Fitting Data and Assessing Goodness-of-fit with Stable Distributions

John P. Nolan
Department of Mathematics and Statistics
American University
Washington, DC 20016 USA
jpnolan@american.edu

June 1999

Abstract

There are now several reliable methods for estimating stable parameters from data. However, little attention has been paid to model verification, i.e. how to assess whether the stable parameters estimated actually do a good job of describing the data. We analyze several heavy tailed data sets and demonstrate diagnostics for assessing both univariate and multivariate stable fits to data sets. Multivariate stable distributions are characterized in terms of one dimensional projections and new results are given for sub-Gaussian stable distributions.

Keywords: Stable distributions, parameter estimation, diagnostics, exploratory data analysis.

1 Introduction

Stable distributions are a rich class of distributions that include the Gaussian and Cauchy distributions in a family that allows skewness and heavy tails. The class was characterized by Paul Lévy (1924) in his study of normalized sums of i.i.d. terms. The general stable distribution is described by four parameters: an index of stability $\alpha \in (0, 2]$, a skewness parameter β, a scale parameter γ and a location parameter δ. The lack of closed formulas for densities and distribution functions for all but a few stable distributions (Gaussian, Cauchy and Lévy) has been a major drawback to the use of stable distributions by practitioners. This paper shows that the computational problems have now been resolved and it is feasible to fit stable models to data and to use diagnostics to assess the goodness of fit.

Stable distributions have been proposed as a model for many types of physical and economic systems. There are several reasons for using a stable distribution to describe a system. The first is where there are solid theoretical reasons for expecting a non-Gaussian stable model, e.g. reflection off a rotating mirror yielding a Cauchy distribution, hitting times for a Brownian motion yielding a Lévy distribution, the gravitational field of stars yielding the Holtsmark distribution; see Feller (1971) for these and other examples. The second reason is the Generalized Central Limit Theorem which states that the only possible non-trivial limits of normalized sums of i.i.d. terms are stable. It has been argued that many observed quantities are the sum of many small terms - the price of a stock, the noise in a communication system, etc. and hence a stable model should be used to describe such systems. The third argument for modeling with stable distributions is empirical: many large data sets exhibit heavy tails and skewness. The strong empirical evidence for these features combined with the Generalized Central Limit Theorem is used by many to justify the use of stable models. Examples in finance and economics are given in Mandelbrot (1963) and (1972), Fama (1965), Embrechts, Klüppelberg, and Mikosch (1997), Cheng and Rachev (1995), McCulloch (1996); in communication systems by Stuck and Kleiner (1974), Zolotarev (1986) and Nikias and Shao (1995). Such data sets are poorly described by a Gaussian model, but possibly can be described by a stable distribution.

Several recent monographs focus on stable models: Zolotarev (1986), Christoph and Wolf (1993), Samorodnitsky and Taqqu (1994), Janicki and Weron (1994), and Nikias and Shao (1995). For the related topic of modeling with heavy tailed distributions, see Embrechts, Klüppelberg and Mikosch (1997) and Adler, Feldman and Taqqu (1998).

Skeptics of stable models recoil from the implicit assumption of infinite variance in a non-Gaussian stable model and have proposed other models for observed heavy tailed and skewed data sets, e.g. mixture models, time varying variances, etc. Of course the same people who argue that the population is inherently bounded and therefore must have a finite variance, routinely use the normal distribution - with unbounded support - as a model for this same population. The variance is but one measure of spread for a distribution, and it is not appropriate for all problems. From an applied point of view, what we generally care about is capturing the shape of a distribution.

We propose that the practitioner approach this dispute as an agnostic. The fact is that until now we have not really been able to compare data sets to a proposed stable model. In some cases there are solid theoretical reasons for believing that a stable model is appropriate; in other cases we will be pragmatic: if a stable distribution describes the data accurately and parsimoniously with four parameters, then we accept it as a model for the observed data.

This paper is organized in the following way. The remainder of this section describes two parameterizations for stable distributions and some basic properties. Section 2 discusses univariate estimation techniques and diagnostics for assessing whether a data set is stable or not. Examples of stable fits for several data sets are given in Section 3. Section 4 focuses on multivariate stable distributions, stressing univariate projections, an approach that we believe simplifies some of the theory and is well suited for estimation and diagnostics. Section 5 is on multivariate estimation and diagnostics. The special case of sub-Gaussian stable distributions is discussed in Section 6. Finally, we give a discussion of our results and information on software in Section 7.

1.1 Parameterizations and basic properties

There are at least half a dozen different parameterizations of stable distributions. All involve different specifications of the characteristic function and all are useful for one reason or another. The parameterization most often used now (see Samorodnitsky and Taqqu (1994)) is the following: $X \sim \mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$ if the characteristic function of X is given by

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

The range of parameters are $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, scale $\gamma > 0$, and location $\delta \in R$. Equation (1) is a slight variation of the (A) parameterization of Zolotarev (1986).

A more useful parameterization in applications is a variation of Zolotarev's (M) parameterization: we will say $X \sim \mathbf{S}(\alpha, \beta, \gamma, \delta; 0)$ if the characteristic function of X is given by

$$E \exp(itX) =$$

$$\begin{cases} \exp\{-\gamma^\alpha|t|^\alpha \left[1 + i\beta(\tan\frac{\pi\alpha}{2})(\operatorname{sign} t)((\gamma|t|)^{1-\alpha} - 1)\right] + i\delta t\} & \alpha \neq 1 \\ \exp\{-\gamma|t| \left[1 + i\beta\frac{2}{\pi}(\operatorname{sign} t)(\ln|t| + \ln\gamma)\right] + i\delta t\} & \alpha = 1. \end{cases}$$

The value of this representation is that the characteristic functions (and hence the corresponding densities and d.f.) are jointly continuous in all four parameters, while the $\mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$ parameterization is not. Accurate numerical calculations of the corresponding densities show that in this representation α and β have a much clear meaning as measures of the heaviness of the tails and skewness parameters, see Figure 1.

We caution the reader that neither the $\mathbf{S}(\alpha, \beta, \gamma, \delta; 0)$ parameterization nor the variables used for scale and location are standard. Most of the literature uses σ for the scale and μ for the location parameter; we are using more neutral variable names γ and δ respectively. Since the scale is never the standard deviation, it is confusing to practioners to use the symbol σ. (When $\alpha < 2$, there is no standard deviation; when $\alpha = 2$, the scale above is chosen so that the standard deviation is $\sqrt{2}\gamma$.) Likewise, the location parameter in the $\mathbf{S}(\alpha, \beta, \gamma, \delta; 0)$ parameterization is not the mean (unless $\alpha > 1$ and $\beta = 0$) and in the $\mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$ parameterization it is not the mean in half the cases (when $\alpha \leq 1$, there is no population mean). We prefer using the $\mathbf{S}(\alpha, \beta, \gamma, \delta; 0)$ parameterization and expressing the mean and variance, when they exist, as functions of the parameters $(\alpha, \beta, \gamma, \delta)$. This is common practice for many of the standard probability distributions, e.g. Gamma, Beta, Weibull, Pareto, t, F, etc.

The parameters α, β and γ have the same meaning for the two parameterizations, while the location parameters δ_0 and δ_1 of the 0 and 1 parameterizations are related by

$$\delta_1 = \begin{cases} \delta_0 - \beta(\tan\frac{\pi\alpha}{2})\gamma & \alpha \neq 1 \\ \delta_0 - \beta\frac{2}{\pi}\gamma\ln\gamma & \alpha = 1. \end{cases} \tag{2}$$

The particular form of the characteristic function was chosen to make the 0 parameterization a location and scale family: if $Y \sim \mathbf{S}(\alpha, \beta, \gamma, \delta; 0)$, then for any $a \neq 0$, b, $aY + b \sim \mathbf{S}(\alpha, (\operatorname{sign} a)\beta, |a|\gamma, a\delta + b; 0)$. We will base the likelihood calculations below on the 0 parameterization because it is the simplest scale-location parameterization which is jointly continuous in all four parameters. Various authors sidestep the discontinuity at $\alpha = 1$ by saying that the probability that $\alpha = 1$ is zero therefore you can ignore it; Buckle (1995) assumes that you know beforehand that either $\alpha < 1$ or $\alpha > 1$ and restricts his prior for α to the appropriate interval. The shape of the data is what we really care about, and that is similar when α is near or at 1; the standard

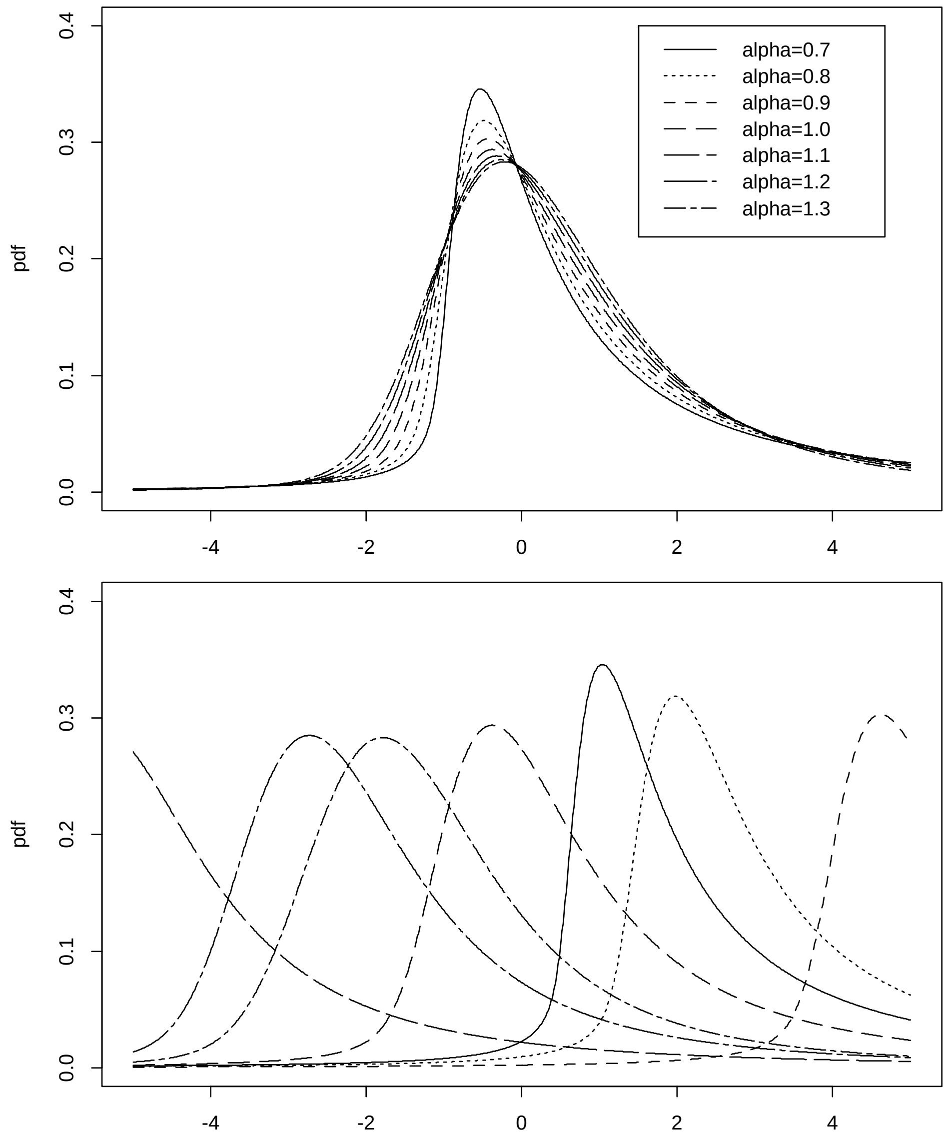

Figure 1: Stable densities in the $\mathbf{S}(\alpha, \beta, \gamma, \delta; 0)$ (top) and $\mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$ (bottom) parameterizations. $\beta = 0.8$, $\gamma = 1$, $\delta = 0$, and α as indicated. In the 1 parameterization, the mode is near 0 for α near 0 or 2, or $\alpha = 1$, but diverges to $+\infty$ as $\alpha \uparrow 1$ and diverges to $-\infty$ as $\alpha \downarrow 1$.

parameterizations simply masks that with a shift. It is preferable to let the data determine what α is and not make assumptions about the parameters, even if α is not near 1. Finally, the use of the 0 parameterization has the technical advantage of reducing the correlation between the parameter estimates, especially when α is near 1. More information on parameterizations, modes of stable densities and generalizations to multivariate stable laws can be found in Nolan (1998).

Basic properties of stable distributions can be found in Samorodnitsky and Taqqu (1994). Some of the prominent properties are: heavy tails that are asymptotically Pareto, possible skewness of the distributions, and smooth unimodal densities with no closed formula. Let $f(x|\alpha,\beta,\gamma,\delta)$ be the density of a $\mathbf{S}(\alpha,\beta,\gamma,\delta;0)$ distribution. Known facts about stable densities in the standard parameterization show that $f(x|\alpha,-\beta,\gamma,\delta) = f(-x|\alpha,\beta,\gamma,-\delta)$ and

$$\text{support}\, f(x|\alpha,\beta,\gamma,\delta) = \begin{cases} [\delta - \gamma\tan\frac{\pi\alpha}{2}, \infty) & \alpha < 1 \text{ and } \beta = 1 \\ (-\infty, \delta + \gamma\tan\frac{\pi\alpha}{2}] & \alpha < 1 \text{ and } \beta = -1 \\ (-\infty, +\infty) & \text{otherwise.} \end{cases}$$

Note that for a totally skewed ($\beta = \pm 1$) distribution when $\alpha < 1$, the finite endpoint of the support goes to $(\operatorname{sign}\beta)\infty$ as $\alpha \uparrow 1$. It can be shown that the mode and most of the distribution stay concentrated near δ, so that only a very small probability is far out on that tail. In fact, the light tail in the totally skewed cases decays faster than Pareto. For algorithms that accurately compute stable densities and cumulatives, see Nolan (1997).

2 Univariate estimation and diagnostics

There are several methods of estimating stable parameters. There are other methods, but we will focus on three general methods, which we describe briefly next. Unpublished simulation results suggest that these are the best three general methods, and that the order of presentation below is from fastest to slowest and from least accurate to most accurate.

The oldest is the quantile/fractile method of Fama and Roll (1971) (symmetric case) and McCulloch (1986) (general case). It uses five sample quantiles ($5^{th}, 25^{th}, 50^{th}, 75^{th}, 95^{th}$ quantiles) and matches these values to the stable distribution with the closest spread pattern.

Since closed forms are known for the characteristic functions of stable laws, several researchers have based estimates on the empirical or sample characteristic function. Press (1972) seems to have been the first to do this.

Several modifications have been made to this approach, see Paulson, Holcomb and Leitch (1975), Feuerverger and McDunnough (1981a) and (1981b), Koutrouvelis (1980) and (1981), Kogon and Williams (1998). The paper of Kogon and Williams incorporates several technical improvements that we highly recommend (using the $\mathbf{S}(\alpha, \beta, \gamma, \delta; 0)$ parameterization and doing an initial scale-location normalization).

Maximum likelihood estimation has been done in certain cases. While not easily accessible, DuMouchel (1971) gives a wealth of information on estimating stable parameters by approximate maximum likelihood at a remarkably early date. See also DuMouchel (1973a), (1973b), (1975) and (1983). For the special case of ML estimation for symmetric stable distributions, see Brorsen and Yang (1990) and McCulloch (1998). The general (not necessarily symmetric case) is described in Nolan (1999), where a fast pre-computed spline approximation to stable densities is used to compute the likelihood and numerically maximize it. Numerical computation of the Fisher information matrix gives confidence bounds for the parameter estimates. While slower than the other methods, the numerical maximum likelihood technique is still quite fast: for a sample of size $n = 1000$, estimates are found within one second on a 300 MHz PC. Since this is asymptotically the minimum variance unbiased estimator, ML estimates are given in the following examples, unless otherwise stated.

2.1 Diagnostics for assessing stability

In principle, it is not surprising that one can fit a data set better with the 4 parameter stable model than with the 2 parameter normal model. The relevant question is whether or not the stable fit actually describes the data well. Any procedure for estimating stable parameters will find a "best fit" by its criteria: the maximum likelihood approach maximizes the likelihood numerically, the quantile methods try to match certain data quantiles with those of stable distributions, the characteristic function based methods fit the empirical characteristic function. All will give some values for parameter estimates, even if the shape of the observed distribution is not similar to the fitted distribution, e.g. the data is multi-modal, has gaps in its support, etc. Therefore, it is important to have some means of assessing whether the resulting fit is reasonable.

The use of a diagnostic depends on what you are planning to do with a data set. For testing residuals from a regression analysis, departures from normality around the center of the distribution are usually not important; outliers are important because they can affect the validity of normal theory conclusions. In the re-insurance field, one is only concerned with extreme events and there one wants to estimate tails of the claim distribution as accurately as possible. In a model of stock prices or exchange rates, one may be interested in the shape of the whole distribution.

While non-Gaussian stable distributions are heavy-tailed distributions, most heavy-tailed distributions are not stable. One can try to fit a heavy-tailed data set with a stable distribution, but it is inappropriate in many cases. As DuMouchel (1983) points out, making a statement about the tails is quite distinct from making a statement about the entire distribution. We amplify this point by an example similar to one used by DuMouchel. Define for $0 < \alpha < 2$, $x_0 > 0$

$$g(x) = g(x|\alpha, x_0) = \begin{cases} c_1 e^{-x^2/2} & |x| \le x_0 \\ c_2 |x|^{-(1+\alpha)} & |x| > x_0, \end{cases}$$

where c_1 and c_2 are chosen to make g continuous and $\int g(x)dx = 1$: $c_1 = c_1(\alpha, x_0) = [\sqrt{2\pi}(2\Phi(x_0)-1)+(2/\alpha)x_0 \exp(x_0^2/2)]^{-1}$, $c_2 = c_1 \exp(-x_0^2/2)x_0^{1+\alpha}$. A random variable X with density g has a normal density in the interval $-x_0 \le x \le x_0$, a Pareto tail, with fraction $p = P(|X| \le x_0) = c_1\sqrt{2\pi}(2\Phi(x_0) - 1)$ in the normal part of the density and $1-p$ on the Pareto tails. For any finite x_0, this density has infinite variance and is in the domain of attraction of a symmetric stable distribution with index of stability α. Suppose we observe a sample of size n from such a distribution and try to fit it with a stable distribution. If $(1 - p)n$ is small, we will likely have few observations from the Pareto part of the distribution and we will not be able to detect the heavy tails. Any reasonable estimation scheme would lead to an $\hat{\alpha} \approx 2$. On the other hand, if $(1-p)n$ is large, then one would get an $\hat{\alpha}$ intermediate between the true α and 2, because the central part of the data is coming from a light tailed density. An incorrect model is being fit to the data, so it is no surprise that we get the "wrong" α. DuMouchel's argument to let the tails speak for themselves is sound, though his suggestion to use the upper 10% of the sample to fit the tail is generally not appropriate, see McCulloch (1997). Fofack and Nolan (1998) show that the asymptotic power decay on a stable tail may take a long time to occur; for an arbitrary distribution, there is no general statement that can be made about what fraction of the tail is appropriate. (For a recent summaries of work on tail estimation, see Beirlant, Vynckier and Teugels (1996), Embrechts, Kluüppelberg and Mikosch (1997) and Reiss and Thomas (1997).)

The diagnostics we are about to discuss are an attempt to detect non-stability. As with any other family of distributions, it is not possible to prove that a given data set is or is not stable. We note that even testing for normality is still an active field of research, e.g. Brown and Hettmansperger (1996). The best we can do is determine whether or not the data are consistent with the hypothesis of stability. These tests will fail if the departure from stability is small or occurs in an unobserved part of the range.

The first step we suggest is to compare estimates from the quantile method, the sample characteristic function method, and maximum likelihood. If they are close, then this supports the idea that the data is stably distributed. If the parameters differ significantly and the sample is large, then this argues that the data is not stably distributed, because the different estimators are all consistent estimators. For the maximum likelihood estimators, we do have asymptotic variances of the estimators and can thus put confidence limits on the parameters. The other estimators do not yet have asymptotic variances worked out, so we cannot make precise statements about differences.

