X|^s - E|X - X'|^s - |x - y|^s, \tag{3.8}$$

where $0 < s < 2$. The kernel $h(x, y)$ is degenerate; that is, $Eh(x, X) = Eh(X, y) = 0$, for all x and y. By the asymptotic theory of V-statistics [60, 20], if $h(x, y)$ is degenerate and $\mathbb{E}h^2(X, X') < \infty$, then $Q_{n,s}$ converges weakly to a weighted sum of independent chi-squared distributions with one degree of freedom.

Property 3.2.1 *If $x_1, \dots, x_n$ are independent and identically distributed as Y, then*

$$EQ_{n,s} = E|Y - Y'|^s + n[2E|Y - X|^s - E|X - X'|^s - E|Y - Y'|^s]. \tag{3.9}$$

Under H_0, X and Y are identically distributed, hence $EQ_{n,s} = E|X - X'|^s$.

Proof Notice that in the notation $E|x_i - X|^s$, the expected value is taken on the probability space of X.

$$\begin{aligned} EQ_{n,s} =& n\left[\frac{2}{n}\sum_{j=1}^n E|Y - X|^s - E|X - X'|^s - \frac{1}{n}\sum_{j\neq k}^n E|Y - Y'|^s\right] \\ =& n\left[2E|Y - X|^s - E|X - X'|^s - \frac{(n-1)n}{n^2}E|Y - Y'|^s\right] \\ =& E|Y - Y'|^s + n\left[2E|Y - X|^s - E|X - X'|^s - E|Y - Y'|^s\right]. \end{aligned}$$

∎

The expression of the energy statistic as in (3.6) has a simple form, but the expectations need to be further evaluated for computational purposes. First, a very useful expression of $|x|^s$ is derived. Let Rex and Imx denote the real part and imaginary part of a complex number x, respectively.

Theorem 3.2.2 *When* $0 < s < 1$,

$$|x|^s = \frac{2}{\pi}\Gamma(1+s)\sin\frac{\pi s}{2}\int_0^\infty (1-\cos tx)t^{-s-1}dt. \tag{3.10}$$

Proof We first establish following equality for $0 < s < 1$:

$$\int_0^\infty (1-e^{-pt})t^{-s-1}dt = p^s\Gamma(1-s)s^{-1}. \tag{3.11}$$

Let $-pt = -x$, hence $t = \frac{x}{p}$. We have

$$\begin{aligned}
\int_0^\infty (1-e^{-pt})t^{-s-1}dt &= \int_0^\infty (1-e^{-x})(x/p)^{-s-1}d(x/p) \\
&= p^s\int_0^\infty (1-e^{-x})x^{-s-1}dx \\
&= p^s\int_0^\infty (1-e^{-x})d\left(\frac{x^{-s}}{-s}\right) \\
&= \frac{p^s}{-s}\left[(1-e^{-x})x^{-s}|_{x=0}^{\infty} - \int_0^\infty x^{-s}e^{-s}dx\right] \\
&= \frac{p^s}{s}\Gamma(1-s).
\end{aligned}$$

The result $(1-e^{-x})x^{-s}|_{x=0}^{\infty} = 0$ follows from the fact that

$$\lim_{x\to 0}\frac{1-e^{-x}}{x^s} = \lim_{x\to 0}\frac{e^{-x}}{sx^{s-1}} = 0.$$

The right hand side of equation (3.11) shows that it is a continuous function of p on the half-plane Re$p > 0$ and we can use the continuity of the expression on the imaginary axis.

By letting $p = -ix$, we can evaluate the real part of (3.11),

$$\mathrm{Re}\int_0^\infty (1-e^{-itx})t^{-s-1}dt = \int_0^\infty \mathrm{Re}(1-e^{-itx})t^{-s-1}dt = \int_0^\infty (1-\cos tx)t^{-s-1}dt. \qquad (3.12)$$

Meanwhile

$$\begin{aligned}\mathrm{Re}\left[(-ix)^s\Gamma(1-s)s^{-1}\right] &= \Gamma(1-s)s^{-1}\mathrm{Re}(-ix)^s \\ &= \begin{cases}\Gamma(1-s)s^{-1}\mathrm{Re}\left[|x|^s e^{i\frac{\pi}{2}s}\right] & \text{if } x \geq 0 \\ \Gamma(1-s)s^{-1}\mathrm{Re}\left[|x|^s e^{-i\frac{\pi}{2}s}\right] & \text{if } x < 0\end{cases} \\ &= \Gamma(1-s)s^{-1}\left|x\right|^s\cos(\pi s/2).\end{aligned}$$

Using the facts that $\Gamma(1-z) = \frac{\pi}{\Gamma(z)\sin(\pi z)}$ and $\Gamma(1+z) = z\Gamma(z)$, we obtain

$$\Gamma(1-s)s^{-1}\left|x\right|^s\cos(\pi s/2) = \left|x\right|^s\frac{\pi s}{\Gamma(1+s)\sin(\pi s)}s^{-1}\cos(\pi s/2) = \frac{\pi\left|x\right|^s}{2\Gamma(1+s)\sin(\frac{\pi s}{2})}. \qquad (3.13)$$

Applying (3.12) together with (3.13), we get

$$|x|^s = \frac{2}{\pi}\Gamma(1+s)\sin\frac{\pi s}{2}\int_0^\infty (1-\cos tx)t^{-s-1}dt.$$

∎

With this expression of $|x|^s$, we immediately obtain that, if $0 < s < 1$ and $s < \alpha$, then

$$E|X|^s = \frac{2}{\pi}\Gamma(1+s)\sin\frac{\pi s}{2}\int_0^\infty (1-\mathrm{Re}\varphi_X(t))t^{-s-1}dt, \qquad (3.14)$$

where $\mathrm{Re}\varphi_X(t)$ is the real part of $\varphi_X(t)$, the characteristic function of X. Notice this a general result for X with any distribution. By Property (1.2.6), $X-X'$ is $S(\alpha,0,0,2)$, hence

$$E\left|X-X'\right|^{s}=\frac{2}{\pi}\Gamma(1+s)\sin\frac{\pi s}{2}\int_{0}^{\infty}(1-\exp(-2t^{\alpha}))t^{-s-1}dt \tag{3.15}$$

$$=\frac{2}{\pi}2^{s/\alpha}\Gamma(1-s/\alpha)\Gamma(s)\sin\frac{\pi s}{2}. \tag{3.16}$$

The integral (3.15) can be easily computed with integration by parts.

The expectation $E|x_j-X|^s$ can be computed by definition

$$E|X-x_j|^s=\int_{-\infty}^{\infty}|x-x_j|^s f(x)dx,$$

only for very few special cases when $f(x)$ has a closed-form expression, while the numerical approximation of $f(x)$ may be slow and inaccurate due to the fact that it is obtained by numerical integration. Generally, the following formula is used, instead; from (3.14)

$$E|x_j-X|^s=\frac{2}{\pi}\Gamma(1+s)\sin\frac{\pi s}{2}\int_{0}^{\infty}(1-\mathrm{Re}\varphi(t))t^{-s-1}dt, \tag{3.17}$$

where $\varphi_j(t)=Ee^{it(X-x_j)}=e^{-itx_j}\varphi_X(t)$ is the characteristic function of $X-x_i$.

By (3.7) and (3.14), when $\alpha\neq 1$,

$$E|X-x_j|^s=\frac{2}{\pi}\Gamma(1+s)\sin\frac{\pi s}{2}\int_{0}^{\infty}\frac{1-e^{-t^{\alpha}}\cos\left(\beta t^{\alpha}\tan\frac{\pi\alpha}{2}-x_j t\right)}{t^{s+1}}dt. \tag{3.18}$$

When $\alpha=1$,

$$E|X-x_j|^s=\frac{2}{\pi}\Gamma(1+s)\sin\frac{\pi s}{2}\int_{0}^{\infty}\frac{1-e^{-t}\cos(\beta\frac{2}{\pi}t\log t+x_j t)}{t^{s+1}}dt. \tag{3.19}$$

In the standard Cauchy case when $\alpha=1$ and $\beta=0$,

$$E|X-x_i|^s=(1+x_i^2)^{s/2}\frac{\cos(s\arctan x_i)}{\cos\frac{\pi s}{2}}. \tag{3.20}$$

However, if x_i is too large, which is very common in samples of heavy-tail distributions, the numerical method for evaluating (3.18) and (3.19) based on quadrature may fail due to the reason that the integrand oscillates too frequently and therefore requires too many subintervals for interpolating the integrand. One solution is to split $(0,\infty)$ into two subintervals and evaluate the integrals separately:

$$\begin{aligned}&\int_0^\infty (1-e^{-t^\alpha}\cos(\omega(t)))t^{-s-1}dt\\ =&\int_0^{t_0} (1-e^{-t^\alpha}\cos(\omega(t)))t^{-s-1}dt+\int_{t_0}^\infty (1-e^{-t^\alpha}\cos(\omega(t)))t^{-s-1}dt,\end{aligned}$$

where

$$\omega(t)=\begin{cases}\beta t^\alpha\tan\frac{\pi\alpha}{2}-x_it & \alpha\neq 1,\\ \beta\frac{2}{\pi}t\log t+x_it & \alpha=1.\end{cases} \tag{3.21}$$

When t_0 is large enough,

$$\int_{t_0}^\infty (1-e^{-t^\alpha}\cos(\omega(t)))t^{-s-1}dt\approx\int_{t_0}^\infty t^{-s-1}dt=\frac{1}{s{t_0}^s},$$

and the error of this approximation is

$$\left|\int_{t_0}^\infty e^{-t^\alpha}\cos(\omega(t))t^{-s-1}dt\right|\leq\int_{t_0}^\infty \left|e^{-t^\alpha}\cos(\omega(t))t^{-s-1}\right|dt<\frac{1}{\alpha t_0^{s+\alpha}e^{t_0^\alpha}}.$$

In practice, $E|x_i-X|^s$ can be approximated by

$$\int_0^{t_0} (1-e^{-t^\alpha}\cos(\omega(t)))t^{-s-1}dt+\frac{1}{s{t_0}^s}.$$

The integral on a finite interval can be numerically computed, and t_0 can be determined by controlling the approximation error of the second part.

Even with the help of the above procedure, the numerical integration of $E|x_i-X|^s$ may converge very slowly or may not reach the given tolerance for x_i with extremely large absolute value, for example, when $|x_i| \geq 10^4$. The following theorem enables us to approximate $E|x_i - X|^s$ with $|x_i|^s$, when $|x_i|^s$ is extremely large.

Theorem 3.2.3 *If $E|x - X|^s < \infty$ for all $x \in R$, then as $x \to \infty$,*

$$\frac{E|x - X|^s}{|x|^s} \to 1. \tag{3.22}$$

Proof If $\varphi(t; x)$ is the characteristic function of $1 - X/x$, then

$$\begin{aligned}\lim_{x\to\infty} E\left|1 - \frac{X}{x}\right|^s &= \lim_{x\to\infty} \frac{2}{\pi}\Gamma(1+s)\sin\frac{\pi}{2}s \int_0^\infty (1 - \mathrm{Re}\varphi(t; x))t^{-s-1}dt \\ &= \frac{2}{\pi}\Gamma(1+s)\sin\frac{\pi}{2}s \int_0^\infty (1 - \mathrm{Re}\lim_{x\to\infty}\varphi(t; x))t^{-s-1}dt.\end{aligned}$$

Notice that

$$\lim_{x\to\infty}\varphi(t; x) = \lim_{x\to\infty} Ee^{it(1-X/x)} = e^{it}\lim_{x\to\infty} Ee^{i\left(-\frac{t}{x}\right)X} = e^{it}.$$

Hence,

$$\lim_{x\to\infty} E\left|1 - \frac{X}{x}\right|^s = \frac{2}{\pi}\Gamma(1+s)\sin\frac{\pi}{2}s \int_0^\infty (1 - \cos t)t^{-s-1}dt = 1$$

∎

A special case of Theorem 3.2.3 can be easily validated on equation (3.20) for the standard Cauchy distribution. Notice that Theorem 3.2.3 does not give the speed of convergence, which actually depends on the distribution of X and the choice of s. In simulations, it was found that the relative approximation error is less than 10^{-5} when $x_i > 2000$ for most of the situations. In this dissertation, the computational formula for $E|x_i - X|^s$ is as follows:

$$E|x_i - X|^s = \begin{cases} \frac{2}{\pi}\Gamma(1+s)\sin\frac{\pi s}{2}\int_0^\infty \frac{1-e^{-t^\alpha}\cos\left(\beta t^\alpha \tan\frac{\pi\alpha}{2} - x_i t\right)}{t^{s+1}}dt & \text{if } x < 2000 \text{ and } \alpha \neq 1, \\ \frac{2}{\pi}\Gamma(1+s)\sin\frac{\pi s}{2}\int_0^\infty \frac{1-e^{-t}\cos(\beta\frac{2}{\pi}t\log t + x_i t)}{t^{s+1}}dt & \text{if } x < 2000 \text{ and } \alpha = 1, \\ |x|^s & \text{if } x \geq 2000. \end{cases} \tag{3.23}$$

Since the exact value of $E|x_i - X|^s$ can be obtained for the standard Cauchy case, the accuracy of the computational formula can be measured, as shown in Figure 3.1. The absolute error and relative error of calculating $E|x_i - X|^s$ are plotted against x_i, where the relative error is defined as $\frac{\text{true value - calculated value}}{\text{true value}}$.

To summarize the above topics in this section, a computational formula was derived for the energy test statistic for testing the simple hypothesis of stability. When $\alpha \neq 1$,

$$\begin{aligned} Q_{n,s} =& \frac{4}{\pi}\Gamma(1+s)\sin\frac{\pi s}{2}\sum_{j=1}^{n}\int_0^\infty \frac{1-e^{-t^\alpha}\cos\left(\beta t^\alpha \tan\frac{\pi\alpha}{2} - x_i t\right)}{t^{s+1}}dt \\ & - n\frac{2^{1+\frac{s}{\alpha}}}{\pi}\Gamma\left(1-\frac{s}{\alpha}\right)\Gamma(s)\sin\frac{\pi s}{2} - \frac{1}{n}\sum_{j,k=1}^{n}|x_j - x_k|^s. \end{aligned} \tag{3.24}$$

When $\alpha = 1$,

$$\begin{aligned} Q_{n,s} =& \frac{4}{\pi}\Gamma(1+s)\sin\frac{\pi s}{2}\sum_{j=1}^{n}\int_0^\infty \frac{1-e^{-t}\cos\left(\beta\frac{2}{\pi}t\log t + x_i t\right)}{t^{s+1}}dt \\ & - n\frac{2^{1+s}}{\pi}\Gamma\left(1-s\right)\Gamma(s)\sin\frac{\pi s}{2} - \frac{1}{n}\sum_{j,k=1}^{n}|x_j - x_k|^s. \end{aligned} \tag{3.25}$$

3.3 Limiting distribution of $Q_{n,s}$ under H_0

The implications of obtaining the asymptotic distribution of the test statistic are twofold. Theoretically, it is important to know that the test statistic converges weakly to some non-degenerate distribution before studying the consistency of the test. Computationally, as the

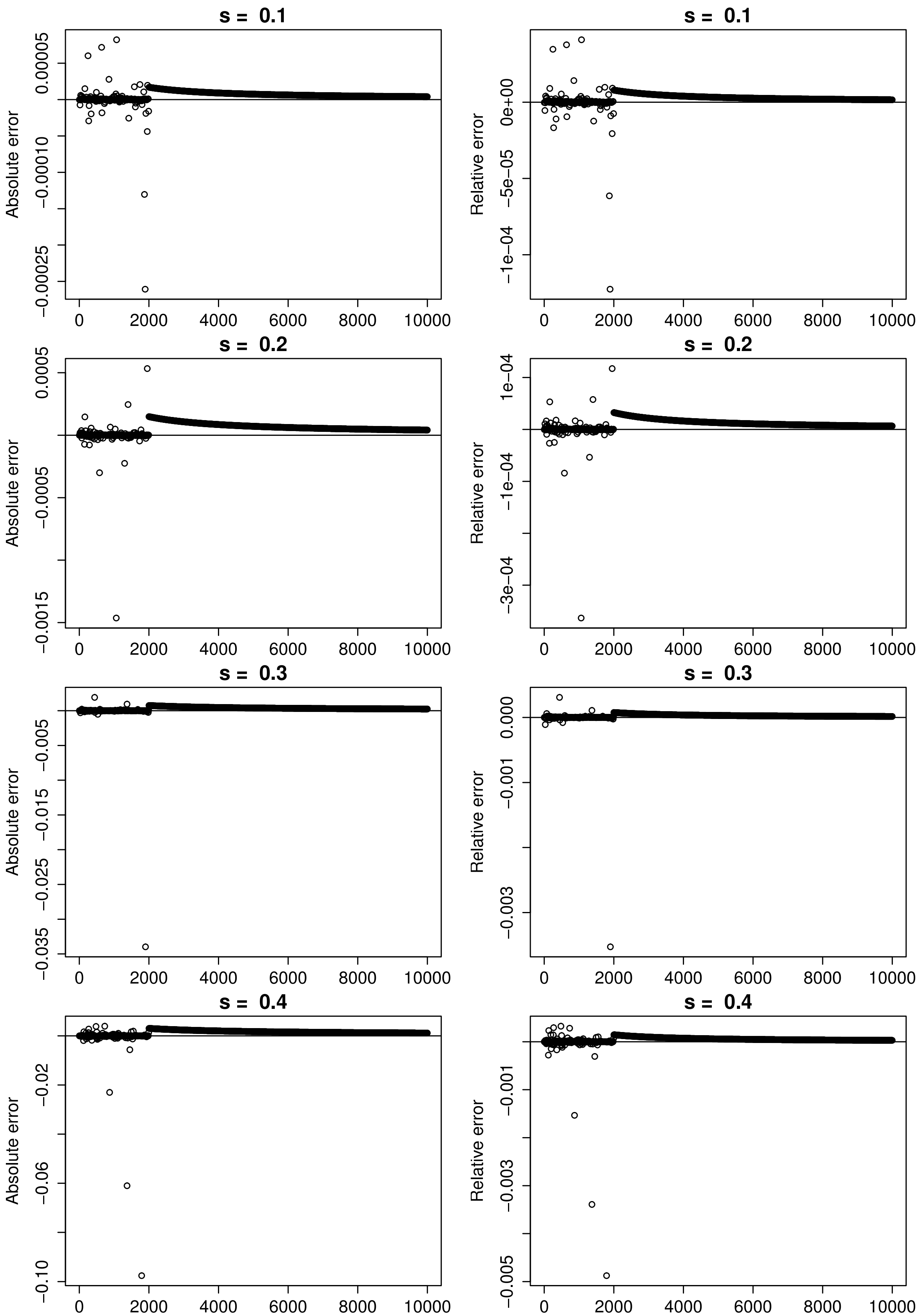

Figure 3.1: Accuracy of the computational formula (3.23) for $E|x - X|^s$ in the case X is standard Cauchy.

number of observations in the test sample increases, the computational cost of simulating the finite-sample distribution of the test statistic will increase dramatically, which may not be acceptable when making real-time decisions, or doing large scale simulations. Therefore, Monte Carlo methods can be adopted for relatively small sample size, and switch to using asymptotic critical values when the sample size is large, provided that the Type I error is controlled.

It has been shown that the statistic $Q_{n,s}/n$ (3.6) is a V-statistic with kernel of degree two and first order degeneracy. With known results in the theory of U-statistics and V-statistics [20, 60], one can obtain the asymptotic distribution of $Q_{n,s}$ and consequently, approximate the critical values for hypothesis testing.

Definition 3.3.1 *Let $h(x_1, \ldots, x_m)$ be a real-valued symmetric measurable kernel, and X_1, $\ldots, X_n$, where $n \geq m$, be independent and identically distributed with distribution function F. A V-statistic is defined as*

$$V_{nm} = \frac{1}{n^m} \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} h(X_{i_1}, \ldots, X_{i_m}). \tag{3.26}$$

The asymptotic distribution of a V-statistic depends on its kernel and the distribution function F, which are summarized and rewritten as follows. The original statement of the theorem and the proof can be found in [50, p.225]. The kernel considered here is symmetric with degree two.

For the kernel $h(x_1, x_2)$, an operator A is defined on $L_2(\mathbb{R}, F)$ by

$$Ag(x) = \int_{-\infty}^{\infty} h(x, y) g(y) dF(y), \ x \in \mathbb{R}, \ g \in L_2,$$

and the associated eigenvalues $\{\lambda_i\}$ are the real numbers satisfying the equation

$$Ag_i(x) = \lambda_i g_i(x), \tag{3.27}$$

where the corresponding $g_i \in L_2$ are called eigenfunctions. The following theorem characterizes the connection between the asymptotic distribution of V_n and its kernel through the eigenvalues of the kernel. Denote $h_1(x_1) = Eh(x_1, X_2)$.

Theorem 3.3.2 *[50, p.225] Let $\{X_i\}$ be independent and identically distributed as F, and $h(x, y)$ a kernel function for which*

1. $h(x, y) = h(y, x)$;
2. $Var_F h_1(X_1) = 0$, *and* $Var_F h(X_1, X_2) > 0$;
3. $E_F h^2(X_1, X_2) < \infty$ *and* $E_F|h(X_1, X_1)| < \infty$;
4. $E_F h(x, X_1)$ is a constant in x.

Put $\mu(F) = E_F h(X_1, X_2)$. If V_n is a V-statistic with kernel $\tilde{h}(x, y) = h(x, y) - \mu(F)$, then

$$nV_n \xrightarrow{d} \sum_{k=1}^{\infty} \lambda_k \chi^2_{1k}, \tag{3.28}$$

where $\chi^2_{1k}(k = 1, 2, \dots)$ are independent chi-square random variables with one degree of freedom, and λ_k's are the eigenvalues of the operator A defined on $L_2(\mathbb{R}, F)$ by

$$Ag(x) = \int_{-\infty}^{\infty} [h(x, y; F) - \mu(F)]g(y)dF(y), \tag{3.29}$$

for $x \in \mathbb{R}$ and $g \in L_2(R, F)$.

As an application of the above theorem, the asymptotic distribution of energy statistic $Q_{n,s}$ is obtained.

