Stable Distributions

Models for Heavy Tailed Data

John P. Nolan
jpnolan@american.edu
Math/Stat Department
American University

Processed January 30, 2018

Dedicated to
Martha, Julia and Erin
and
Anne Zemitus Nolan (1919-2016)

Contents

Part I

UNIVARIATE STABLE DISTRIBUTIONS

1

Basic Properties of Univariate Stable Distributions

Stable distributions are a rich class of probability distributions that allow skewness and heavy tails and have many intriguing mathematical properties. The class was characterized by Paul Lévy in his study of sums of independent identically distributed terms in the 1920's. The lack of closed formulas for densities and distribution functions for all but a few stable distributions (Gaussian, Cauchy and Lévy, see Figure 1.1), has been a major drawback to the use of stable distributions by practitioners. There are now reliable computer programs to compute stable densities, distribution functions and quantiles. With these programs, it is possible to use stable models in a variety of practical problems.

This book describes the basic facts about univariate and multivariate stable distributions, with an emphasis on practical applications. Part I focuses on univariate stable laws. This chapter describes basic properties of univariate stable distributions. Chapter 2 gives examples of stable laws arising in different problems. Chapter 3 gives proofs of the results in this chapter, as well as more technical details about stable distributions. Chapter 4 describes methods of fitting stable models to data. This structure is continued in Part II, which concerns multivariate stable laws. Chapters 5, 6, and 8 give basic facts about multivariate stable distributions, proofs and technical results, and estimation respectively. Part III is about stable regression, stable times series, and general stable processes. At the end of the book, Part IV describes related distributions and the appendices give tables of stable quantiles, modes and asymptotic standard deviations of maximum likelihood estimators of stable parameters.

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) and Uchaikin and Zolotarev (1999) for these and other examples. The second reason is the Generalized Central Limit Theorem which states that the only possible non-trivial limit of normalized sums of independent identically distributed terms is stable. It is argued that some 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), Fama (1965), Samuelson (1967), Roll (1970), Embrechts et al. (1997), Rachev and Mittnik (2000), 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 can be well described by a stable distribution.

Several recent monographs focus on stable models: Zolotarev (1986), Uchaikin and Zolotarev (1999), Christoph and Wolf (1992), Samorodnitsky and Taqqu (1994), Janicki and Weron (1994), and Nikias and Shao (1995). The related topic of modeling with the extremes of data and heavy tailed distributions is discussed in Embrechts et al. (1997), Adler et al. (1998), and in Reiss and Thomas (2001).

1.1 Definition of stable

An important property of normal or Gaussian random variables is that the sum of two of them is itself a normal random variable. One consequence of this is that if X is normal, then for X_1 and X_2 independent copies of X and any positive constants a and b,

$$aX_1 + bX_2 \stackrel{d}{=} cX + d, \tag{1.1}$$

for some positive c and some $d \in \mathbb{R}$. (The symbol $\stackrel{d}{=}$ means equality in distribution, i.e. both expressions have the same probability law.) In words, equation (1.1) says that the shape of X is preserved (up to scale and shift) under addition. This book is about the class of distributions with this property.

Definition 1.1 A random variable X is *stable* or *stable in the broad sense* if for X_1 and X_2 independent copies of X and any positive constants a and b, (1.1) holds for some positive c and some $d \in \mathbb{R}$. The random variable is *strictly stable* or *stable in the narrow sense* if (1.1) holds with $d = 0$ for all choices of a and b. A random variable is *symmetric stable* if it is stable and symmetrically distributed around 0, e.g. $X \stackrel{d}{=} -X$.

The addition rule for independent normal random variables says that the mean of the sum is the sum of the means and the variance of the sum is the sum of the variances. Suppose $X \sim \mathbf{N}(\mu, \sigma^2)$, then the terms on the left hand side above are $\mathbf{N}(a\mu, (a\sigma)^2)$ and $\mathbf{N}(b\mu, (b\sigma)^2)$ respectively, while the right hand side is $\mathbf{N}(c\mu + d, (c\sigma)^2)$. By the addition rule one must have $c^2 = a^2 + b^2$ and $d = (a + b - c)\mu$. Expressions for c and d in the general stable case are given below.

The word stable is used because the shape is stable or unchanged under sums of the type (1.1). Some authors use the phrase *sum stable* to emphasize the fact that (1.1) is about a sum and to distinguish between these distributions and max-stable, min-stable, multiplication stable and geometric stable distributions (see Chapter 13). Also, some older literature used slightly different terms: stable was originally used for what we now call strictly stable, *quasi-stable* was reserved for what we now call stable.

Two random variables X and Y are said to be of the same *type* if there exist constants $A > 0$ and $B \in \mathbb{R}$ with $X \stackrel{d}{=} AY + B$. The definition of stability can be restated as $aX_1 + bX_2$ has the same type as X.

There are three cases where one can write down closed form expressions for the density and verify directly that they are stable - normal, Cauchy and Lévy distributions. The parameters α and β mentioned below are defined in Section 1.3.

Example 1.2 Normal or Gaussian distributions. $X \sim \mathbf{N}(\mu, \sigma^2)$ if it has a density

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < \infty.$$

The cumulative distribution function, for which there is no closed form expression, is $F(x) = P(X \le x) = \Phi((x-\mu)/\sigma)$, where $\Phi(z) =$ probability that a standard normal r.v. is less than or equal z. Problem 1.1 shows a Gaussian distribution is stable with parameters $\alpha = 2$, $\beta = 0$. □

Example 1.3 Cauchy distributions. $X \sim \text{Cauchy}(\gamma, \delta)$ if it has density

$$f(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x-\delta)^2} \quad -\infty < x < \infty.$$

These are also called Lorentz distributions in physics. Problem 1.2 shows a Cauchy distribution is stable with parameters $\alpha = 1$, $\beta = 0$ and Problem 1.3 gives the d.f. of a Cauchy distribution. □

Example 1.4 Lévy distributions. $X \sim \text{Lévy}(\gamma, \delta)$ if it has density

$$f(x) = \sqrt{\frac{\gamma}{2\pi}} \frac{1}{(x-\delta)^{3/2}} \exp\left(-\frac{\gamma}{2(x-\delta)}\right), \quad \delta < x < \infty.$$

Note that some authors use the term Lévy distribution for all sum stable laws; we shall only use it for this particular distribution. Problem 1.4 shows a Lévy distribution is stable with parameters $\alpha = 1/2$, $\beta = 1$ and Problem 1.5 gives the d.f. of a Lévy distribution. □

Figure 1.1 shows a plot of these three densities. Both normal distributions and Cauchy distributions are symmetric, bell-shaped curves. The main qualitative distinction between them is that the Cauchy distribution has much heavier tails, see Table 1.1. In particular, there is a tiny amount of probability above 3 for the normal distribution, but a significant amount above 3 for a Cauchy. In a sample of data from these two distributions, there will be (on average) approximately 100 times more values above 3 in the Cauchy case than in the normal case. This is the reason stable distributions are called heavy tailed. In contrast to the normal and Cauchy distributions, the Lévy distribution is highly skewed, with all of the probability concentrated on $x > 0$, and it has even heavier tails than the Cauchy.

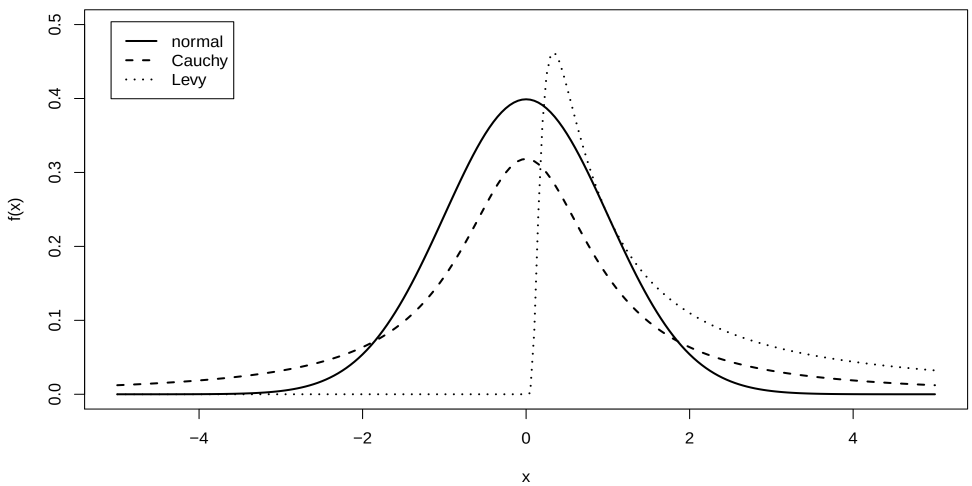

Figure 1.1: Graphs of standardized normal $\mathbf{N}(0,1)$, Cauchy(1,0) and Lévy(1,0) densities.

c	$P(X > c)$		
	Normal	Cauchy	Lévy
0	0.5000	0.5000	1.0000
1	0.1587	0.2500	0.6827
2	0.0228	0.1476	0.5205
3	0.001347	0.1024	0.4363
4	0.00003167	0.0780	0.3829
5	0.0000002866	0.0628	0.3453

Table 1.1: Comparison of tail probabilities for standard normal, Cauchy and Lévy distributions.