The next step is to do a smoothed density plot of the data. If there are clear multiple modes or gaps in the support, then the data can't come from a stable distribution. For density plots, we smoothed the data with a Gaussian kernel with standard deviation given by a "width" parameter. We found that the commonly suggested width of 2(inter-quartile range)$n^{-1/3}$ works reasonably when the tails of the data are not too heavy, say $\alpha > 1.5$, but works poorly for heavier tailed data. For such cases, we used trial and error to find a width parameter that was a small as possible without showing oscillations from individual data points. The density plots give a good sense of whether the fit matches the data near the mode of the distribution, but generally is uninformative on the tails where both the fitted density and the smoothed data density are small. If the smoothed density is plausibly stable, proceed with a stable fit and compare the fitted distribution with the data using q-q and p-p plots.

We note a practical problem with q-q plots for heavy tailed data. While using q-q plots to compare simulated stable data sets with the exact corresponding cumulative d.f., we routinely had two problems with extreme values: (1) most of the data is visually compressed to a small region and (2) on the tails there seems to be an unacceptably large amount of fluctuation around the theoretical straight line. For heavy tailed stable distributions, we should expect such fluctuations: if $X_{(i)}$ is the i^{th} order statistic from an i.i.d. stable sample of size n, $p = (i - \frac{1}{2})/n$ and x_p is the p^{th} percentile, then for n large, the distribution of $X_{(i)}$ is approximately normal with $\mathrm{E}X_{(i)} = x_p$ and $\mathrm{Var}(X_{(i)}) = p(1-p)/nf(x_p)^2$, e.g. page 91 of Ferguson (1996). These formulas are used to show pointwise 95% confidence bounds around the expected value in the figures below. In sum, q-q plots may appear non-linear on the tails, even when the data set is stable.

Because q-q plots have some problems, we also recommend using a modified p-p plot. Standard p-p plots tend to emphasize behavior around the mode of the distribution, where they have more variation, and necessarily pinch the curve near the tails. In Michael (1983), a "stabilized" p-p plot was defined that eliminates this non-uniformity by using a transformation. (The word stabilized refers to making the variance in the p-p plot uniform, and has nothing to do with stable distributions. To stress this, we will use the phrase "variance stabilized" p-p plot.) The result is better than the regular p-p plot for detecting a poor fit near the extremes.

Finally, we tried comparing distribution functions, but did not find it very helpful. Because of the curvature in the distribution functions, it is hard to compare the fitted and empirical d.f. visually, especially on the tails.

3 Applications

3.1 Simulated stable data

Three data sets were generated, each of size $n = 1,000$ using the method of Chambers, Mallows and Stuck (1976). The first example used $(\alpha, \beta, \gamma, \delta) = (0.7, 0.5, 1, 0)$ and the ML parameter estimates with naive 95% confidence intervals are respectively 0.746 ± 0.062, 0.447 ± 0.075, 0.929 ± 0.116, and -0.005 ± 0.081. The second example used $(\alpha, \beta, \gamma, \delta) = (1.3, 0.5, 1, 0)$ and the ML parameter estimates with naive 95% confidence intervals are respectively 1.347 ± 0.092, 0.485 ± 0.129, 0.998 ± 0.072, and 0.027 ± 0.105. The third example used $(\alpha, \beta, \gamma, \delta) = (1.8, 0.5, 1, 0)$ and the ML parameter estimates with naive 95% confidence intervals are respectively 1.850 ± 0.084, 0.374 ± 0.324, 1.026 ± 0.054, and 0.013 ± 0.108. Diagnostics are shown in Figures 2, 3, and 4 respectively.

The problems with q-q plots mentioned above are clear for $\alpha = 0.7$, persist with $\alpha = 1.3$, and are minor for $\alpha = 1.8$. To show how a normal distribution describes these data sets, the density plots also show a normal fit, where the sample mean and variance are used as the parameters. When $\alpha = 0.7$, the sample variance is so large that the normal fit appears to be a flat line. When $\alpha = 1.3$, the normal fit is far from the data and stable fit. When $\alpha = 1.8$, the normal fit still differs noticeably from the data and stable fit.

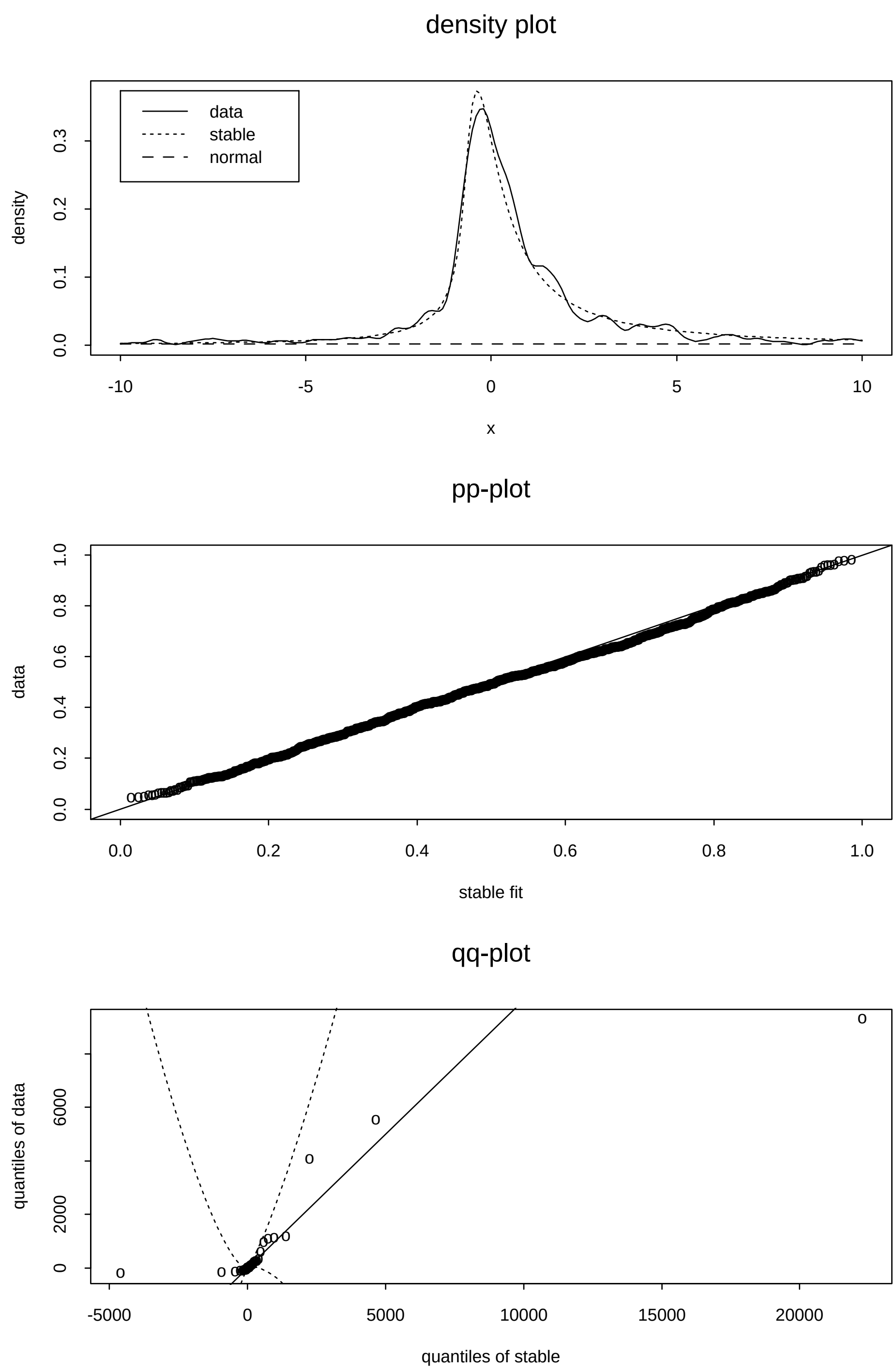

Figure 2: Simulated stable data set with n=1000, $\alpha = 0.7$, $\beta = 0.5$,$\gamma = 1$, $\delta = 0$.

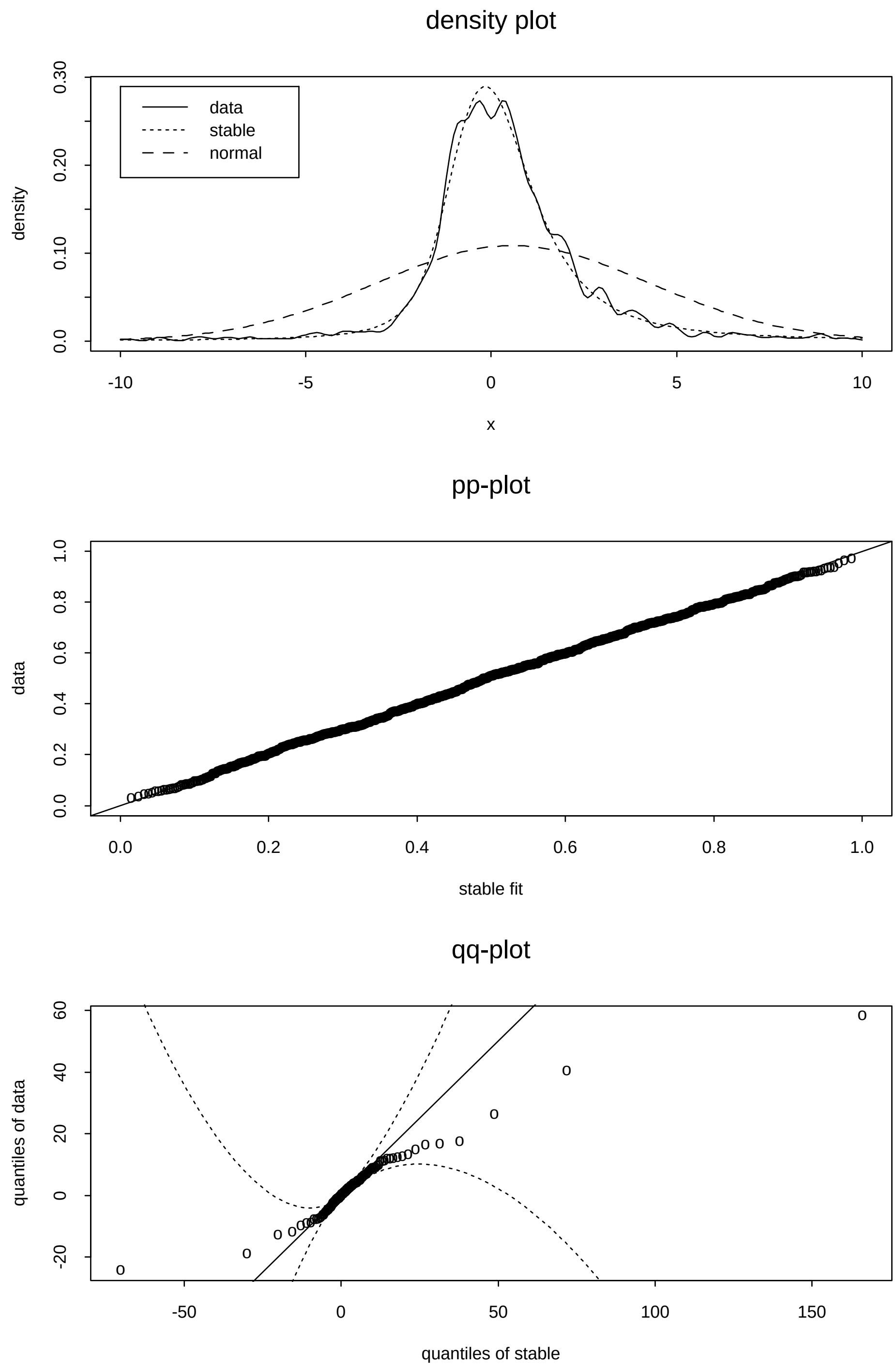

Figure 3: Simulated stable data set with n=1000, $\alpha = 1.3$, $\beta = 0.5$,$\gamma = 1$, $\delta = 0$.

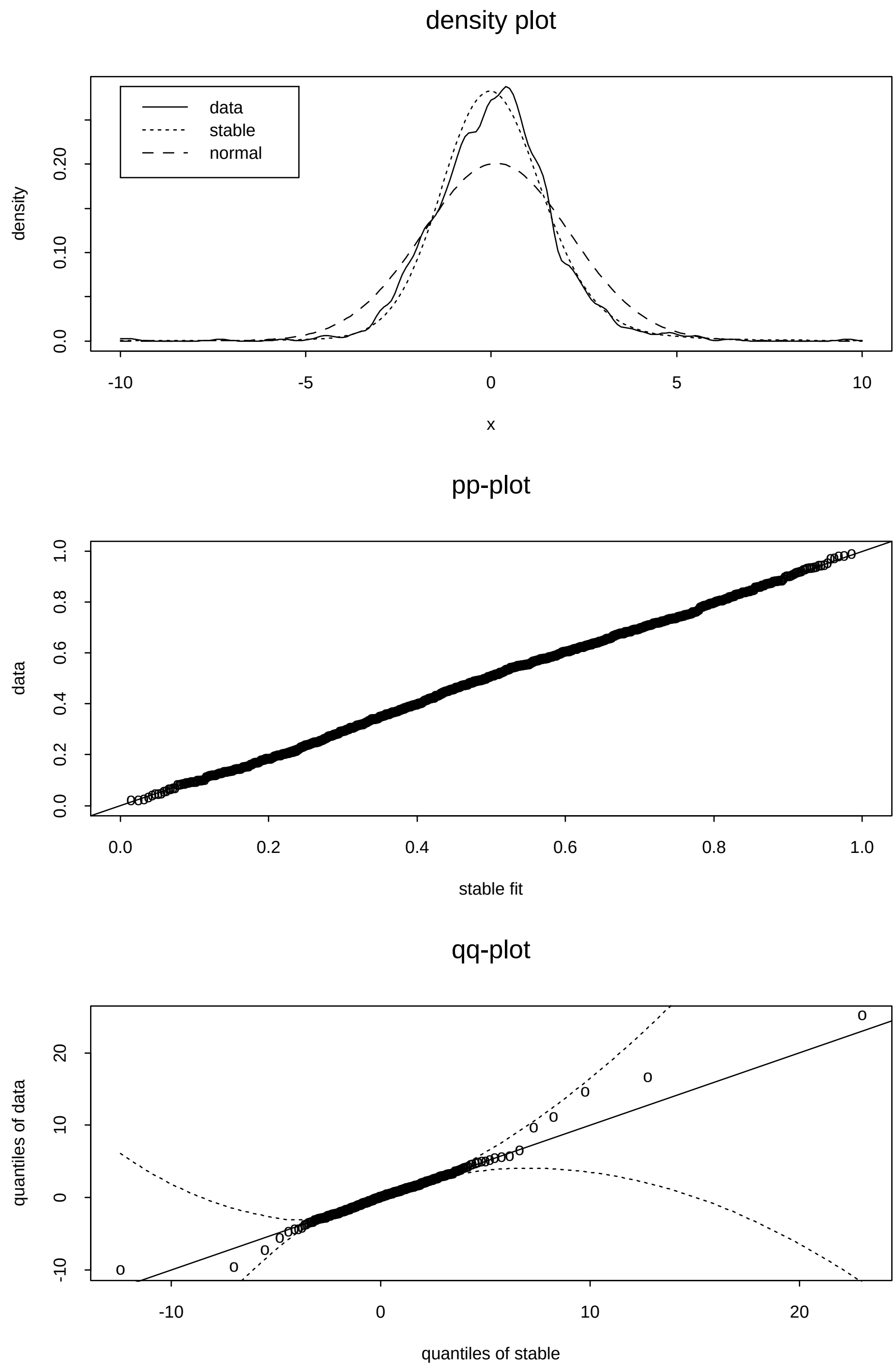

Figure 4: Simulated stable data set with n=1000, $\alpha = 1.8$, $\beta = 0.5$,$\gamma = 1$, $\delta = 0$.

country	α	β	γ	δ
Australia	1.479 ± 0.047	0.033 ± 0.080	0.00413 ± 0.00013	-0.00015 ± 0.00022
Austria	1.559 ± 0.047	-0.119 ± 0.092	0.00285 ± 0.00009	0.00014 ± 0.00015
Belgium	1.473 ± 0.047	-0.061 ± 0.080	0.00306 ± 0.00010	0.00009 ± 0.00016
Canada	1.574 ± 0.047	-0.051 ± 0.093	0.00379 ± 0.00012	0.00004 ± 0.00020
Denmark	1.545 ± 0.047	-0.119 ± 0.090	0.00272 ± 0.00008	0.00022 ± 0.00014
France	1.438 ± 0.047	-0.146 ± 0.078	0.00245 ± 0.00008	0.00028 ± 0.00013
Germany	1.495 ± 0.047	-0.182 ± 0.085	0.00244 ± 0.00008	0.00019 ± 0.00013
Italy	1.441 ± 0.046	-0.043 ± 0.076	0.00266 ± 0.00009	0.00017 ± 0.00014
Japan	1.511 ± 0.047	-0.148 ± 0.086	0.00368 ± 0.00012	0.00013 ± 0.00019
Netherlands	1.467 ± 0.047	-0.167 ± 0.081	0.00244 ± 0.00008	0.00016 ± 0.00013
Norway	1.533 ± 0.047	-0.070 ± 0.088	0.00253 ± 0.00008	0.00005 ± 0.00013
Spain	1.512 ± 0.047	-0.007 ± 0.083	0.00268 ± 0.00008	0.00012 ± 0.00014
Sweden	1.517 ± 0.047	-0.081 ± 0.085	0.00256 ± 0.00008	0.00006 ± 0.00013
Switzerland	1.599 ± 0.047	-0.179 ± 0.100	0.00295 ± 0.00009	0.00014 ± 0.00016
United States	1.530 ± 0.047	-0.088 ± 0.088	0.00376 ± 0.00012	0.00009 ± 0.00020

Table 1: Exchange rate analysis. Maximum likelihood parameter estimates and 95% confidence intervals with sample size of $n = 4274$.

3.2 Exchange rate data

Daily exchange rate data for 15 different currencies were recorded (in U.K. pounds) over a 16 year period (2 January 1980 to 21 May 1996). The data was transformed by $y_t = \ln(x_{t+1}/x_t)$, giving $n = 4,274$ data values. The transformed data was fit with a stable distribution; results are shown in Table 1.

Figure 5 shows smoothed density, q-q plot and variance stabilized p-p plot for the German mark data set. The data sets are clearly not normal: the heavy tails in the data causes the sample variance to be large, and the normal fit poorly describes both the center and the tails of the distribution. The q-q plot shows that the extreme tails of the data set are lighter than the stable model. The horizontal line segment at the center of the p-p graph is from the days where the exchange rate was unchanged on successive days. As another measure of non-normality, the ratio of the stable fit log likelihood to the normal log likelihood was computed for each currency. The ratio of the log likelihoods for the ML stable fit to the normal fit were computed and the values ranged from 113 to 1041.

Plots for the other currencies were similar, showing that the stable fit does a reasonable job of describing the exchange rate data. We note in passing that the currency with the heaviest tails ($\hat{\alpha} = 1.441$) was the Italian lire, while the one with the lightest tails ($\hat{\alpha} = 1.530$) was the Swiss Franc.