Theorem 3.3.3 *If $s < \frac{\alpha}{2}$, then under H_0, $Q_{n,s}$ converges weakly to a weighted sum of iid chi-squared random variables with one degree of freedom; that is,*

$$Q_{n,s} \xrightarrow{d} \sum_{k=1}^{\infty} \lambda_k \chi^2_{1k}, \tag{3.30}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots > 0$ *are solutions to the eigenvalue problem (3.27).*

Proof Since $s < \frac{\alpha}{2}$, by Proposition 3.1.2, the kernel $h(x,y) = E|x-X|^s + E|y-X|^s - E|X-X'|^s - |x-y|^s$ exists for all real-valued x and y. Under H_0, X_1 and X_2 are independent and identical copies of X, hence $E[h(X_1,X_2)] = 2E|X_1-X|^s - E|X_1-X_2|^s - E|X-X'|^s = 0$;. Therefore, we just need to check that the conditions in Theorem 3.3.2 are satisfied for the kernel $h(x,y)$. The kernel $h(x,y)$ has following properties under H_0, which are easy to check:

1. It is obvious that $h(x,y) = h(y,x)$;

2. $h_1(x_1) = Eh(x_1,X_2) \equiv 0$, hence $Var(h_1(X_1)) = 0$; $Var(h(X_1,X_2)) = Var(|X_1 - X_2|^s) > 0$;

3. $E\{h(X_1,X_1)\} = E|X_1 - X|^s < \infty$;

4. When $s < \frac{\alpha}{2}$, $Eh^2(X_1,X_2) = Var[h(X_1,X_2)] = Var(|X_1-X_2|^s) = E|X_1-X_2|^{2s} - (E|X_1-X_2|^s)^2 < \infty$ by Proposition 3.1.2.

To obtain the critical values of $Q_{n,s}$, the distribution of $\sum_{k=1}^{\infty} \lambda_k \chi^2_{1k}$, an infinite sum, is approximated with $\sum_{k=1}^{N} \tilde{\lambda}_k \chi^2_{1k}$, a finite sum, by first numerically solving the eigenvalue problem (3.27). Let $x = u/(1-u^2)$ and $y = v/(1-v^2)$, so that the equation(3.27) becomes

$$\int_{-1}^{1} h(x,y) f(y) g(y) \frac{1+v^2}{(1-v^2)^2} dv = \lambda g(x),$$

where $f(y)$ is the density function of the null distribution. The discretized version of the problem is

$$\sum_{j=1}^{N} h(x_i,x_j) f(x_j) g(x_j) \frac{1+v_j^2}{(1-v_j^2)^2} w_j = \tilde{\lambda} g(x_i).$$

Here $v_1, \ldots, v_N$ are equally spaced points on interval $[-1,1]$, and $x_i = v_i/(1-v_i^2)$. The w_j's are corresponding coefficients of certain quadrature method. This type of treatment for integral equation is called Nyström method [51]. Since both endpoints -1 and 1 are singular points, midpoint quadrature rule can be used.

Note that when the null hypothesis is true, $EQ_{n,s} = E|X - X'|^s = \sum_{k=1}^{\infty} \lambda_i$, thus one can check the convergence of the approximation by comparing $E|X - X'|^s$ and $\sum_{k=1}^{N} \tilde{\lambda}_k \chi_{1k}^2$.

After the approximate eigenvalues are obtained, one can find the approximate p-value by Imhof's method [21]. The R [43] package **CompQuadForm** [11] provides implementation of this method.

3.4 Consistency of the energy test for simple hypotheses

For a goodness-of-fit test of fitting a stable distribution model, consistency should be one of the desirable properties. It is desired that the goodness-of-fit test can eventually reject all alternative hypotheses as more data are collected to increase the power of the goodness-of-test, which is the case when the test is consistent.

Theorem 3.4.1 *Let $F_0 = S(\alpha, \beta, \gamma, \delta)$, and let $q_{n,\xi} = \inf\{x : P(Q_{n,s} > x) \leq \xi\}$ under $H_0 : F = F_0$, for fixed $0 < \xi < 1$. The energy test for simple hypothesis of stability is consistent for all alternatives such that $E|X|^s < \infty$ for some $s > 0$; that is, as $n \to \infty$,*

$$P_F(Q_{n,s} > q_{n,\xi}) \to 1. \tag{3.31}$$

Proof Since $Q_{n,s}$ is continuous, the condition $q_{n,\xi} = \inf\{x : P(Q_{n,s} > x) \leq \xi\}$ is equivalent to $P(Q_{n,s} > q_{n,\xi}) = \xi$. First, we need to show that for fixed α and s, the set $\{q_{n,\xi}\}$ is bounded. Notice that under H_0,

$$
\begin{aligned}
E(Q_{n,s}) =& E\left\{2\sum_{j=1}^{n} E|X_j - X|^s - nE|X - X'|^s - \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}|X_j - X_k|^s\right\} \\
=& 2nE|X - X'|^s - nE|X - X'|^s - \frac{n^2 - n}{n}E|X - X'|^s \\
=& E|X - X'|^s < \infty.
\end{aligned}
$$

Because $Q_{n,s} \geq 0$ for all values of $x_1, \ldots, x_n$,

$$
\begin{aligned}
E(Q_{n,s}) = \int_0^{\infty} P(Q_{n,s} > x)dx > \int_0^{q_{n,\xi}} P(Q_{n,s} > x)dx \\
> \int_0^{q_{n,\xi}} P(Q_{n,s} > q_{n,\xi})dx = \xi q_{n,\xi},
\end{aligned}
$$

therefore,

$$
q_{n,\xi} < \frac{E|X - X'|^s}{\xi}, \text{ for all } n.
$$

On the other hand, let $\hat{f}(t)$ and $\hat{g}(t)$ be the characteristic functions of the alternative distribution and null distribution, respectively. If $\hat{f}_n(t)$ is the empirical characteristic function of a random sample $X_1, \ldots, X_n$, then $\hat{f}_n(t)$ converges to $\hat{f}(t)$ almost surely.

Suppose the alternative distribution from which $X_1, \ldots, X_n$ are sampled is different from the null distribution. Then there exists at least one t_0 such that $\hat{f}(t_0) \neq \hat{g}(t_0)$. Obviously, $t_0 \neq 0$. Because characteristic functions are uniformly continuous, we can find a neighborhood $(a, b)$of t_0, such that $t_0 \in (a, b)$, $0 \notin (a, b)$ and for all $t \in (a, b)$, $\hat{f}(t) \neq \hat{g}(t)$.

Because

$$
\frac{|\hat{f}_n(t) - \hat{g}(t)|^2}{|t|^{s+1}} \leq \frac{(|\hat{f}_n(t)| + |\hat{g}(t)|)^2}{|t|^{s+1}} \leq \frac{4}{|t|^{s+1}}
$$

and

$$\int_a^b \frac{4}{|t|^{s+1}}dt = \left\{ \begin{array}{ll} \frac{4}{s}(a^{-s} - b^{-s}) & \text{if } a > 0 \\ \frac{4}{s}((-b)^{-s} - (-a)^{-s}) & \text{if } b < 0 \end{array} \right\} < \infty,$$

by The Dominated Convergence Theorem, we have,

$$\begin{aligned} \lim_{n\to\infty} \frac{Q_{n,s}}{n} &= \liminf_{n\to\infty} \frac{1}{C(s)} \int_{-\infty}^{\infty} \frac{|\hat{f}_n(t) - \hat{g}(t)|^2}{|t|^{s+1}}dt \\ &\geq \liminf_{n\to\infty} \frac{1}{C(s)} \int_a^b \frac{|\hat{f}_n(t) - \hat{g}(t)|^2}{|t|^{s+1}}dt \\ &= \frac{1}{C(s)} \int_a^b \frac{|\liminf_{n\to\infty} \hat{f}_n(t) - \hat{g}(t)|^2}{|t|^{s+1}}dt, \end{aligned}$$

where

$$C(s) = \frac{2\pi^{\frac{1}{2}}\Gamma(1 - \frac{s}{2})}{s2^s\Gamma(\frac{1+s}{2})}.$$

Because $\hat{f}_n(t)$ converges to $\hat{f}(t)$ almost surely, we have, almost surely,

$$\lim_{n\to\infty} \frac{Q_{n,s}}{n} \geq \frac{1}{C(s)} \int_a^b \frac{|\hat{f}(t) - \hat{g}(t)|^2}{|t|^{s+1}}dt > 0.$$

Therefore, $P_F(Q_{n,s} > q_{n,\xi}) > P_F(Q_{n,s} > \frac{E|X-X'|^s}{\xi}) \to 1$, as $n \to \infty$. ∎

3.5 Implementation of energy test

To the author's knowledge, there are no existing general methods for deriving the exact distribution of the finite-sample test statistic $Q_{n,s}$. This situation is similar to other EDF-based or ECF-based goodness-of-fit tests reviewed in Chapter 2. Usually, tables of critical values are provided based on simulation. However, with the advance of modern computer technology, fast computation enables us to take advantage of bootstrap methods to simulate the distribution of the test statistic on the fly. Compared with using a table of critical values, bootstrap methods are not restricted by the parameters of the test, such as the sample size n and exponent s for energy test. More importantly, simulation methods can provide a good approximation to the p-value, which is not feasible using tables of critical values. The computational cost of bootstrap methods are acceptable when the sample size is relatively small. On the other hand, when the sample size becomes relatively large, it is questionable to what extent the effort of creating an accurate table of critical values will be compensated by the advantage of an exact level test. If it can be shown that the achieved Type I error rate is close to the nominal significance level, then it is reasonable to use the results in Section 3.3 to either create a table of critical values or compute the approximate p-value directly.

Therefore, we propose to adopt two different schemes to implement the energy test. For small samples ($n < 100$), the test is implemented by parametric bootstrap. For large samples ($n > 300$), the test is implemented by numerical approximation to the asymptotic distributions. For sample sizes between 100 and 300, the choice of method may depend on the trade-off between the computational resources, the control of Type I error and other advantages of each scheme.

An energy test for stable distributions can be implemented with parametric bootstrap as follows:

1. For the test sample $x_1, \ldots, x_n$, fix $s < \frac{\alpha}{2}$, and calculate the test statistic $Q_{n,s}$;

2. Generate m replications of random samples of size n from $S(\alpha, \beta; 1)$, denoted as $\{x_{i1}, \ldots, x_{in}\}$, for $i = 1, \ldots, m$;

3. Compute the test statistic $Q_{n,s}^{(i)}$ for each sample;

4. The empirical p-value is calculated as the proportion of $Q_{n,s}^{(i)}$ which are greater than $Q_{n,s}$.

An energy test by the asymptotic distribution method can be implemented as follows:

1. For the test sample $x_1, \ldots, x_n$, fix $s < \frac{\alpha}{2}$, and calculate the test statistic $Q_{n,s}$;

2. Approximate the eigenvalues of the kernel function $h(x, y)$ by Nyström's method (see page 48);

3. Apply Imhof's method [21] to get the approximate p-value.

In each simulation, the number of rejections (the empirical/approximate p-value is less than the nominal significance level ξ) were counted. It should be expected that the proportion of rejections are very close to the nominal level ξ, if the random sample in each replication is generated from the null distribution.

CHAPTER 4

SIMULATION STUDY

In this chapter, results of extensive simulation studies are reported to investigate the empirical type I error and power of the energy test. Two examples will be discussed: Cauchy and symmetric stable distribution $S(1.5, 0)$. All the simulation studies are implemented in R environment [43]. It is noted that numbers in tables of empirical Type I error and empirical power are percentages.

4.1 Test of Cauchy distribution

In this section, the simulation study results of the energy test are presented when the null distribution is the standard Cauchy distribution. Let $X_1, \ldots, X_n$ be independent and identically distributed random variables, and $x_1, \ldots, x_n$ be a sample of observations correspondingly. Without loss of generality, one can consider testing standard Cauchy distribution. Then the energy goodness-of-fit test statistic (3.6) for the standard Cauchy distribution is

$$Q_{n,s} = 2\sum_{j=1}^{n} \frac{(1+x_j^2)^{s/2}\cos(s\arctan x_j)}{\cos\frac{\pi s}{2}} - \frac{n2^s}{\cos\frac{\pi s}{2}} - \frac{1}{n}\sum_{j,k=1}^{n} |x_j - x_k|^s. \tag{4.1}$$

The test statistic $Q_{n,s}/n = V_n$ is a V-statistic whose kernel is

$$\begin{aligned}h(x_1, x_2) =& E|x_1 - X|^s + E|x_2 - X|^s - E|X - X'|^s - |x_1 - x_2|^s \\ =& \frac{(1 + x_1^2)^{s/2} \cos(s \arctan x_1)}{\cos \frac{\pi s}{2}} + \frac{(1 + x_2^2)^{s/2} \cos(s \arctan x_2)}{\cos \frac{\pi s}{2}} - \frac{2^s}{\cos \frac{\pi s}{2}} - |x_1 - x_2|^s.\end{aligned}$$

The asymptotic distribution of $Q_{n,s}$ then can be described as follows,

$$Q_{n,s} \xrightarrow{d} \sum_{j=1}^{\infty} \lambda_j \chi^2_{1j},$$

where $\{\lambda_j\}$ are eigenvalues of the operator A,

$$Ag(x) = \int_{-\infty}^{\infty} h(x, y) g(y) f(y) dy,$$

and $f(y) = 1/(\pi(1 + y^2))$ is the density function of standard Cauchy distribution. The eigenvalue problem

$$\int_{-\infty}^{\infty} h(x, y) g(y) f(y) dy = \lambda g(x),$$

by transformation $y = u/(1 - u^2)$, can be rewritten as

$$\int_{-1}^{1} h(x, y) f(y) g(y) \frac{1 + u^2}{(1 - u^2)^2} du = \lambda g(x).$$

By numerically approximating the integral, we get an equation in $g(x)$

$$\sum_{j=1}^{q} h(x, y_j) f(y_j) g(y_j) \frac{1 + u_j^2}{(1 - u_j^2)^2} \omega_j = \tilde{\lambda} g(x),$$

where $y_j = u_j/(1 - u_j^2)$, and $u_1, \ldots, u_q$ are so-called collocation points on interval $[-1, 1]$. $\omega_1, \ldots, \omega_q$ are the weights of specific quadrature rule. The equation, when evaluating at

Table 4.1: Asymptotic critical values for testing standard Cauchy, at significance level $\xi = 0.10$. N denotes the number of collocation points used in Nyström's method, and s denotes the exponent in energy statistic 3.6.

s	$N = 500$	$N = 1000$	$N = 2000$
0.1	1.263888	1.253216	1.247987
0.2	1.526356	1.521186	1.518994
0.3	1.853511	1.851547	1.851242
0.4	2.275369	2.277855	2.280661
0.5	2.841562	2.853597	2.863627

$x = y_1, \ldots, y_q$, can be approximated by the linear equation system

$$\sum_{j=1}^{q} h(y_i, y_j) f(y_j) g(y_j) \frac{1 + u_j^2}{(1 - u_j^2)^2} \omega_j = \tilde{\lambda} g(y_i), \text{ for } i = 1, \ldots, q.$$

Since we are only interested in the eigenvalue $\tilde{\lambda}$, rather than the unknown function g, the problem is reduced to finding the eigenvalues of matrix $\tilde{A} = (\tilde{a}_{ij})$, where

$$\tilde{a}_{ij} = h(y_i, y_j) f(y_j) \frac{1 + u_j^2}{(1 - u_j^2)^2} \omega_j.$$

This method of solving integral equations is a type of Nyström's method [51]. Since both -1 and 1 are integrable singular points, midpoint rule will be used to determine u_j's and the weights ω_j's. The critical values for significance level $\xi = 0.10$ are shown in Table 4.1. More collocation points are recommended when making a table of critical values, where high accuracy is preferred. Less points may be used when lower accuracy is required for approximating critical values and fast decisions need to be made. In the table, we used $N = 500, 1000,$ and 2000 collocation points in Nyström's method to approximate the asymptotic distribution of the energy statistic under the null hypothesis.

4.1.1 Alternative distributions

Various alternatives to the standard Cauchy distribution are considered, including those that are often proposed to model financial asset returns. They are all standard with location parameter 0 and scale parameter 1.

Mixtures We consider the mixtures of normal and Cauchy distributions in the form $pN(0,1)+(1-p)C(0,1)$, for $p \in \{0.1, 0.3, 0.7, 0.9\}$. In a table of empirical power, for example, $N10C90$ is used to denote when $p = 0.10$.

Student's t The standard Student's t distribution density function with n degrees of freedom is given by

$$f(x;n) = \frac{n^{-\frac{1}{2}}}{B(\frac{n}{2}, \frac{1}{2})} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}. \tag{4.2}$$

where $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$ is the beta function. We consider Student's t distribution with degree of freedom $n \in \{2, 3, 4, 5, 6, 7, 10\}$.

Stable We consider standard symmetric stable distributions under S_1 parameterization, with varying values of α.

Laplace The Laplace distribution is a symmetric continuous probability distribution with exponential tails. The density function of standard Laplace distribution is given by

$$f(x) = \frac{1}{2}e^{-|x|}. \tag{4.3}$$

Gumbel The Gumbel distribution is used in extreme value theory to predict the chance of rare events. We consider the standard Gumbel distribution, whose density function is given by

$$f(x) = e^{-x}e^{-e^{-x}},\ x > 0. \tag{4.4}$$

4.1.2 Other goodness-of-fit tests for the standard Cauchy distribution used for comparison

The simulation study results of three other tests, including the Kolmogorov-Smirnov (K-S) test, the Anderson-Darling (A-D) test and the Gürtler-Henze (G-H) test, are reported and their performances are compared with energy test. Let $x_1, \ldots, x_n$ be a random sample of size n to be tested for standard Cauchy distribution, and $x_{(1)}, \ldots, x_{(n)}$ be the order statistic of the sample. Let $u_{(i)} = \frac{1}{\pi}\arctan(x_{(i)}) + \frac{1}{2}$ be the sample transformed with the cumulative distribution function of standard Cauchy distribution.

In the simulation study, the Kolmogorov-Smirnov test statistic is computed as

$$D_n = \max_{1\le j\le n} \max\left\{\frac{j}{n} - u_{(j)}, u_{(j)} - \frac{j-1}{n}\right\}.$$

An exact p-value is computed if the sample size is less than 100, as described in [31]. Otherwise, the asymptotic distribution is used to compute the p-value. The asymptotic distribution function of K-S statistic D_n under null hypothesis [31] is

$$\lim_{n\to\infty} \Pr(\sqrt{n}D_n \le x) = 1 - 2\sum_{j=1}^{\infty}(-1)^{j-1}e^{-2j^2x^2} = \frac{\sqrt{2\pi}}{x}\sum_{j=1}^{\infty} e^{-(2j-1)^2\pi^2/(8x^2)}.$$

The K-S tests are implemented using the **ks.test** function in the **stats** [43] package in **R**.

The Anderson-Darling test statistic is computed as

$$A_n = -n - \frac{1}{n}\left[\sum_{j=1}^{n}(2j-1)(\log u_{(j)} + \log(1 - u_{(n-j+1)}))\right].$$

The p-values of the A-D test are obtained based on asymptotic distribution of the test statistic. As with the energy statistic, the A-D test statistic converges weakly to a weighted sum of independent squared standard normal random variables. More efficient methods were developed for computing the tail probability of its asymptotic distribution [30]. Marsaglia's method [30] is adopted in the simulation study to compute the p-value of the A-D test, and implemented using the **ad.test** function in the **ADGofTest** [4] package in **R**.

The Gürtler-Henze test statistic is computed as

$$D_{n,\kappa} = \frac{2}{n}\sum_{j,k=1}^{n}\frac{\kappa}{\kappa^2+(x_j-x_k)^2} - 4\sum_{j=1}^{n}\frac{1+\kappa}{(1+\kappa)^2+x_j^2} + \frac{2n}{2+\kappa}.$$

In testing the simple hypothesis of the standard Cauchy distribution, we use $\kappa = 5$ in the G-H test, because it was found in [19, 33] that the test had the best overall performance when $\kappa = 5$. The critical values for significance level $\xi = 0.10$ were generated by simulation with 10^5 replications, and are provided in Table 4.2.

Table 4.2: Critical values for the Gürtler-Henze test generated by simulation with 10^5 replications, significance level $\xi = 0.10$, $\kappa = 5$.

	20	50	100	200
$\kappa = 5$	0.212449	0.211303	0.209615	0.210239

4.1.3 Results of implementation with asymptotic critical values

Simulation studies are carried out to evaluate the performance of the energy goodness-of-fit test. For each combination of the alternative distribution, sample size n and exponent s, 10^5 replicates of the test statistic are generated, and compared with the asymptotic critical value calculated by Nyström's method with 2000 collocation point (see Table 4.1, significance level $\xi = 0.10$). The percentages of rejections are reported as the empirical power of the test against corresponding alternative distribution in Table 4.3 to Table 4.7.

We can see that for most values of s, the actual type I error rate is very close to the nominal significance level ($\xi = 0.10$) when the sample size is > 50. Surprisingly, the actual type I error rate is also well controlled when s is small, even when the sample size is as small as 20. However, when s gets larger, the test is no longer consistent with the nominal level, though the error rate is very close to the nominal level. The special case here is when

Table 4.3: Empirical power of energy test in testing the standard Cauchy distribution, using asymptotic critical values

Alternative	$s = 0.1$				$s = 0.2$				$s = 0.3$				$s = 0.4$				$s = 0.5$			
Sample size n	20	50	100	200	20	50	100	200	20	50	100	200	20	50	100	200	20	50	100	200
Cauchy	11	10	9	10	11	10	11	10	13	10	11	10	13	12	11	10	13	12	12	12
N0C100	10	10	10	10	11	11	10	10	12	11	10	11	12	12	11	11	13	13	12	12
N10C90	9	9	10	13	10	9	11	13	10	10	11	13	10	10	11	13	11	11	11	13
N30C70	8	16	27	52	9	15	28	55	9	15	28	55	8	15	28	56	8	14	27	54
N50C50	12	32	67	96	12	35	68	97	11	34	71	97	10	32	71	97	8	29	68	96
N70C30	19	65	97	100	19	66	97	100	18	67	98	100	15	66	98	100	12	63	98	100
N90C10	33	94	100	100	33	96	100	100	30	96	100	100	27	96	100	100	21	96	100	100
N100C0	42	100	100	100	41	100	100	100	41	100	100	100	36	100	100	100	27	100	100	100
normal	42	100	100	100	43	100	100	100	41	100	100	100	36	100	100	100	28	100	100	100
$t(2)$	10	30	68	98	10	32	73	99	9	33	76	99	7	32	78	100	5	28	78	100
$t(3)$	16	58	96	100	16	62	98	100	14	65	99	100	12	65	99	100	9	63	99	100
$t(4)$	21	74	100	100	20	79	100	100	18	82	100	100	15	83	100	100	11	82	100	100
$t(5)$	23	84	100	100	24	88	100	100	22	90	100	100	18	91	100	100	13	90	100	100
$t(6)$	26	89	100	100	26	92	100	100	25	94	100	100	21	94	100	100	15	95	100	100
$t(7)$	28	92	100	100	28	95	100	100	27	96	100	100	23	97	100	100	17	97	100	100
$t(10)$	31	96	100	100	31	98	100	100	29	99	100	100	26	99	100	100	20	99	100	100
stable(0.5,0)	76	97	100	100	81	98	100	100	84	99	100	100	86	99	100	100	88	99	100	100
stable(0.8,0)	22	28	40	61	26	33	47	68	30	39	53	73	34	43	57	77	34	46	61	81
stable(1.2,0)	8	10	14	27	8	10	16	30	7	10	15	31	7	8	14	32	6	8	14	32
stable(1.5,0)	10	23	55	94	9	24	60	97	8	23	63	97	6	22	65	98	5	19	65	99
stable(1.8,0)	13	47	95	100	13	50	97	100	11	51	98	100	9	51	99	100	6	48	99	100
Laplace	21	70	99	100	20	75	100	100	19	77	100	100	16	78	100	100	12	76	100	100
Gumbel	45	100	100	100	47	100	100	100	46	100	100	100	44	100	100	100	38	100	100	100

$s = \frac{\alpha}{2} = \frac{1}{2}$, the condition for convergence to asymptotic distribution is not satisfied, and Nyström's method may fail to produce the correct asymptotic distribution. Therefore, $s = \frac{\alpha}{2}$ is not recommended when choosing exponent in energy statistic.