General stable distributions allow for varying degrees of tail heaviness and varying degrees of skewness.

Other than the normal distribution, the Cauchy distribution, the Lévy distribution, and the reflection of the Lévy distribution, there are no known closed form expressions for general stable densities and it is unlikely that any other stable distributions have closed forms for their densities. Zolotarev (1986) (pg. 155-158) shows that in a few cases stable densities or distribution functions are expressible in terms of certain special functions. This may seem to doom the use of stable models in practice, but recall that there is no closed formula for the normal cumulative distribution function. There are tables and accurate computer algorithms for the standard normal distribution function, and people routinely use those values in normal models. We now have computer programs to compute quantities of interest for stable distributions, so it is possible to use them in practical problems.

1.2 Other definitions of stability

There are other equivalent definitions of stable random variables. Two are stated here, the proofs of the equivalence of these definitions are given in Section 3.1.

Definition 1.5 Non-degenerate X is *stable* if and only if for all $n > 1$, there exist constants $c_n > 0$ and $d_n \in \mathbb{R}$ such that

$$X_1 + \cdots + X_n \stackrel{d}{=} c_n X + d_n,$$

where $X_1, \ldots, X_n$ are independent, identical copies of X. X is *strictly stable* if and only if $d_n = 0$ for all n.

Section 3.1 shows that the only possible choice for the scaling constants is $c_n = n^{1/\alpha}$ for some $\alpha \in (0,2]$. Both the original definition of stable and the one above use distributional properties of X, yet another distributional characterization is given by the Generalized Central Limit Theorem, Theorem 1.20. While useful, these conditions do not give a concrete way of parameterizing stable distributions. The most concrete way to describe all possible stable distributions is through the characteristic function or Fourier transform. (For a random variable X with distribution function $F(x)$, the characteristic function is defined by $\phi(u) = E\exp(iuX) = \int_{-\infty}^{\infty} \exp(iux) dF(x)$. The function $\phi(u)$ completely determines the distribution of X and has many useful mathematical properties, see Appendix A.) The sign function is used below, it is defined as

$$\operatorname{sign} u = \begin{cases} -1 & u < 0 \\ 0 & u = 0 \\ 1 & u > 0. \end{cases}$$

In the expression below for the $\alpha = 1$ case, $0 \cdot \log 0$ is always interpreted as $\lim_{x \downarrow 0} x \log x = 0$.

Definition 1.6 A random variable X is *stable* if and only if $X \stackrel{d}{=} aZ + b$, where $0 < \alpha \le 2$, $-1 \le \beta \le 1$, $a \ne 0$, $b \in \mathbb{R}$ and Z is a random variable with characteristic function

$$E\exp(iuZ) = \begin{cases} \exp\left(-|u|^\alpha \left[1 - i\beta \tan \frac{\pi\alpha}{2} (\operatorname{sign} u)\right]\right) & \alpha \ne 1 \\ \exp\left(-|u| \left[1 + i\beta \frac{2}{\pi} (\operatorname{sign} u) \log |u|\right]\right) & \alpha = 1. \end{cases} \tag{1.2}$$

These distributions are symmetric around zero when $\beta = 0$ and $b = 0$, in which case the characteristic function of aZ has the simpler form

$$\phi(u) = e^{-a^\alpha |u|^\alpha}.$$

Problems 1.1, 1.2 and 1.4 show that a $\mathbf{N}(\mu, \sigma^2)$ distribution is stable with $(\alpha = 2, \beta = 0, a = \sigma/\sqrt{2}, b = \mu)$, a Cauchy$(\gamma, \delta)$ distribution is stable with $(\alpha = 1, \beta = 0, a = \gamma, b = \delta)$ and a Lévy(γ, δ) distribution is stable with $(\alpha = 1/2, \beta = 1, a = \gamma, b = \delta)$.

1.3 Parameterizations of stable laws

Definition 1.6 shows that a general stable distribution requires four parameters to describe: an *index of stability* or *characteristic exponent* $\alpha \in (0,2]$, a skewness parameter $\beta \in [-1,1]$, a scale parameter and a location parameter. We will use γ for the scale parameter and δ for the location parameter to avoid confusion with the symbols σ and μ, which will be used exclusively for the standard deviation and mean. The parameters are restricted to the range $\alpha \in (0,2]$, $\beta \in [-1,1]$, $\gamma \geq 0$ and $\delta \in \mathbb{R}$. Generally $\gamma > 0$, although $\gamma = 0$ will sometimes be used to denote a degenerate distribution concentrated at δ when it simplifies the statement of a result. Since α and β determine the form of the distribution, they may be considered shape parameters.

There are multiple parameterizations for stable laws and much confusion has been caused by these different parameterizations. The variety of parameterizations is caused by a combination of historical evolution, plus the numerous problems that have been analyzed using specialized forms of the stable distributions. There are good reasons to use different parameterizations in different situations. If numerical work or fitting data is required, then one parameterization is preferable. If simple algebraic properties of the distribution are desired, then another is preferred. If one wants to study the analytic properties of strictly stable laws, then yet another is useful. This section will describe three parameterizations; in Section 3.4 eight others are described.

In most of the recent literature, the notation $\mathbf{S}_\alpha(\sigma,\beta,\mu)$ is used for the class of stable laws. We will use a modified notation of the form $\mathbf{S}(\alpha,\beta,\gamma,\delta;k)$ for three reasons. First, the usual notation singles out α as different and fixed. In statistical applications, all four parameters $(\alpha,\beta,\gamma,\delta)$ are unknown and need to be estimated; the new notation emphasizes this. Second, the scale parameter is not the standard deviation (even in the Gaussian case), and the location parameter is not generally the mean. So we use the neutral symbols γ for the scale (not σ) and δ for the location (not μ). And third, there should be a clear distinction between the different parameterizations; the integer k does that. Users of stable distributions need to state clearly what parameterization they are using, this notation makes it explicit.

Definition 1.7 A random variable X is $\mathbf{S}(\alpha,\beta,\gamma,\delta;0)$ if

$$X \overset{d}{=} \begin{cases} \gamma(Z - \beta \tan \frac{\pi\alpha}{2}) + \delta & \alpha \neq 1 \\ \gamma Z + \delta & \alpha = 1 \end{cases}, \tag{1.3}$$

where $Z = Z(\alpha,\beta)$ is given by (1.2). X has characteristic function

$$E\exp(iuX) = \begin{cases} \exp\left(-\gamma^\alpha |u|^\alpha \left[1 + i\beta (\tan \frac{\pi\alpha}{2})(\operatorname{sign} u)(|\gamma u|^{1-\alpha} - 1)\right] + i\delta u\right) & \alpha \neq 1 \\ \exp\left(-\gamma |u| \left[1 + i\beta \frac{2}{\pi} (\operatorname{sign} u) \log(\gamma |u|)\right] + i\delta u\right) & \alpha = 1. \end{cases} \tag{1.4}$$

When the distribution is standardized, i.e. scale $\gamma = 1$, and location $\delta = 0$, the symbol $\mathbf{S}(\alpha,\beta;0)$ will be used as an abbreviation for $\mathbf{S}(\alpha,\beta,1,0;0)$.