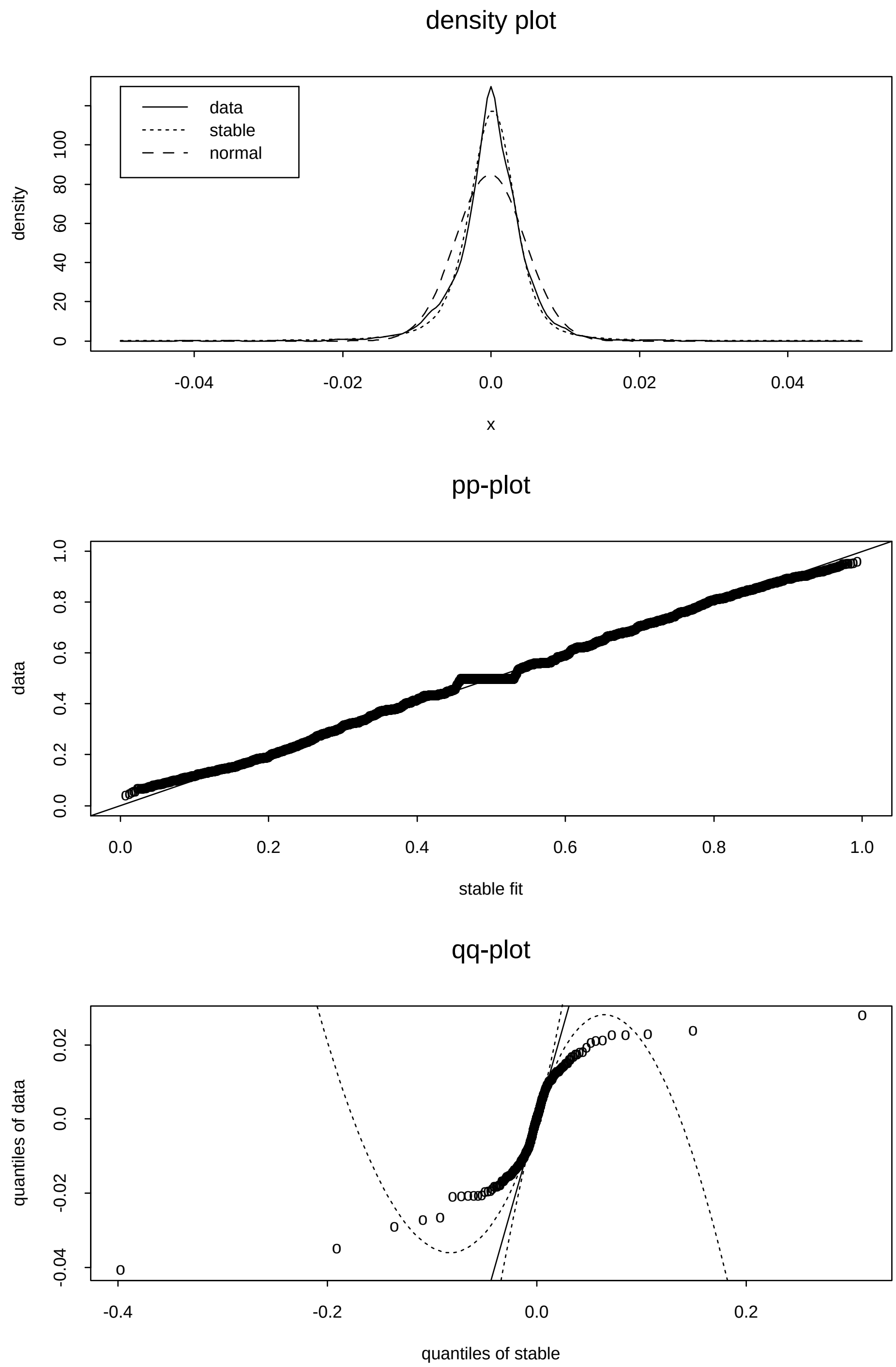

Figure 5: Density, variance stabilized p-p and q-q plots for the German mark exchange rate data, n=4274.

method	α	β	γ	δ
quantile	1.996	1.000	.008579	-.003445
MCMC	1.650	.768	.007900	-.003187
ML	$1.518 \pm .422$	$.743 \pm .651$	$.006828 \pm .001931$	$-.003064 \pm .003359$

Table 2: Abbey National share price parameter estimates, $n = 49$.

3.3 CRSP stock prices

McCulloch (1997) analyzed forty years (January 1953 - December 1992) of monthly stock price data from the Center for Research in Security Prices (CRSP). The data set consists of 480 values of the CRSP value-weighted stock index, including dividends, and adjusted for inflation. The quantile estimates were $\hat{\alpha} = 1.965$, $\hat{\beta} = -1$, $\hat{\gamma} = 2.755$ and $\hat{\delta} = 0.896$. McCulloch used ML with symmetric stable distributions to fit this data and obtained $\hat{\alpha} = 1.845$, $\hat{\beta} = 0$, $\hat{\gamma} = 2.712$ and $\hat{\delta} = 0.673$. Our ML estimates with naive 95% confidence intervals are $\hat{\alpha} = 1.855 \pm 0.110$, $\hat{\beta} = -0.558 \pm 0.615$, $\hat{\gamma} = 2.711 \pm 0.213$ and $\hat{\delta} = 0.871 \pm 0.424$. The diagnostics in Figure 6 show a close fit.

We note that the confidence interval for $\hat{\alpha}$ is close to the upper bound of 2 for α and the one for $\hat{\beta}$ is large and extends beyond the lower bound of -1, so the asymptotic normality of these parameters has not been achieved and the naive confidence intervals should not be strictly believed.

3.4 Abbey National share price

Buckle (1995) listed a small data set of stock price data. The price for Abbey National shares was recorded for the period 31 July 1991 through 8 October 1991. The return was defined as $(x_{t+1}/x_t) - 1$, yielding $n = 49$ data points, which were fit with a stable distribution. In the Monte Carlo Markov chain (MCMC) approach used by Buckle, the means of the posterior distributions were given. Table 2 lists these MCMC parameter estimates (transformed to the 0 parameterization), the quantile estimates, and the ML estimates with naive 95% confidence intervals.

The quantile method fit is essentially a normal distribution with $\alpha = 1.996$, yet highly skewed. This is likely caused by the small sample size: with $n = 49$, the 5^{th} percentile is found by interpolating between the second and third data point. It is hard to detect heavy tails when there is virtually no tail. The MCMC method and ML method reach similar estimates. We

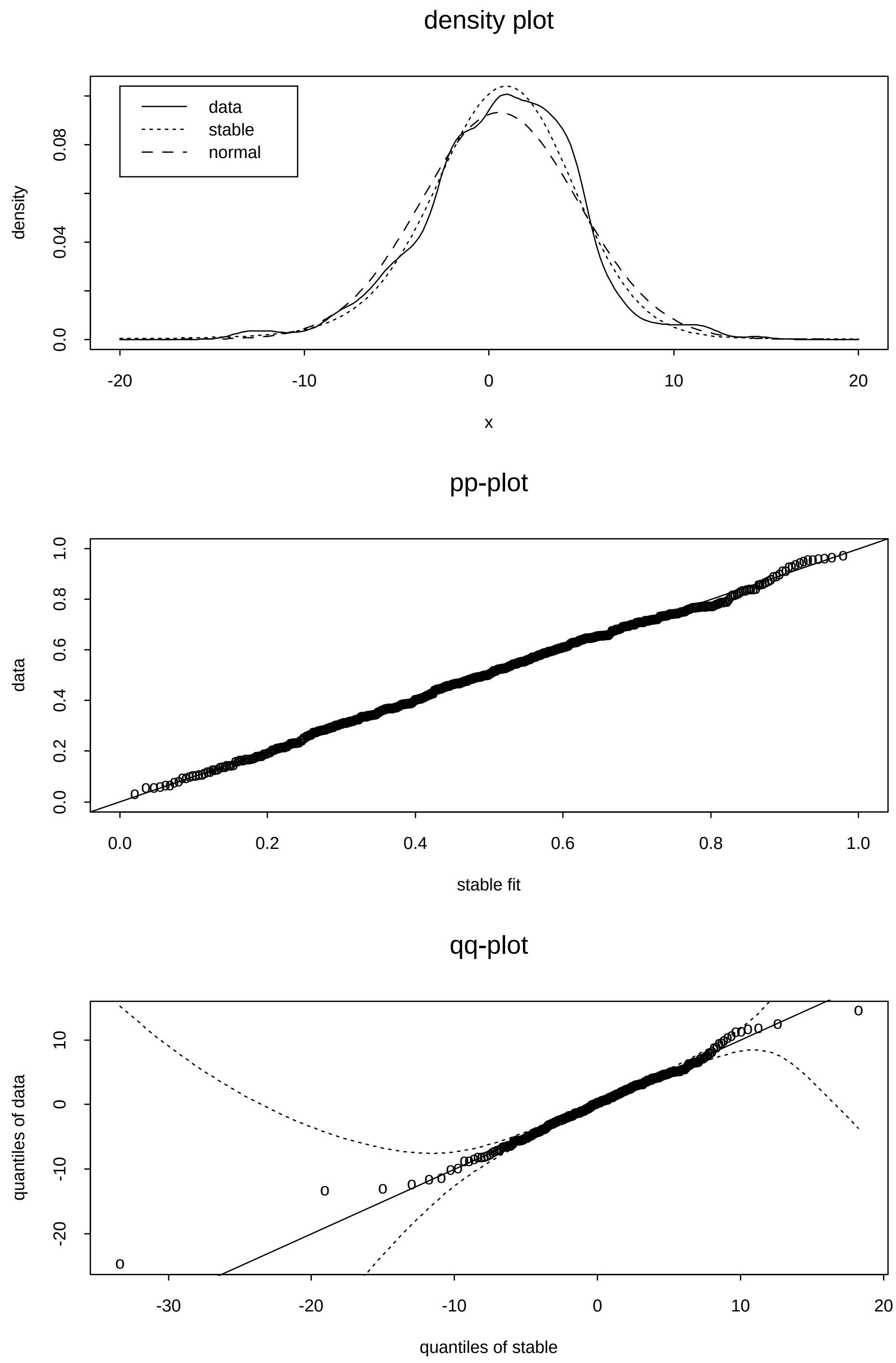

Figure 6: Density, variance stabilized p-p plot and q-q plots for the CRSP stock price data, n=480.

tried the diagnostics on this data set and got mixed results, see Figure 7. The p-p plot shows that the data are concentrated on a small set of values and it is not clear how good a stable model is for this small data set. In particular, healthy skepticism is called for when making statements about tail probabilities unless a large data set is available to verify stable behavior.

We note that the density plot here changed radically as the smoothing width parameter changed. Since the data is concentrated at a small number of discrete points, a small smoothing parameter gives a density with multiple spikes. We compromised on an intermediate value that showed location of spikes, but smoothed things out.

3.5 Radar noise

This is a very large data set with $n = 320,000$ pairs of data points. The two coordinates are the in-phase and quadrature components of sea clutter radar noise. We focus on the in-phase component only here; see Section 5 for an analysis of the bivariate distribution. The parameter estimates are $\hat{\alpha} = 1.7966 \pm .0048$, $\hat{\beta} = .0054 \pm .0173$, $\hat{\gamma} = .4402 \pm .0013$ and $\hat{\delta} = -.00060 \pm .00247$. (The quantile based estimators are $\hat{\alpha} = 1.7042$, $\hat{\beta} = .0058$, $\hat{\gamma} = .3981$ and $\hat{\delta} = -.00040$.) With this large sample size, the confidence intervals for the ML parameter estimates are very small. Again the correct question is not how tight the parameter estimates are, but whether or not the fit accurately describes the data. The plots in Figure 8 show a very close stable fit,even far out on the tails. Because 320,000 data points add little to the plots, we actually show thinned q-q and p-p plots with 1,000 values.

3.6 Ocean wave power

Pierce (1997) proposed using positive α-stable distributions to model inherently positive quantities such as energy or power. One example he uses is the power in ocean waves, which is proportional to the square of the wave height. Pierce lists a National Oceanographic and Atmospheric Administration (NOAA) web site where hourly wave data can be downloaded. We downloaded the same data set, edited out invalid numbers (99.00) and had 8084 values for the wave height variable WVHT. Pierce compared the data with an α=0.75, $\beta = 1$ distribution (it is not indicated how these values are obtained). Our analysis gave quantile estimates of $\hat{\alpha} = 1.139255$, $\hat{\beta} = 1$, $\hat{\gamma} = 0.813324$ and $\hat{\delta} = 0.841235$; the ML estimates with naive 95% confidence intervals are $\hat{\alpha}$=0.800 $\pm$ 0.0177, $\hat{\beta}$=1 $\pm$ 0, $\hat{\gamma}$=0.566 $\pm$ 0.018 and $\hat{\delta}$=0.965 $\pm$ 0.021. The fact that we get very different estimates of α is an indication that

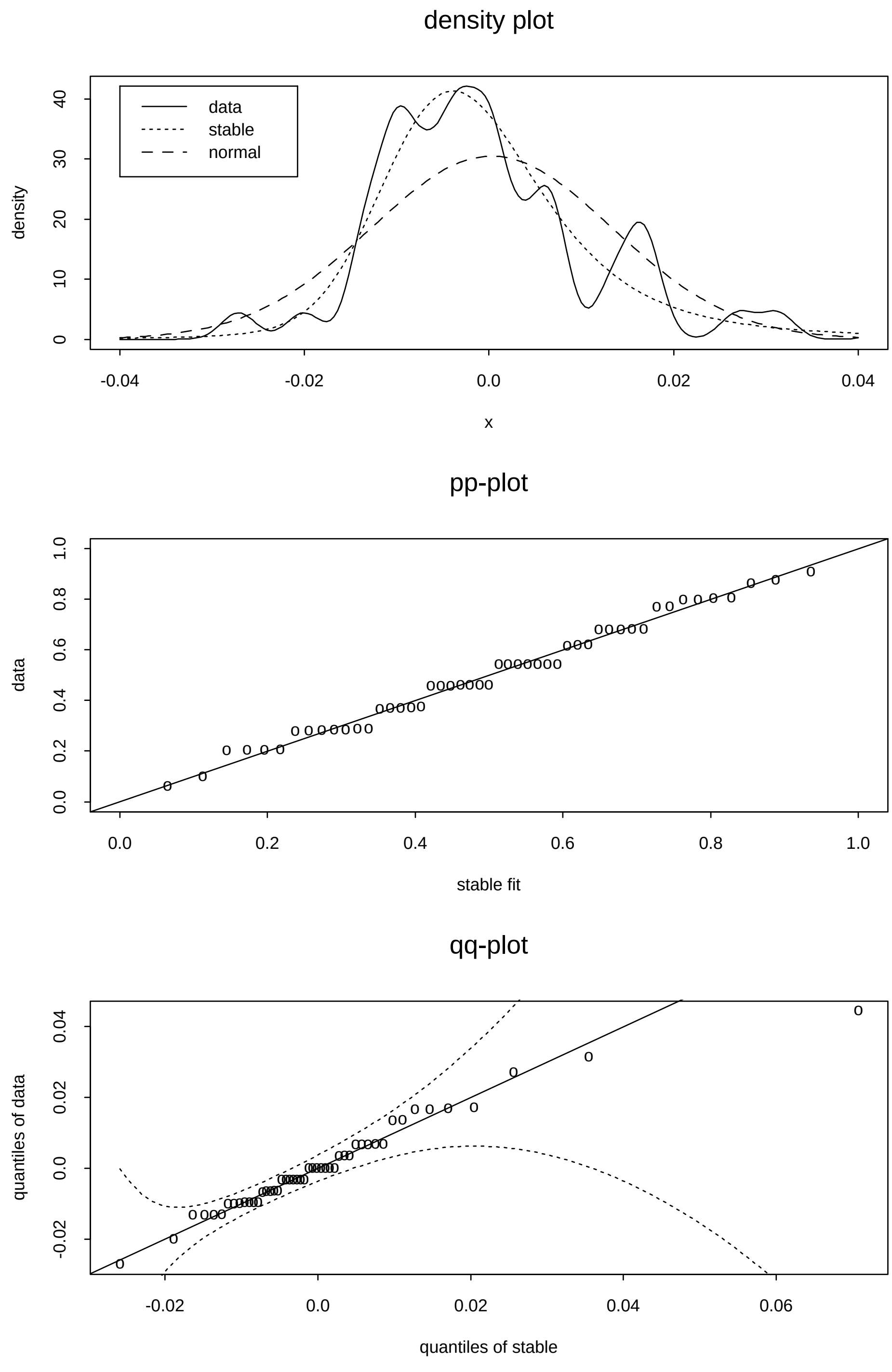

Figure 7: Density, variance stabilized p-p, and q-q plots and densities for Abbey share price data, n=49.

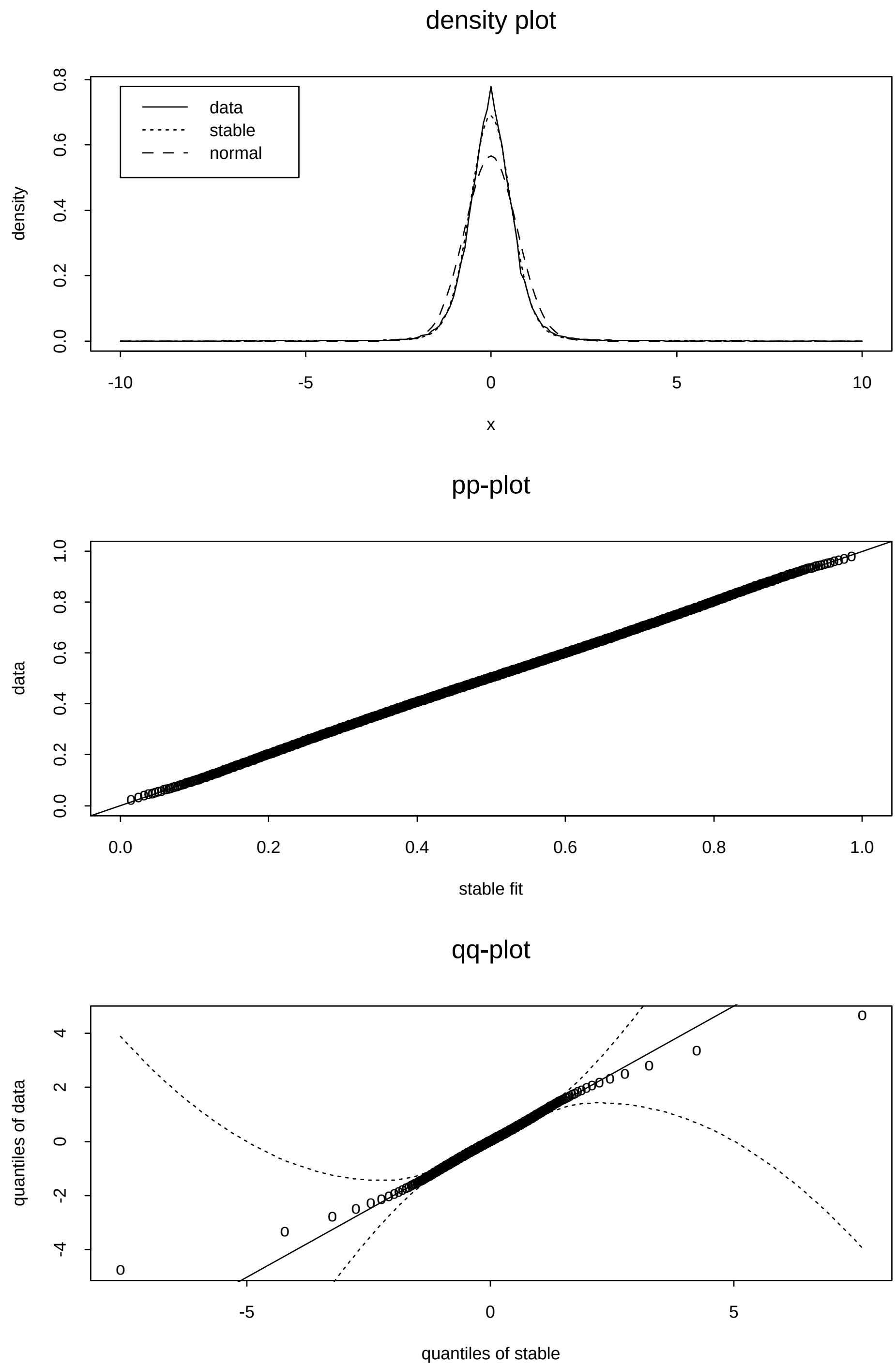

Figure 8: Density, variance stabilized p-p, and q-q plots for the in-phase component of sea clutter radar noise, n=320000.

the data set is not stable. The diagnostics in Figure 9 support this idea. The plots show a reasonable fit around the mode, but a poor fit on both tails. While it is possible that the power in waves is stably distributed, but that measurement of extremes (both high and low) of wave height are inaccurate, we would reject a stable model in this problem.