Table 4.4: Empirical power of testing standard Cauchy. Energy tests use asymptotic critical values. $n = 20$.

Alternative	$s = 0.1$ [a]	$s = 0.2$	$s = 0.3$	$s = 0.4$	$s = 0.5$	G-H	A-D	K-S
Cauchy	11	11	13	13	13	10	10	10
N10C90	9	10	10	10	11	8	9	9
N30C70	8	9	9	8	8	6	8	9
N50C50	12	12	11	10	8	8	8	9
N70C30	19	19	18	15	12	18	9	11
N90C10	33	33	30	27	21	38	13	11
normal	42	43	41	36	28	57	16	12
$t(2)$	10	10	9	7	5	9	8	9
$t(3)$	16	16	14	12	9	16	8	10
$t(4)$	21	20	18	15	11	23	10	10
$t(5)$	23	24	22	18	13	27	11	11
$t(6)$	26	26	25	21	15	33	11	11
$t(7)$	28	28	27	23	17	34	12	11
$t(10)$	31	31	29	26	20	40	14	11
stable(0.5,0)	76	81	84	86	88	68	54	20
stable(0.8,0)	22	26	30	34	34	23	17	11
stable(1.2,0)	8	8	7	7	6	6	8	9
stable(1.5,0)	10	9	8	6	5	7	8	10
stable(1.8,0)	13	13	11	9	6	11	8	10
Laplace	21	20	19	16	12	21	10	12
Gumbel	45	47	46	44	38	66	48	43

[a] $s = 0.1$ through $s = 0.5$ are energy tests with corresponding exponent s.

4.1.4 Results of implementation with parametric bootstrap

Empirical p-values are calculated as described in Section 3.5, for each simulated sample of size n from the alternative distribution, with 200 replicates of random sample of size n generated from the standard Cauchy distribution. This process is repeated 5000 times for each combination of exponent s and sample size n. The percentage of times when the empirical p-value is less than the significance level $\xi = 0.10$ is called the percentage of

Table 4.5: Empirical power of testing standard Cauchy. Energy tests use asymptotic critical values. $n = 50$.

Alternative	$s = 0.1$ [a]	$s = 0.2$	$s = 0.3$	$s = 0.4$	$s = 0.5$	G-H	A-D	K-S
Cauchy	10	10	10	12	12	10	10	10
N10C90	9	9	10	10	11	9	9	10
N30C70	16	15	15	15	14	16	10	11
N50C50	32	35	34	32	29	39	18	14
N70C30	65	66	67	66	63	77	37	21
N90C10	94	96	96	96	96	99	74	37
normal	100	100	100	100	100	100	93	48
$t(2)$	30	32	33	32	28	47	17	12
$t(3)$	58	62	65	65	63	85	33	18
$t(4)$	74	79	82	83	82	96	46	21
$t(5)$	84	88	90	91	90	99	57	25
$t(6)$	89	92	94	94	95	100	63	28
$t(7)$	92	95	96	97	97	100	68	30
$t(10)$	96	98	99	99	99	100	77	34
stable(0.5,0)	97	98	99	99	99	93	84	35
stable(0.8,0)	28	33	39	43	46	32	19	12
stable(1.2,0)	10	10	10	8	8	10	9	10
stable(1.5,0)	23	24	23	22	19	33	12	10
stable(1.8,0)	47	50	51	51	48	72	17	11
Laplace	70	75	77	78	76	94	46	24
Gumbel	100	100	100	100	100	100	99	98

[a] $s = 0.1$ through $s = 0.5$ are energy tests with corresponding exponent s.

Table 4.6: Empirical power of testing standard Cauchy. Energy tests use asymptotic critical values. $n = 100$.

Alternative	$s = 0.1$ [a]	$s = 0.2$	$s = 0.3$	$s = 0.4$	$s = 0.5$	G-H	A-D	K-S
Cauchy	9	11	11	11	12	10	10	9
N10C90	10	11	11	11	11	10	9	9
N30C70	27	28	28	28	27	31	17	13
N50C50	67	68	71	71	68	76	43	24
N70C30	97	97	98	98	98	99	85	54
N90C10	100	100	100	100	100	100	100	92
normal	100	100	100	100	100	100	100	99
$t(2)$	68	73	76	78	78	89	44	19
$t(3)$	96	98	99	99	99	100	85	37
$t(4)$	100	100	100	100	100	100	97	55
$t(5)$	100	100	100	100	100	100	99	67
$t(6)$	100	100	100	100	100	100	100	75
$t(7)$	100	100	100	100	100	100	100	82
$t(10)$	100	100	100	100	100	100	100	90
stable(0.5,0)	100	100	100	100	100	99	99	62
stable(0.8,0)	40	47	53	57	61	44	26	12
stable(1.2,0)	14	16	15	14	14	19	11	9
stable(1.5,0)	55	60	63	65	65	74	24	11
stable(1.8,0)	95	97	98	99	99	100	55	14
Laplace	99	100	100	100	100	100	95	50
Gumbel	100	100	100	100	100	100	100	100

[a] $s = 0.1$ through $s = 0.5$ are energy tests with corresponding exponent s.

Table 4.7: Empirical power of testing standard Cauchy. Energy tests use asymptotic critical values. $n = 200$.

Alternative	$s = 0.1$ [a]	$s = 0.2$	$s = 0.3$	$s = 0.4$	$s = 0.5$	G-H	A-D	K-S
Cauchy	10	10	10	10	12	11	10	9
N10C90	13	13	13	13	13	14	11	10
N30C70	52	55	55	56	54	61	33	21
N50C50	96	97	97	97	96	98	84	59
N70C30	100	100	100	100	100	100	100	97
N90C10	100	100	100	100	100	100	100	100
normal	100	100	100	100	100	100	100	100
$t(2)$	98	99	99	100	100	100	91	46
$t(3)$	100	100	100	100	100	100	100	90
$t(4)$	100	100	100	100	100	100	100	99
$t(5)$	100	100	100	100	100	100	100	100
$t(6)$	100	100	100	100	100	100	100	100
$t(7)$	100	100	100	100	100	100	100	100
$t(10)$	100	100	100	100	100	100	100	100
stable(0.5,0)	100	100	100	100	100	100	100	95
stable(0.8,0)	61	68	73	77	81	67	38	15
stable(1.2,0)	27	30	31	32	32	40	15	10
stable(1.5,0)	94	97	97	98	99	99	67	21
stable(1.8,0)	100	100	100	100	100	100	100	66
Laplace	100	100	100	100	100	100	100	96
Gumbel	100	100	100	100	100	100	100	100

[a] $s = 0.1$ through $s = 0.5$ are energy tests with corresponding exponent s.

rejections, or empirical power if the alternative distribution is other than standard Cauchy. The small bootstrap sample was used because of the huge computation time required for parametric bootstrap. In the mean time, since the empirical p-value is actually an unbiased estimator of the true p-value, using a large number of simulation replicates, the empirical power can give a satisfactory estimate of actual test power. In general, when only one test decision is required, a large bootstrap sample should be used. The standard error of the empirical power obtained by 5000 replications is $\sqrt{\frac{p(1-p)}{5000}} \leq \frac{0.5}{\sqrt{5000}} = 0.0071$, where p is the probability that a random test sample generated from alternative distribution is rejected by energy test implemented with 200-replicate bootstrap. Therefore, the empirical results are only accurate up to two decimal places, and it is reasonable to display the results as percentages.

From the tables of empirical power, it is apparent that the power of energy test increases very fast as the sample size increases. The null distribution, standard Cauchy, can be regarded as a special case of normal-Cauchy mixture defined on page 49 when $p = 1$, or a special case of Student's t distribution of one degree of freedom, or a special case of stable distribution with $\alpha = 1$ and $\beta = 0$. The results show a distinct trend that the test power increases as the alternative distribution becomes "more" different from the null distribution.

It is noted that for testing the standard Cauchy distribution, the energy test and the Gürtler-Henze test are generally much more powerful than traditional goodness-of-fit tests such as the Kolmogorov-Smirnov test or the Anderson-Darling test, regardless of the sample size and the alternative distribution. The power of the energy test and the Gürtler-Henze test are comparable for most alternatives. While the Gürtler-Henze test is more powerful in detecting alternatives with lighter tails such as Student's t and stable distributions with $\alpha > 1$, the energy test is more sensitive in identifying alternatives with heavier tails such as stable distributions with $\alpha < 1$.

Table 4.8: Empirical power of energy test for the standard Cauchy distribution, implemented with bootstrap, $n = 20$.

Alternative	$s = 0.1$	$s = 0.2$	$s = 0.3$	$s = 0.4$
Cauchy	10	10	10	10
stable(0.5,0)	76	81	83	84
stable(0.8,0)	22	25	27	28
stable(1.2,0)	9	7	6	5
stable(1.5,0)	9	8	7	5
stable(1.8,0)	13	13	9	6
stable(2.0,0)	15	15	12	8
normal	42	38	34	27
$t(2)$	10	9	7	5
$t(3)$	16	14	11	9
$t(4)$	21	17	15	12
$t(5)$	24	23	18	14
Laplace	21	19	16	12
Gumbel	46	45	43	37

Table 4.9: Empirical power of energy test for the standard Cauchy distribution, implemented with bootstrap, $n = 50$.

Alternative	$s = 0.1$	$s = 0.2$	$s = 0.3$	$s = 0.4$
Cauchy	10	9	10	10
stable(0.5,0)	97	98	98	99
stable(0.8,0)	29	32	36	40
stable(1.2,0)	11	10	9	8
stable(1.5,0)	24	25	22	19
stable(1.8,0)	47	50	50	46
stable(2.0,0)	67	71	72	70
normal	100	100	100	100
$t(2)$	31	31	30	28
$t(3)$	57	61	61	59
$t(4)$	75	77	78	77
$t(5)$	85	87	89	89
Laplace	71	73	74	73
Gumbel	100	100	100	100

Table 4.10: Empirical power of energy test for the standard Cauchy distribution, implemented with bootstrap, $n = 100$.

Alternative	$s = 0.1$	$s = 0.2$	$s = 0.3$	$s = 0.4$
Cauchy	10	11	10	10
stable(0.5,0)	100	100	100	100
stable(0.8,0)	40	45	50	53
stable(1.2,0)	16	16	16	13
stable(1.5,0)	57	60	62	61
stable(1.8,0)	95	97	98	98
stable(2.0,0)	100	100	100	100
normal	100	100	100	100
$t(2)$	66	72	74	76
$t(3)$	97	98	99	99
$t(4)$	100	100	100	100
$t(5)$	100	100	100	100
Laplace	99	100	100	100
Gumbel	100	100	100	100

Table 4.11: Empirical power of energy test for the standard Cauchy distribution, implemented with bootstrap, $n = 200$.

Alternative	$s = 0.1$	$s = 0.2$	$s = 0.3$	$s = 0.4$
Cauchy	10	10	10	9
stable(0.5,0)	100	100	100	100
stable(0.8,0)	62	67	73	76
stable(1.2,0)	28	30	32	31
stable(1.5,0)	95	97	98	98
stable(1.8,0)	100	100	100	100
stable(2.0,0)	100	100	100	100
normal	100	100	100	100
$t(2)$	98	99	99	100
$t(3)$	100	100	100	100
$t(4)$	100	100	100	100
$t(5)$	100	100	100	100
Laplace	100	100	100	100
Gumbel	100	100	100	100

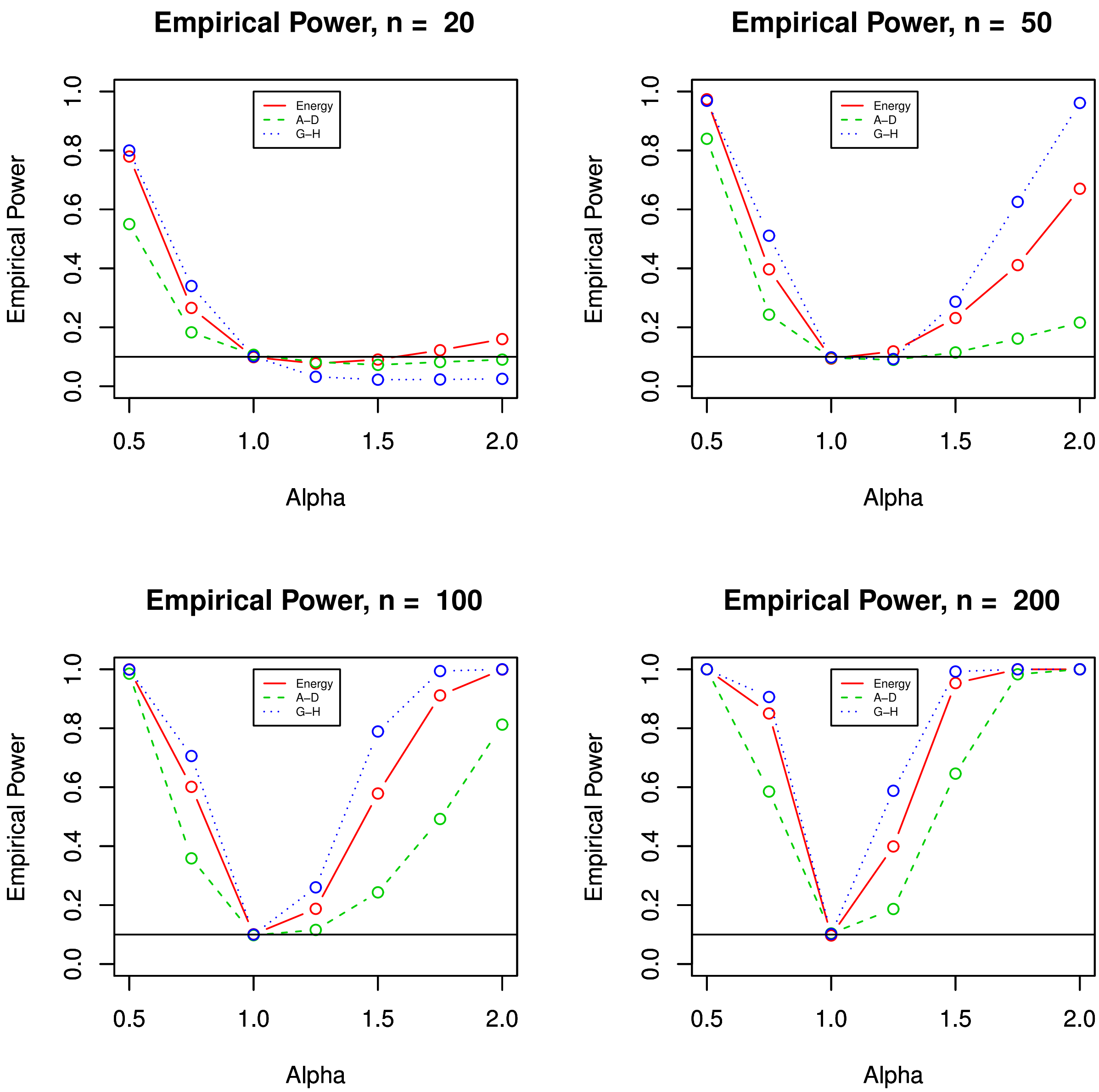

Figure 4.1: Empirical power of testing standard Cauchy against $S(\alpha, 0)$ with varying tail index α, implemented with bootstrap.

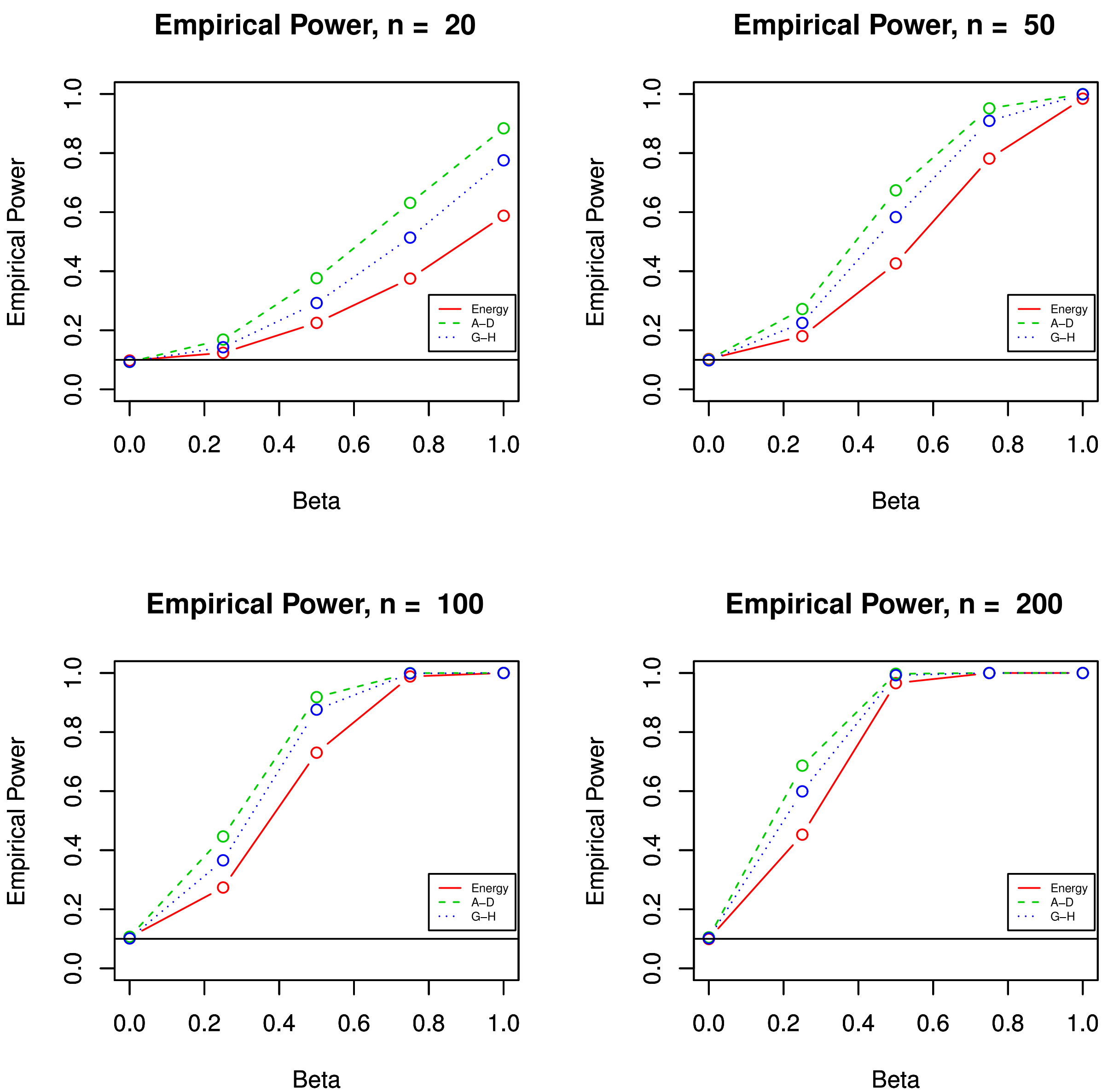

Figure 4.2: Empirical power of testing standard Cauchy against $S(1, \beta)$ with varying skewness parameter β, implemented with bootstrap.

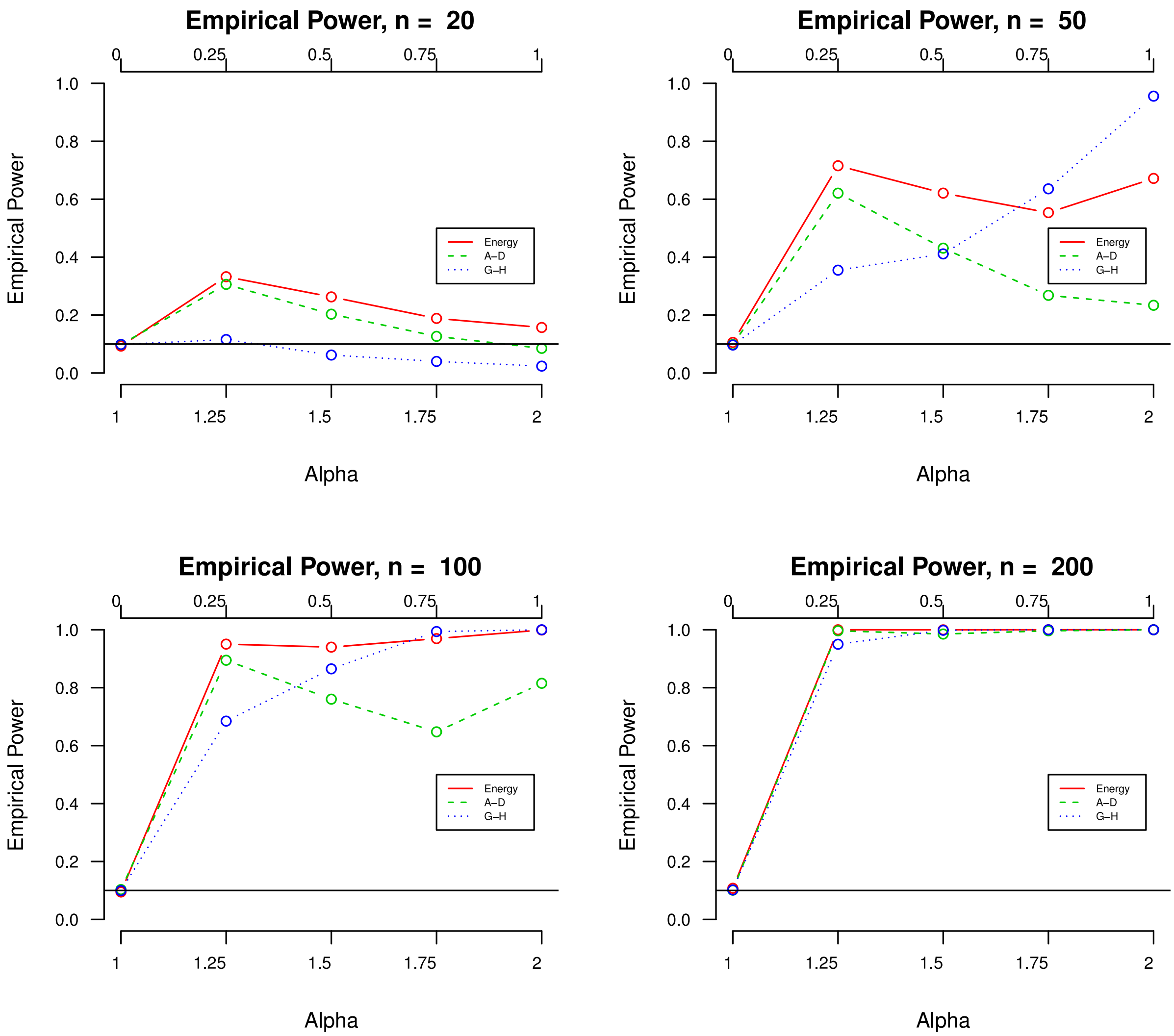

Figure 4.3: Empirical power of testing standard Cauchy against $S(1, \beta)$ with varying tail index α and skewness parameter β simultaneously, implemented with bootstrap.