Definition 1.8 A random variable X is $\mathbf{S}(\alpha,\beta,\gamma,\delta;1)$ if

$$X \overset{d}{=} \begin{cases} \gamma Z + \delta & \alpha \neq 1 \\ \gamma Z + (\delta + \beta \frac{2}{\pi} \gamma \log \gamma) & \alpha = 1, \end{cases} \tag{1.5}$$

where $Z = Z(\alpha,\beta)$ is given by (1.2). X has characteristic function

$$E\exp(iuX) = \begin{cases} \exp\left(-\gamma^\alpha |u|^\alpha \left[1 - i\beta (\tan \frac{\pi\alpha}{2})(\operatorname{sign} u)\right] + i\delta u\right) & \alpha \neq 1 \\ \exp\left(-\gamma |u| \left[1 + i\beta \frac{2}{\pi} (\operatorname{sign} u) \log |u|\right] + i\delta u\right) & \alpha = 1. \end{cases} \tag{1.6}$$

When the distribution is standardized, i.e. scale $\gamma = 1$, and location $\delta = 0$, the symbol $\mathbf{S}(\alpha,\beta;1)$ will be used as an abbreviation for $\mathbf{S}(\alpha,\beta,1,0;1)$.

Above we defined the general stable law in the 0-parameterization and 1-parameterization in terms of a standardized $Z \sim \mathbf{S}(\alpha,\beta;1)$. Alternatively, we could start with $Z_0 \sim \mathbf{S}(\alpha,\beta;0)$, in which case

$$\gamma Z_0 + \delta \sim \mathbf{S}(\alpha,\beta,\gamma,\delta;0)$$

and

$$\begin{cases} \gamma Z_0 + \delta + \beta\gamma \tan\frac{\pi\alpha}{2} & \alpha \neq 1 \\ \gamma Z_0 + \delta + \beta\frac{2}{\pi}\gamma\log\gamma & \alpha = 1 \end{cases} \sim \mathbf{S}(\alpha,\beta,\gamma,\delta;1).$$

Since the density of Z_0 is continuous with respect to x, α, and β, this makes it clear how the 1-parameterization is not continuous as $\alpha \to 1$ (because of the $\tan(\pi\alpha/2)$ term) and not a scale location family when $\alpha = 1$ (because the $\gamma\log\gamma$ term). Note that if $\beta = 0$, then the 0- and 1-parameterizations are identical, but when $\beta \neq 0$ the asymmetry factor (the imaginary term in the characteristic function) becomes an issue. The symbol $\mathbf{S\alpha S}$ is used as an abbreviation for symmetric α-stable. When a scale parameter is used, $\mathbf{S\alpha S}(\gamma) = \mathbf{S}(\alpha,0,\gamma,0;0) = \mathbf{S}(\alpha,0,\gamma,0;1)$.

The different parameterizations have caused repeated misunderstandings. Hall (1981a) describes a "comedy of errors" caused by parameterization choices. The most common mistake concerns the sign of the skewness parameter when $\alpha = 1$. Zolotarev (1986) briskly switches between half a dozen parameterizations. Another example is the stable random number generator of Chambers et al. (1976) which has two arguments: α and β. Most users expect to get a $\mathbf{S}(\alpha,\beta;1)$ result, however, the routine actually returns random variates with a $\mathbf{S}(\alpha,\beta;0)$ distribution. One book even excludes the cases $\beta \neq 0$ when $\alpha = 1$.

In principle, any choice of scale and location is as good as any other choice. We recommend using the $\mathbf{S}(\alpha,\beta,\gamma,\delta;0)$ parameterization for numerical work and statistical inference with stable distributions: it has the simplest form for the characteristic function that is continuous in all parameters. See Figure 1.2 for plots of stable densities in the 0-parameterization. It lets α and β determine the shape of the distribution, while γ and δ determine scale and location in the standard way: if $X \sim \mathbf{S}(\alpha,\beta,\gamma,\delta;0)$, then $(X-\delta)/\gamma \sim \mathbf{S}(\alpha,\beta,1,0;0)$. This is not true for the $\mathbf{S}(\alpha,\beta,\gamma,\delta;1)$ parameterization when $\alpha = 1$.

On the other hand, if one is primarily interested in a simple form for the characteristic function and nice algebraic properties, the $\mathbf{S}(\alpha,\beta,\gamma,\delta;1)$ parameterization is favored. Because of these properties, this is the most common parameterization in use and we will generally use it when we are proving facts about stable distributions. The main practical disadvantage of the $\mathbf{S}(\alpha,\beta,\gamma,\delta;1)$ parameterization is that the location of the mode is unbounded in any neighborhood of $\alpha = 1$: if $X \sim \mathbf{S}(\alpha,\beta,\gamma,\delta;1)$ and $\beta > 0$, then the mode of X tends to $+\infty$ as $\alpha \uparrow 1$ and tends to $-\infty$ as $\alpha \downarrow 1$. Moreover, the $\mathbf{S}(\alpha,\beta,\gamma,\delta;1)$ parameterization does not have the intuitive properties desirable in applications (continuity of the distributions as the parameters vary, a scale and location family, etc.). See Figure 1.3 for densities in the 1-parameterization and Section 3.2.2 for more information on modes.

When $\alpha = 2$, a $\mathbf{S}(2,0,\gamma,\delta;0) = \mathbf{S}(2,0,\gamma,\delta;1)$ distribution is normal with mean δ, but the standard deviation is not γ. Because of the way the characteristic function is defined above, $\mathbf{S}(2,0,\gamma,\delta;0) = \mathrm{N}(\delta, 2\gamma^2)$, so the normal standard deviation $\sigma = \sqrt{2}\gamma$. This fact is a frequent source of confusion when one tries to compare stable quantiles when $\alpha = 2$ to normal quantiles. This complication is not inherent in the properties of stable laws; it is

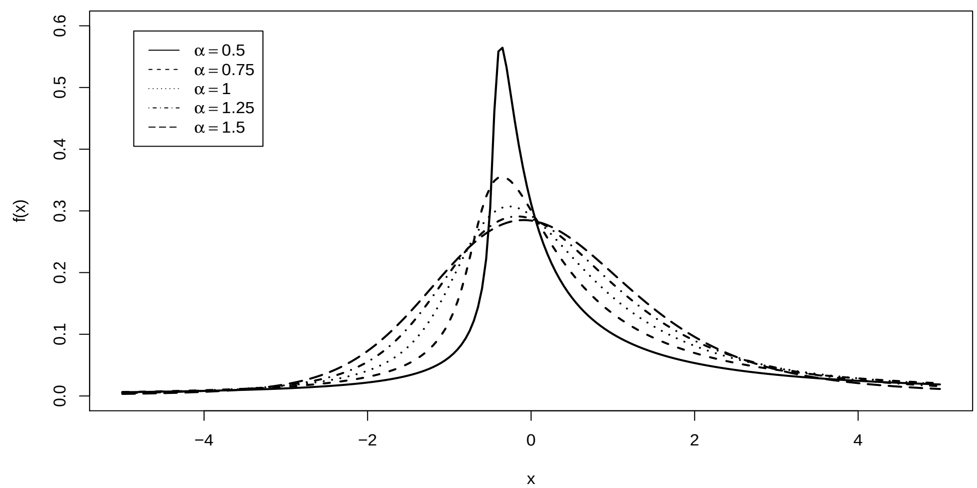

Figure 1.2: Stable densities in the $\mathbf{S}(\alpha, 0.5, 1, 0; 0)$ parameterization, $\alpha = 0.5, 0.75, 1, 1.25, 1.5$.

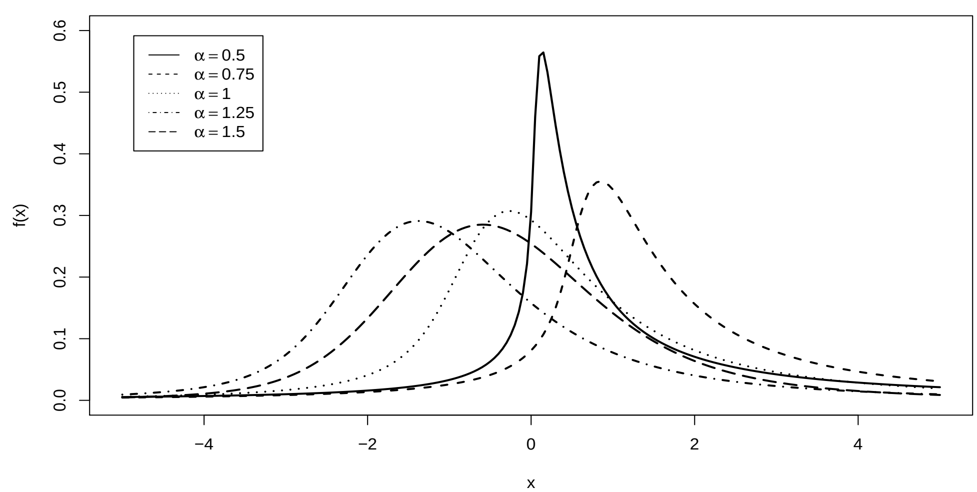

Figure 1.3: Stable densities in the $\mathbf{S}(\alpha, 0.5, 1, 0; 1)$ parameterization, $\alpha = 0.5, 0.75, 1, 1.25, 1.5$.