3.7 Simulated non-stable data

We simulated several data sets that were not stable and used our diagnostics to assess the fit with a stable model. The first data set is a mixture of two Cauchy distributions with different modes: $\alpha = 1$, $\beta = 0$, $\gamma = 1$, with $\delta = 5$ for 100 data points and then $\delta = -5$ for another 100 data points. Both the p-p plot and the density plot in Figure 10 show the bimodality, so a stable model is clearly not appropriate. Still, it is instructive to see what happens if we fit these data with a stable model. The maximum likelihood estimates are $\alpha = 2$, $\gamma = 3.867$, and $\delta_0 = -.395$ (β is meaningless in the normal case). Apparently the likelihood for this data set is dominated by the central terms and is maximized by taking a normal curve with large variance. Even though this is a heavy tailed data set, the use of an inappropriate stable model leads to a light tailed fit! This is an example where the numerical maximum likelihood method might get trapped in a local minimum, centered on one of the modes, and lead to a "wrong" answer. In this case, there is no right answer, but simple diagnostics show the data is not stably distributed.

The next example is a simulated data set consisting of a mixture of 9,000 Gaussian random variables with scale 1 and 1,000 Gaussian random variables with scale 10, a "contaminated" normal mixture. The mixture has heavier tails than a pure normal, so one might try to fit it with a stable distribution. However, what we would really like to do is detect that it is not a stably distributed data set. The ML estimates of the parameters are $\alpha = 1.346 \pm .030$ and $\gamma = 1.048 \pm .033$ (the search was restricted to $\beta = 0$, $\delta = 0$). Here the confidence intervals are small because the sample size of $n = 10,000$, not because we have a good fit. The density plot in Figure 11 shows the smoothed data density and the stable fit. The curves show a systematic difference that indicates departure from a stable distribution. It is interesting to note that in this example, the percentile estimate of α is 1.535, quite different from the ML estimate. This is another indication that the data is not stable: if the distribution is stable, then all consistent estimators of the parameters should be close when there is a large sample.

Next, we generated a data set that was the sum of 1,000 independent normal and Cauchy variables, a data set that is not stable. The ML estimates

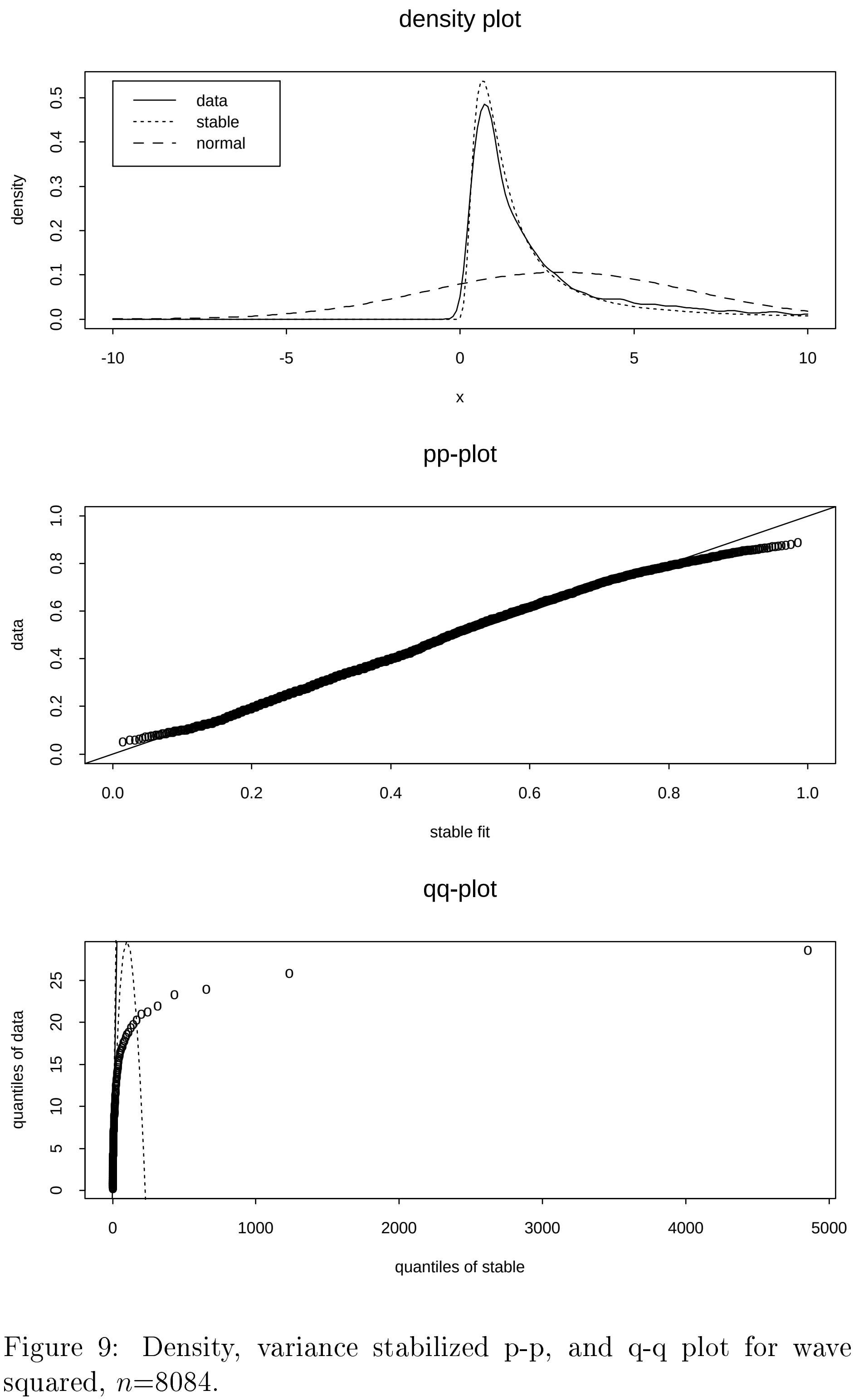

Figure 9: Density, variance stabilized p-p, and q-q plot for wave height squared, n=8084.

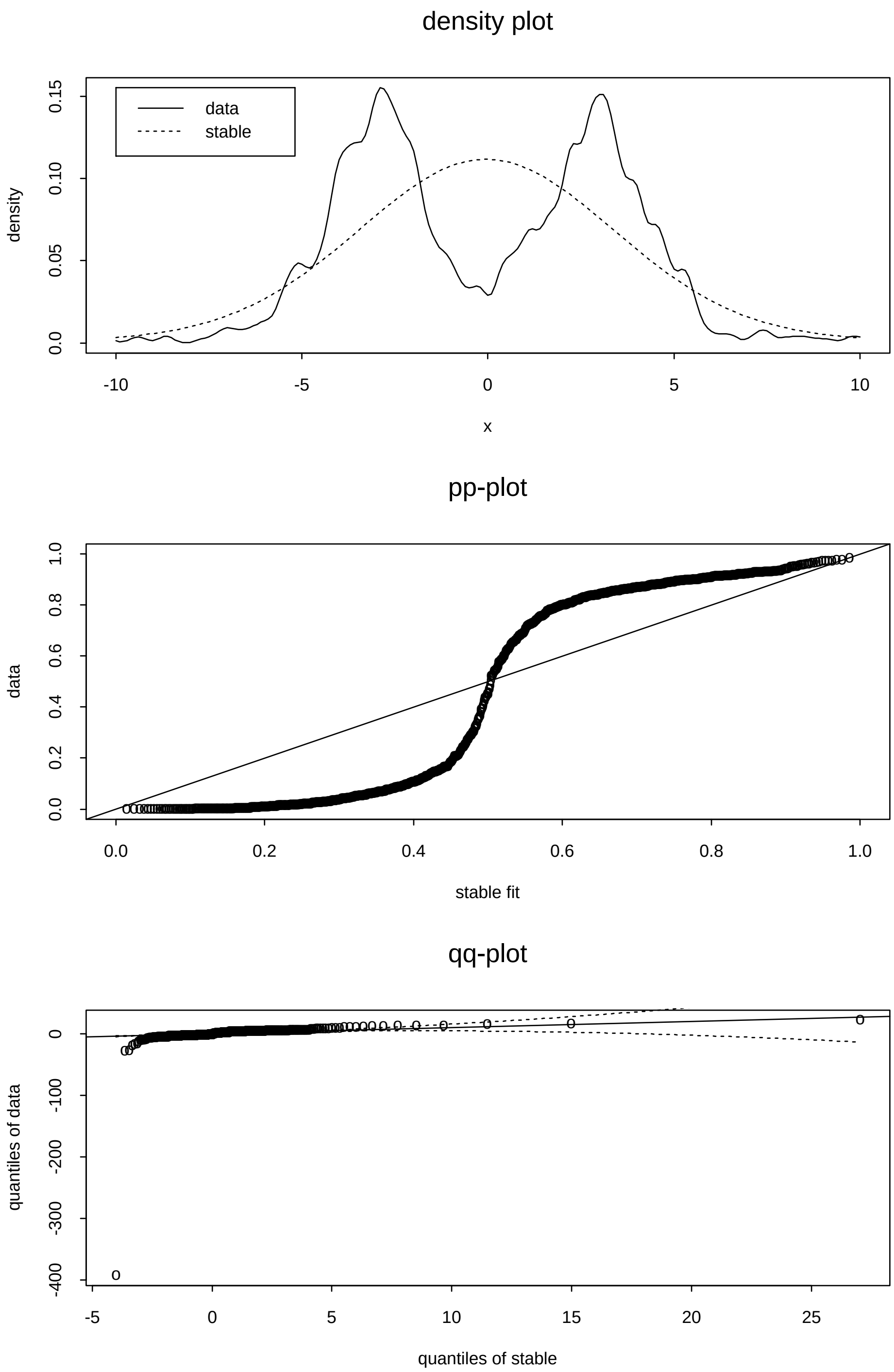

Figure 10: Simulated bimodal mixture data with heavy tails, n=200.

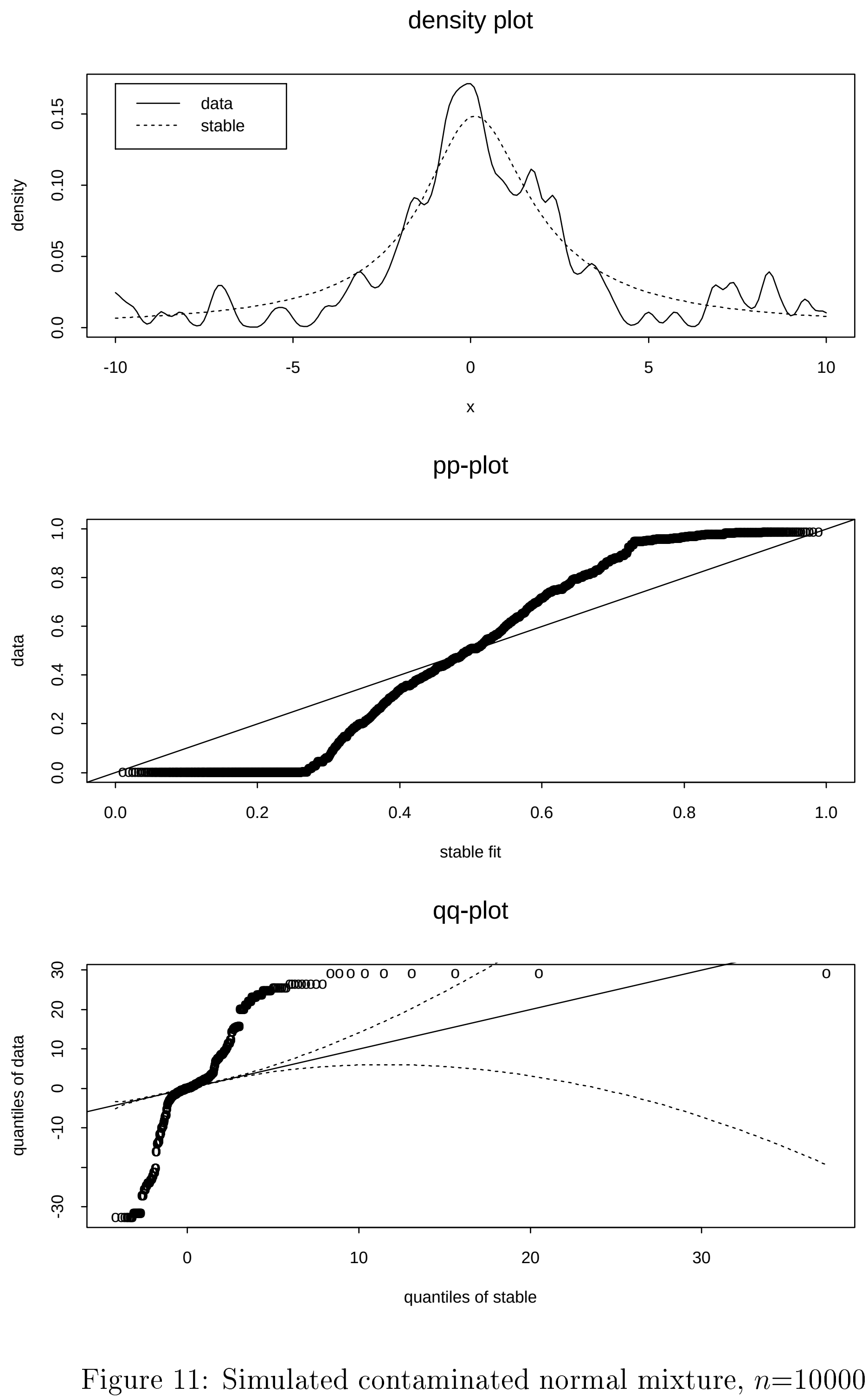

Figure 11: Simulated contaminated normal mixture, n=10000.

of the parameters are $\hat{\alpha} = 1.20$, $\hat{\beta} = 0.030$, $\hat{\gamma} = 1.344$, $\hat{\delta} = -0.089$. Figure 12 shows the diagnostics are not good at detecting non-stability here. Our only intuition about why this is true is that the true ch. f. $\exp(-u^2/2 - |u|)$ is close to the fitted ch. f. $\exp(-|1.344u|^{1.20})$ for small and moderate u values, so the distributions are close in an L^1 and L^∞ sense. Consequently, it will be hard to discriminate between this kind of distribution and a stable one from data, unless there is a massive data set.

We briefly mention two other experiments we did. In one experiment, 10000 variables were generated from a Pareto distribution ($F(x) = 1 - x^{-1-\alpha}$; $x > 1$) with $\alpha = 1.5$. The quantile and ML estimates of α were 1.23 and 0.9 respectively, β was essentially 1. This shows that a stable fit to a data set with genuine Pareto tails will give poor estimates of the tail index. In the second experiment 10000 Gamma(2) variates were generated and fit with a stable distribution. The quantile and ML estimates of α were 1.98 and 1.80 respectively, β was essentially 1. This shows that the light tails of the Gamma distribution lead to estimates of α close to the Gaussian case, but the skewed nature of the data showed up in the estimate of β. While not shown, the diagnostics did show the non-stability of these data sets.

4 Multivariate stable distributions

A formal definition for random vector $\mathbf{X} = (X_1, X_2, \ldots, X_d)$ to be stable is given in Section 2.1 of Samorodnitsky and Taqqu (1994). The "jointly stable" is sometimes used to stress the fact that the definition forces all the components X_j to be univariate α-stable with one α. This follows from the following theorem and justifies the term α-stable random vector.

Theorem 1 *(i) Let* $\mathbf{X}$ *be a stable random vector. Then every one dimensional projection* $\mathbf{u} \cdot \mathbf{X} = \sum u_i X_i$ *is one dimensional stable random variable with the same index* α *for every* $\mathbf{u}$.

(ii) Conversely, suppose $\mathbf{X}$ *is a random vector with the property that every one-dimensional projection* $\mathbf{u} \cdot \mathbf{X}$ *is one dimensional stable, e.g.* $\mathbf{u} \cdot \mathbf{X} \sim \mathbf{S}(\alpha(\mathbf{u}), \beta(\mathbf{u}), \gamma(\mathbf{u}), \delta(\mathbf{u}); 1)$. *Then there is one* α *that is the index of all projections, i.e.* $\alpha(\mathbf{u}) = \alpha$ *is constant. If* $\alpha \geq 1$, *then* $\mathbf{X}$ *is stable. If* $\alpha < 1$ *and the location parameter function* $\delta(\mathbf{u})$ *and the vector of location parameters* $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_d)$ *of the components* $X_1, X_2, \ldots, X_d$ *(all in the 1 parameterization) are related by*

$$\delta(\mathbf{u}) = \mathbf{u} \cdot \boldsymbol{\delta}, \tag{3}$$

then $\mathbf{X}$ *is stable.*

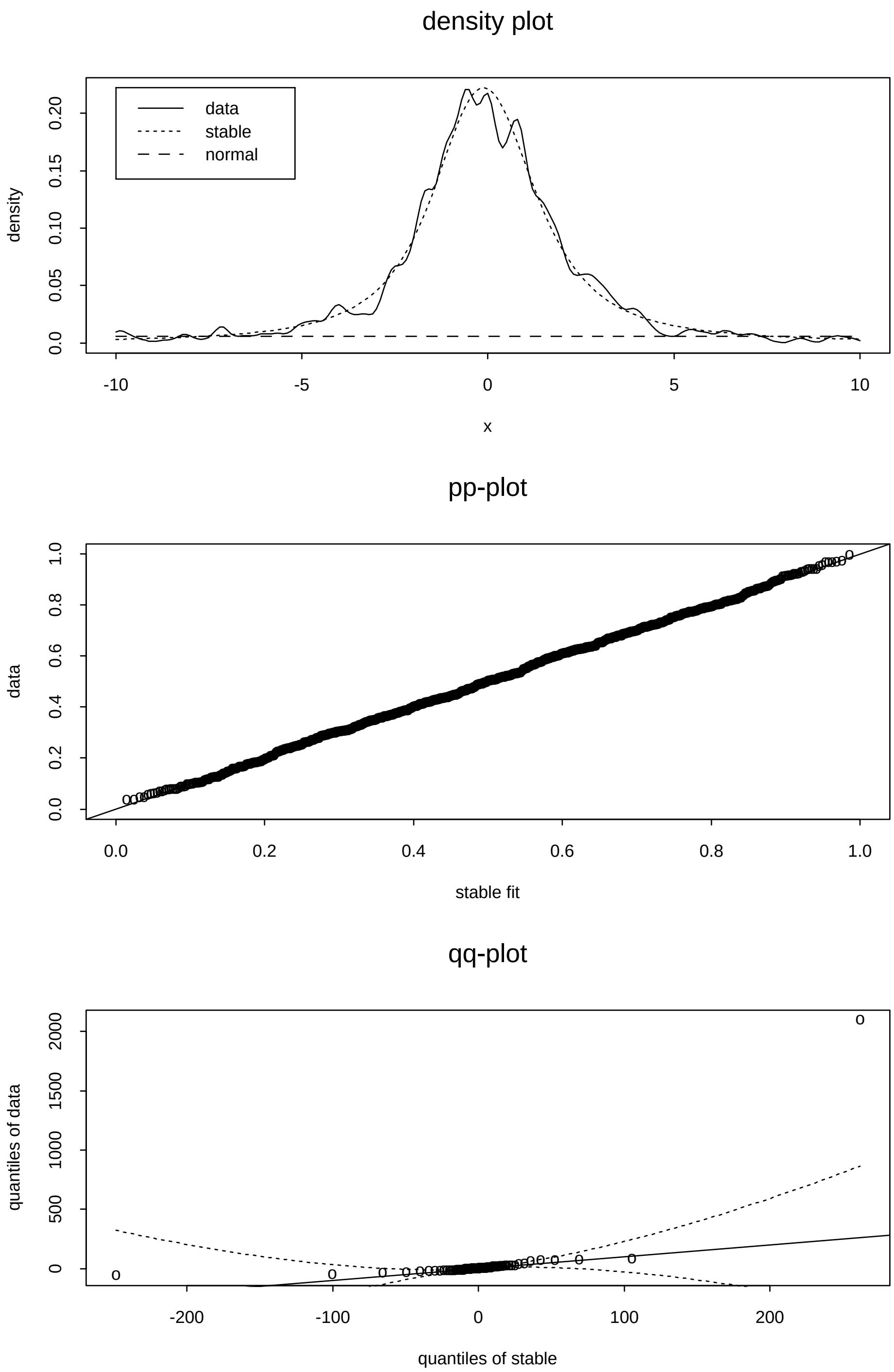

Figure 12: Sum of normal and Cauchy data.