4.2 Test of Symmetric stable distribution $S(1.5, 0)$

The test statistic (3.6) when the null hypothesis is symmetric stable distribution has the following form:

$$Q_{n,s} = 2\sum_{j=1}^{n}\int_{0}^{\infty}\frac{1-e^{-t^{\alpha}}\cos(x_j t)}{t^{1+s}}dt - \frac{n}{\pi}2^{\frac{s+\alpha}{\alpha}}\Gamma(1-\frac{s}{\alpha})\Gamma(s)\sin\frac{\pi}{2}s - \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}|x_j - x_k|^s.$$

4.2.1 Results of implementation with asymptotic critical values

In this section, we will show the simulation results of the energy test implemented with asymptotic critical values. With the same scheme as in the Cauchy case, we obtained the asymptotic critical values for significance level $\xi = 0.10$ and different values of s, as illustrated by Table 4.12. In the table, N is the number of collocation points in Nyström's method.

Table 4.12: Asymptotic critical values of energy test for testing S(1.5,0), significance level $\xi = 0.10$. N denotes the number of collocation points, and s denotes the exponent in energy statistic.

s	N=**100**	N=**500**	N=**1000**	N=**2000**
0.1	1.288338	1.212503	1.202177	1.197269
0.2	1.435955	1.392295	1.387822	1.386023
0.3	1.620209	1.591604	1.589238	1.588731
0.4	1.837044	1.815455	1.814061	1.814344
0.5	2.091995	2.073112	2.072245	2.073322
0.6	2.395514	2.376382	2.375914	2.378132
0.7	2.762651	2.740838	2.740881	2.744964

We used $N = 100, 500, 1000$ and 2000 collocation points, respectively, in Nyström's method of computing eigenvalues of the kernel function. For each critical value, a simulation study was carried out to investigate the type I error rate with 10^4 replications. The standard error of each Type I error reported is $\sqrt{\frac{p(1-p)}{10^4}} < \frac{0.5}{10^2} = 0.005$, where p is the probability that the null hypothesis is rejected in each replication.

Results are shown in Tables 4.13 to 4.16. One can see that when asymptotic critical values calculated with 2000-point Nyström's method are used, the type I error is well controlled for small and medium-sized value of s, even when the sample size is very small. However, it is not recommended to use combination of large values of s with small sample sizes.

Table 4.13: Type I error of energy test for $S(1.5, 0)$ using asymptotic critical values with 100 collocation points.

$s\backslash n$	20	50	100	200
0.1	4	7	10	11
0.2	4	9	10	11
0.3	4	9	11	10
0.4	4	9	11	11
0.5	8	9	10	11
0.6	7	10	10	10
0.7	8	10	12	10

Table 4.14: Type I error of energy test for $S(1.5, 0)$ using asymptotic critical values with 500 collocation points.

$s\backslash n$	20	50	100	200
0.1	8	8	8	8
0.2	10	10	9	10
0.3	10	10	10	10
0.4	11	10	10	10
0.5	11	11	10	10
0.6	12	11	11	11
0.7	12	12	11	11

A simulation study was implemented for the empirical power (percentage of rejections) of the energy test of stable distributions against a collection of alternative distributions, using the asymptotic critical value obtained with 2000 collocation point. The simulation results were summarized in Table 4.17 to Table 4.20. Figure 4.4 to Figure 4.11 better illustrate the comparison between the performance of the energy test, the Anderson-Darling (A-D) test and the Matsui-Takemura (M-T) test. Figure 4.4 to Figure 4.7 illustrated the empirical power of energy test for null distribution $S(1.5, 0)$ against alternatives $S(\alpha, 0)$ with varying tail index α, implemented with asymptotic critical values. Figure 4.8 to Figure 4.11 also illustrated the better performance of the energy test and the M-T test when the alternatives

Table 4.15: Type I error of energy test for $S(1.5, 0)$ using asymptotic critical values with 1000 collocation points.

$s\backslash n$	20	50	100	200
0.1	9	9	9	9
0.2	10	10	10	9
0.3	10	11	10	10
0.4	11	10	10	10
0.5	11	10	11	10
0.6	12	11	10	10
0.7	11	11	10	11

Table 4.16: Type I error of energy test for $S(1.5, 0)$ using asymptotic critical values with 2000 collocation points.

$s\backslash n$	20	50	100	200	1000
0.1	10	9	10	10	10
0.2	10	10	10	10	10
0.3	10	10	10	10	10
0.4	10	10	10	10	10
0.5	11	11	11	10	10
0.6	12	11	10	10	10
0.7	12	11	10	11	10

were Student's t distribution with different degrees of freedom. It was shown that for all sample sizes studied, the two tests based on empirical characteristic functions, the energy test and the M-T test, outperformed the Anderson-Darling test. Similar to what was found in testing the standard Cauchy distribution, the power of the energy test and the M-T test are comparable for most alternatives. While the M-T test is more powerful in detecting alternatives with lighter tails, the energy test is more sensitive in identifying alternatives with heavier tails such as stable distributions with small tail index α.

Future simulation study need to be done to find the optimal value of s. Based on current simulation results, it is recommended to choose s to be about $\alpha/3$.

Table 4.17: Empirical power of energy test for $S(1.5, 0)$, sample size $n = 20$, implemented with asymptotic critical values.

Alternative	$s = 0.1$	$s = 0.2$	$s = 0.3$	$s = 0.4$	$s = 0.5$	$s = 0.6$	$s = 0.7$
Cauchy	30	34	38	43	45	49	51
stable(1.1,0)	22	26	28	32	35	38	39
stable(1.2,0)	17	18	21	23	25	27	29
stable(1.3,0)	13	15	15	16	19	20	22
stable(1.4,0)	11	12	12	13	14	15	16
stable(1.5,0)	10	10	10	10	12	11	11
stable(1.6,0)	8	9	9	9	9	10	10
stable(1.7,0)	9	9	9	8	9	9	8
stable(1.8,0)	9	9	9	8	8	7	7
stable(1.9,0)	9	9	9	8	8	8	7
Normal	26	26	26	24	22	19	16
$t(2)$	11	11	10	11	10	9	9
$t(3)$	11	11	10	10	10	9	7
$t(4)$	14	13	13	12	11	10	8
$t(5)$	15	15	14	14	13	11	9
$t(10)$	20	20	18	18	16	14	13
Laplace	17	17	15	15	13	11	10

4.2.2 Results of implementation with parametric bootstrap

This section reports the simulation results of the energy test for symmetric stable distribution $S(1.5, 0)$ implemented with bootstrap. Due to the computational demands of the bootstrap

Table 4.18: Empirical power of energy test for $S(1.5, 0)$, sample size $n = 50$, implemented with asymptotic critical values.

Alternative	$s = 0.1$	$s = 0.2$	$s = 0.3$	$s = 0.4$	$s = 0.5$	$s = 0.6$	$s = 0.7$
Cauchy	47	53	57	62	67	70	73
stable(1.1,0)	30	35	40	44	48	51	56
stable(1.2,0)	20	23	25	28	32	35	37
stable(1.3,0)	14	16	17	19	20	22	24
stable(1.4,0)	11	12	13	14	14	14	16
stable(1.5,0)	9	10	10	11	11	10	11
stable(1.6,0)	9	9	10	9	9	9	9
stable(1.7,0)	10	9	10	10	9	9	8
stable(1.8,0)	10	11	10	10	9	9	8
stable(1.9,0)	11	11	11	11	11	9	9
Normal	73	76	77	78	78	78	75
$t(2)$	11	11	10	11	10	9	9
$t(3)$	23	25	25	23	21	20	18
$t(4)$	32	33	32	31	31	29	26
$t(5)$	38	39	40	38	39	36	33
$t(10)$	53	56	58	57	57	55	53
Laplace	39	39	38	36	35	33	29

Table 4.19: Empirical power of energy test for $S(1.5, 0)$, sample size $n = 100$, implemented with asymptotic critical values.

Alternative	$s = 0.1$	$s = 0.2$	$s = 0.3$	$s = 0.4$	$s = 0.5$	$s = 0.6$	$s = 0.7$
Cauchy	72	77	82	85	88	89	91
stable(1.1,0)	47	53	57	63	67	71	74
stable(1.2,0)	27	32	36	39	43	46	51
stable(1.3,0)	16	18	21	23	24	28	30
stable(1.4,0)	11	12	13	14	14	16	17
stable(1.5,0)	10	10	10	10	10	10	11
stable(1.6,0)	9	10	10	10	9	10	9
stable(1.7,0)	11	11	11	10	10	10	11
stable(1.8,0)	12	13	12	13	13	12	11
stable(1.9,0)	16	16	16	16	16	15	14
Normal	99	99	100	100	100	100	100
$t(2)$	11	11	10	11	10	9	9
$t(3)$	45	46	46	46	44	43	41
$t(4)$	61	63	63	65	65	63	62
$t(5)$	72	75	77	77	77	78	76
$t(10)$	91	93	94	95	95	95	95
Laplace	69	70	69	69	68	66	64

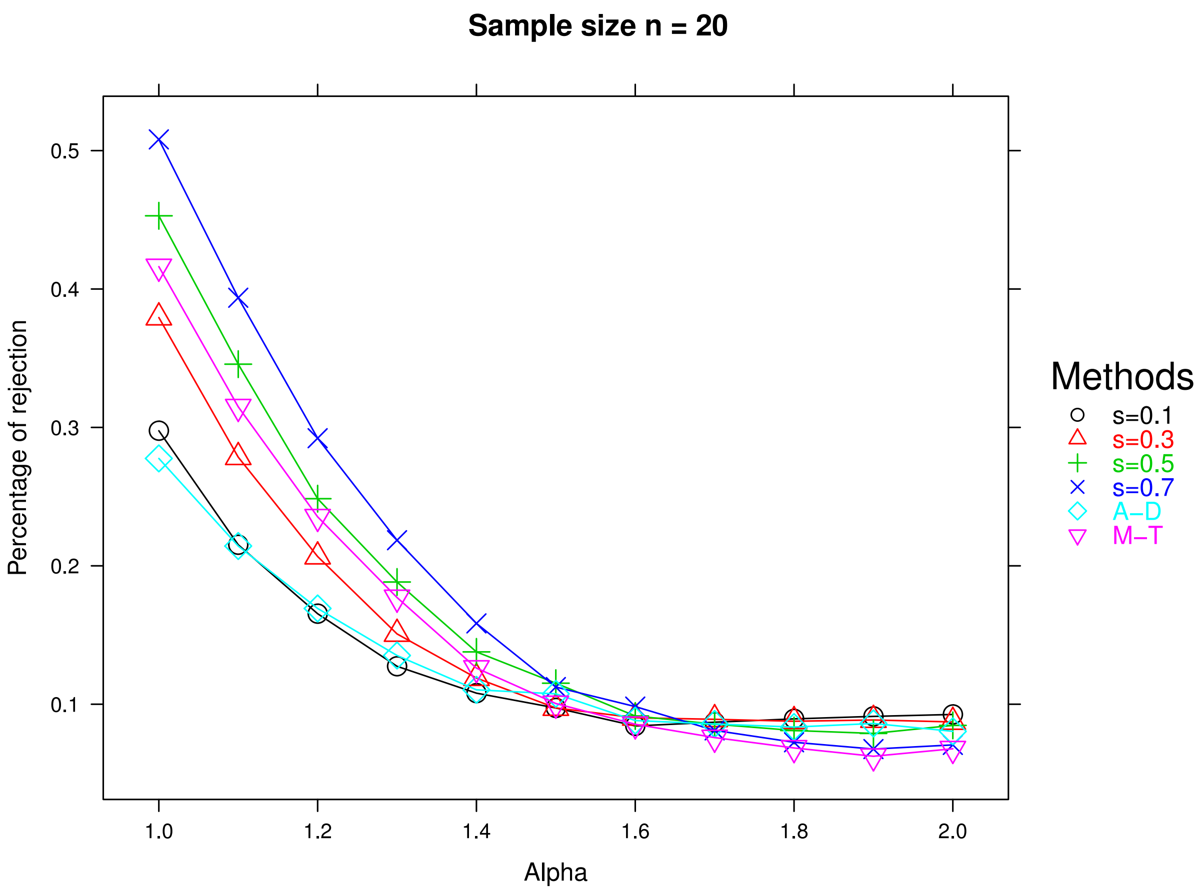

Figure 4.4: Empirical power of energy test for null distribution $S(1.5, 0)$ against $S(\alpha, 0)$ with varying tail index α, implemented with asymptotic critical values. Sample size $n = 20$.

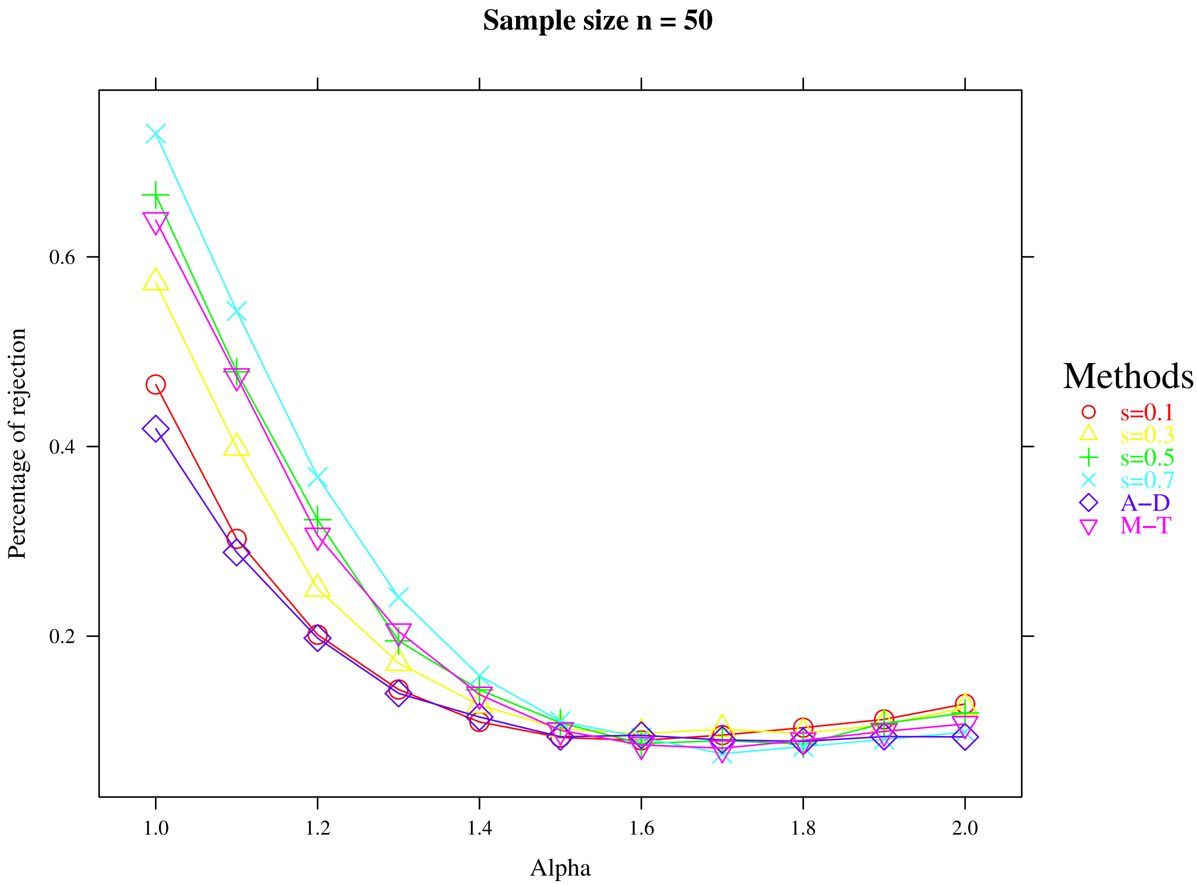

Figure 4.5: Empirical power of energy test for null distribution $S(1.5, 0)$ against $S(\alpha, 0)$ with varying tail index α, implemented with asymptotic critical values. Sample size $n = 50$.

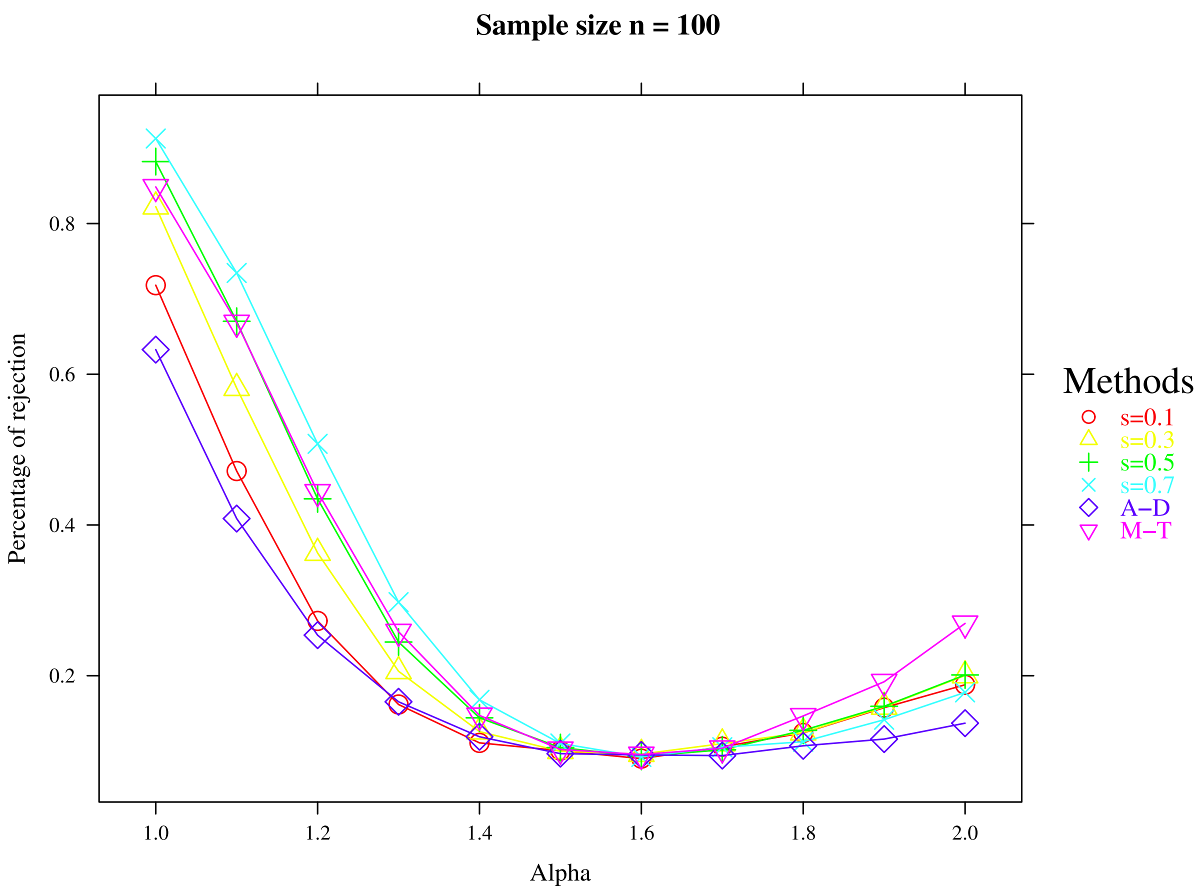

Figure 4.6: Empirical power of energy test for null distribution $S(1.5, 0)$ against $S(\alpha, 0)$ with varying tail index α, implemented with asymptotic critical values. Sample size $n = 100$.