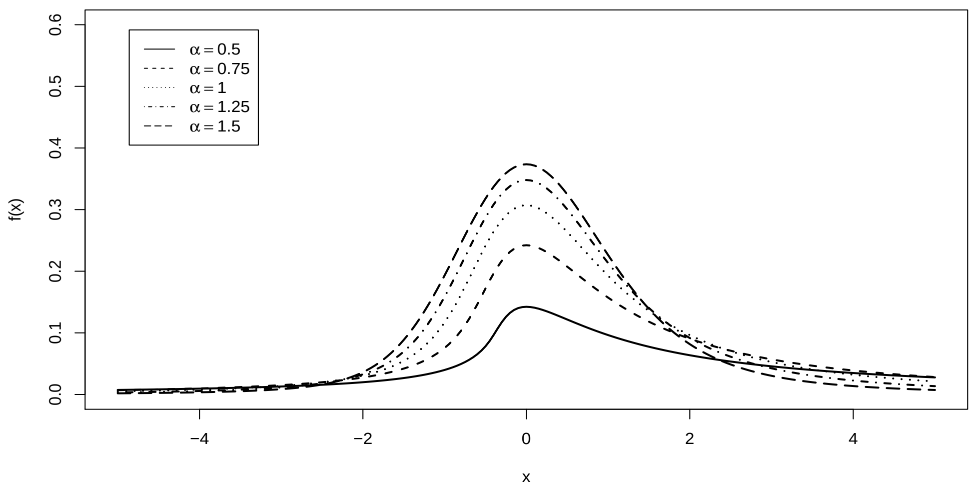

Figure 1.4: Stable densities in the $\mathbf{S}(\alpha, 0.5; 2)$ parameterization, $\alpha = 0.5, 0.75, 1, 1.25, 1.5$.

a consequence of the way the parameterization has been chosen. The 2-parameterization mentioned below rescales to avoid this problem, but the above scaling is standard in the literature. Also, when $\alpha = 2$, β is irrelevant because then the factor $\tan(\pi\alpha/2) = 0$. While you can allow any $\beta \in [-1, 1]$, it is customary take $\beta = 0$ when $\alpha = 2$; this emphasizes that the normal distribution is always symmetric.

Since multiple parameterizations are used for stable distributions, it is perhaps worthwhile to ask if there is another parameterization where the scale and location parameter have a more intuitive meaning. Section 3.4 defines the $\mathbf{S}(\alpha, \beta, \gamma, \delta; 2)$ parameterization so that the location parameter is at the mode and the scale parameter agrees with the standard scale parameters in the Gaussian and Cauchy cases. While technically more cumbersome, this parameterization may be the most intuitive for applications. In particular, it is useful in signal processing and in linear regression problems when there is skewness. Figure 1.4 shows plots of the densities in this parameterization.

A stable distribution can be represented in any one of these or other parameterizations. For completeness, Section 3.4 lists eleven different parameterizations that can be used, and the relationships of these to each other. We will generally use the $\mathbf{S}(\alpha, \beta, \gamma, \delta; 0)$ and $\mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$ parameterizations in what follows to avoid (or at least limit) confusion. In these two parameterizations, α, β and the scale γ are always the same, but the location parameters will have different values. The notation $X \sim \mathbf{S}(\alpha, \beta, \gamma, \delta_k; k)$ for $k = 0, 1$ will be shorthand for $X \sim \mathbf{S}(\alpha, \beta, \gamma, \delta_0; 0)$ and $X \sim \mathbf{S}(\alpha, \beta, \gamma, \delta_1; 1)$ simultaneously. In this case, the parameters are related by (see Problem 1.9)

$$\delta_0 = \begin{cases} \delta_1 + \beta\gamma\tan\frac{\pi\alpha}{2} & \alpha \neq 1 \\ \delta_1 + \beta\frac{2}{\pi}\gamma\log\gamma & \alpha = 1 \end{cases} \qquad \delta_1 = \begin{cases} \delta_0 - \beta\gamma\tan\frac{\pi\alpha}{2} & \alpha \neq 1 \\ \delta_0 - \beta\frac{2}{\pi}\gamma\log\gamma & \alpha = 1 \end{cases} \tag{1.7}$$

In particular, note that in (1.2), $Z(\alpha,\beta) \sim \mathbf{S}(\alpha,\beta,1,\beta\tan\frac{\pi\alpha}{2};0) = \mathbf{S}(\alpha,\beta,1,0;1)$ when $\alpha \neq 1$ and $Z(1,\beta) \sim \mathbf{S}(1,\beta,1,0;0) = \mathbf{S}(1,\beta,1,0;1)$ when $\alpha = 1$.

1.4 Densities and distribution functions

While there are no explicit formulas for general stable densities, a lot is known about their theoretical properties. The most basic fact is the following.

Theorem 1.9 *All (non-degenerate) stable distributions are continuous distributions with an infinitely differentiable density.*

To distinguish between the densities and cumulative distribution functions in different parameterizations, $f(x|\alpha,\beta,\gamma,\delta;k)$ will denote the density and $F(x|\alpha,\beta,\gamma,\delta;k)$ will denote the d.f. of a $\mathbf{S}(\alpha,\beta,\gamma,\delta;k)$ distribution. When the distribution is standardized, i.e. scale $\gamma = 1$, and location $\delta = 0$, $f(x|\alpha,\beta;k)$ will be used for the density, and $F(x|\alpha,\beta;k)$ will be used for the d.f..

Since all stable distributions are shifts and scales of some $Z \sim \mathbf{S}(\alpha,\beta;0)$, we will focus on those distributions here. The computer program STABLE, using algorithms described in Section 3.3, was used to compute the probability density functions (pdf) and (cumulative) distribution functions (d.f.) below to illustrate the range of shapes of these distributions.

Stable densities are supported on either the whole real line or a half line. The latter situation can only occur when $\alpha < 1$ and ($\beta = +1$ or $\beta = -1$). Precise limits are given by the following lemma.

Lemma 1.10 *The support of a stable distribution in the different parameterizations is*

$$\text{support}\, f(x|\alpha,\beta,\gamma,\delta;0) = \begin{cases} [\delta - \gamma\tan\frac{\pi\alpha}{2}, \infty) & \alpha < 1 \textit{ and } \beta = 1 \\ (-\infty, \delta + \gamma\tan\frac{\pi\alpha}{2}] & \alpha < 1 \textit{ and } \beta = -1 \\ (-\infty,+\infty) & \textit{otherwise} \end{cases}$$

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

The constant $\tan\frac{\pi\alpha}{2}$ appears frequently when working with stable distributions, so it is worth recording its behavior. As $\alpha \uparrow 1$, $\tan\frac{\pi\alpha}{2} \uparrow +\infty$, the expression is undefined at $\alpha = 1$, and when $\alpha \downarrow 1$, $\tan\frac{\pi\alpha}{2} \downarrow -\infty$. This essential discontinuity at $\alpha = 1$ is sometimes a nuisance when working with stable distributions, but here it is natural: if $|\beta| = 1$ then as $\alpha \uparrow 1$, the support in Lemma 1.10 grows to $\mathbb{R}$ in a natural way.

Another basic fact about stable distributions is the reflection property.

Proposition 1.11 ***Reflection Property.*** *For any α and β, $Z \sim \mathbf{S}(\alpha,\beta;k)$, $k = 0,1,2$*

$$Z(\alpha,-\beta) \stackrel{d}{=} -Z(\alpha,\beta).$$

Thus the density and distribution function of a $Z(\alpha,\beta)$ random variable satisfy $f(x|\alpha,\beta;k) = f(-x|\alpha,-\beta;k)$ and $F(x|\alpha,\beta;k) = 1 - F(-x|\alpha,-\beta;k)$.

More generally, if $X \sim \mathbf{S}(\alpha,\beta,\gamma,\delta;k)$, then $-X \sim \mathbf{S}(\alpha,-\beta,\gamma,-\delta;k)$, so $f(x|\alpha,\beta,\gamma,\delta;k) = f(-x|\alpha,-\beta,\gamma,-\delta;k)$ and $F(x|\alpha,\beta,\gamma,\delta;k) = 1 - F(-x|\alpha,-\beta,\gamma,-\delta;k)$.