Proof The first part is Theorem 2.1.2 of Samorodnitsky and Taqqu (1994). The second part is Theorem 2.1.5(c) of Samorodnitsky and Taqqu (1994) when $\alpha \geq 1$. It remains to show that if $\alpha < 1$ and (3) holds, then $\mathbf{X}$ is stable. To show this, define $\mathbf{Y} = \mathbf{X} - \boldsymbol{\delta}$. Then $\mathbf{Y}$ has the property that every one dimensional projection is strictly stable: $\mathbf{u} \cdot \mathbf{Y} = \mathbf{u} \cdot \mathbf{X} - \mathbf{u} \cdot \boldsymbol{\delta}$ has location parameter (in the 1 parameterization) $\delta(\mathbf{u}) - \mathbf{u} \cdot \boldsymbol{\delta} = 0$. Thus for every $\mathbf{u}$, $\mathbf{u} \cdot \mathbf{Y}$ is strictly stable, so by Theorem 2.1.5(a) of Samorodnitsky and Taqqu (1994), $\mathbf{Y}$ is strictly stable. The definition of stable shows that any shift of a stable r. vector is stable, so $\mathbf{X} = \mathbf{Y} + \boldsymbol{\delta}$ is also stable. □

We note that (3) holds automatically when $\alpha > 1$, so the condition is only required when $\alpha < 1$. Furthermore, (3) is necessary when $\alpha \neq 1$, so it cannot be dropped. Section 2.2 of Samorodnitsky and Taqqu (1994) gives an example due to David J. Marcus where $\alpha < 1$ and all one dimensional projections are stable, but (3) fails and $\mathbf{X}$ is not jointly stable.

One advantage of the preceding Theorem is that it gives a way of parameterizing multivariate stable distributions in terms of one dimensional projections. For any vector $\mathbf{u} \in \mathbf{R}^d$, $\mathbf{u} \cdot \mathbf{X} \sim \mathbf{S}(\alpha, \beta(\mathbf{u}), \gamma(\mathbf{u}), \delta(\mathbf{u}); k)$, $k = 0, 1$. Thus we know the (univariate) characteristic function of $\mathbf{u} \cdot \mathbf{X}$ for every $\mathbf{u}$, and hence the joint characteristic function of $\mathbf{X}$. Therefore α and the functions $\beta(\cdot)$, $\gamma(\cdot)$ and $\delta(\cdot)$ completely characterize the joint distribution. In fact, knowing these functions on the sphere $\mathbf{S}^d = \{\mathbf{u} \in \mathbf{R}^d : |\mathbf{u}| = 1\}$ characterizes the distribution.

The functions $\beta(\cdot)$, $\gamma(\cdot)$ and $\delta(\cdot)$ must satisfy certain regularity conditions. The standard way of describing multivariate stable distributions is in terms of a finite measure Λ on the sphere $\mathbf{S}^d$, called the spectral measure. The following result is due to Feldheim (1937), Section 2.3 of Samorodnitsky and Taqqu (1994) contains a proof. It is typical to use the spectral measure to describe the joint characteristic function, we find it more convenient to relate it to the functions $\beta(\cdot)$, $\gamma(\cdot)$, and $\delta(\cdot)$.

Theorem 2 *Let $\mathbf{X} = (X_1, \ldots, X_d)$ be jointly stable, say*

$$\mathbf{u} \cdot \mathbf{X} \sim \mathbf{S}(\alpha, \beta(\mathbf{u}), \gamma(\mathbf{u}), \delta(\mathbf{u}); k), \quad k = 0, 1.$$

Then there exists a finite measure Λ on $\mathbf{S}^d$ and a location vector $\boldsymbol{\delta} \in \mathbf{R}^d$ with

$$\gamma(\mathbf{u}) = \left(\int_{\mathbf{S}^d} |\mathbf{u} \cdot \mathbf{s}|^\alpha \Lambda(d\mathbf{s}) \right)^{1/\alpha}$$

$$
\begin{aligned}
\beta(\mathbf{u}) &= \frac{\int_{\mathbf{S}^d} |\mathbf{u}\cdot\mathbf{s}|^\alpha \text{sign}\,(\mathbf{u}\cdot\mathbf{s})\Lambda(d\mathbf{s})}{\int_{\mathbf{S}^d} |\mathbf{u}\cdot\mathbf{s}|^\alpha \Lambda(d\mathbf{s})} \\
\delta(\mathbf{u}) &= \begin{cases}
\boldsymbol{\delta}\cdot\mathbf{u} & k=1, \alpha\neq 1 \\
\boldsymbol{\delta}\cdot\mathbf{u} - \frac{2}{\pi}\int_{\mathbf{S}^d}(\mathbf{u}\cdot\mathbf{s})\ln|\mathbf{u}\cdot\mathbf{s}|\Lambda(d\mathbf{s}) & k=1, \alpha=1 \\
\boldsymbol{\delta}\cdot\mathbf{u} + (\tan\frac{\pi\alpha}{2})\beta(\mathbf{u})\gamma(\mathbf{u}) & k=0, \alpha\neq 1 \\
\boldsymbol{\delta}\cdot\mathbf{u} - \frac{2}{\pi}\int_{\mathbf{S}^d}(\mathbf{u}\cdot\mathbf{s})\ln(\mathbf{u}\cdot\mathbf{s})\Lambda(d\mathbf{s}) & \\
\qquad +\frac{2}{\pi}\beta(\mathbf{u})\gamma(\mathbf{u})\ln\gamma(\mathbf{u}) & k=0, \alpha=1.
\end{cases}
\end{aligned}
$$

Proof Theorem 2.3.1 and Example 2.3.4 of Samorodnitsky and Taqqu (1994) give the formulas for $\gamma(\mathbf{u})$, $\beta(\mathbf{u})$ and the formulas for $\delta(\mathbf{u})$ when $k=1$. For $\delta(\mathbf{u})$ when $k=0$, use (2). We note that the expressions can be rewritten when $k=0, \alpha\neq 1$ as

$$
\begin{aligned}
\delta_0(\mathbf{u}) &= \delta_1(\mathbf{u}) + \tan\frac{\pi\alpha}{2}\beta(\mathbf{u})\gamma(\mathbf{u}) \\
&= \delta_1(\mathbf{u}) + \tan\frac{\pi\alpha}{2}\left(\int_{\mathbf{S}^d}|\mathbf{u}\cdot\mathbf{s}|^\alpha \text{sign}\,(\mathbf{u}\cdot\mathbf{s})\Lambda(d\mathbf{s})/\gamma^\alpha(\mathbf{u})\right)\gamma(\mathbf{u}) \\
&= \boldsymbol{\delta}\cdot\mathbf{u} + \tan\frac{\pi\alpha}{2}\int_{\mathbf{S}^d}|\mathbf{u}\cdot\mathbf{s}|^\alpha \text{sign}\,(\mathbf{u}\cdot\mathbf{s})\Lambda(d\mathbf{s})\left(\int_{\mathbf{S}^d}|\mathbf{u}\cdot\mathbf{s}|^\alpha\Lambda(d\mathbf{s})\right)^{(1/\alpha)-1}
\end{aligned}
$$

Likewise, when $k=0, \alpha=1$,

$$
\begin{aligned}
\delta_0(\mathbf{u}) &= \delta_1(\mathbf{u}) + \frac{2}{\pi}\beta(\mathbf{u})\gamma(\mathbf{u})\ln\gamma(\mathbf{u}) \\
&= \boldsymbol{\delta}\cdot\mathbf{u} + \frac{2}{\pi}\int_{\mathbf{S}^d}\mathbf{u}\cdot\mathbf{s}\left[\ln\int_{\mathbf{S}^d}|\mathbf{u}\cdot\mathbf{s}|\Lambda(d\mathbf{s}) - \ln|\mathbf{u}\cdot\mathbf{s}|\right]\Lambda(d\mathbf{s})
\end{aligned}
$$

□

It is possible for $\mathbf{X}$ to be non-degenerate, but singular. For example, $\mathbf{X} = (X_1, 0)$ is formally a two dimensional stable distribution if X_1 is univariate stable, but it is supported on the one dimensional subspace $\mathbf{R}\times\{0\}$. In what follows, we will always assume that $\mathbf{X}$ is d-dimensional. It can be shown that the following are equivalent:

- $\mathbf{X}$ is nonsingular.
- $\gamma(\mathbf{u}) > 0$ for all $\mathbf{u}\in\mathbf{R}^d$.
- span support$(\Lambda) = \mathbf{R}^d$.

5 Multivariate estimation

There are several methods of estimating for multivariate stable distributions; all involve some estimate of α and some discrete estimate of the spectral measure $\hat{\Lambda} = \sum_{k=1}^{m} \lambda_k 1_{\{\mathbf{s}_k\}}$, $\mathbf{s}_k \in \mathbf{S}^d$. If you know the distribution is isotropic (radially symmetric), then you can adapt the univariate fractional lower order moment method to d-dimensions using problem 4, pg. 44 of Nikias and Shao (1995) to estimate α and then the (constant) spectral measure. In general one should let the data speak for itself, and see if the spectral measure Λ is constant. Rachev and Xin (1993) and Cheng and Rachev (1995) use the fact that the directional tail behavior of multivariate stable distributions is Pareto, and base an estimate of Λ on this. Nolan, Panorska and McCulloch (1996) define two other estimates of Λ, one based on the joint empirical/sample ch. f. and one based on the one dimensional projections of the data.

We note that an estimate of Λ is preferred to an estimate of covariation, since covariation is only a partial measure of the dependence in a bivariate stable distribution, but Λ gives complete information. The covariation of the components of $\mathbf{X} = (X_1, X_2)$ can be obtained from Λ by $[X_1, X_2] = \int_{\mathbf{S}^d} t_1 t_2 |t_2|^{\alpha-2} \Lambda(d\mathbf{t})$.

Another advantage of Theorem 1 is that it gives a way of assessing whether a multivariate data set is stable by looking at just one dimensional projections of the data. Fit projections in multiple directions using the univariate techniques described above, and see if they are well described by a univariate stable fit. If so, and if the α's are the same for every direction (and if $\alpha < 1$, the location parameters satisfy (3)), then a multivariate stable model is appropriate. We will illustrate this in examples below.

For the purposes of comparing two multivariate stable distributions, the parameters $(\alpha, \beta(\mathbf{u}), \gamma(\mathbf{u}), \delta(\mathbf{u}))$ are more useful than Λ itself. This is because the distribution of $\mathbf{X}$ depends more on how Λ distributes mass around the sphere than exactly on the measure. Two spectral measures can be far away in the traditional total variation norm (e.g. one can be discrete and the other continuous), but their corresponding directional scale functions and densities can be very close. Indeed, Theorem 2 shows that the only way Λ enters into the joint distribution is through the parameter functions.

The diagnostics suggested are:

- Project the data in a variety of directions $\mathbf{t}$ and use the univariate diagnostics described in Section 2 on each of those distributions. Bad fits in any direction indicate that the data is not stable.

- For each direction $\mathbf{t}$, estimate the parameter functions $\alpha(\mathbf{t})$, $\beta(\mathbf{t})$, $\gamma(\mathbf{t})$, $\delta(\mathbf{t})$ by ML estimation. The plot of $\alpha(\mathbf{t})$ should be a constant, significant departures from this indicate that the data has different decay rates in different directions. (Note that $\gamma(\mathbf{t})$ will be a constant iff the distribution is isotropic.)

- Assess the goodness-of-fit by computing a discrete $\hat{\Lambda}$ by one of the methods above. Substitute the discrete $\hat{\Lambda}$ in Theorem 2 to compute parameter functions. If it differs from the one obtained above by projection, then either the data is not jointly stable, or not enough points were chosen in the discrete spectral measure approximation.

These techniques are illustrated next. Each bivariate data set will have two pages of graphs. The first is series of smoothed density, q-q plot and variance stabilized p-p plot for projections in 8 different directions: $\pi/2$, $\pi/3$, $\pi/4$, $\pi/6$,0,$-\pi/6$, $-\pi/4$, $-\pi/3$. (Because $(-\mathbf{u})\cdot\mathbf{x}) = -\mathbf{u}\cdot\mathbf{X}$, projections in the left half plane are reflections of those in the right half plane). The second page will show the discrete estimate of the spectral measure (with $m = 100$ evenly spaced point masses) in polar form, a cumulative plot of the spectral measure in rectangular form, and then four plots for the parameter estimates $(\alpha(\mathbf{t}), \beta(\mathbf{t}), \gamma(\mathbf{t}), \delta(\mathbf{t}))$. Also on the $\alpha(\mathbf{t})$ plot is a horizontal line showing the average value of all the estimated indices which is taken as the estimate of the common α that should come from a jointly stable distribution. The plots of $\beta(\mathbf{t})$ and $\gamma(\mathbf{t})$ also show the skewness and scale functions computed from the estimated spectral measure and Theorem 2. All three of these curves should be close to the separately estimated directional parameters.

Some details on these plots. The polar plots of the spectral measure show a unit circle and lines connecting the points (θ_j, r_j), where $\theta_j = 2\pi(j-1)/m$ and $r_j = 1 + (\lambda_j/\lambda_{max})$, where $\lambda_{max} = \max_j \lambda_j$. The polar plots are spiky, because we are estimating a discrete object. What should be looked at is the overall spread of mass, not specific spikes in the plot. In cases where the spectral measure is really smooth, it may be appropriate to smooth these plots out to better show it's true nature. In cases where the measure is discrete, i.e. the independent case, then one wants to emphasize the spikes. So there is no satisfactory general solution and we just plot the raw data.

Finally, most graphing programs will set vertical scale so that the data fills the graph. This emphasizes minor fluctuations in the data that are not of practical significance. In the graphs below, the vertical scales for the parameter functions $\alpha(\mathbf{t})$, $\beta(\mathbf{t})$, $\gamma(\mathbf{t})$ are respectively [0,2], [-1,1], and $[0, 1.2 \times \max \gamma(\mathbf{t})]$. These bounds show how the functions vary over their possible

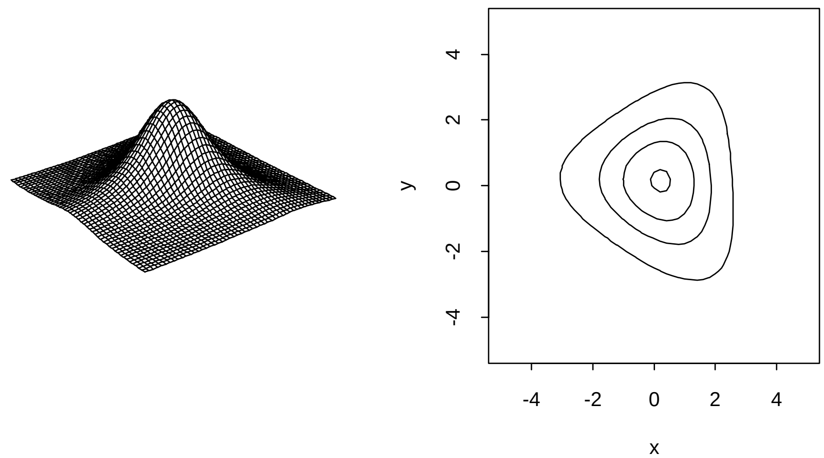

Figure 13: Density surface and level curves for "triangle" example.

range. For $\delta(\mathbf{t})$, we used the bounds $[-1.2\times\max|\delta(\mathbf{t})|, 1.2\times\max|\delta(\mathbf{t})|]$, which visually exagerates the changes in $\delta(\mathbf{t})$. A scale that depends on $\max\gamma(\mathbf{t})$ may be more appropriate.

5.1 Simulated "triangle" stable data

This data set is simulated from a known spectral measure using the method of Modarres and Nolan (1994). A value of $\alpha = 1.2$ and a discrete spectral measure having three unit point masses, distributed at angles $\pi/3$, π and $-\pi/3$. A plot of the density surface and level curves are given in Figure 13. The triangular spread of the spectral measure shows up in the triangular shape of the level curves.

The simulated data set has $n = 5,000$ points. Figure 14 shows EDA plots for 8 projections. The rightmost column of plots has text which shows the angle of projection (upper left corner) and the estimated parameter values for those projections (lower right corner).

Figure 15 shows the estimated spectral measure and parameter functions. The upper left polar plot shows the estimated spectral measure is mostly concentrated at the correct three points, cumulative spectral measure shows the same info in a different format. The plot of $\hat{\alpha}(\mathbf{t})$ shows an essentially flat curve near the correct $\alpha = 1.2$. The plot of $\hat{\beta}(\mathbf{t})$ and $\hat{\gamma}(\mathbf{t})$ show how skewness and scale vary as projection angle varies.

5.2 Simulated sub-Gaussian stable data

Here $n = 5000$ data points were simulated from a sub-Gaussian distribution with $\alpha = 1.5$ and

$$R = \begin{pmatrix} 1.0 & 0.7 \\ 0.7 & 1.0 \end{pmatrix} \tag{4}$$

Figure 16 shows projection diagnostics, Figure 17 shows the estimation results. See Section 6 for more info on detecting the sub-Gaussian behavior.

5.3 Simulated non-stable data

A data set was generated with independent stable components of different α's, so the data set is not jointly stable. The first coordinate is Gaussian ($\alpha_1 = 2$) and the second is Cauchy ($\alpha_2 = 1$), there were $n = 1000$ data points.

While the projection diagnostics in Figure 18 are individually plausible, they do not fit together in a reasonable way for a jointly stable distribution. (The third row of this plot is similar to Figure 12.) The non joint stability is best detected from the plot of $\hat{\alpha}(\mathbf{t})$ in Figure 19, which varies from 2 to 1, far from a flat curve that occurs in the jointly stable case.

5.4 Bivariate radar data

The radar sea clutter data was analyzed as a bivariate data set with $n = 320,000$ pairs of (in-phase,quadrature) values. The one dimensional projections in different directions are indistinguishable from Figure 8, so they are not shown. They give strong support to an underlying stable distribution. In Figure 20, the spectral measure and the parameter functions are shown. The index function is essentially constant ($1.698 \leq \alpha(\mathbf{t}) \leq 1.709$), showing that the tails die off at the same rate in all directions. Also, the scale function is essentially constant ($.3975 \leq \gamma(\mathbf{t}) \leq .3993$) and skewness function is essentially 0, strongly supporting an isotropic/radially symmetric fit. The cumulative spectral measure is close to a uniform measure.

5.5 Foreign exchange rates

In Section 3, foreign exchange rates were analyzed individually. Here we will examine the joint distribution of the German mark and the Japanese yen.