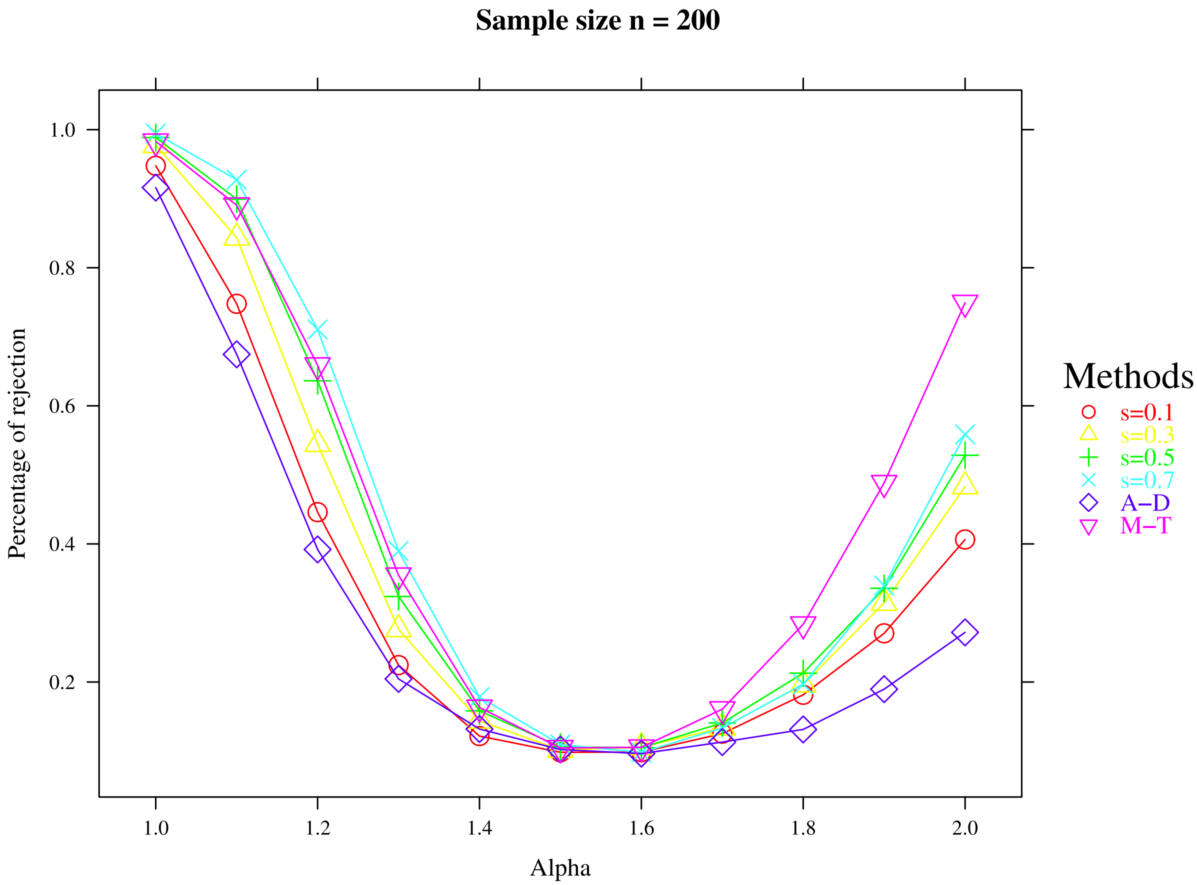

Figure 4.7: Empirical power of energy test for null distribution $S(1.5, 0)$ against $S(\alpha, 0)$ with varying tail index α, implemented with asymptotic critical values. Sample size $n = 200$.

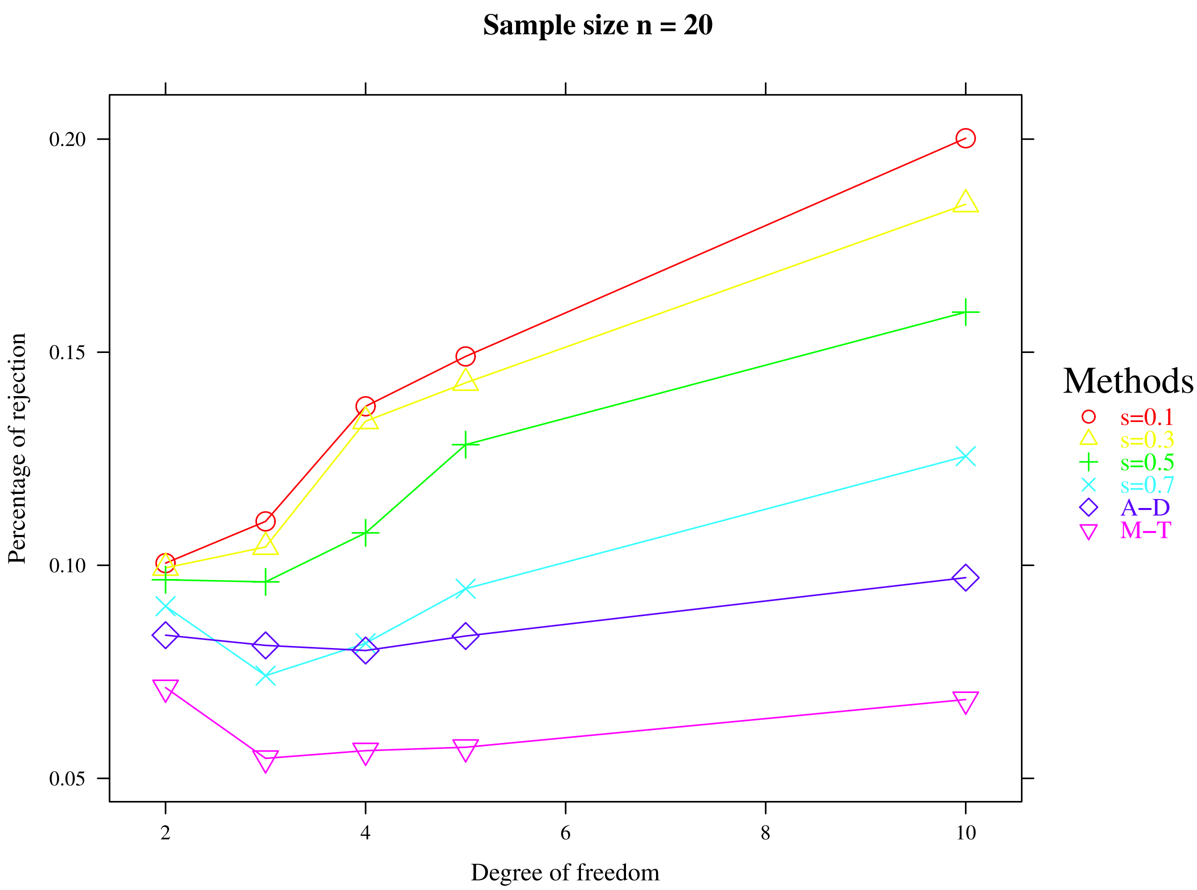

Figure 4.8: Empirical power of energy test for null distribution $S(1.5, 0)$ against Student's t distribution with varying degree of freedom, implemented with asymptotic critical values. Sample size $n = 20$.

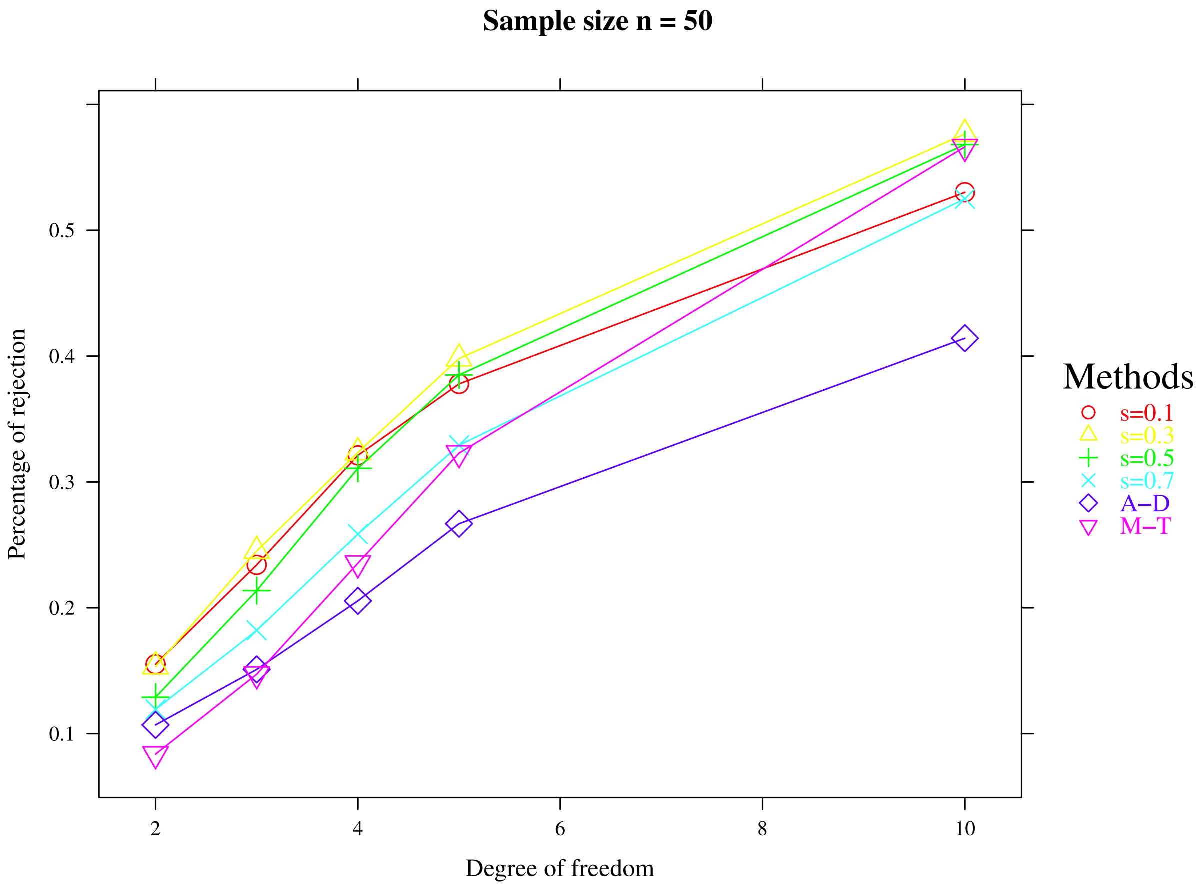

Figure 4.9: Empirical power of energy test for null distribution $S(1.5, 0)$ against Student's t distribution with varying degree of freedom, implemented with asymptotic critical values. Sample size $n = 50$.

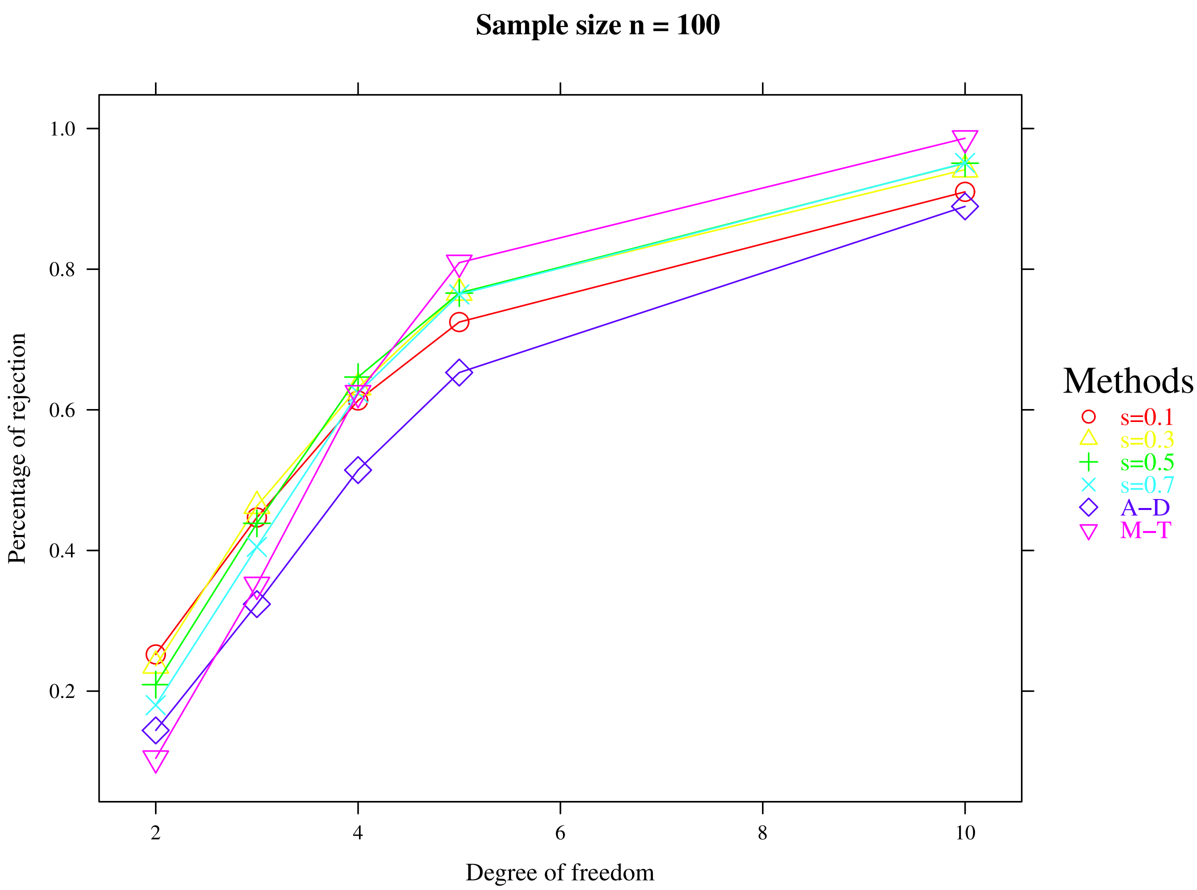

Figure 4.10: Empirical power of energy test for null distribution $S(1.5, 0)$ against Student's t distribution with varying degree of freedom, implemented with asymptotic critical values. Sample size $n = 100$.

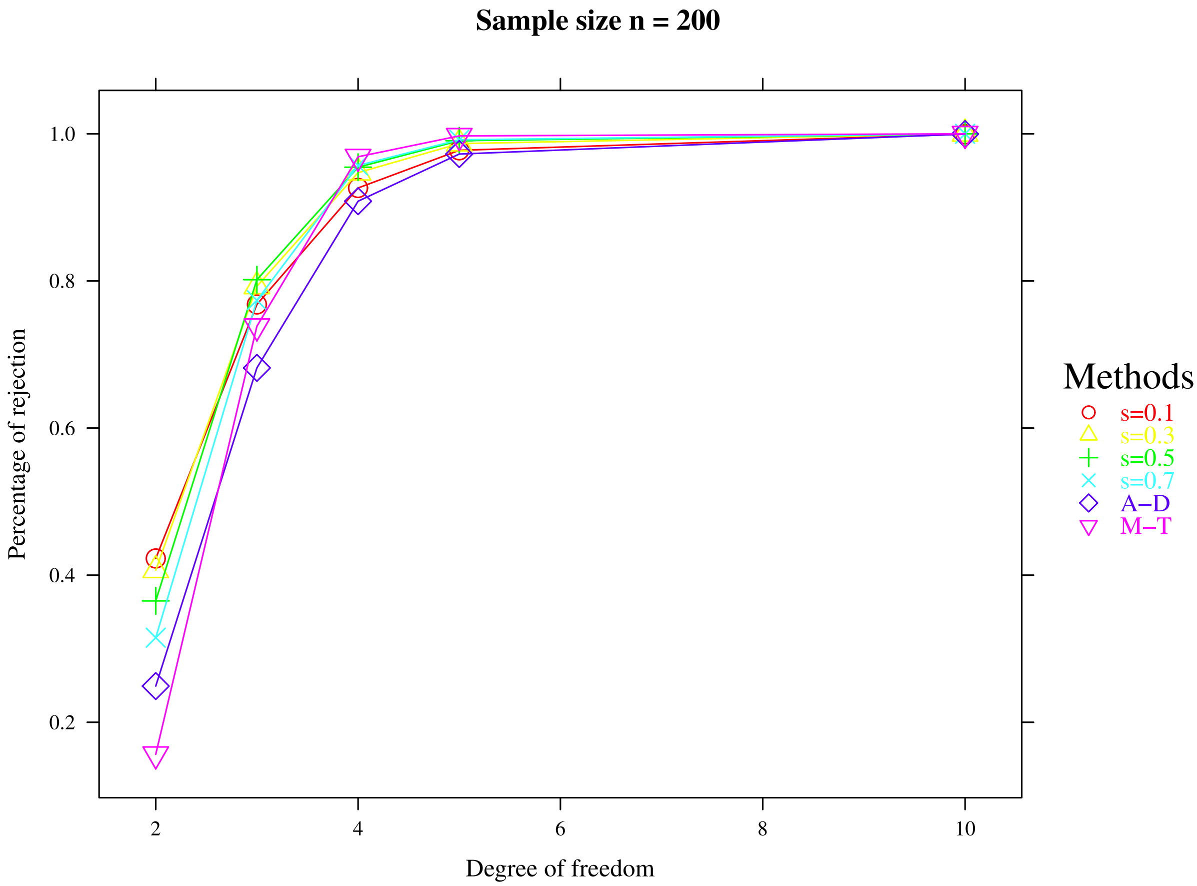

Figure 4.11: Empirical power of energy test for null distribution $S(1.5, 0)$ against Student's t distribution with varying degree of freedom, implemented with asymptotic critical values. Sample size $n = 200$.

Table 4.20: Empirical power of energy test for $S(1.5, 0)$, sample size $n = 200$, implemented with asymptotic critical values.

Alternative	$s = 0.1$	$s = 0.2$	$s = 0.3$	$s = 0.4$	$s = 0.5$	$s = 0.6$	$s = 0.7$
Cauchy	95	97	98	99	99	99	99
stable(1.1,0)	75	80	85	88	90	91	93
stable(1.2,0)	45	50	54	60	64	67	71
stable(1.3,0)	22	25	28	30	32	36	39
stable(1.4,0)	12	14	14	14	16	16	18
stable(1.5,0)	10	10	10	10	11	10	11
stable(1.6,0)	10	11	11	10	10	10	10
stable(1.7,0)	13	14	13	14	14	14	13
stable(1.8,0)	18	21	20	21	21	20	20
stable(1.9,0)	27	30	31	33	34	34	34
Normal	100	100	100	100	100	100	100
$t(2)$	11	11	10	11	10	9	9
$t(3)$	77	79	79	79	80	79	77
$t(4)$	93	94	95	95	95	95	96
$t(5)$	98	98	99	99	99	99	99
$t(10)$	100	100	100	100	100	100	100
Laplace	95	96	96	96	96	96	95

implementation, the simulation study is done only for $s = \frac{1}{8}$. Energy test's performance is compared with that of the Anderson-Darling (A-D) test.

Figure 4.12 shows the empirical power of the tests when the alternative is Student's t distribution with different degree of freedom. The horizontal axis represents degree of freedom in log scale.

Figure 4.13 shows the empirical power of the tests when the alternative is symmetric stable distribution with different values of α.

The empirical results suggest that the energy test is more powerful than the Anderson-Darling test in detecting distributions with lighter and heavier tails than $S(1.5, 0)$.

4.3 Test of asymmetric stable distribution $S(1.8, 0.5)$

In this section, the simulation study results are presented when the null distribution is asymmetric stable distribution $S(1.8, 0.5)$. The alternative distributions considered include

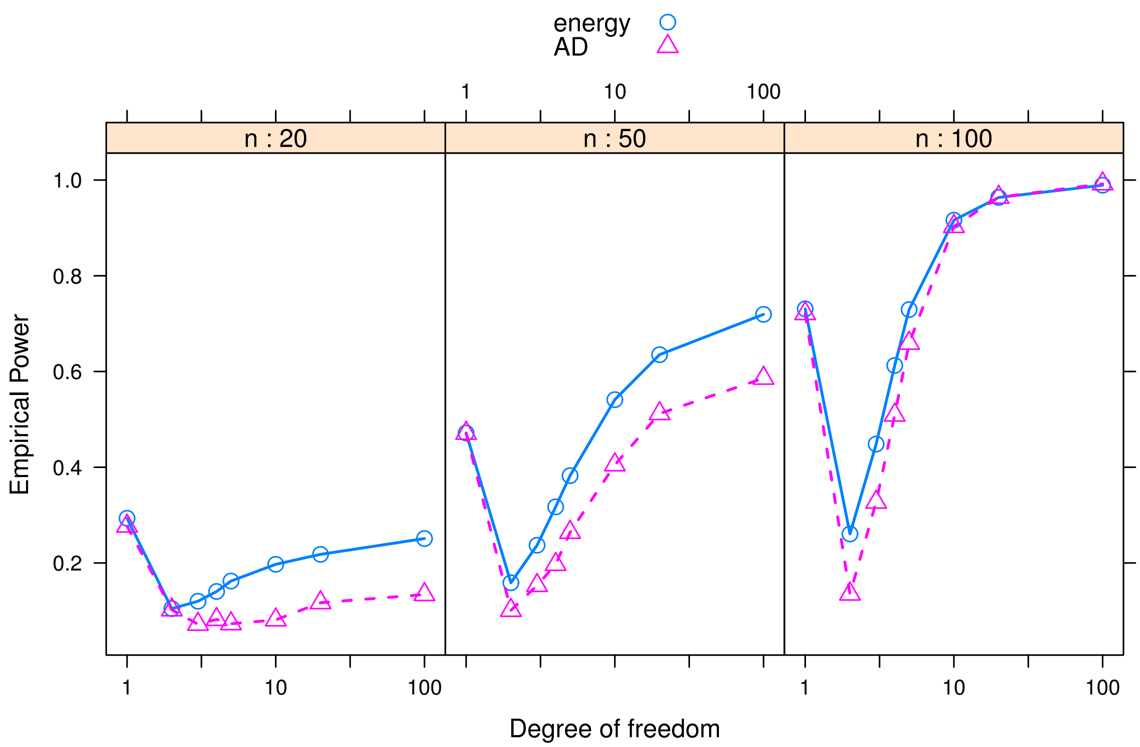

Figure 4.12: Empirical power of testing $S(1.5, 0)$ against Student's t distribution with varying degree of freedom, implemented with bootstrap.