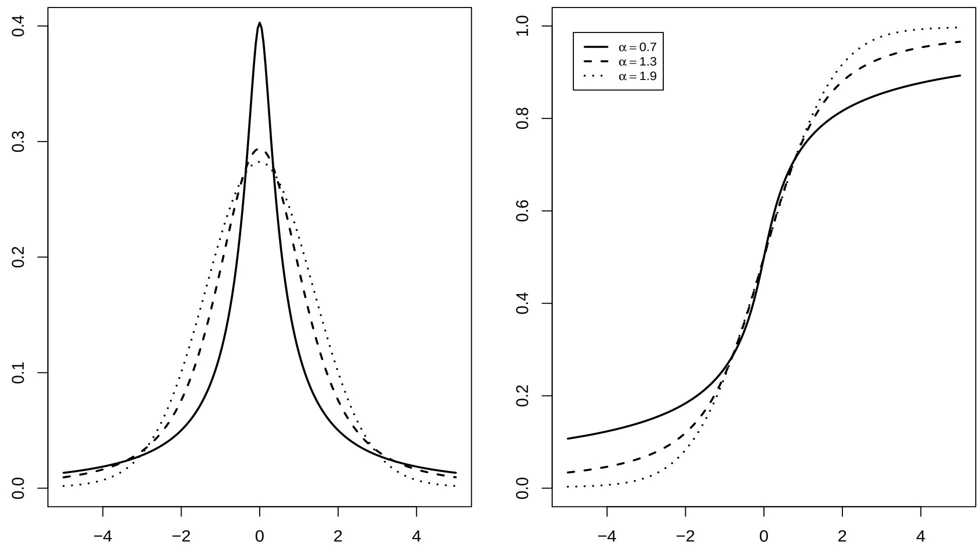

Figure 1.5: Symmetric stable densities and cumulative distribution functions for $Z \sim \mathbf{S}(\alpha, 0; 0)$, $\alpha = 0.7, 1.3, 1.9$.

First consider the case when $\beta = 0$. In this case, the reflection property says $f(x|\alpha, 0; k) = f(-x|\alpha, 0; k)$, so the density and d.f. are symmetric around 0. Figure 1.5 shows the bell-shaped density of symmetric stable distributions. As α decreases, three things occur to the density: the peak gets higher, the region flanking the peak get lower, and the tails get heavier. The d.f. plot shows how as α decreases, the tail probabilities increase.

If $\beta > 0$, then the distribution is skewed with the right tail of the distribution heavier than the left tail: $P(X > x) > P(X < -x)$ for large $x > 0$. (Here and later, statements about the tail of a distribution will always refer to large $|x|$, nothing is implied about $|x|$ small.) When $\beta = 1$, we say the stable distribution is *totally skewed to the right*. By the reflection property, the behavior of the $\beta < 0$ cases are reflections of the $\beta > 0$ ones, with left tail being heavier. When $\beta = -1$, the distribution is *totally skewed to the left*.

When $\alpha = 2$, the distribution is a (non-standardized) normal distribution. Note that $\tan \frac{\pi\alpha}{2} = 0$ in (1.2) so the characteristic function is real and hence the distribution is always symmetric, no matter what the value of β. In symbols, $Z(2, \beta) \stackrel{d}{=} Z(2, 0)$. In general, as $\alpha \uparrow 2$, all stable distributions get closer and closer to being symmetric and β becomes less meaningful in applications (and harder to estimate accurately).

Figure 1.6 shows the density and d.f. when $\alpha = 1.9$, with varying β, and there is little visible difference as β varies. As α decreases, the effect of β becomes more pronounced: the left tail gets lighter and lighter for $\beta > 0$, see Figure 1.7 ($\alpha = 1.3$), Figure 1.8 ($\alpha = 0.7$), and Figure 1.9 ($\alpha = 0.2$). The last figure shows that when α approaches 0, the density gets extremely high at the peak, and the d.f. gets closer and closer to a degenerate distribution (see Section 3.2 for more information on this topic). As Lemma 1.10 shows, the light tail actually is 0 after some point when $\alpha < 1$ and $|\beta| = 1$.

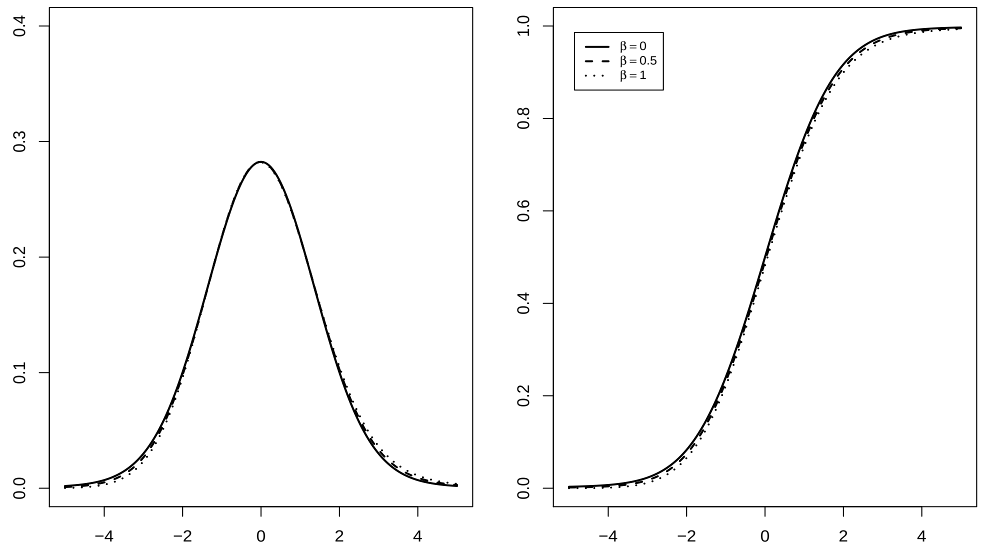

Figure 1.6: Stable densities and cumulative distribution functions for $Z \sim \mathbf{S}(1.9, \beta; 0)$, $\beta = 0, 0.5, 1$.

Finally, all stable densities are unimodal, but there is no known formula for the location of the mode. However, the mode of a $Z \sim \mathbf{S}(\alpha, \beta; 0)$ distribution, denoted by $m(\alpha, \beta)$, has been numerically computed. The values of $m(\alpha, \beta)$ are shown for $\beta \geq 0$ in Figure 1.10 and a table of modes is given in Appendix C. By the reflection property, $m(\alpha, -\beta) = -m(\alpha, \beta)$. Numerically, it is also observed that $P(Z > m(\alpha, \beta)) > P(Z < m(\alpha, \beta))$ (more mass to the right of the mode) when $\beta > 0$, $P(Z > m(\alpha, \beta)) = P(Z < m(\alpha, \beta)) = 1/2$ when $\beta = 0$, and by reflection $P(Z > m(\alpha, \beta)) < P(Z < m(\alpha, \beta))$ when $\beta < 0$ (more mass to the left of the mode). Note that these statements are all in the 0-parameterization, not the 1-parameterization. See Section 3.2.2 for more information about modes.

1.5 Tail probabilities, moments and quantiles

When $\alpha = 2$, the normal distribution has well understood asymptotic tail properties. Here we give a brief discussion of the tails of non-Gaussian ($\alpha < 2$) stable laws, see Section 3.5 for more information. For $\alpha < 2$, stable distributions have one tail (when $\alpha < 1$ and $\beta = \pm 1$) or both tails (all other cases) that are asymptotically power laws with heavy tails. The statement $h(x) \sim g(x)$ as $x \to a$ means $\lim_{x \to a} h(x)/g(x) = 1$.