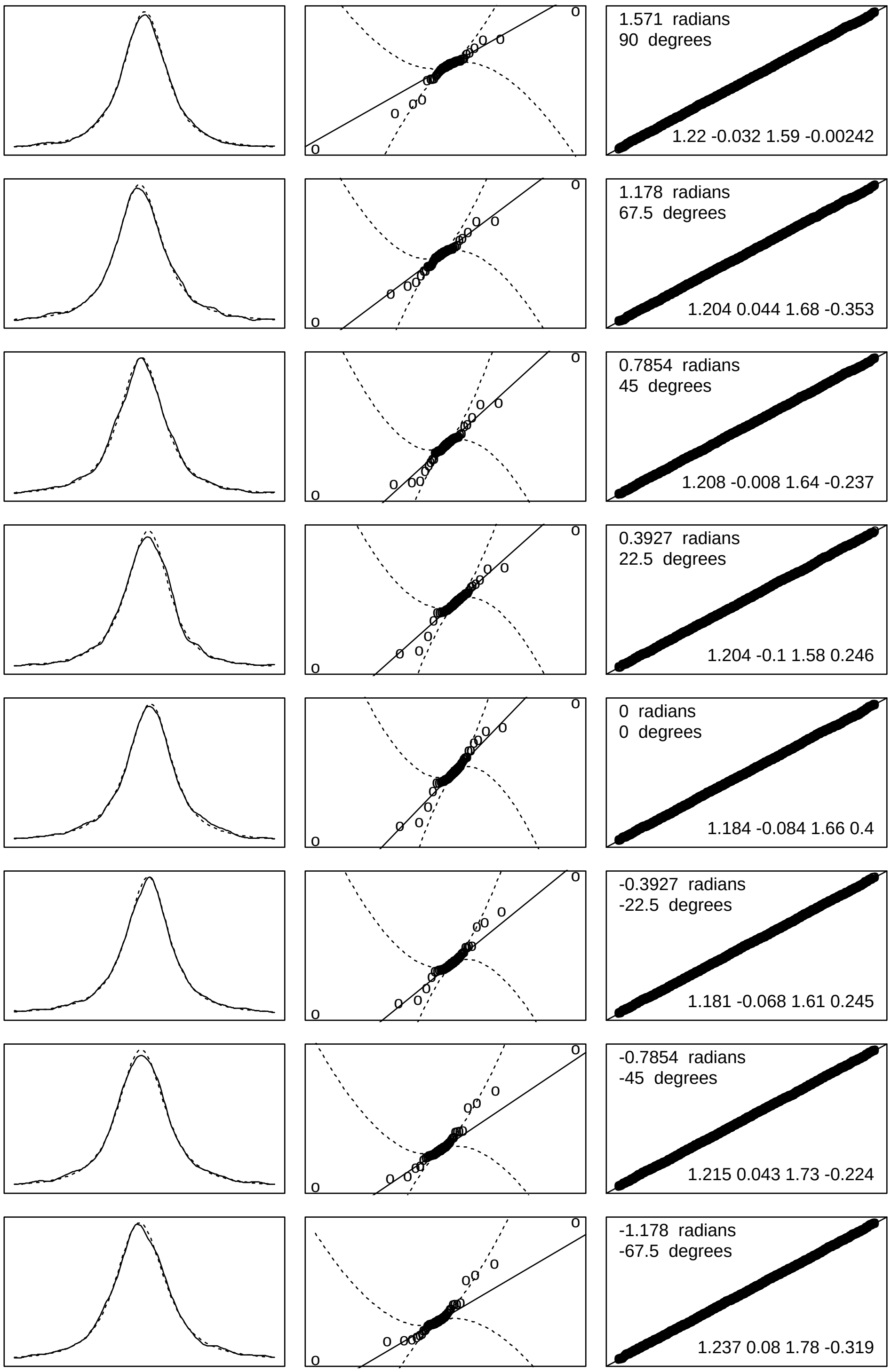

Figure 14: Projections diagnostics for the simulated “triangle” data set.

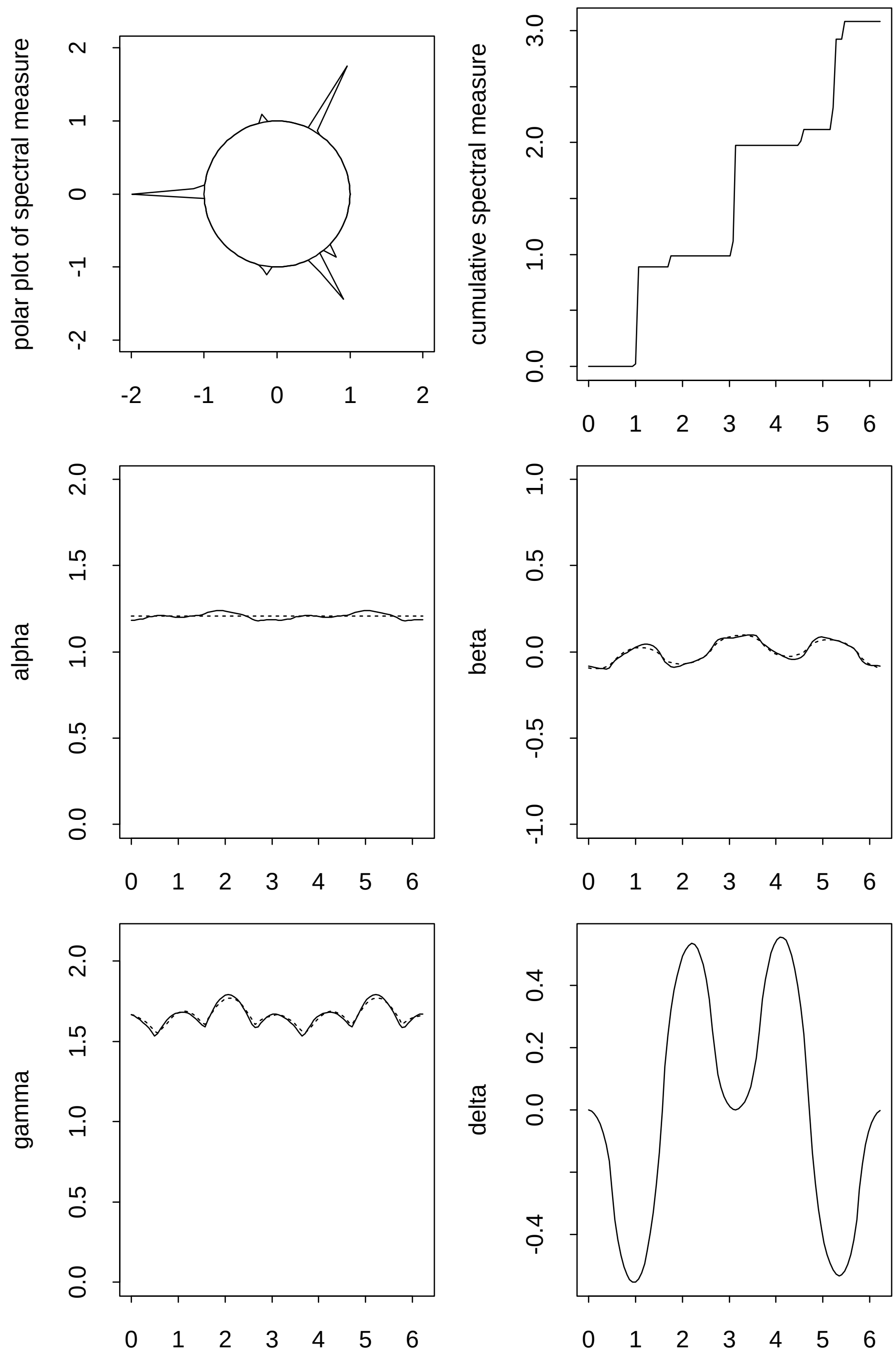

Figure 15: Estimation results for simulated “triangle” data set.

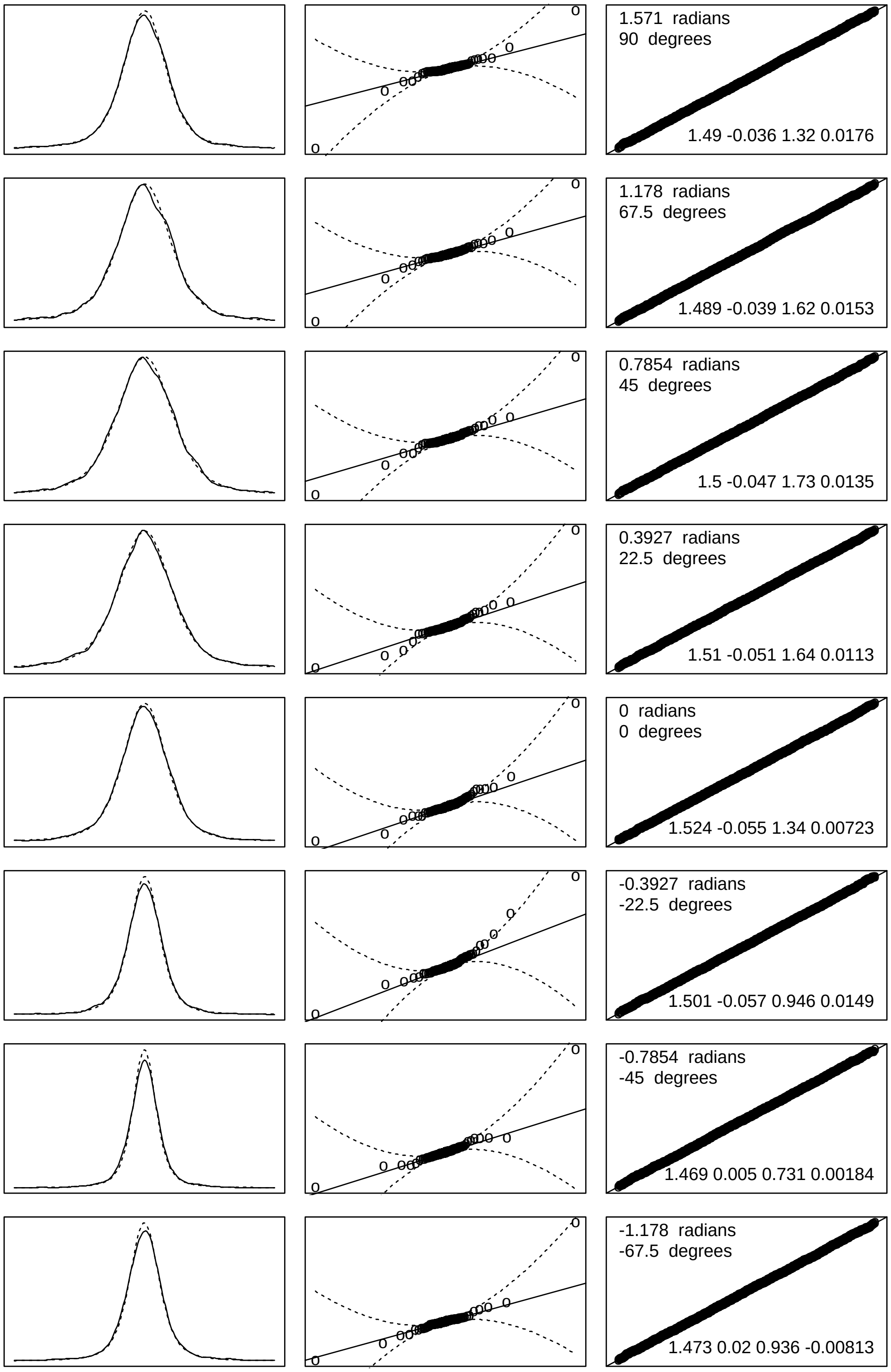

Figure 16: Projections diagnostics for the simulated sub-Gaussian data set.

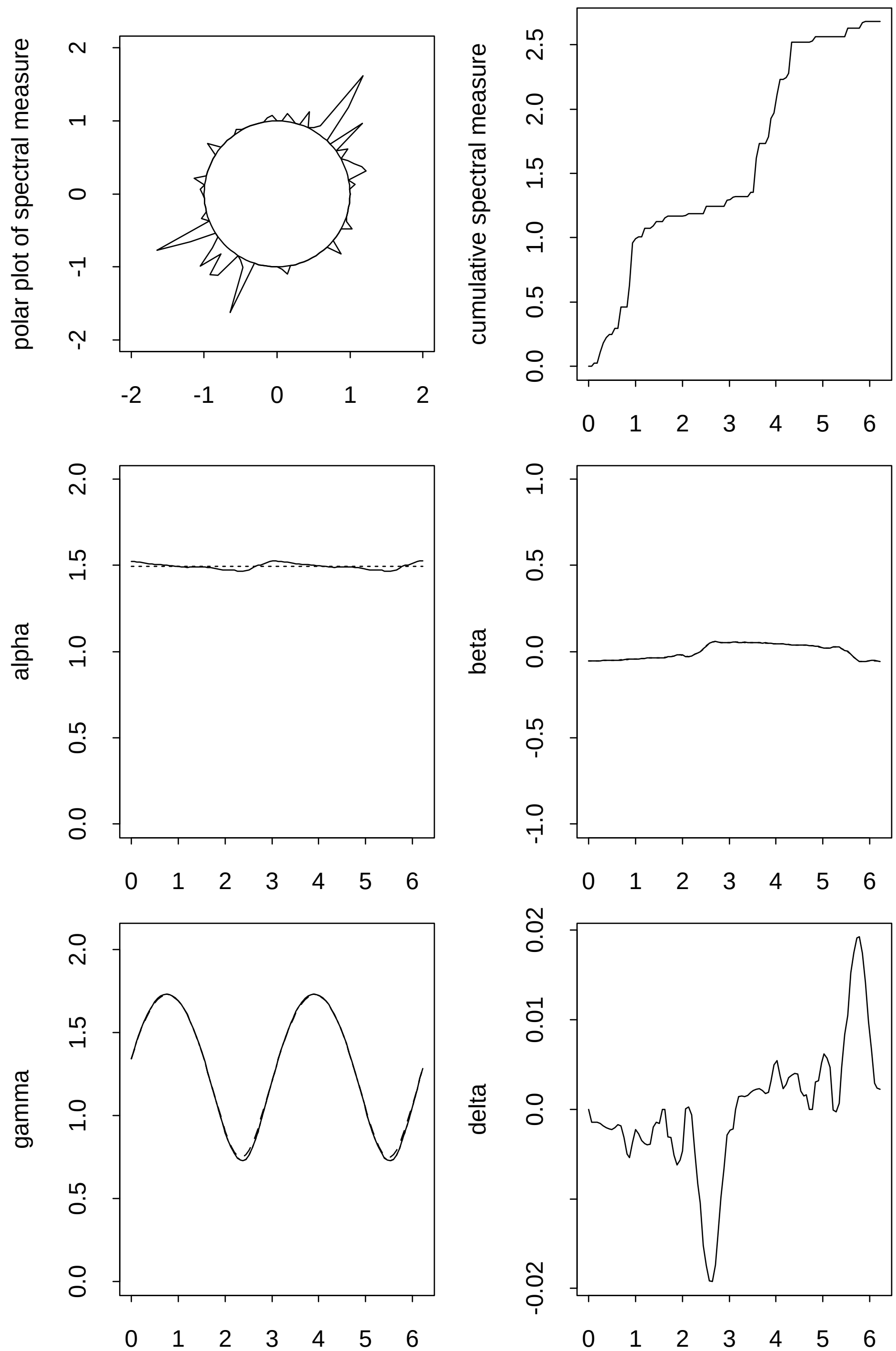

Figure 17: Estimation results for simulated sub-Gaussian data set.

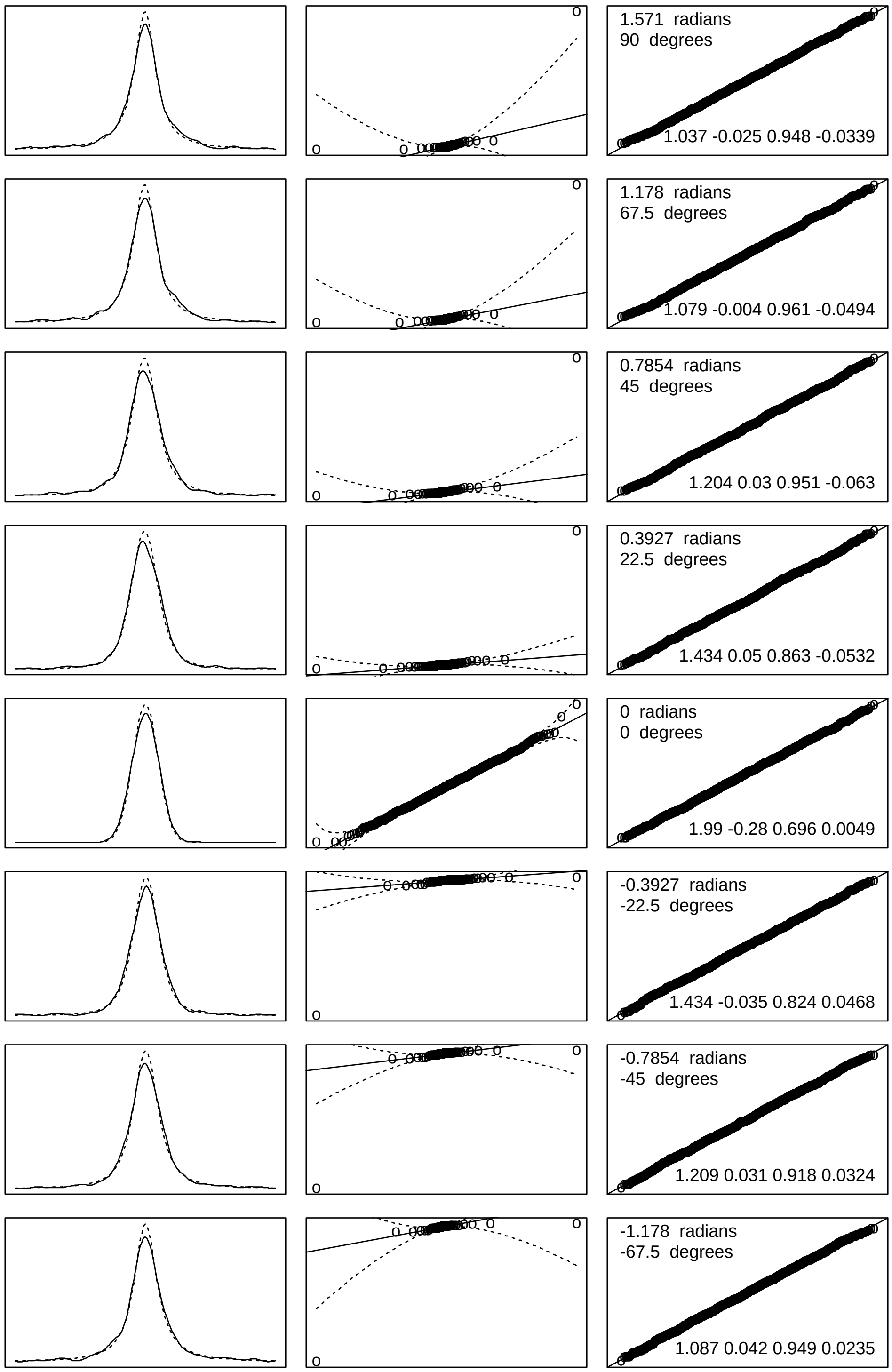

Figure 18: Projections diagnostics for the simulated Gaussian and Cauchy data set.