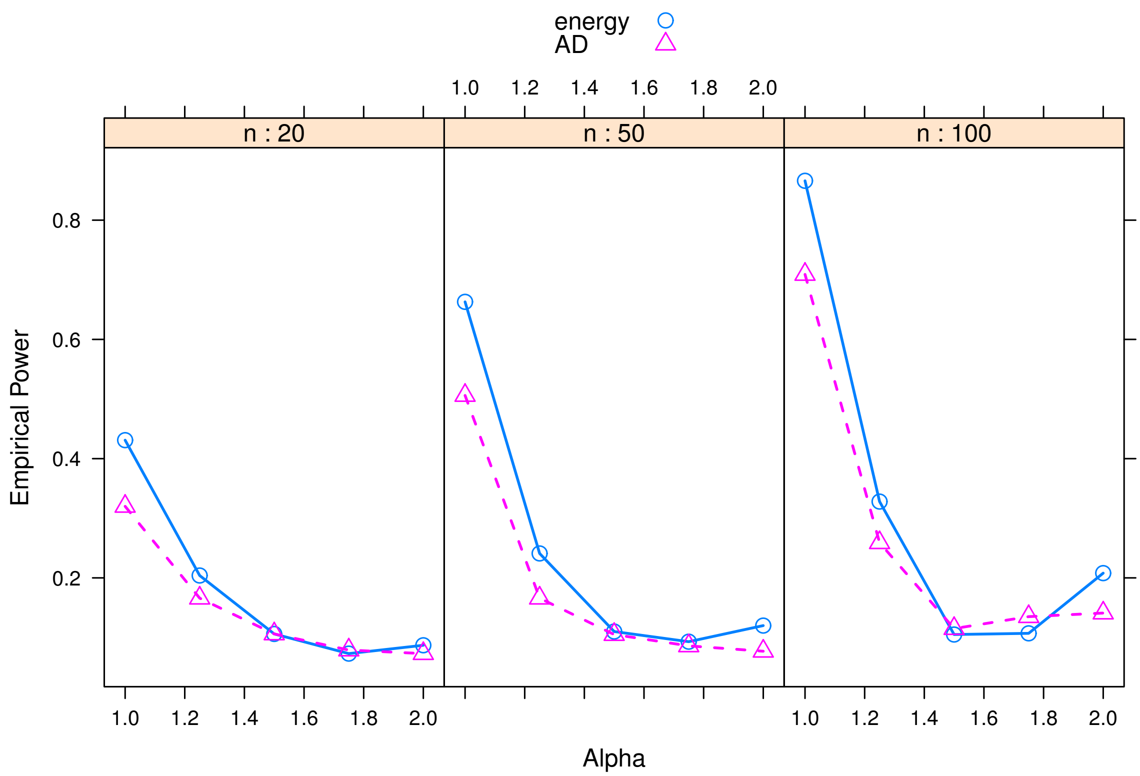

Figure 4.13: Empirical power of testing $S(1.5, 0)$ against symmetric stable distribution with varying α, implemented with bootstrap.

stable distribution with varying skewness parameter β and Student's t distribution with varying degree of freedom. Simulation study was done only for energy tests implemented by asymptotic critical values.

The asymptotic critical values were obtained using Nyström's method with 2000 collocation points, which are shown in Table 4.21, where s is the exponent in the energy statistic.

Table 4.21: Asymptotic critical values for testing $S(1.8, 0.5)$

$s = 0.3$	$s = 0.6$	$s = 0.9$
1.531216	2.153583	3.065132

The empirical power results of the energy test for null distribution $S(1.8, 0.5)$ are summarized in Table 4.22. Figure 4.14 to Figure 4.21 better illustrate the comparison between the performance of the energy test and the Anderson-Darling (A-D) test. Figure 4.14 to Figure 4.17 illustrated the empirical power of energy test for null distribution $S(1.8, 0.5)$ against alternatives $S(1.8, \beta)$ with varying skewness parameter β, implemented with asymptotic critical values. It was shown that for all sample sizes studied, the energy test had higher power than the Anderson-Darling test, and the difference between the two tests may vary by different alternatives. Figure 4.18 to Figure 4.21 also illustrated that the energy test outperformed the Anderson-Darling test when the alternatives were Student's t distribution with different degrees of freedom. In addition, it is noted that the choosing of exponent s in the energy test did not affect the test power, when $s < \alpha/2$ and the alternatives were stable distribution with the same tail index α as the null distribution but with different skewness parameter β. However, when the alternatives are Student's t distribution, the exponent $s = 0.3$ lead to better empirical power results than $s = 0.6$ and $s = 0.9$. Future simulation study need to be done to find the optimal value of s. Based on current simulation results, it is recommended to choose s to be about $\alpha/3$.

Table 4.22: Empirical power of energy test for $S(1.8, 0.5)$, implemented with asymptotic critical values.

	$s = 0.3$				$s = 0.6$				$s = 0.9$			
Alternative \ n	20	50	100	200	20	50	100	200	20	50	100	200
$S(1.1, 0.5)$	100	100	100	100	100	100	100	100	100	100	100	100
$S(1.2, 0.5)$	97	100	100	100	98	100	100	100	98	100	100	100
$S(1.3, 0.5)$	70	96	100	100	74	97	100	100	75	97	100	100
$S(1.4, 0.5)$	41	70	92	100	45	74	94	100	48	75	95	100
$S(1.5, 0.5)$	24	39	60	86	27	43	64	89	30	45	66	90
$S(1.6, 0.5)$	15	20	29	46	18	24	32	49	20	25	33	50
$S(1.7, 0.5)$	12	13	15	18	13	13	15	17	14	14	15	19
$S(1.9, 0.5)$	10	11	12	14	10	11	12	14	9	10	11	13
$S(2.0, 0.5)$	11	13	17	26	10	13	16	25	9	11	14	22
$S(1.8, -1.0)$	23	39	65	90	23	39	62	89	21	35	58	86
$S(1.8, -0.8)$	20	33	54	81	19	32	53	79	18	29	47	76
$S(1.8, -0.5)$	16	24	38	61	16	23	36	59	15	21	33	55
$S(1.8, -0.2)$	12	17	24	36	12	17	22	36	13	16	21	32
$S(1.8, 0)$	11	14	17	25	12	13	17	23	12	13	16	21
$S(1.8, 0.2)$	10	12	13	15	11	11	12	14	12	12	13	14
$S(1.8, 0.5)$	10	10	10	10	10	10	10	10	11	10	11	10
$S(1.8, 0.8)$	10	11	13	15	10	12	13	15	12	12	13	14
$S(1.8, 1.0)$	12	14	17	23	12	13	17	23	12	14	16	22
$t(2)$	16	27	46	77	16	26	44	74	19	25	39	67
$t(3)$	15	27	52	82	13	25	45	77	11	18	35	66
$t(4)$	15	34	62	91	13	29	57	89	10	22	47	84
$t(5)$	17	38	70	96	13	34	67	95	11	27	58	92
$t(6)$	18	43	76	97	15	39	74	97	10	31	66	96
$t(7)$	18	47	80	99	16	43	78	99	11	33	72	98

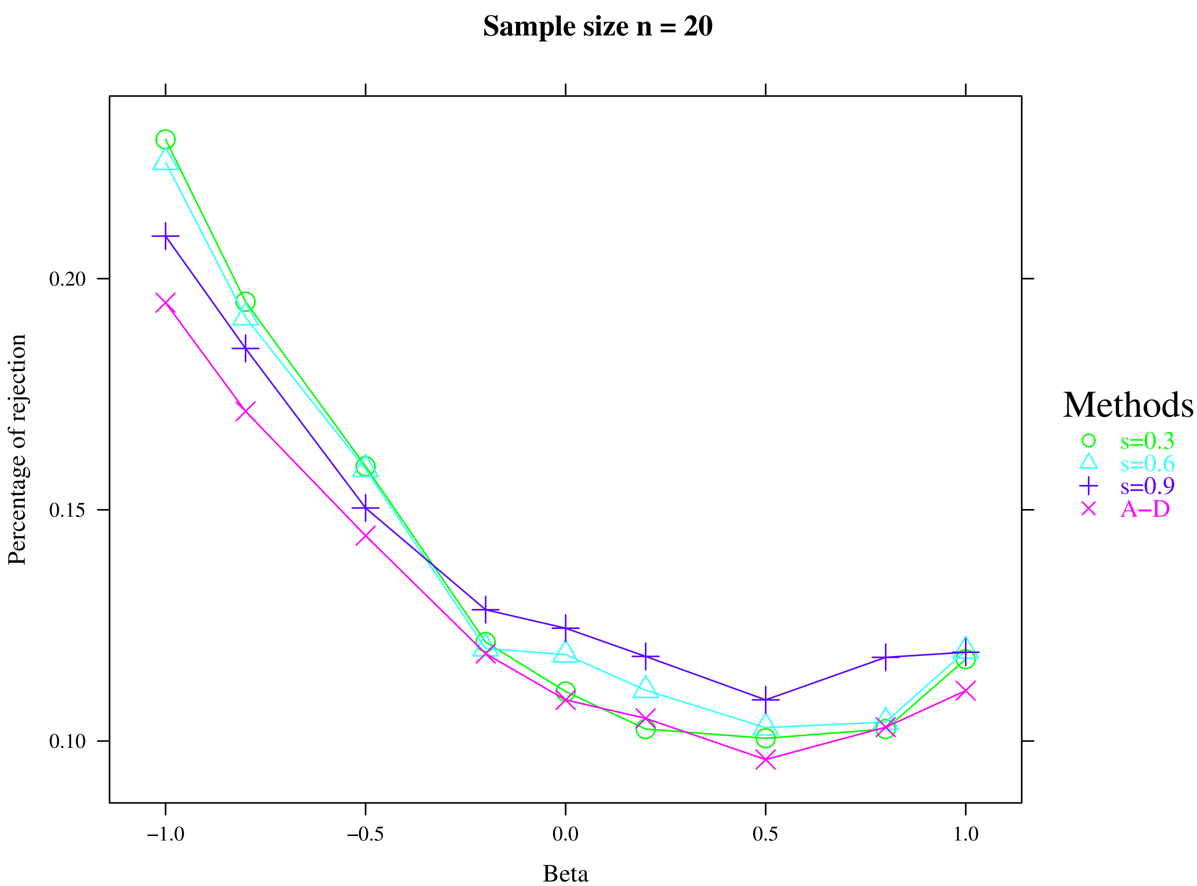

Figure 4.14: Empirical power of energy test for null distribution $S(1.8, 0.5)$ against $S(1.8, \beta)$ with varying skewness parameter β, implemented with asymptotic critical values. Sample size $n = 20$.

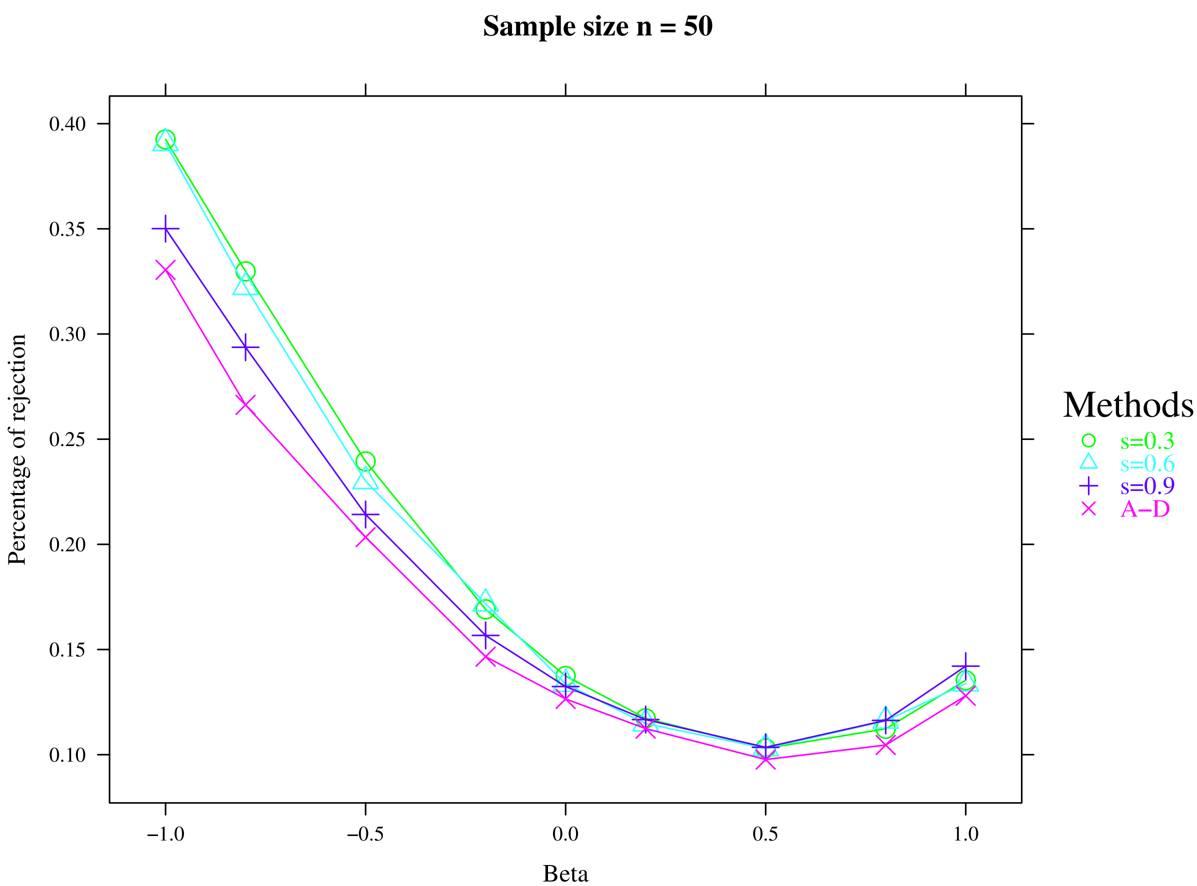

Figure 4.15: Empirical power of energy test for null distribution $S(1.8, 0.5)$ against $S(1.8, \beta)$ with varying skewness parameter β, implemented with asymptotic critical values. Sample size $n = 50$.

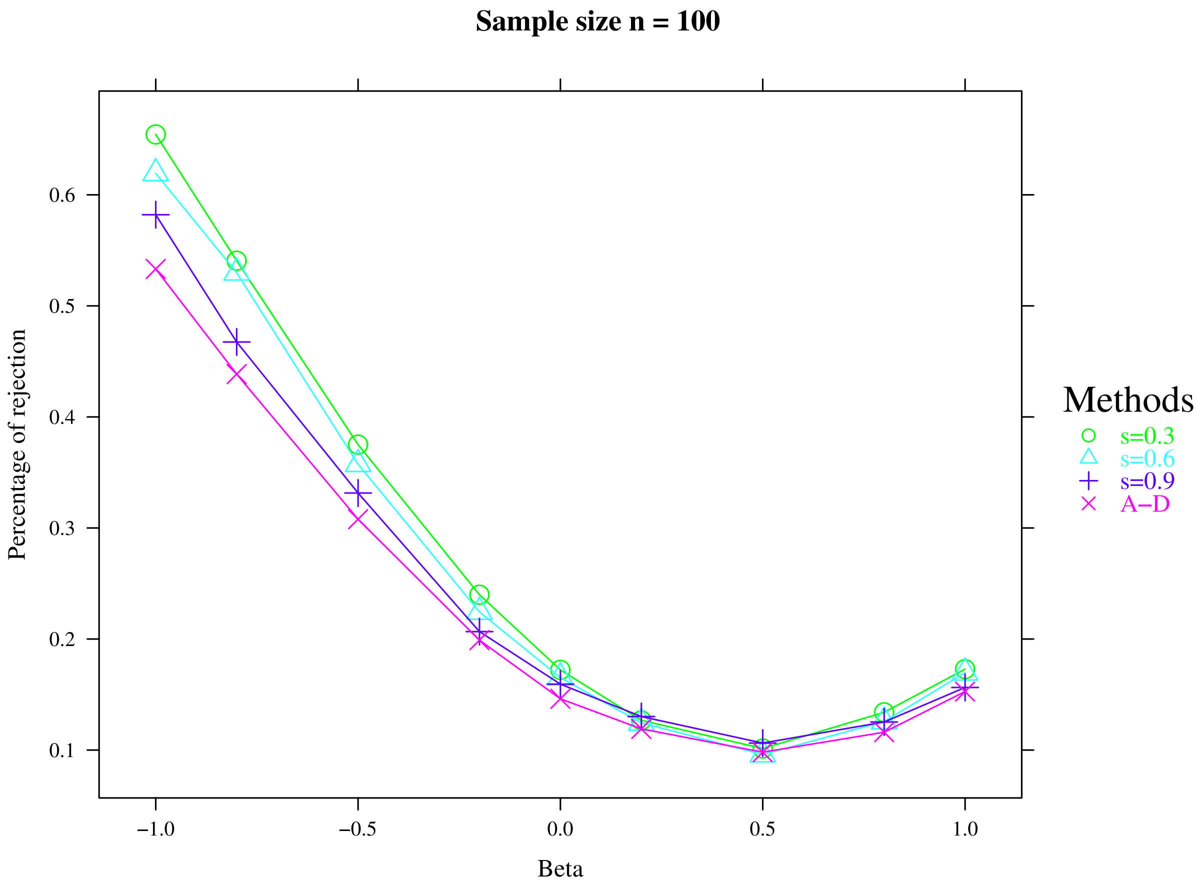

Figure 4.16: Empirical power of energy test for null distribution $S(1.8, 0.5)$ against $S(1.8, \beta)$ with varying skewness parameter β, implemented with asymptotic critical values. Sample size $n = 100$.

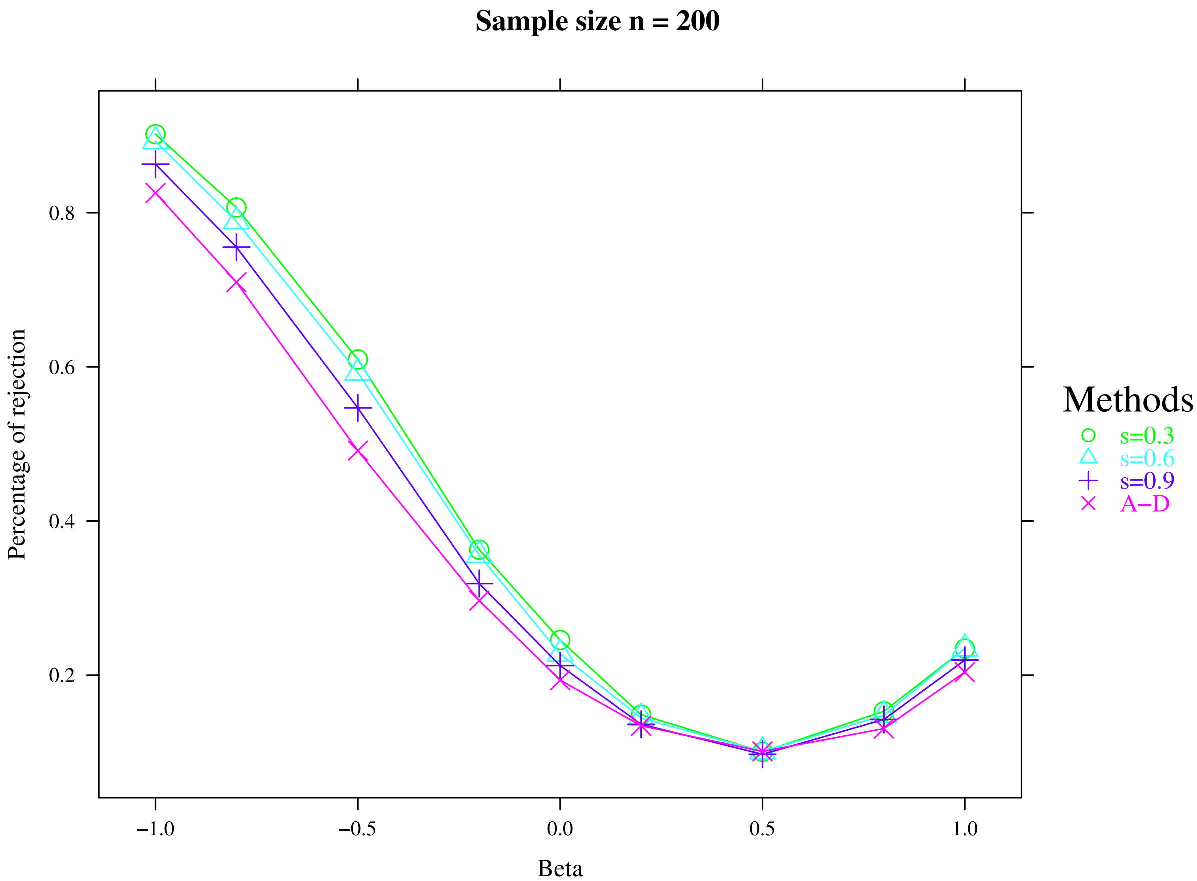

Figure 4.17: Empirical power of energy test for null distribution $S(1.8, 0.5)$ against $S(1.8, \beta)$ with varying skewness parameter β, implemented with asymptotic critical values. Sample size $n = 200$.

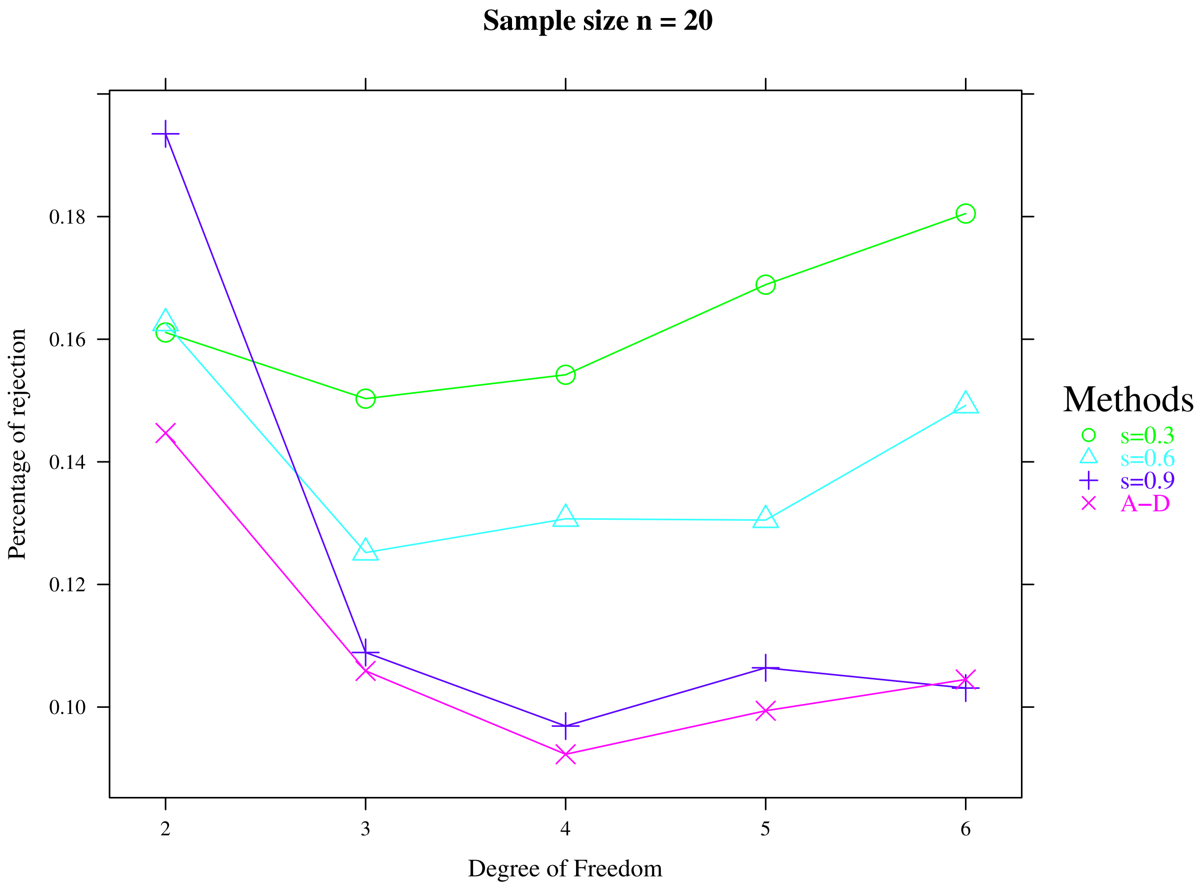

Figure 4.18: Empirical power of energy test for null distribution $S(1.8, 0.5)$ against Student's t distribution with varying degree of freedom, implemented with asymptotic critical values. Sample size $n = 20$.