Theorem 1.12 ***Tail approximation.*** *Let* $X \sim \mathbf{S}(\alpha, \beta, \gamma, \delta; 0)$ *with* $0 < \alpha < 2$, $-1 < \beta \leq 1$. *Then as* $x \to \infty$,

$$
\begin{aligned}
P(X > x) &\sim \gamma^\alpha c_\alpha (1+\beta) x^{-\alpha} \\
f(x|\alpha, \beta, \gamma, \delta; 0) &\sim \alpha \gamma^\alpha c_\alpha (1+\beta) x^{-(\alpha+1)}
\end{aligned}
$$

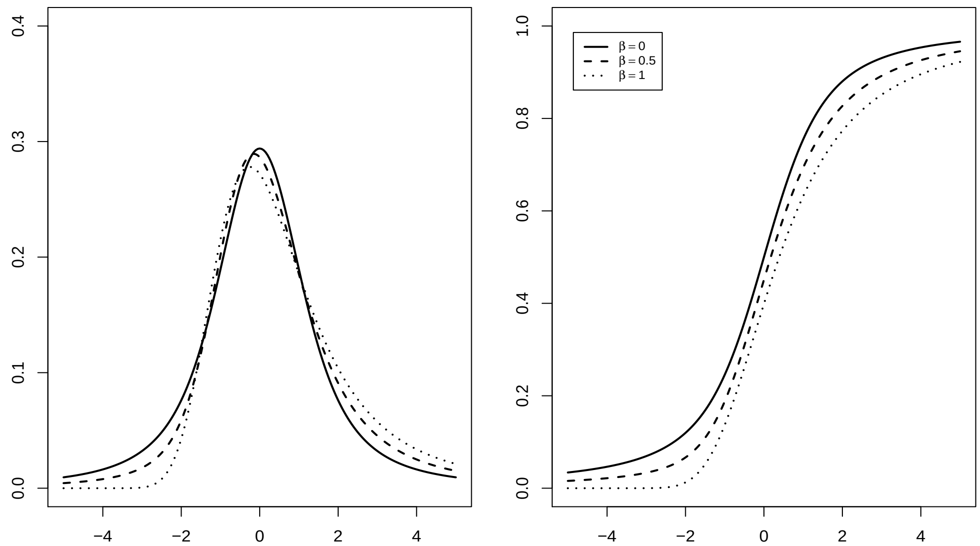

Figure 1.7: Stable densities and cumulative distribution functions for $Z \sim \mathbf{S}(1.3, \beta; 0)$, $\beta = 0, 0.5, 1$.

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

$$\begin{aligned} P(X < -x) &\sim \gamma^\alpha c_\alpha (1-\beta) x^{-\alpha} \\ f(-x|\alpha,\beta,\gamma,\delta;0) &\sim \alpha\gamma^\alpha c_\alpha (1-\beta) x^{-(\alpha+1)}. \end{aligned}$$

For all $\alpha < 2$ and $-1 < \beta < 1$, both tail probabilities and densities are asymptotically power laws. When $\beta = -1$, the right tail of the distribution is not asymptotically a power law; likewise when $\beta = 1$, the left tail of the distribution is not asymptotically a power law. The point at which the tail approximation becomes useful is a complicated issue, it depends on both the parameterization and the parameters $(\alpha, \beta, \gamma, \delta)$. See Section 3.5 for more information on both of these issues.

Pareto distributions (see Problem 1.10) are a class of probability laws with upper tail probabilities given exactly by the right hand side of Theorem 1.12. The term *stable Paretian laws* is used to distinguish between the fast decay of the Gaussian law and the Pareto like tail behavior in the $\alpha < 2$ case.

One consequence of heavy tails is that not all moments exist. In most statistical problems, the first moment EX and variance $\text{Var}(X) = E(X^2) - (EX)^2$ are routinely used to describe a distribution. However, these are not generally useful for heavy tailed distributions, because the integral expressions for these expectations may diverge. In their place, it is sometimes useful to use fractional absolute moments: $E|X|^p = \int_{-\infty}^{\infty} |x|^p f(x) dx$, where p is any real number. Some review on moments and fractional moments is given in Appendix A. Problem 1.11 shows that for $0 < \alpha < 2$, $E|X|^p$ is finite for $0 < p < \alpha$, and that $E|X|^p = +\infty$ for $p \ge \alpha$. Explicit formulas for moments of strictly stable laws are given in Section 3.6.

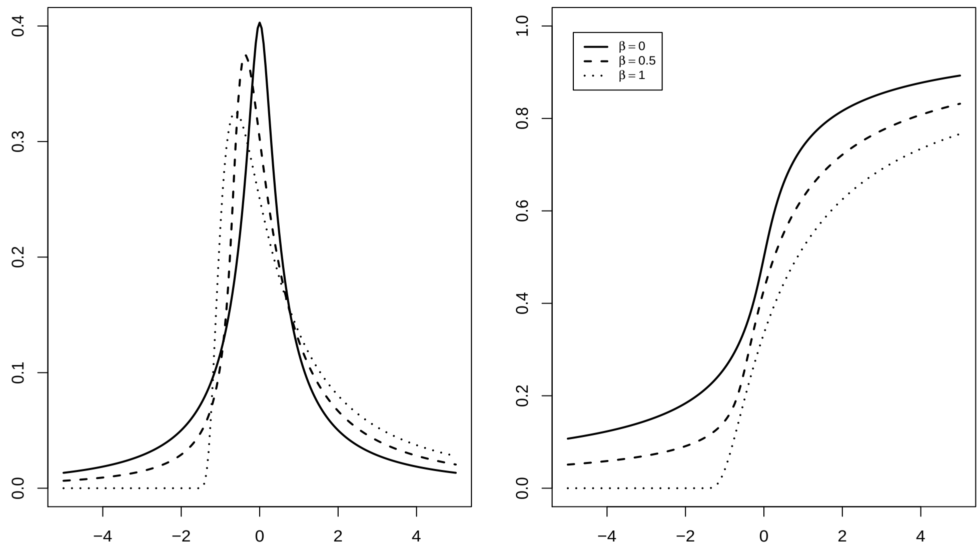

Figure 1.8: Stable densities and cumulative distribution functions for $Z \sim \mathbf{S}(0.7, \beta; 0)$, $\beta = 0, 0.5, 1$.

Thus, when $0 < \alpha < 2$, $E|X|^2 = EX^2 = +\infty$ and stable distributions do not have finite second moments or variances. This fact causes some to immediately dismiss stable distributions as being irrelevant to any practical problem. Section 2.13 discusses this in more detail. When $1 < \alpha \leq 2$, $E|X| < \infty$ and the mean of X is given below. On the other hand, when $\alpha \leq 1$, $E|X| = +\infty$, so means are undefined.

Proposition 1.13 *When $1 < \alpha \leq 2$, the mean of $X \sim \mathbf{S}(\alpha, \beta, \gamma, \delta_k; k)$ for $k = 0, 1$ is*

$$\mu = EX = \delta_1 = \delta_0 - \beta\gamma \tan \tfrac{\pi\alpha}{2}.$$

Consider what happens to the mean of $X \sim \mathbf{S}(\alpha, \beta; 0)$ as $\alpha \downarrow 1$. Even though the mode of the distribution stays close to 0, it has a mean of $\mu = -\beta \tan \frac{\pi\alpha}{2}$. When $\beta = 0$, the distribution is symmetric and the mean is always 0. When $\beta > 0$, the mean tends to $+\infty$ because while both tails are getting heavier, the right tail is heavier than the left. By reflection, the $\beta < 0$ case has $\mu \downarrow -\infty$. Finally, when α reaches 1, the tails are too heavy for the integral $EX = \int_{-\infty}^{\infty} x f(x) dx$ to converge. In contrast, a $\mathbf{S}(\alpha, \beta; 1)$ distribution keeps the mean at 0 by shifting the whole distribution by an increasing amount as $\alpha \downarrow 1$. For example, when $1 < \alpha < 2$, Theorem 3.16 shows that $F(0|\alpha, 1; 1) = 1/\alpha$, which converges up to 1 as $\alpha \downarrow 1$. In these cases, most of the probability is to the left of zero, and only a tiny amount is to the right of zero, yet the mean is still zero because of the very slow decay of the right tail. The behavior is essentially the same for any $\beta > 0$. A $\mathbf{S}(\alpha, \beta; 2)$ distribution keeps the mode exactly at 0, and the mean as a function of (α, β) is continuous, like the mean of a $\mathbf{S}(\alpha, \beta; 0)$ distribution.