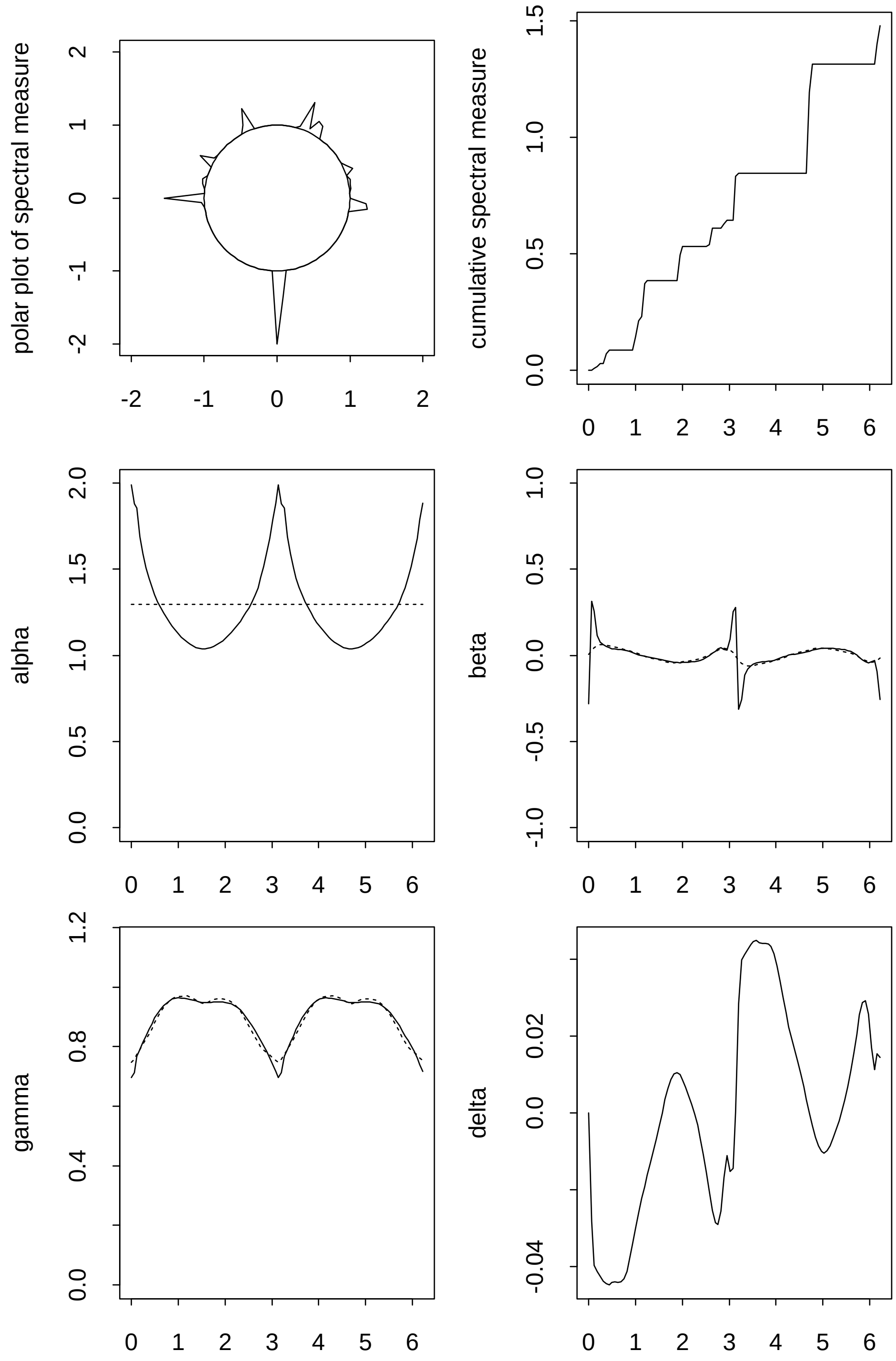

Figure 19: Estimation results for simulated Gaussian and Cauchy data set.

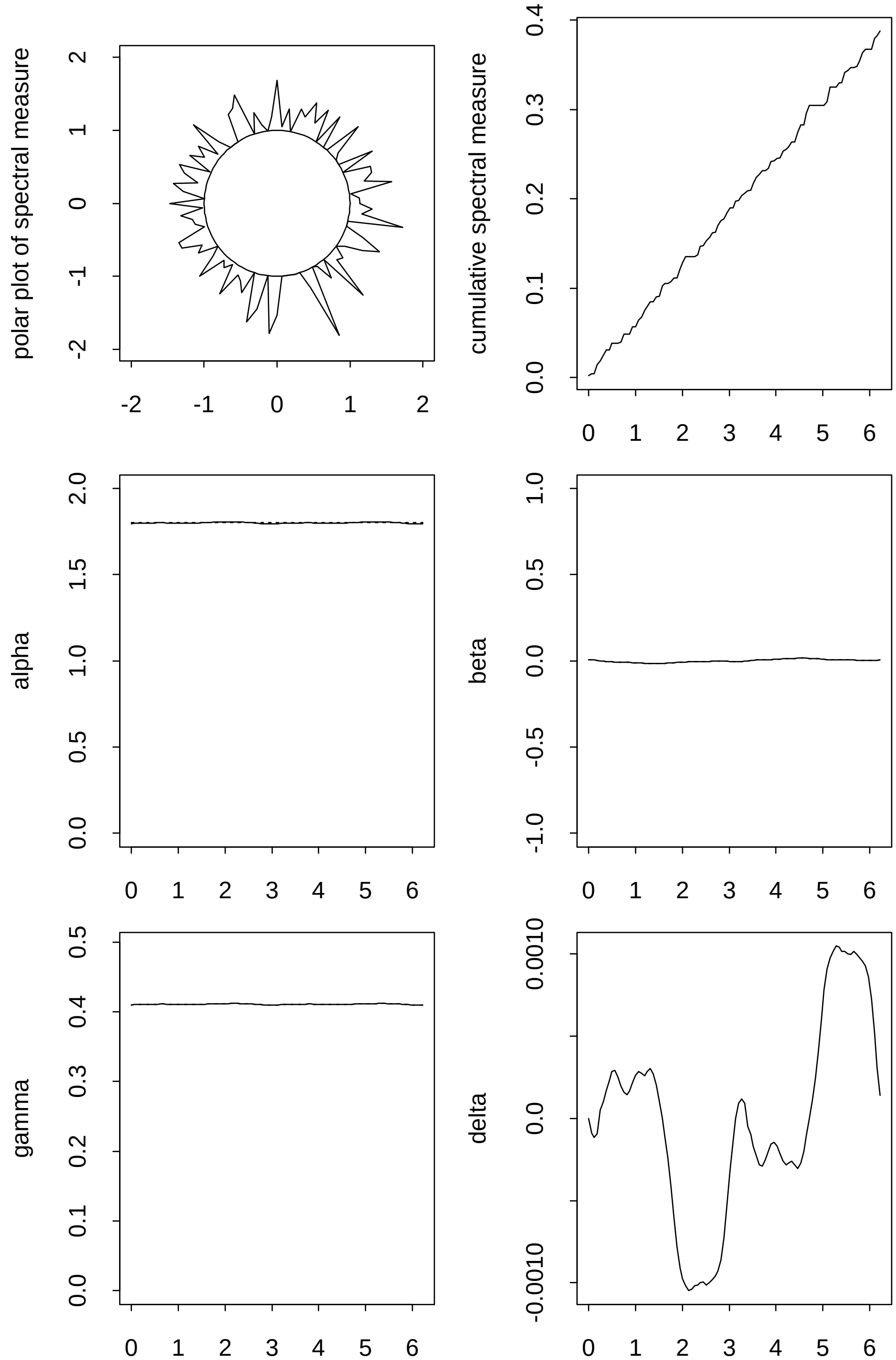

Figure 20: Bivariate stable fit to sea clutter data.

The diagnostics in Figure 21 shows that a sequence of projections are similar to Figure 5, in fact the fifth row of Figure 21 is exactly the same as Figure 5. Except on the extreme tails, the stable fit does a good job of describing the data.

The projection functions $\alpha(\mathbf{t})$, $\beta(\mathbf{t})$, $\gamma(\mathbf{t})$, and $\delta(\mathbf{t})$ were estimated and used to compute an estimate of the spectral measure. The results are shown in Figure 22. The fitted spectral measure was used to plot an estimate of the fitted bivariate density using the program described in Nolan and Rajput (1995), shown in Figure 23. The spread of the spectral measure is spiky, and masks a pattern that is more obvious in the density surface: the approximate elliptical contours of the fitted density. This suggests modeling the data by a sub-Gaussian stable distribution, a topic discussed in the next section.

6 Sub-Gaussian distributions

Since the radar clutter data is well described by a radially symmetric stable distribution and the foreign exchange data seems to be approximately elliptically contoured, there may be interest in categorizing such stable distributions. The main practical advantage to this is that all d-dimensional elliptically contoured stable distributions are parameterized by α and a symmetric, positive definite $d \times d$ matrix. Since the matrix is symmetric, there are a total of $1 + d(d+1)/2$ parameters. This is quite different from the general stable case, which involves an infinite dimensional spectral measure. Even a discrete approximating measure involves a much larger number of terms: if a "polar grid" is used with each of the angles divided up evenly with k subintervals, then there are k^{d-1} point masses to be estimated.

Let $\mathbf{X}$ be an non-singular symmetric α-stable random vector. The following are equivalent:

- $\mathbf{X}$ is elliptically contoured around the origin.
- $\mathbf{X}$ is sub-Gaussian, i.e. $\mathbf{X} \overset{d}{=} A^{1/2}\mathbf{G}$, where $A \sim \mathbf{S}(\alpha, 1, \gamma, 0; 1)$ and $\mathbf{G} \sim N(0, R_{\mathbf{G}})$.
- The characteristic function is $E \exp(i\mathbf{u} \cdot \mathbf{X}) = \exp(-(\mathbf{u}R\mathbf{u}^T)^{\alpha/2})$, for some symmetric, positive definite matrix R.

Computing densities for the elliptically contoured distributions is a one dimensional problem: given the R matrix,

$$f(\mathbf{x}) = (1/\det(B))g_\alpha(|B\mathbf{x}|^{1/2}),$$

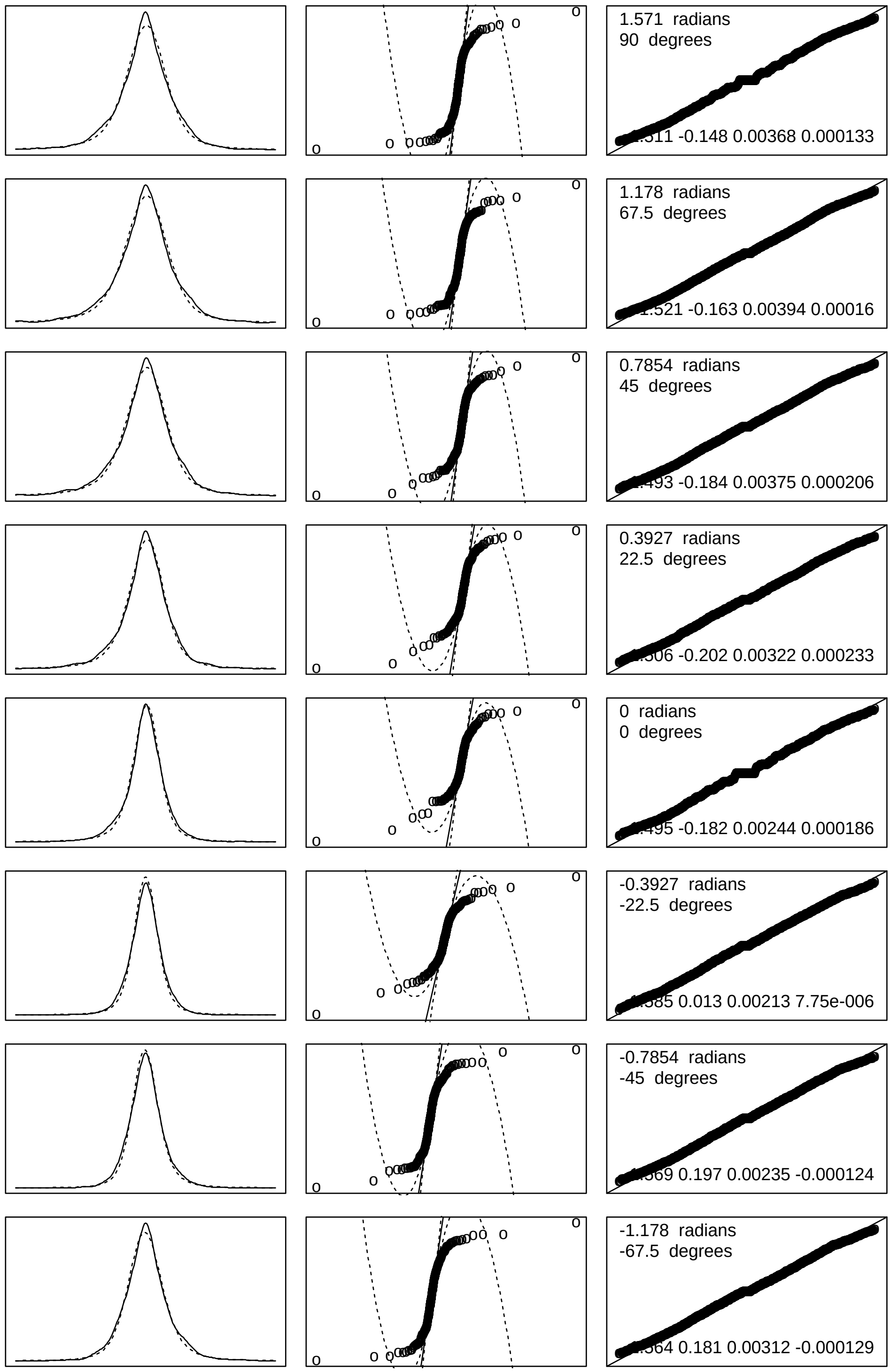

Figure 21: Projection diagnostics for the German mark and Japanese yen exchange rates.

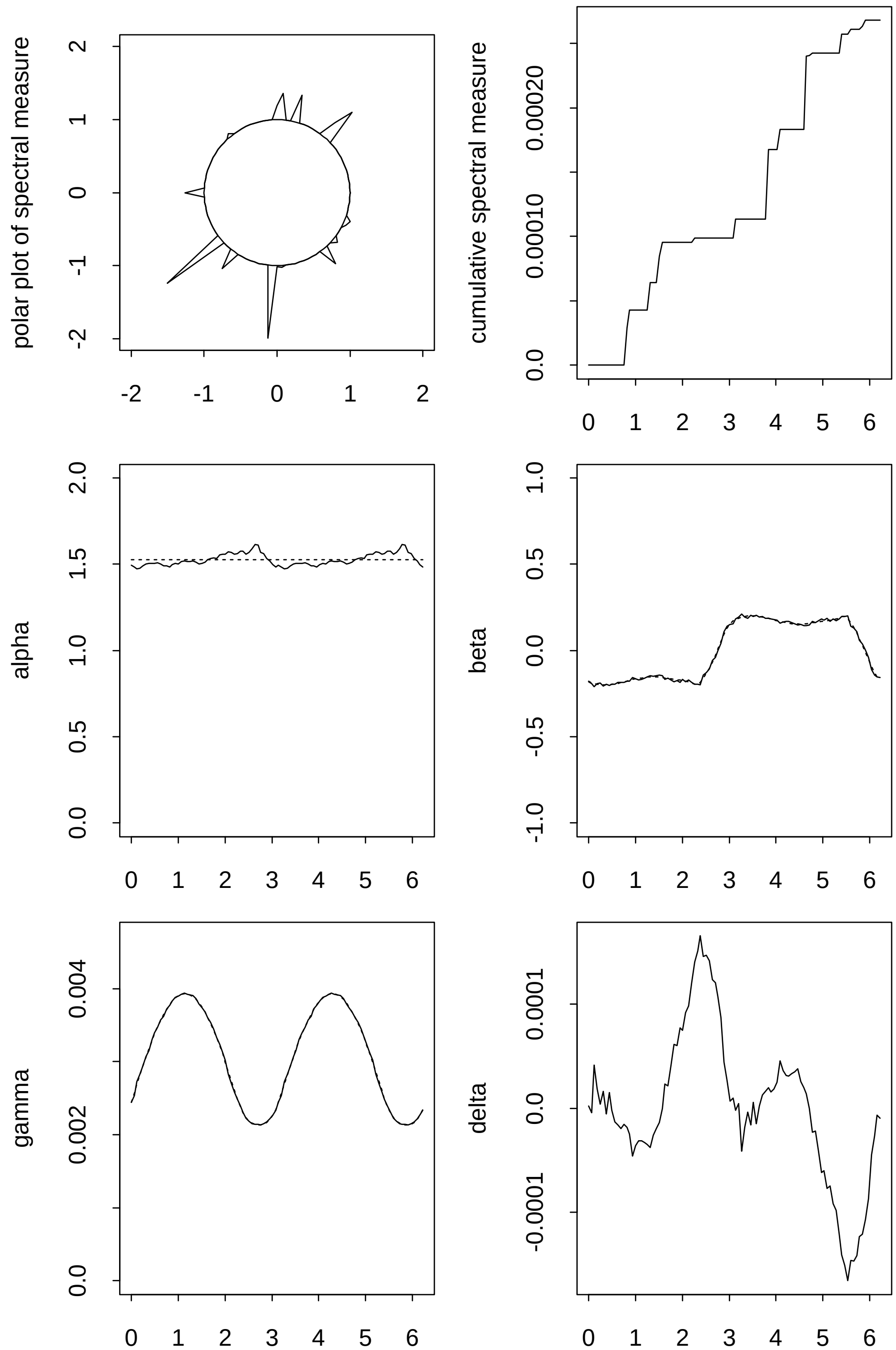

Figure 22: Estimation results for the German mark and Japanese yen exchange rates.

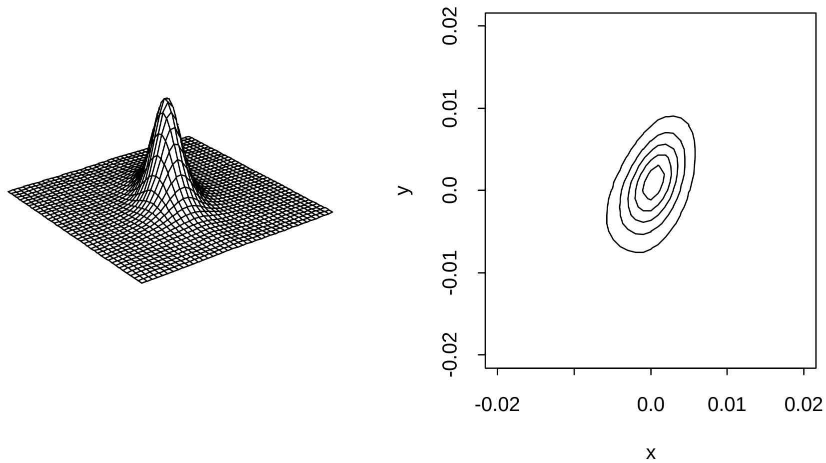

Figure 23: Estimated density surface and level curves for a bivariate stable fit to the German mark and Japanese yen exchange rates.

where

$$g_\alpha(v) = c \int_0^\infty J_0(|\mathbf{x}|u) u^{d-1} e^{-u^\alpha} du,$$

$R = BB^T$, and $J_0(\cdot)$ is the 0^{th} order Bessel function of the first kind. Likewise, computing probabilities of the natural ellipses is a one dimensional problem once $g_\alpha(\cdot)$ is known.

We next describe ways of assessing a d-dimensional data set to see if it is approximately sub-Gaussian and then estimating the parameters of a sub-Gaussian vector.

First perform a one dimensional stable fit to each coordinate of the data using one of the methods described above, to get estimates $\hat{\boldsymbol{\theta}}_i = (\hat{\alpha}_i, \hat{\beta}_i, \hat{\gamma}_i, \hat{\delta}_i)$. If the α_i's are significantly different, then the data is not jointly α-stable, so it cannot be sub-Gaussian. Likewise, if the β_i's are not all close to 0, then the distribution is not symmetric and it cannot be sub-Gaussian.

If the α_i's are all close, form a pooled estimate of $\alpha = (\sum_{i=1}^d \alpha_i)/d =$ average of the indices of each component. Then shift the data by $\hat{\boldsymbol{\delta}} = (\hat{\delta}_1, \hat{\delta}_2, \ldots, \hat{\delta}_d)$ so the distribution is centered at the origin.

Next, test for sub-Gaussian behavior. This can be accomplished by examining two dimensional projections because of the following result.

Theorem 3 *Let* $\mathbf{X}$ *be a d-dimensional sub-Gaussian α-stable random vector.*

Then every two dimensional projection ($\mathbf{a}_1, \mathbf{a}_2 \in \mathbf{R}^d$)

$$\mathbf{Y} = (Y_1, Y_2) = (\mathbf{a}_1 \cdot \mathbf{X}, \mathbf{a}_2 \cdot \mathbf{X}) \tag{5}$$

is a 2-dimensional sub-Gaussian α-stable random vector.

Conversely, suppose $\mathbf{X}$ is a d-dimensional α-stable random vector with the property that every two dimensional projection of form (5) is non-singular sub-Gaussian. Then $\mathbf{X}$ is non-singular sub-Gaussian α-stable.