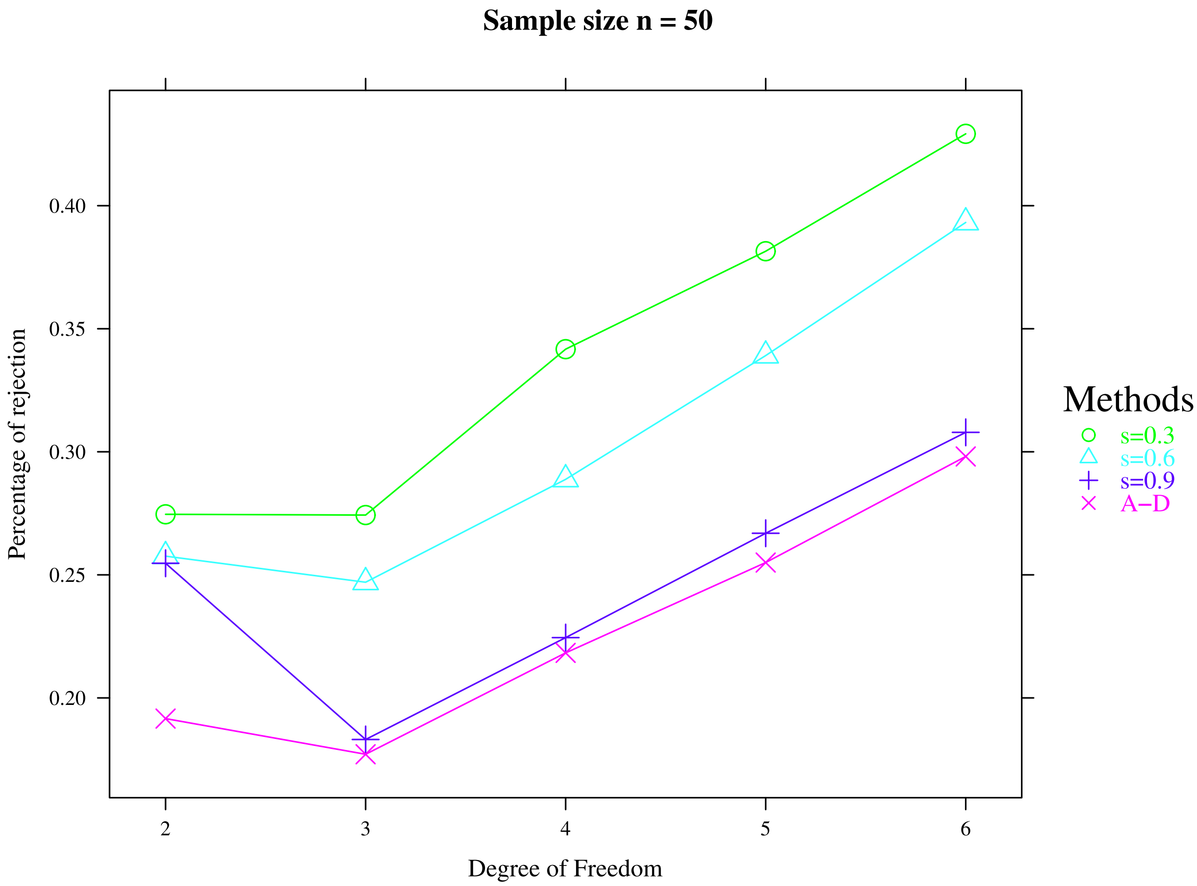

Figure 4.19: Empirical power of energy test for null distribution $S(1.8, 0.5)$ against Student's t distribution with varying degree of freedom, implemented with asymptotic critical values. Sample size $n = 50$.

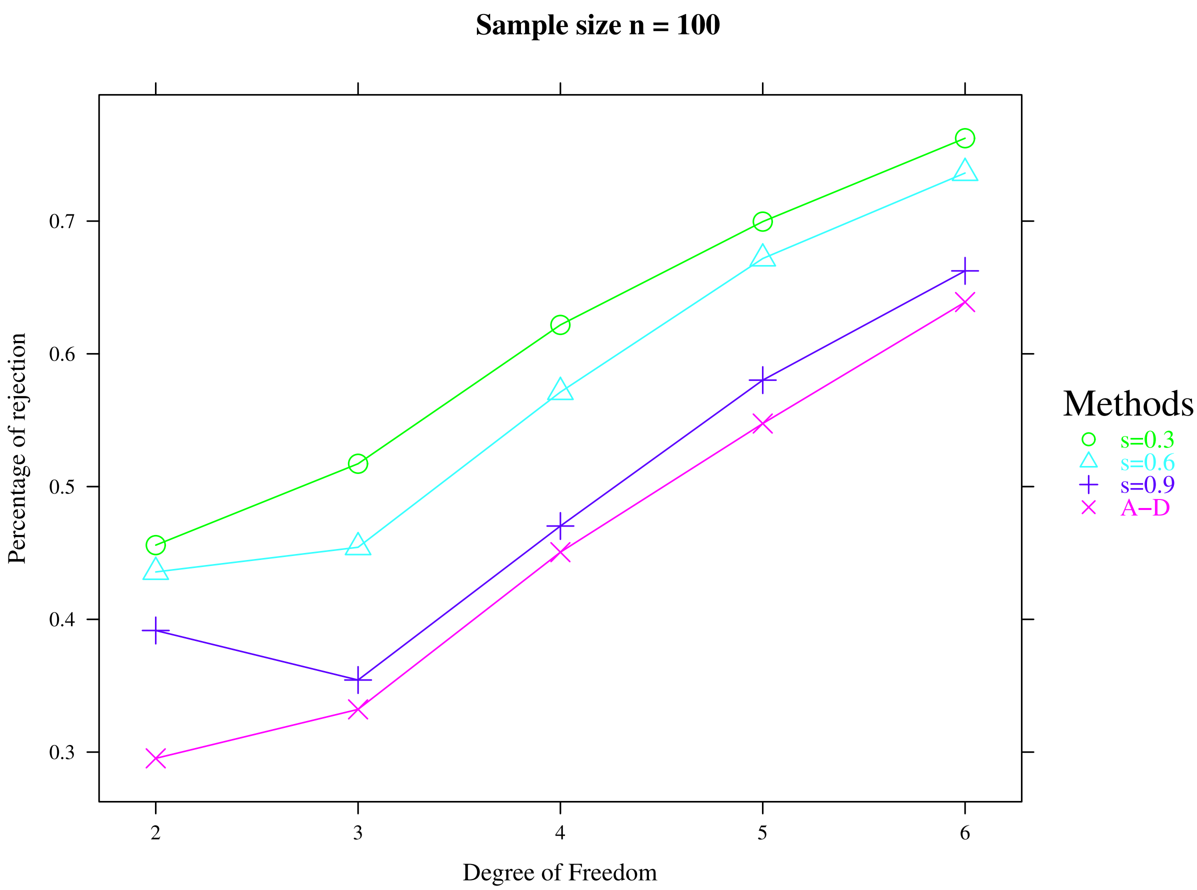

Figure 4.20: Empirical power of energy test for null distribution $S(1.8, 0.5)$ against Student's t distribution with varying degree of freedom, implemented with asymptotic critical values. Sample size $n = 100$.

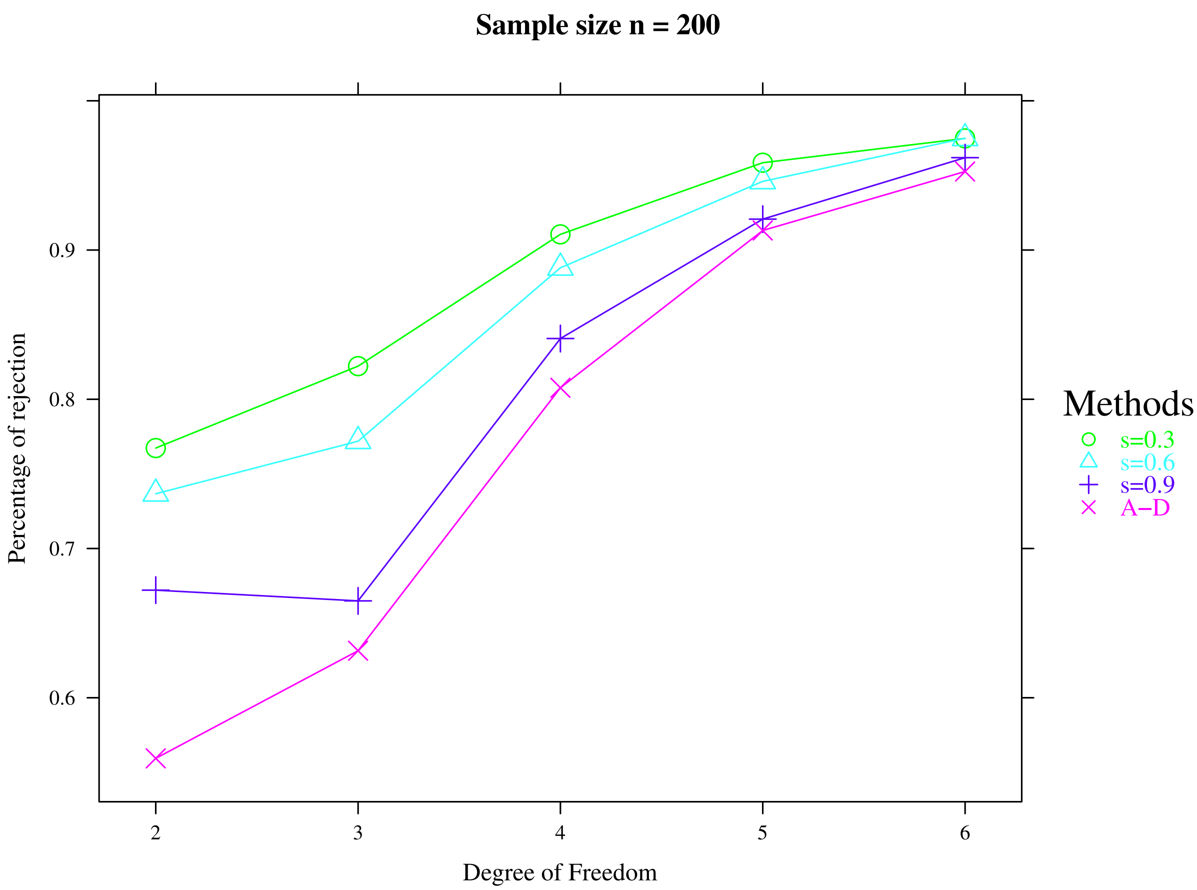

Figure 4.21: Empirical power of energy test for null distribution $S(1.8, 0.5)$ against Student's t distribution with varying degree of freedom, implemented with asymptotic critical values. Sample size $n = 200$.

4.4 Application to financial data

Financial returns, such as bonds, stocks, funds, and options, usually exhibit patterns which stable distributions seem able to capture, such as heavy tails and skewness. Complex methods have been developed to model the change of financial assets returns. In this chapter, we will not try to adopt very sophisticated methods to model financial data because it is not the purpose of this study. Instead, a stable distribution is fitted to the daily logarithmic returns, and then tested the goodness-of-fit for a simple hypothesis. The historical prices of these financial assets can be obtained from Yahoo Finance or Google Finance.

The historical price of Bank of America (BAC) is used as an illustrative example. The daily closing price of BAC was collected from Yahoo Finance from January 1, 2007 to December 31, 2010, which includes 1008 trading days. Let P_t be the closing price of BAC at day t, for $t = 1, \ldots, n$. The logarithmic asset return at day t is defined as $\log\left(\frac{P_t}{P_{t-1}}\right) = \log(P_t) - \log(P_{t-1})$. After the transformation, there are 1007 data points in the sample. The histogram is plotted in Figure 4.22.

The **stableFit** function in R package **fBasics** [62] and the executable program **stable.exe** provided on Nolan's web page `http://academic2.american.edu/~jpnolan/stable/stable.html` can be applied to get the Maximum Likelihood estimates (MLE) of four parameters. In practice, both give very similar estimates. Nolan's program is much faster than the R function **stableFit**, but cannot be implemented automatically by calling a function. Nolan's program was used in this study to get the MLE's.

For BAC data, the MLE of the parameters, energy test statistic $Q_{n,s}$ and p-value based on its asymptotic distribution are listed in Table 4.23.

Table 4.23: MLE of BAC historical price data

Ticker	Sample size	α	β	γ	δ	$Q_{n,s}$	p-value
BAC	1007	1.2539	-0.0270	0.0164	-0.0022	1.383098	0.019034

Figure 4.23 illustrates how well the data is fitted. The green curve is the density function

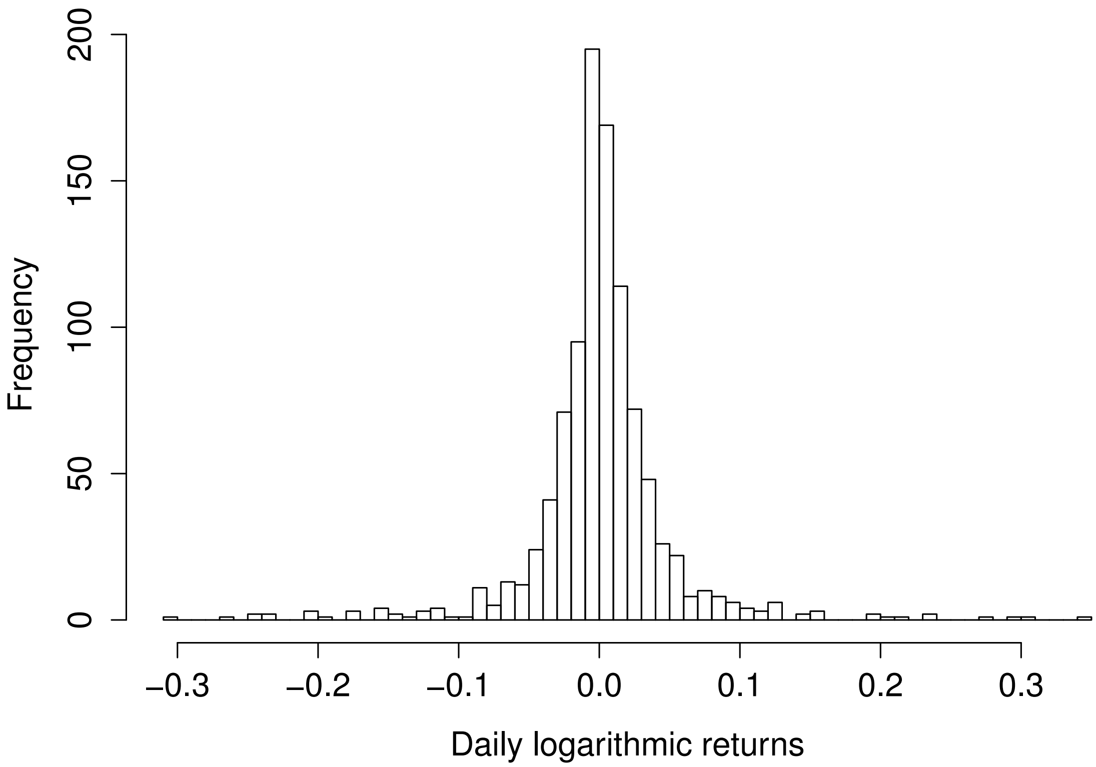

Figure 4.22: The distribution of daily logarithmic returns of Bank of America stock

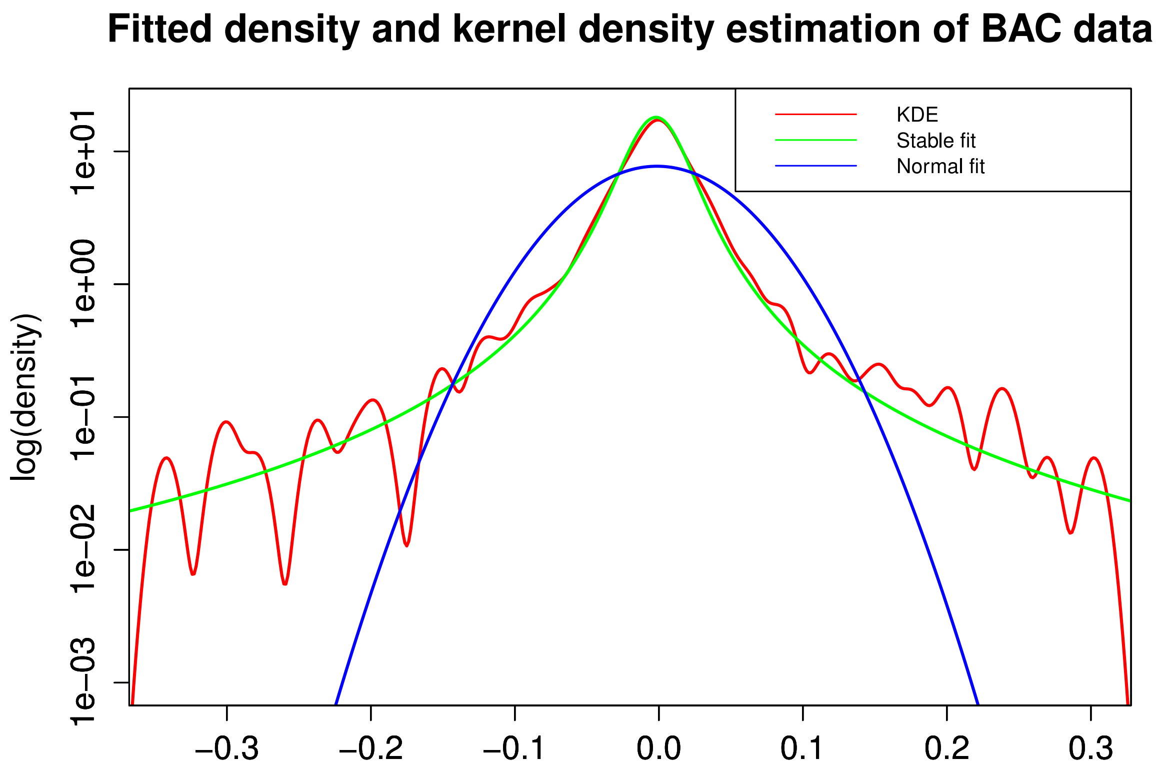

Figure 4.23: Fitted density function on BAC data

of the stable distribution with parameters equal MLE, the red curve represents the kernel density estimation at evenly spaced points, and the blue curve represents the density function of normal distribution with parameters equal MLE. It can be seen that the density fits well with data near the mode, but the tail is poorly fitted. The data shows a heavier tail than the stable distribution fitted by MLE.

Due to the problem formulation of testing the simple hypothesis, a conclusion may be drawn, after the energy test rejects the null hypothesis, that the assumption of random sample from distribution with specified parameters fails. However, there are multiple reasons that possibly explain why the null hypothesis gets rejected. First, it may be that the observations are iid, but are not sampled from a stable distribution family, and hence it may be worthwhile trying model the data with other type of distributions. Second, the parameters may be mis-specified because of incorrect estimation. It could also be due to the fact that the data points are not independent or identically distributed, as dependence and heterogeneity are very common in financial returns data. More complicated methods such as generalized autoregressive conditional heteroscedasticity (GARCH) models can be applied to capture the dependence and heterogeneity, and model the residuals by stable distributions [40].

CHAPTER 5

GOODNESS-OF-FIT TESTS FOR COMPOSITE CAUCHY HYPOTHESIS

It is often interesting to test whether a random sample is from some family of distributions, for example, a location-scale family. In this case, the parameters of the distribution F are not fully specified. A common way of handling this type of goodness-of-fit problem is to first estimate the unknown parameters with some "nice" estimators, then assume the parameters are fully specified by substituting the unknown parameters with their estimators, and then use the same test statistic as in simple hypothesis testing. These types of tests are called goodness-of-fit tests for a composite hypothesis, or goodness-of-fit tests with estimated parameters. Test for composite hypothesis are more complicated than those for simple hypotheses, because the distribution of the test statistic now depends not only on the underlying distribution of the random sample, but also the way unknown parameters are estimated. It is much more difficult to derive the asymptotic distribution of the test statistic, and the critical value tables of simple hypothesis cannot be used for the composite cases [8, 10, 9]. Nevertheless, the test can still be implemented by parametric bootstrap in general. In this chapter, the asymptotic distribution of the energy statistic for testing composite Cauchy hypothesis is derived when the parameters are estimated by maximum likelihood estimators.

5.1 Maximum likelihood estimators of Cauchy distribution

Let $x_1, \ldots, x_n$ be a random sample. The log-likelihood function is given by

$$L(\gamma, \delta; \mathbf{x}) = n \log \gamma - \sum_{j=1}^{n} \log(\gamma^2 + (x_j - \delta)^2) - n \log \pi. \tag{5.1}$$

By differentiating L with respect to γ and δ, we get the likelihood equations

$$\sum_{j=1}^{n} \frac{x_j - \delta}{\gamma^2 + (x_j - \delta)^2} = 0, \tag{5.2}$$

$$\sum_{j=1}^{n} \frac{\gamma^2}{\gamma^2 + (x_j - \delta)^2} = \frac{n}{2}. \tag{5.3}$$

The maximum likelihood estimators (MLE) can be obtained by solving the above equations numerically.