Note that the skewness parameter β is not the same thing as the classical skewness parameter. The latter is undefined for every non-Gaussian stable distribution because neither the

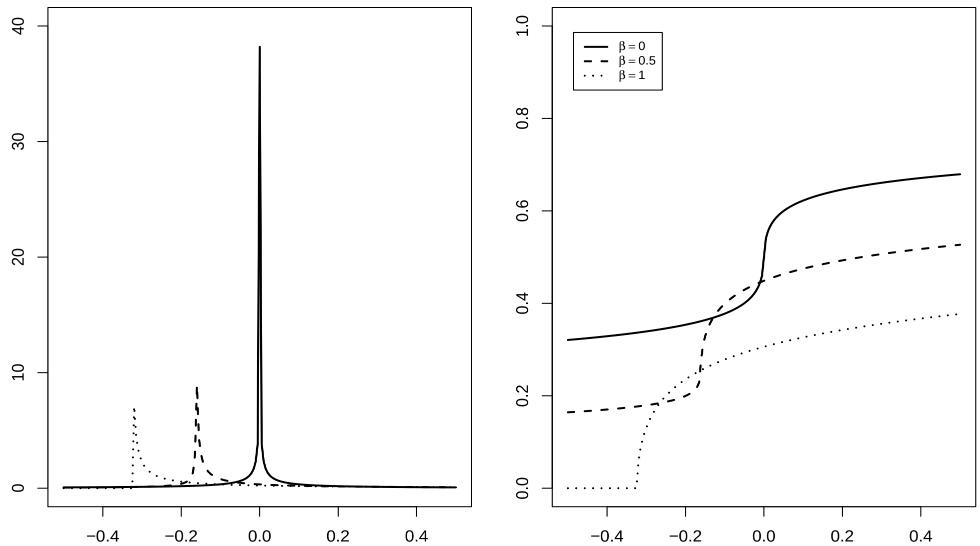

Figure 1.9: Stable densities and cumulative distribution functions for $Z \sim \mathbf{S}(0.2, \beta; 0)$, $\beta = 0, 0.5, 1$. Note that both the horizontal and vertical scales are very different from Figures 1.6 - 1.8.

third moment or the variance exist. Likewise, the kurtosis is undefined, because the fourth moment is undefined for every non-Gaussian stable distribution.

It is sometimes useful to consider non-integer moments of stable distributions. In Section 3.6 it will be shown that for $0 < p < \alpha$, the p-th absolute moment exists: $E|X|^p < \infty$. Such moments are sometimes called *fractional lower order moments* (FLOM). When X is strictly stable there is an explicit form for such moments. Such moments can be used as a measure of dispersion of a stable distribution, and are used in some estimation schemes.

Tables of standard normal quantiles or percentiles are given in most basic probability and statistic books. Let z_λ be the λ^{th} quantile, i.e. the z value for which the standard normal distribution has lower tail probability λ, i.e. $P(Z < z_\lambda) = \lambda$. The value $z_{0.975} = 1.96$ is commonly used: for $X \sim \mathbf{N}(\mu, \sigma^2)$, the 0.025^{th} quantile is $\mu - 1.96\sigma$ and the 0.975^{th} quantile is $\mu + 1.96\sigma$. Quantiles are used to quantify risk. For example, in a Gaussian/normal model for the price of an asset, the interval from $\mu - 1.96\sigma$ to $\mu + 1.96\sigma$ contains 95% of the distribution of the asset price.

Quantiles of the standard stable distributions are used in the same way. The difficulty is that there are different quantiles for every value of α and β. The symbol $z_\lambda(\alpha, \beta)$ will be used for the λ^{th} quantile of a $\mathbf{S}(\alpha, \beta; 0)$ distribution: $P(Z < z_\lambda(\alpha, \beta)) = \lambda$. The easiest way to find these values is to use the program STABLE. Less accurately, one can use the tabulated values in Appendix B and interpolate on the α and β values. Appendix B shows selected quantiles for α=0.1, 0.2,..., 1.9, 1.95, 1.99, 2.0 and β=0, 0.1, 0.2,..., 0.9, 1. (Reflection can be used for negative beta: by Proposition 1.11, $z_\lambda(\alpha, \beta) = z_{1-\lambda}(\alpha, -\beta)$).

We caution the reader about two ways that stable quantiles are different from normal quantiles. First, if the distribution is not symmetric, i.e. $\beta \neq 0$, then the quantiles are not

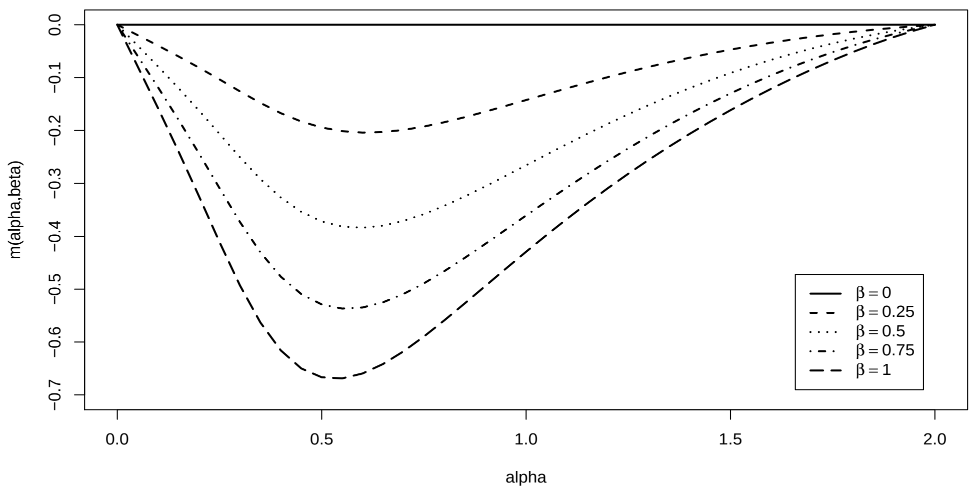

Figure 1.10: The location of the mode of a $\mathbf{S}(\alpha,\beta;0)$ density.

symmetric. Second, the way the quantiles scale depend on what parameterization is being used. In the $\mathbf{S}(\alpha,\beta,\gamma,\delta;0)$ parameterization, it is straightforward; in other parameterizations one has to either convert to the $\mathbf{S}(\alpha,\beta,\gamma,\delta;0)$ parameterization using (1.7), or scale and shift according to the definition of each parameterization. These issues are illustrated in the following examples.

Example 1.14 Find the 5^{th} and 95^{th} quantiles for $X \sim \mathbf{S}(1.3,0.5,2,7;0)$. From Appendix B, the 5^{th} quantile is $z_{0.05}(1.3,0.5) = -2.355$ and the 95^{th} quantile is $z_{0.95}(1.3,0.5) = +5.333$ for a standardized $\mathbf{S}(1.3,0.5,1,0;0)$ distribution. So the corresponding quantiles for X are $\delta - 2.355\gamma = 2.289$ and $\delta + 5.333\gamma = 17.666$. □

Example 1.15 If $X \sim \mathbf{S}(1.3,0.5,2,7;1)$, then using the previous example, the 5^{th} and 95^{th} quantiles are $\gamma(-2.355) + (\delta + \beta\gamma \tan\frac{\pi\alpha}{2}) = 0.327$ and $\gamma(5.333) + (\delta + \beta\gamma \tan\frac{\pi\alpha}{2}) = 15.704$. Alternatively, $\mathbf{S}(1.3,0.5,2,7;1) = \mathbf{S}(1.3,0.5,2,5.037;0)$ by (1.7), so the 5^{th} and 95^{th} quantiles are $2(-2.355) + 5.037 = 0.327$ and $2(5.333) + 5.037 = 15.704$. □

1.6 Sums of stable random variables

A basic property of stable laws is that sums of α-stable random variables are α-stable. In the independent case, the exact parameters of the sums are given below. As always, the results depend on the parameterization used. In these results it is essential that the summands all have the same α, as Problem 1.12 shows that otherwise the sum will not be stable. Section 13.9 discusses this issue briefly. When the summands are dependent, the sum is stable but the precise statement is more difficult and depends on the exact dependence structure; this is explained in Section 5.5.

Proposition 1.16 *The* $\mathbf{S}(\alpha,\beta,\gamma,\delta;0)$ *parameterization has the following properties.*
(a) If $X \sim \mathbf{S}(\alpha,\beta,\gamma,\delta;0)$*, then for any* $a \neq 0$*,* $b \in \mathbb{R}$*,*

$$aX + b \sim \mathbf{S}(\alpha, (\operatorname{sign} a)\beta, |a|\gamma, a\delta + b; 0).$$

(b) The characteristic functions, densities and distribution functions are jointly continuous in all four parameters $(\alpha,\gamma,\beta,\delta)$ *and in x.*
(c) If $X_1 \sim \mathbf{S}(\alpha,\beta_1,\gamma_1,\delta_1;0)$ *and* $X_2 \sim \mathbf{S}(\alpha,\beta_2,\gamma_2,\delta_2;0)$ *are independent, then* $X_1 + X_2 \sim \mathbf{S}(\alpha,\beta,\gamma,\delta;0)$ *where*

$$\beta = \frac{\beta_1\gamma_1^\alpha + \beta_2\gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha}, \quad \gamma^\alpha = \gamma_1^\alpha + \gamma_2^\alpha,$$

$$\delta = \begin{cases} \delta_1 + \delta_2 + (\tan\frac{\pi\alpha}{2})\,[\beta\gamma - \beta_1\gamma_1 - \beta_2\gamma_2] & \alpha \neq 1 \\ \delta_1 + \delta_2 + \frac{2}{\pi}[\beta\gamma\log\gamma - \beta_1\gamma_1\log\gamma_1 - \beta_2\gamma_2\log\gamma_2] & \alpha = 1. \end{cases}$$

The formula $\gamma^\alpha = \gamma_1^\alpha + \gamma_2^\alpha$ in (c) is the generalization of the rule for adding variances of independent random variables: $\sigma^2 = \sigma_1^2 + \sigma_2^2$. It holds for both parameterizations. Note that one adds the α^{th} power of the scale parameters, not the scale parameters themselves.