Proof If $\mathbf{X}$ is sub-Gaussian α-stable, then there is a symmetric positive definite matrix R such that $\phi(\mathbf{u}) = E\exp(i\mathbf{u}\cdot\mathbf{X}) = \exp(-(\mathbf{u}R\mathbf{u}^T)^{\alpha/2})$. So

$$\begin{aligned} E\exp(i(v_1, v_2)\cdot(Y_1, Y_2)) &= E\exp(i(v_1\mathbf{a}_1 + v_2\mathbf{a}_2)\cdot\mathbf{X}) \\ &= \exp(-((v_1\mathbf{a}_1 + v_2\mathbf{a}_2)R(v_1\mathbf{a}_1 + v_2\mathbf{a}_2)^T)^{\alpha/2}) \\ &= \exp(-[(v_1\mathbf{a}_1)R(v_1\mathbf{a}_1)^T + (v_1\mathbf{a}_1)R(v_2\mathbf{a}_2)^T \\ &\qquad +(v_2\mathbf{a}_2)R(v_1\mathbf{a}_1)^T + (v_2\mathbf{a}_2)R(v_2\mathbf{a}_2)^T]^{\alpha/2}) \\ &= \exp(-[(v_1, v_2)S(v_1, v_2)^T]^{\alpha/2}), \end{aligned}$$

for $s_{11} = \mathbf{a}_1 R\mathbf{a}_1^T$, $s_{22} = \mathbf{a}_2 R\mathbf{a}_2^T$, $s_{12} = s_{21} = \mathbf{a}_1 R\mathbf{a}_2^T$. Hence $\mathbf{Y}$ is sub-Gaussian.

Conversely, suppose every two dimensional projection is non-singular sub-Gaussian α-stable. Then $\phi(\mathbf{u}) > 0$, so $f(\mathbf{u}) := (-\ln\phi(\mathbf{u}))^{2/\alpha}$ is well defined. The assumption says every two dimensional projection of f is a non-singular, positive definite, symmetric quadratic form: $f(v_1\mathbf{u}_1 + v_2\mathbf{u}_2) = c_1v_1^2 + c_2v_2^2 + c_3v_1v_2$. The next lemma shows that $f(\mathbf{u}) = \mathbf{u}R\mathbf{u}^T$ for some symmetric, positive definite $d \times d$ matrix R. □

The following lemma was proved with the help of J. Hakim and D. Kalman.

Lemma 1 *Let $f(\mathbf{u})$ be a real valued function on $\mathbf{R}^d$ such that every two dimensional projection is a symmetric positive definite quadratic form on $\mathbf{R}^2$. Then f is a symmetric positive definite quadratic form on $\mathbf{R}^d$.*

Proof The hypothesis says $f(v_1\mathbf{u}_1 + v_2\mathbf{u}_2) = c_1v_1^2 + c_2v_2^2 + c_3v_1v_2$, for some constants $c_i = c_i(\mathbf{u}_1, \mathbf{u}_2)$. Substituting $(v_1, v_2) = (1, 0)$ shows $c_1 = f(\mathbf{u}_1)$, substituting $(v_1, v_2) = (0, 1)$ shows $c_2 = f(\mathbf{u}_2)$, and substituting $(v_1, v_2) = (1, 1)$ shows $c_3 = f(\mathbf{u}_1 + \mathbf{u}_2) - f(\mathbf{u}_1) - f(\mathbf{u}_2)$. Hence

$$f(v_1\mathbf{u}_1 + v_2\mathbf{u}_2) = f(\mathbf{u}_1)v_1^2 + f(\mathbf{u}_2)v_2^2 + [f(\mathbf{u}_1 + \mathbf{u}_2) - f(\mathbf{u}_1) - f(\mathbf{u}_2)]v_1v_2 \tag{6}$$

Let $\mathbf{e}_1 = (1,0,0,\ldots,0)$, $\mathbf{e}_2 = (0,1,0,\ldots,0)$, $\ldots$, $\mathbf{e}_1 = (0,0,0,\ldots,1)$ be the standard unit basis vectors. Define the $d \times d$ symmetric matrix $R = [r_{ij}]$ with entries $r_{ii} = f(\mathbf{e}_i)$ and $r_{ij} = r_{ji} = (1/2)[f(\mathbf{e}_i + \mathbf{e}_j) - f(\mathbf{e}_i) - f(\mathbf{e}_j)]$. Then (6) says

$$f(v_1\mathbf{e}_i + v_2\mathbf{e}_j) = r_{ii}v_1^2 + r_{jj}v_2^2 + 2r_{ij}v_1v_2 \tag{7}$$

We first show the case $d = 3$: suppose $\mathbf{u} = (u_1, u_2, u_3)$. If $u_2 \neq 0$, then write $\mathbf{u} = u_2(u_1/u_2, 1/2, 0) - u_2(0, -1/2, -u_3/u_2)$ and apply (6) to show

$$\begin{aligned} f(\mathbf{u}) &= f(u_2(u_1/u_2, 1/2, 0) - u_2(0, -1/2, -u_3/u_2)) \\ &= u_2^2 f(u_1/u_2, 1/2, 0) + (-u_2)^2 f(0, -1/2, -u_3/u_2) + u_2(-u_2) \\ &\quad \times [f(u_1/u_2, 0, -u_3/u_2) - f(u_1/u_2, 1/2, 0) - f(0, -1/2, -u_3/u_2)] \\ &= u_2^2\,[2f(u_1/u_2, 1/2, 0) + 2f(0, -1/2, -u_3/u_2) - f(u_1/u_2, 0, -u_3/u_2)] \end{aligned}$$

Next apply (7) to each of the terms above to get

$$\begin{aligned} f(\mathbf{u}) &= u_2^2 \Big[2\left((u_1/u_2)^2 r_{11} + (1/2)^2 r_{22} + 2(u_1/u_2)(1/2) r_{12}\right) \\ &\quad +2\left((-1/2)^2 r_{22} + (-u_3/u_2)^2 r_{33} + 2(-1/2)(-u_3/u_2) r_{23}\right) \\ &\quad - \left((u_1/u_2)^2 r_{11} + (-u_3/u_2)^2 r_{33} + 2(u_1/u_2)(-u_3/u_2) r_{13}\right)\Big] \\ &= r_{11}u_1^2 + r_{22}u_2^2 + r_{33}u_3^2 + 2u_1u_2r_{12} + 2u_1u_3r_{13} + 2u_2u_3r_{23} = \mathbf{u}R\mathbf{u}^T. \end{aligned}$$

If $u_2 = 0$, then apply (7) to $\mathbf{u} = u_1(1,0,0) + u_3(0,0,1)$ to show that the form above is valid in that case.

For $d > 3$, use induction with the above argument applied to $\mathbf{u} = (u_1, u_2, \ldots, u_d) = u_2(u_1/u_2, 1/2, u_3/u_2, \ldots, u_{d-1}/u_2, 0) - u_2(0, -1/2, 0, 0, \ldots, 0, -u_d/u_2)$. □

Estimating the $d(d+1)/2$ parameters (upper triangular part of) R can be done in at least two ways. For the first method, set $r_{ii} = \gamma_i^2$, i.e. the square of the scale parameter of the i-th coordinate. Then estimate r_{ij} by analyzing the pair (X_i, X_j) and take $r_{ij} = (\gamma^2(1,1) - r_{ii} - r_{jj})/2$, where $\gamma(1,1)$ is the scale parameter of $(1,1) \cdot (X_i, X_j) = X_i + X_j$. This involves estimating $d + d(d-1)/2 = d(d+1)/2$ one dimensional scale parameters.

For the second method, note that if $\mathbf{X}$ is α-stable sub-Gaussian, then $E\exp(i\mathbf{u}\cdot\mathbf{X}) = \exp(-(\mathbf{u}R\mathbf{u}^T)^{\alpha/2})$, so

$$[-\ln E\exp(i\mathbf{u}\cdot\mathbf{X})]^{2/\alpha} = \mathbf{u}R\mathbf{u}^T = \sum_i u_i^2 r_{ii} + 2\sum_{i<j} u_iu_jr_{ij}.$$

This is a linear function of the r_{ij}'s which can be estimated by regression. This method may be more accurate because it uses multiple directions, whereas the first method uses only three directions: (1,0), (0,1) and (1,1). If a two dimensional fit has already been done, then one has already estimated $\gamma(\mathbf{u})$ on a circular grid. Note that $\mathbf{u}R\mathbf{u}^T = \gamma^2(\mathbf{u})$ is the square of the scale parameter in the direction $\mathbf{u}$. Sample estimates of $\gamma^2(\mathbf{u})$ on a grid of $\mathbf{u}$ points can be used on the left hand side above.

In both methods, checks should be made to test that the resulting matrix R is positive definite.

The first method was used to estimate the matrix R for three of the data sets considered above. For the simulated sub-Gaussian data set, R was given by a multiple of (4), and the estimated matrix was

$$\hat{R} = 1.776 \begin{pmatrix} 1.016 & 0.686 \\ 0.686 & 0.985 \end{pmatrix}.$$

The plot of $\gamma(\mathbf{t})$ shown in the lower left corner of Figure 22 also shows $\sqrt{\mathbf{t}\hat{R}\mathbf{t}^T}$ as a dashed line. It is virtually indistinguishable from the curve of $\gamma(\mathbf{t})$, supporting the idea that a sub-Gaussian stable fit does a good job of fitting the data.

For the radar data set

$$\hat{R} = \begin{pmatrix} 0.16829 & 0.00041 \\ 0.00041 & 0.16897 \end{pmatrix}.$$

For the German mark/Japanese yen exchange rate data,

$$\hat{R} = 10^{-6} \begin{pmatrix} 5.9552 & 4.0783 \\ 4.0783 & 13.9861 \end{pmatrix}.$$

7 Discussion

We have shown that estimation of general stable parameters is now feasible. The diagnostics show that several large data sets with heavy tails are well described by stable distributions. We also showed that stable models are not a panacea - not all heavy tailed data sets can be well described by stable distributions.

In practice, the decision to use a stable model should be based on the purpose of the model. In cases where a large data set shows close agreement with a stable fit, confident statements can be made about the population. In other cases, one should clearly not use a stable model. In intermediate cases, one could tentatively use a stable model as a descriptive method of summarizing the general shape of the distribution, but not try to make statements about tail probabilities. In such problems, it may actually be better to use the quantile parameter estimates rather than ML estimates, because the former tries to match the shape of the empirical distribution and ignores the top and bottom 5% of the data.

We have not considered parameters that vary with time, mixture models, etc. While we do not do so here, it is straightforward to use an information criteria like AIC to compare a stable model to mixture models or GARCH models for a data set. It seems likely that certain problems, e.g. the radar sea clutter problem, have physical explanations that make a stationary model plausible. Other problems, particularly economic time series, may very well have time varying parameters that reflect changes in the underlying conditions for that series. We cannot resolve this issue here. Our main purpose is to make stable models a practical tool that can be used and evaluated by the scientific community.

In multivariate problems where the dimension is large, it will be very difficult to model with a stable distribution unless there is some special structure. If some components are independent, then they should be separated out and analyzed alone. If the dependent components are sub-Gaussian, then Section 6 discusses how to jointly analyze them. In general stable case, one may try to group the components into smaller dependent groups, estimate within groups, and then try to characterize dependence between groups. We are not aware of work on this topic.

The program STABLE for univariate data is available on the Web at `http://www.cas.american.edu/~jpnolan`. MVSTABLE, a similar program for multivariate stable is under development.

Acknowledgments

The data sets analyzed above were graciously provided by C. Klüppelberg (exchange rate data), P. Tsakalides (radar data) and J. H. McCulloch (CRSP stock data). R. Jernigan provided discussion and references on EDA techniques.

References

[1] Abdul-Hamid, H. and Nolan, J. P. (1998) Multivariate stable densities as functions of one dimensional projections. *J. Multivar. Analysis*, **67**, 80-89.

[2] Adler, R., Feldman, R. and Taqqu, M. (eds.) (1998), *A Practical Guide to Heavy Tailed Data*, Birkhäuser, Boston, MA.

[3] Akgiray, V. and Lamoureux, C. G. (1989), Estimation of stable law parameters: a comparative study, *J. Bus. Econ. Stat* **7**:85-93.

[4] Beirlant, J., Vynckier, P. and Teugels, J. L. (1996), Tail index estimation, Pareto quantile plots and regression diagnostics, *JASA* **91**:1659-1667.

[5] Brant, R. (1984), Approximate likelihood and probability calculations based on transforms, *Ann. Stat* **12**:989-1005.

[6] Brorsen, B. W. and Yang, S. R. (1990), Maximum likelihood estimates of symmetric stable distribution parameters, *Commun. Stat. Simula.* **19**:1459-1464.

[7] Brown, B. M. and Hettmansperger, T. P. (1996), Normal scores, normal plots and tests for normality, *JASA*, **91**:1668-1675.

[8] Buckle, D. J. (1995), Bayesian inference for stable distributions, *JASA* **90**:605-613.

[9] Byczkowski, T., Nolan, J. P. and Rajput, B. (1993). Approximation of multidimensional stable densities. *J. Multvar. Anal.* **46**, 13-31.

[10] Chambers, J. M., Mallows, C. and Stuck, B. W. (1976) A method for simulating stable random variables. *JASA* **71**, 340-344.

[11] Cheng, B.N., and Rachev, S.T. (1995), Multivariate stable securities in financial markets, *Mathematical Finance* (to appear).

[12] Christoph, G. and Wolf, W. (1993), *Convergence Theorems with a Stable Limit Law*, Akademie Verlag, Berlin.

[13] DuMouchel, W. H. (1971), *Stable Distributions in Statistical Inference*, Ph.D. dissertation, Department of Statistics, Yale University.

[14] DuMouchel, W. H. (1973a), On the asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution, *Ann. Stat.* **1**:948-957.

[15] DuMouchel, W. H. (1973b), Stable distributions in statistical inference. 1: Symmetric stable distributions compared to other symmetric long-tailed distributions, *JASA* **68**:469-477.

[16] DuMouchel, W. H. (1975), Stable distributions in statistical inference. 2: Information from stably distributed samples, *JASA* **70**:386- 393.

[17] DuMouchel, W. H. (1983), Estimating the stable index α in order to measure tail thickness: A critique, *Ann. Stat.* **11**:1019-1031.

[18] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997), *Modelling Extremal Events with Emphasis on Insurance and Finance*, Springer Verlag, Berlin.

[19] Fama, E. and Roll, R. (1971). Parameter estimates for symmetric stable distributions. *J. Amer. Stat. Assoc.* **66**:331-338.

[20] Feldheim, E. (1937) Etude de la stabilité des lois de probabilité, PhD thesis, Thése de la Faculté des Sciences de Paris.

[21] Feller, W. (1975), *An Introduction to Probability Theory and It's Applications, Vol. II*, Second Edition, John Wiley and Sons, NY, NY.

[22] Feuerverger, A. and McDunnough, P. (1981a), On the efficiency of empirical characteristic function procedures, *J. Roy. Stat. Soc, ser. B* **43**:20-27.

[23] Feuerverger, A. and McDunnough, P. (1981b), On efficient inference in symmetric stable laws and processes, *Statistics and Related Topics*, ed. by M. Csörgö, Dawson, D. A., Rao, N. J. K. and Saleh, A. K., 109-122.

[24] Ferguson, T. S. (1996) *A Course in Large Sample Theory*, Chapman and Hall, NY, NY.

[25] Fofack, H. and Nolan, J. P. (1998), Tail behavior, modes and other characteristics of stable distributions. To appear in *Extremes*.

[26] Hoffmann-Jørgensen, J. (1994), Stable densities, *Theor. Prob. Appl.* **38**:350-355.

[27] Janicki, A. and Weron, A. (1994), *Simulation and Chaotic Behavior of α-Stable Stochastic Processes*, Marcel Dekker, NY, NY.

[28] Kogon, S. M. and Williams, D. B. (1998), Characteristic function based estimation of stable parameters, In Adler, R., Feldman, R. and Taqqu, M. (eds.) *A Practical Guide to Heavy Tailed Data*, Birkhäuser, Boston, MA.

[29] Koutrouvelis, I. A. (1980), Regression type estimation of the parameters of stable laws, *JASA* **75**: 918-928.

[30] Koutrouvelis, I. A. (1981), An iterative procedure for the estimation of the parameters of stable laws, *Commun. Statist. Simula* **10**:17-28.

[31] Lévy, P. (1924), Théorie de l'addition des variables aléatoires. *Bulletin de la Société de France* **52**:49-85.

[32] Ma, X. and Nikias, C. L. (1995). On blind channel identification for impulsive signal environments, *Proc. ICASSP'95.*

[33] McCulloch, J. H. (1986), Simple consistent estimators of stable distribution parameters, *Commun. Statist. Simula.* **15**: 1109-1136.

[34] McCulloch, J. H. (1996), Financial applications of stable distributions, *Statistical Methods in Finance, Handbook of Statistics, Vol. 14*, Maddala, G. S. and Rao, C. R. (eds.) North-Holland, NY.

[35] McCulloch, J. H. (1997) Measuring Tail Thickness to Estimate the Stable Index Alpha: A Critique *J. Bus. Econ. Stat.*, **15**: 74-81.

[36] McCulloch, J. H. (1998), Linear regression with stable disturbances, In Adler, R., Feldman, R. and Taqqu, M. (eds.) *A Practical Guide to Heavy Tailed Data*, Birkhäuser, Boston, MA.

[37] Michael, J. R. (1983), The stabilized probability plot, *Biometrika* **70**:11-17.

[38] Modarres, R. and Nolan, J. P. (1994). A Method for Simulating Stable Random Vectors, *Computational Statistics*, **9**, 11-19.

[39] Nikias, C. L. and Shao, M. (1995), *Signal Processing with Alpha-Stable Distributions and Applications*, Wiley, NY, NY.

[40] Nolan, J. P. (1997), Numerical computation of stable densities and distribution functions, *Comm. in Stat. - Stochastic Models*, **13**: 759-774.

[41] Nolan, J. P. (1998), Parameterizations and modes of stable distributions, *Stat. and Prob. Letters*, **38**, 187-195.

[42] Nolan, J. P., Panorska, A. K. and McCulloch, J. H. (1996), Estimation of stable spectral measures, Preprint.

[43] Nolan, J. P. and Panorska, A. K. (1997), Data analysis for heavy tailed multivariate samples, *Comm. in Stat. - Stochastic Models*, **13**: 687-702.

[44] Nolan, J. P. and Rajput, B. (1995). Calculation of multidimensional stable densities. *Commun. Statist.-Simula.* **24:** 551-566.

[45] Paulson, A. S., Holcomb, E. W., and Leitch, R. (1975), The estimation of the parameters of the stable laws, *Biometrika* **62**: 163-170.

[46] Pierce, R. D. (1997) Application of the positive α stable distribution, *1997 IEEE Signal Processing Workshop on Higher-Order Statistics*, to appear.

[47] Press, S. J. (1972), Estimation in univariate and multivariate stable distributions, *JASA*, **67**: 842-846.

[48] Rachev S.T., and Xin, H. (1993), Test for Association of Random Variables in the Domain of Attraction of Multivariate Stable Law, *Probability and Mathematical Statistics*, **14**:125 - 141.

[49] Reiss, R.-D. and Thomas, M. (1997) *Statistical Analysis of Extreme Values*, Birkhäuser, Boston, MA.

[50] Samorodnitsky, G. and Taqqu, M. S. (1994), *Stable Non-Gaussian Random Processes.* Chapman and Hall, NY, NY.

[51] Stephens, M. A. (1981). Kolmogorov-Smirnov-type tests of fit, in *Encycl. of Statistical Sciences*, Kotz, S. and Johnson, N. L., Wiley, NY, NY.

[52] Stuck, B. W. and Kleiner, B. (1974), A statistical analysis of telephone noise, *Bell Syst. Tech. J.* **53**:1263-1320.

[53] Tsihrintzis, G. A. and Nikias, C. L. (1996). Fast estimation of the parameters of alpha-stable impulsive interference. *IEEE Trans. on Sig. Proc.* **44**:1492-1503.

[54] Zolotarev, V. M. (1986), *One-dimensional Stable Distributions, Amer. Math. Soc. Transl. of Math. Monographs, Vol. 65.* Amer. Math. Soc., Providence, RI. (Transl. of the original 1983 Russian)