5.2 Test of composite Cauchy hypothesis

Let $y_j = \frac{x_j - \hat{\delta}}{\hat{\gamma}}$, where $\hat{\delta}$ and $\hat{\gamma}$ are consistent and equivariant estimators of location parameter δ and scale parameter γ of Cauchy distribution, and suppose X follows standard Cauchy distribution. The goodness-of-fit test statistic for composite hypothesis that the sampled distribution is Cauchy has the form

$$Q_{n,s} = n\left(\frac{2}{n}\sum_{i=1}^{n} E|y_i - X|^s - E|X - X'|^s - \frac{1}{n^2}\sum_{i,j=1}^{n}|y_i - y_j|^s\right) \tag{5.4}$$

$$= 2\sum_{j=1}^{n}\frac{(1+y_j^2)^{s/2}\cos(s\arctan y_j)}{\cos\frac{\pi s}{2}} - \frac{n2^s}{\cos\frac{\pi s}{2}} - \frac{1}{n}\sum_{j,k=1}^{n}|y_j - y_k|^s. \tag{5.5}$$

We can get an alternative expression of $Q_{n,s}$ that is very useful for deriving its asymptotic distribution:

$$\begin{aligned} Q_{n,s} &= \int_{-\infty}^{\infty}\left|\frac{1}{n}\sum\exp(it\frac{x_j-\hat{\delta}}{\hat{\gamma}}) - e^{-|t|}\right|^2 \frac{|t|^{-s-1}}{C}dt \\ &= \int_{-\infty}^{\infty}\left|\frac{1}{n}\sum\exp(it(x_j-\hat{\delta})) - e^{-|\hat{\gamma}t|}\right|^2 \frac{|\hat{\gamma}t|^{-s-1}}{C}\hat{\gamma}dt \\ &= \int_{-\infty}^{\infty}|e^{it\hat{\delta}}|^2\left|\frac{1}{n}\sum\exp(itx_j) - e^{-|\hat{\gamma}t|+it\hat{\delta}}\right|^2 \frac{|\hat{\gamma}t|^{-s-1}}{C}\hat{\gamma}dt \\ &= \int_{-\infty}^{\infty}|\hat{Z}_n(t)|^2\frac{|t|^{-s-1}}{|\hat{\gamma}|^s C}dt, \end{aligned} \tag{5.6}$$

where

$$\hat{Z}_n(t) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left\{\cos(tx_j) + i\sin(tx_j) - e^{-|\hat{\gamma}t|}\left(\cos(t\hat{\delta}) + i\sin(t\hat{\delta})\right)\right\}, \tag{5.7}$$

and

$$C = \frac{2\pi^{\frac{1}{2}}\Gamma(1-\frac{s}{2})}{s2^s\Gamma(\frac{1+s}{2})},$$

as in (3.5). It is noted that both [19] and [33] gave the wrong expression of $\hat{Z}_n(t)$, and [34] corrected it.

Hence, $\hat{Z}_n(t)$ can be rewritten as

$$\hat{Z}_n(t) = \int k(x,t)d\{\sqrt{n}(F_n(x) - F(x;\hat{\theta}_n))\}, \tag{5.8}$$

where $F_n(x)$ is the empirical distribution function and $F(x;\hat{\theta}_n)$ is the cumulative distribution function with parameter vector θ substituted by its estimator $\hat{\theta}_n$. So $\hat{Z}_n(t)$ corresponds to the empirical characteristic function process, where $k(x,t) = \cos(tx) + i\sin(tx)$ is the kernel of the kernel transformed empirical process $\hat{Z}_n(t)$. Let $\theta = (\gamma, \delta)$ and $\hat{\theta} = (\hat{\gamma}, \hat{\delta})$ be the maximum likelihood estimator of θ. Matsui and Takemura [32, 33] have done extensive work in deriving the asymptotic distribution of the empirical process $\hat{Z}_n(t)$ when the Cauchy distribution parameters are estimated by MLE. Their result, Theorem 2.1 in [33], is stated as follows:

Theorem 5.2.1 *If $X_1, \ldots, X_n$ are iid standard Cauchy random variables and $\hat{Z}_n(t)$ is defined as in (5.7) then $\hat{Z}_n(t) \xrightarrow{d} Z(t)$ in $C(\mathbb{R})$, where $C(\mathbb{R})$ is the Fréchet space of continuous functions on real line $\mathbb{R}$, and Z is a zero mean Gaussian process with covariance function*

$$\operatorname{cov}(Z(u), Z(v)) = \Gamma(u,v) = e^{-|u-v|} - \{1 + 2(uv + |uv|)\}e^{-|u|-|v|}. \tag{5.9}$$

Matsui and Takemura also studied the case when parameters are estimated by EISE, or equivariant integrated squared error estimator. However, according to their numerical results, EISE adds much computational cost to the goodness-of-fit test, but limited value to power of the test, so this case will not be discussed in this dissertation. Notice that the kernel $\Gamma(u,v)$ is symmetric about origin; that is, $\Gamma(-u,-v) = \Gamma(u,v)$, and $\Gamma(u,v) = 0$ if $uv < 0$. Hence $Z(u)$ and $Z(-u)$ are independent and identically distributed Gaussian random variables.

Similar to Theorem 2.2 in [33], we can derive the asymptotic distribution of $Q_{n,s}$ by its alternative expression (5.6) and the result in Theorem (5.2.1).

Theorem 5.2.2 *If $X_1, \ldots, X_n$ are iid standard Cauchy random variables and $\hat{Z}_n(t)$ is defined as (5.7) and $Q_{n,s}$ is defined as (5.6), then*

$$Q_{n,s} = \int_{-\infty}^{\infty} |\hat{Z}_n(t)|^2 \frac{|t|^{-s-1}}{|\hat{\gamma}|^s C} dt \xrightarrow{d} Q_s := \int_{-\infty}^{\infty} Z(t)^2 \frac{|t|^{-s-1}}{C} dt. \tag{5.10}$$

Proof The proof is similar to that of Theorem 2.2 in [19].

Asymptotic critical values can be obtained after we derive the distribution of Q_s with help from theory of integral equations. Notice $Q_s = \int_0^\infty Z(t)^2 \frac{|t|^{-s-1}}{C} dt + \int^\infty Z(-t)^2 \frac{|t|^{-s-1}}{C} dt$, where $Z(t)$ and $Z(-t)$ are independent and identically distributed, hence we can only consider the empirical process $z(t)$ on $t \in [0, \infty)$. Mercer's theorem enables us to express a symmetric positive-definite function as a sum of a convergent sequence of product functions.

Theorem 5.2.3 *If $K(x, y)$ is the kernel of a positive self-adjoint operator on $L^2[0, 1]$ and suppose that $K(x, y)$ is continuous in both variables, then*

$$K(x, y) = \sum_{j=1} \lambda_j \phi_j(x) \phi_j(y), \ \lambda_1 \geq \lambda_2 \geq \cdots > 0, \tag{5.11}$$

where λ_j is an eigenvalue and ϕ_j is the corresponding orthonormal eigenfunction of the integral equation

$$\int_0^1 K(x, y) \phi(y) dy = \lambda \phi(x). \tag{5.12}$$

The series (5.11) converge uniformly and absolutely to $K(x, y)$.

Since Mercer's theorem only applies to kernels defined on compact spaces, we need to transform the kernel to $[0, 1]^2$ while keeping all the eigenvalues invariant. However, after the transformation, we need to deal with kernels that are not continues at the point $(1, 1)$, the following version of Mercer's theorem by Hammerstein is required, as in [2] and [33].

Theorem 5.2.4 *Let $K(x, y)$ be continuous in the unit square $[0, 1]^2$ except possibly at the corners of the square; let $\partial K(x, y)/\partial x$ be continuous in the interior of both triangles in which the square is divided by the line between $(0, 0)$ and $(1, 1)$, and let the partial derivative be bounded in the domain $\epsilon \leq x \leq 1 - \epsilon$ and $0 \leq y \leq 1$ for each $\epsilon > 0$. The series on the right of (5.11) converges uniformly to $K(x, y)$ in every domain in the interior of the unit square*

$[0,1]^2$.

With the above theorem, the following theorem can be stated.

Theorem 5.2.5 *If Q_s is defined as in (5.10), and $x = \frac{u}{1+u}$ and $y = \frac{v}{1+v}$, then Q_s admits a representation*

$$Q_s \stackrel{d}{=} \sum_{j=1}^{\infty} \lambda_j (Z_{2j-1}^2 + Z_{2j}^2), \tag{5.13}$$

where λ_j's are eigenvalues of kernel function defined on $[0,1]^2$

$$K(x,y) = \Gamma(u(x), v(y)) \frac{|u(x)v(y)|^{\frac{-s-1}{2}}}{C} \sqrt{u'(x)v'(y)} \tag{5.14}$$

and $Z_j, j = 1, 2, \ldots$ are independent standard normal random variables.

Proof Notice $Q_s = \int_0^\infty Z(t)^2 \frac{|t|^{-s-1}}{C} dt + \int^\infty Z(-t)^2 \frac{|t|^{-s-1}}{C} dt$, where $Z(t)$ and $Z(-t)$ are independent and identically distributed, hence we can only consider the empirical process $z(t)$ on $t \in [0, \infty)$.

$$\int_0^\infty Z(u)^2 \frac{|u|^{-s-1}}{C} du = \int_0^1 Z(u(x))^2 \frac{|u(x)|^{-s-1}}{C} u'(x) dx. \tag{5.15}$$

Let $Y(x) = Z(u(x)) \sqrt{\frac{|u(x)|^{-s-1}}{C} u'(x)}$ be a zero-mean Gaussian process. Its covariance function is

$$\begin{aligned}
K(x,y) =& \operatorname{cov}(Y(x), Y(y)) \\
=& \operatorname{cov}\left((Z(u(x)) \sqrt{\frac{|u(x)|^{-s-1}}{C} u'(x)}, Z(v(y)) \sqrt{\frac{|v(y)|^{-s-1}}{C} v'(y)} \right) \\
=& \Gamma(u(x), v(y)) \frac{|u(x)v(y)|^{\frac{-s-1}{2}}}{C} \sqrt{u'(x)v'(y)} \\
=& \Gamma(u(x), v(y)) \left(\frac{x}{1-x} \right)^{\frac{-s-1}{2}} \left(\frac{y}{1-y} \right)^{\frac{-s-1}{2}} \frac{1}{C(1-x)(1-y)}.
\end{aligned}$$

Now $K(x,y)$ is a symmetric kernel function defined on $[0,1]^2$, and by Theorem 5.2.5,

$$K(x,y) = \sum_{j=1}^{\infty} \lambda_j \phi_j(x)\phi_j(y).$$

Let $W(x) = \sum_{j=1}^{\infty} \sqrt{\lambda_j}\phi_j(x)Z_j$. It is easy to see that $W(x)$ is a Gaussian process with $E(W(x)) = 0$ and

$$\begin{aligned} &\operatorname{cov}(W(x), W(y)) = E(W(x)W(y)) \\ =&E\left(\sum_{j=1}^{\infty} \lambda_j \phi_j(x)\phi_j(y)Z_j^2\right) = \sum_{j=1}^{\infty} \lambda_j \phi_j(x)\phi_j(y) = K(x,y), \end{aligned}$$

so $W(x)$ defines the same Gaussian process as $Y(x)$, and therefore

$$\int_0^1 Y(x)^2 dx \stackrel{d}{=} \int_0^1 W(x)^2 dx = \int_0^1 \sum_{j=1}^{\infty} \sqrt{\lambda_j}\phi_j(x)Z_j dx = \sum_{j=1}^{\infty} \lambda_j Z_j^2.$$

∎

CHAPTER 6

SUMMARY

In this dissertation, a new goodness-of-fit test for stable distributions is proposed based on the energy distance. The test statistic is a V-statistic with second degree degenerate kernel, which can also be expressed as a weighted L^2-norm of the difference between the two characteristic functions, with a suitable weight function. Both theory of V-statistics and empirical process theory are applied to study the asymptotic distribution of the test statistic. For testing a simple hypothesis of a stable distribution in the general case, the computational formula of the energy statistic is provided and consistency of the test is proved. For testing a composite hypothesis of Cauchy distribution, an asymptotic results are developed when parameters are estimated by maximum likelihood estimators.

Because of the difficulty deriving the exact distribution of the finite-sample test statistic, parametric bootstrap is recommended to implement the test for small and medium-sized samples (sample size less than 200). For large samples, asymptotic critical values can be used. Simulation results are presented for testing the standard Cauchy distribution and the standard symmetric stable distribution with $\alpha = 1.5$. Overall, the empirical power results presented demonstrate that the new test is more powerful than existing methods including the Kolmogorov-Smirnov test and the Anderson-Darling test. Compared to tests based on empirical characteristic functions, such as the Gütler-Henze test and the Matsui-Takemura test, the energy test is slightly more powerful when the alternatives have heavier tails than the null distribution, and slightly less powerful in testing lighter-tailed alternatives.

Some future research directions include developing asymptotic results for testing composite hypothesis of stable distribution in the general case, extending the energy test to multivariate stable distribution, and more simulation studies on the power of energy test against broader classes of alternative distributions.

BIBLIOGRAPHY

[1] V. Akgiray and C.G. Lamoureux. Estimation of stable-law parameters: A comparative study. *Journal of Business & Economic Statistics*, 7(1):85–93, 1989.

[2] T W Anderson and D A Darling. Asymptotic theory of certain" goodness of fit" criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 23(2):193–212, 1952.

[3] R. W. Arad. Parameter estimation for symmetric stable distribution. *International Economic Review*, 21(1):209–220, 1980.

[4] C. J. Bellosta. *ADGofTest: Anderson-Darling GoF test*, 2011. R package version 0.3.

[5] R. C. Blattberg and N. J. Gonedes. A comparison of the stable and student distributions as statistical models for stock prices. *The Journal of Business*, 47(2):244–280, 1974.

[6] R. F. Brcich, D. R. Iskander, and A. M. Zoubir. The stability test for symmetric alpha-stable distributions. *IEEE Transactions on Signal Processing*, 53(3):977–986, March 2005.

[7] J. M. Chambers and C. L. Mallows. A method for simulating stable random variables. *Journal of the American Statistical Association*, 71(354):340– 344, 1976.

[8] R. B. D'Agostino and M. A. Stephens. *Goodness-of-fit Techniques*. CRC Press, 1986.

[9] A. DasGupta. *Asymptotic theory of statistics and probability*. Springer Verlag, New York, 2008.

[10] E. Del Barrio, P. Deheuvels, and S. A. Geer. *Lectures on Empirical Processes: Theory and Statistical Applications*. European Mathematical Society, 2007.

[11] P. Duchesne and L. P. de Micheaux. Computing the distribution of quadratic forms: Further comparisons between the Liu-Tang-Zhang approximation and exact methods. *Computational Statistics and Data Analysis*, 54:858–862, 2010.

[12] W. H. DuMouchel. On the asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution. *The Annals of Statistics*, 1(5):948–957, 1973.

[13] W. H. DuMouchel. Estimating the stable index alpha in order to measure tail thickness: a critique. *The Annals of Statistics*, 11(4):1019–1031, 1983.

[14] E. F. Fama. The behavior of stock-market prices. *The Journal of Business*, 38(1):34–105, 1965.

[15] E. F. Fama and R. Roll. Some properties of symmetric stable distributions. *Journal of the American Statistical Association*, 63(323):817–836, 1968.

[16] E. F. Fama and R. Roll. Parameter Estimates for Symmetric Stable Distributions. *Journal of the American Statistical Association*, 66(334):331, June 1971.

[17] Z. Fan. Parameter Estimation of Stable Distributions. *Communications in Statistics: Theory and Methods*, 35(2):245–255, March 2006.

[18] Z. Fan. Minimum-Distance Estimator for Stable Exponent. *Communications in Statistics - Theory and Methods*, 38(4):511–528, January 2009.

[19] N. Gurtler and N. Henze. Goodness-of-fit tests for the Cauchy distribution based on the empirical characteristic function. *Annals of the Institute of Statistical Mathematics*, 52(2):267–286, 2000.

[20] W. Hoeffding. A class of statistics with asymptotically normal distribution. *The Annals of Mathematical Statistics*, 19(3):293–325, 1948.

[21] J. P. Imhof. Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48(3/4):419–426, 1961.

[22] I. A. Koutrouvelis. Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association*, 21:918–928, 1980.

[23] I. A. Koutrouvelis. An iterative procedure for the estimation of the parameters of stable laws. *Communications in Statistics - Simulation and Computation*, 10(1):17–28, 1981.

[24] A. H. L. Lau, H. S. Lau, and J. R. Wingender. The Distribution of Stock Returns: New Evidence against the Stable Model. *Journal of Business & Economic Statistics*, 8(2):217–223, 1990.

[25] E. L. Lehmann. *Elements of Large-sample Theory*. Springer Verlag, 1999.

[26] S. M. Liu and B. W. Brorsen. Maximum likelihood estimation of a GARCH-stable model. *Journal of Applied Econometrics*, 10(3):273–285, 1995.

[27] B. Mandelbrot. The stable Paretian income distribution when the apparent exponent is near two. *International Economic Review*, 4(1):111–115, 1963.

[28] B. Mandelbrot. The variation of certain speculative prices. *The Journal of Business*, 36(4):394–419, 1963.

[29] B. Mandelbrot and H. M. Taylor. On the distribution of stock price differences. *Operations research*, pages 1057–1062, 1967.

[30] G. Marsaglia and J. Marsaglia. Evaluating the Anderson-Darling Distribution. *Journal of Statistical Software*, 9(2):1–5, 2004.

[31] G. Marsaglia, W. W. Tsang, and J. Wang. Evaluating Kolmogorov's Distribution. *Journal of Statistical Software*, 8(18):1–4, 2003.

[32] M. Matsui and A. Takemura. 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. Technical report, Center for International Research on the Japanese Economy, University of Tokyo, 2003.

[33] M. Matsui and A. Takemura. Empirical characteristic function approach to goodness-of-fit tests for the Cauchy distribution with parameters estimated by MLE or EISE. *Annals of the Institute of Statistical Mathematics*, 57(1):183–199, 2005.

[34] M. Matsui and A. Takemura. Goodness-of-fit tests for symmetric stable distributions - Empirical characteristic function approach. *Test*, 17(3):546–566, 2008.

[35] J. H. McCulloch. Simple consistent estimators of stable distribution parameters. *Communications in Statistics-Simulation and Computation*, 15(4):1109–1136, 1986.

[36] J. P. Nolan. Numerical calculation of stable densities and distribution functions. *Communications in Statistics-Stochastic Models*, 13(4):759–774, 1997.

[37] J. P. Nolan. Fitting data and assessing goodness-of-fit with stable distributions. In *Proceedings of the Conference on Applications of Heavy Tailed Distributions in Economics, Engineering and Statistics*, American University, Washington, DC, 1999.

[38] J. P. Nolan. Maximum likelihood estimation and diagnostics for stable distributions. *Levy Processes: Theory and Applications*, pages 379–400, 2001.

[39] J. P. Nolan. *Stable Distributions - Models for Heavy Tailed Data*. Birkhäuser, Boston, 2010.

[40] A. K. Panorska, S. Mittnik, and S. T. Rachev. Stable GARCH models for financial time series. *Applied Mathematics Letters*, 8(5):33–37, 1995.

[41] A. S. Paulson, E. W. Holcomb, and R. A. Leitch. The estimation of the parameters of the stable laws. *Biometrika*, 62(1):163, 1975.

[42] S. J. Press. Estimation in univariate and multivariate stable distributions. *Journal of the American Statistical Association*, pages 842–846, 1972.

[43] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011.

[44] M. L. Rizzo. *A new rotation invariant goodness-of-fit test*. PhD thesis, Bowling Green State University, 2002.

[45] M. L. Rizzo. New Goodness-of-Fit Tests for Pareto Distributions. *ASTIN Bulletin*, 39(2):691–715, 2009.

[46] M. L. Rizzo and G. J. Szekely. DISCO Analysis: A Nonparametric Extension of Analysis of Variance. *Annals of Applied Statistics*, 4(2):1034–1055, 2010.

[47] P. A. Samuelson. Efficient portfolio selection for Pareto-Levy investments. *Journal of Financial and Quantitative Analysis*, 2(2):107–122, 1967.

[48] E. M. Saniga and J. C. Hayya. Simple goodness-of-fit tests for symmetric stable distributions. *Journal of Financial and Quantitative Analysis*, 12(02):276–289, 1977.

[49] E. M. Saniga and J. A. Miles. Power of some standard goodness-of-fit tests of normality against asymmetric stable alternatives. *Journal of the American Statistical Association*, 74(368):861–865, 1979.

[50] R. J. Serfling. *Approximation Theorems of Mathematical Statistics*. Wiley-Interscience, New York, 1980.

[51] I. H. Sloan. Quadrature methods for integral equations of the second kind over infinite intervals. *Mathematics of Computation*, 36(154):511–523, 1981.

[52] M. A. Stephens. EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69(347):730–737, September 1974.

[53] G. J. Szekely. Technical Report 03-05: E-statistics: Energy of Statistical Samples. *Department of Mathematics and Statistics, Bowling Green State University*, 2000.

[54] G. J. Szekely and M. L. Rizzo. Testing for equal distributions in high dimension. *InterStat*, 5, 2004.

[55] G. J. Szekely and M. L. Rizzo. A new test for multivariate normality. *Journal of Multivariate Analysis*, 93(1):58–80, 2005.

[56] G. J. Szekely and M. L. Rizzo. Hierarchical Clustering via Joint Between-Within Distances: Extending Ward's Minimum Variance Method. *Journal of Classification*, 22(2):151–183, 2005.

[57] G. J. Szekely and M. L. Rizzo. Brownian Distance Covariance. *Annals of Applied Statistics*, 3(4):1236–1265, 2009.

[58] G. J. Szekely, M. L. Rizzo, and N. K. Bakirov. Measuring and testing dependence by correlation of distances. *The Annals of Statistics*, 35(6):2769–2794, December 2007.

[59] V. V. Uchaikin and V. M. Zolotarev. *Chance and Stability: Stable Distributions and Their Applications*. VSP, 1999.

[60] R. von Mises. On the asymptotic distribution of differentiable statistical functions. *The annals of mathematical statistics*, 18(3):309–348, 1947.

[61] R. Weron. On the Chambers-Mallows-Stuck method for simulating skewed stable random variables. *Statistics & Probability Letters*, 28(2):165–171, June 1996.

[62] Diethelm Wuertz, Rmetrics Core team members, uses code builtin from the following R contributed packages: gmm from Pierre Chauss, gld from Robert King, gss from Chong Gu, nortest from Juergen Gross, HyperbolicDist from David Scott, sandwich from Thomas Lumley, Achim Zeileis, and and fortran/C code from Kersti Aas. *fBasics: Rmetrics - Markets and Basic Statistics*, 2010. R package version 2110.79.

[63] V. M. Zolotarev. *One-dimensional Stable Distributions*, volume 65. American Mathematical Society, Providence, 1986.