Proposition 1.17 *The* $\mathbf{S}(\alpha,\beta,\gamma,\delta;1)$ *parameterization has the following properties.*
(a) If $X \sim \mathbf{S}(\alpha,\beta,\gamma,\delta;1)$*, then for any* $a \neq 0$*,* $b \in \mathbb{R}$*,*

$$aX + b \sim \begin{cases} \mathbf{S}(\alpha, (\operatorname{sign} a)\beta, |a|\gamma, a\delta + b; 1) & \alpha \neq 1 \\ \mathbf{S}(1, (\operatorname{sign} a)\beta, |a|\gamma, a\delta + b - \frac{2}{\pi}\beta\gamma a\log|a|; 1) & \alpha = 1. \end{cases}$$

(b) The characteristic functions, densities and distribution functions are continuous away from $\alpha = 1$*, but discontinuous in any neighborhood of* $\alpha = 1$*.*

(c) If $X_1 \sim \mathbf{S}(\alpha,\beta_1,\gamma_1,\delta_1;1)$ *and* $X_2 \sim \mathbf{S}(\alpha,\beta_2,\gamma_2,\delta_2;1)$ *are independent, then* $X_1 + X_2 \sim \mathbf{S}(\alpha,\beta,\gamma,\delta;1)$ *where*

$$\beta = \frac{\beta_1\gamma_1^\alpha + \beta_2\gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha}, \quad \gamma^\alpha = \gamma_1^\alpha + \gamma_2^\alpha, \quad \delta = \delta_1 + \delta_2.$$

The corresponding results for the $\mathbf{S}(\alpha,\beta,\gamma,\delta;2)$ parameterization are given in Proposition 3.43.

Part (a) of the above results shows that γ and δ are standard scale and location parameters in the $\mathbf{S}(\alpha,\beta,\gamma,\delta;0)$ parameterization, but not in the $\mathbf{S}(\alpha,\beta,\gamma,\delta;1)$ parameterization when $\alpha = 1$. In contrast, part (b) shows that the location parameter δ of a sum is the sum of the location parameters $\delta_1 + \delta_2$ only in the $\mathbf{S}(\alpha,\beta,\gamma,\delta;1)$ parameterization. Unfortunately there is no parameterization that has both properties.

In the symmetric case, i.e. $\beta_1 = \beta_2 = 0$, both the previous propositions are simpler to state: if $X_1 \sim \mathbf{S}(\alpha,0,\gamma_1,\delta_1;k)$ and $X_2 \sim \mathbf{S}(\alpha,0,\gamma_2,\delta_2;k)$ (with $k = 0$ or $k = 1$), then $X_1 + X_2 \sim \mathbf{S}(\alpha,0,\gamma,\delta;k)$ with $\gamma^\alpha = \gamma_1^\alpha + \gamma_2^\alpha$ and $\delta = \delta_1 + \delta_2$. This is exactly like the normal case: if $X_1 \sim \mathrm{N}(\mu_1,\sigma_1^2)$ and $X_2 \sim \mathrm{N}(\mu_2,\sigma_2^2)$ and independent, then $X_1 + X_2 \sim \mathrm{N}(\mu,\sigma^2)$ where $\sigma^2 = \sigma_1^2 + \sigma_2^2$ and $\mu = \mu_1 + \mu_2$.

By induction (see Problem 1.13), one gets formulas for sums of n stable random variables: for $X_j \sim \mathbf{S}(\alpha, \beta_j, \gamma_j, \delta_j; k)$, $j = 1, 2, \ldots, n$ independent and arbitrary $w_1, \ldots, w_n$, the sum

$$w_1X_1 + w_2X_2 + \cdots + w_nX_n \sim \mathbf{S}(\alpha, \beta, \gamma, \delta; k) \tag{1.8}$$

where

$$\begin{aligned}
\gamma^\alpha &= \sum_{j=1}^{n} |w_j\gamma_j|^\alpha \\
\beta &= \frac{\sum_{j=1}^{n} \beta_j (\operatorname{sign} w_j)|w_j\gamma_j|^\alpha}{\gamma^\alpha} \\
\delta &= \begin{cases} \sum_j w_j\delta_j + \tan\frac{\pi\alpha}{2}\left(\beta\gamma - \sum_j \beta_j w_j \gamma_j\right) & k=0, \alpha \neq 1 \\ \sum_j w_j\delta_j + \frac{2}{\pi}\left(\beta\gamma\log\gamma - \sum_j \beta_j w_j\gamma_j \log|w_j\gamma_j|\right) & k=0, \alpha = 1 \\ \sum_j w_j\delta_j & k=1, \alpha\neq 1 \\ \sum_j w_j\delta_j - \frac{2}{\pi}\sum_j \beta_j w_j\gamma_j\log|w_j| & k=1, \alpha=1\ . \end{cases}
\end{aligned}$$

Note that if $\beta_j = 0$ for all j, then $\beta = 0$ and $\delta = \sum_j w_j\delta_j$. An important case is the *scaling property* for stable random variables: when the terms are independent and identically distributed, say $X_j \sim \mathbf{S}(\alpha, \beta, \gamma, \delta; k)$, then

$$X_1 + \cdots + X_n \sim \mathbf{S}\left(\alpha, \beta, n^{1/\alpha}\gamma, \delta_n; k\right) \tag{1.9}$$

where

$$\delta_n = \begin{cases} n\delta + \gamma\beta\tan\frac{\pi\alpha}{2}(n^{1/\alpha} - n) & k=0, \alpha\neq 1 \\ n\delta + \gamma\beta\frac{2}{\pi}n\log n & k=0, \alpha=1 \\ n\delta & k=1\ . \end{cases}$$

This is a restatement of Definition 1.5: the shape of the sum of n terms is the same as the original shape. We stress that no other distribution has this property.

With the above properties of linear combinations of stable random variables, we can characterize strict stability.

Proposition 1.18 *Let $X \sim \mathbf{S}(\alpha, \beta, \gamma, \delta_k; k)$ for $k = 0, 1$.*
(a) If $\alpha \neq 1$, then X is strictly stable if and only if $\delta_1 = \delta_0 - \beta\gamma\tan\frac{\pi\alpha}{2} = 0$.
(b) If $\alpha = 1$, then X is strictly stable if and only if $\beta = 0$.

Here there is an essential difference between the $\alpha = 1$ case and all other cases. When $\alpha = 1$, only the symmetric case is strictly stable and in that case the location parameter δ can be anything. In contrast, when $\alpha \neq 1$, any β can be strictly stable, as long as the location parameter is chosen correctly. This can be rephrased as follows: any stable distribution with $\alpha \neq 1$ can be made strictly stable by shifting; when $\alpha = 1$, a symmetric stable distribution with any shift is strictly stable and no shift can make a nonsymmetric 1-stable distribution strictly stable.

In addition to the basic properties described above, there are other linear and nonlinear properties of stable random variables given in Section 3.8.

1.7 Simulation

In this section U, U_1, U_2 will be used to denote independent Uniform(0,1) random variables. For a few special cases, there are simple ways to generate stable random variables.

For the normal case, Problem 1.15 shows

$$\begin{aligned} X_1 &= \mu + \sigma\sqrt{-2\log U_1}\cos 2\pi U_2 \\ X_2 &= \mu + \sigma\sqrt{-2\log U_1}\sin 2\pi U_2 \end{aligned} \tag{1.10}$$

give two independent $\mathbf{N}(\mu, \sigma^2)$ random variables. This is known as the Box-Muller algorithm.

For the Cauchy case, Problem 1.16 shows

$$X = \gamma \tan(\pi(U - 1/2)) + \delta \tag{1.11}$$

is Cauchy(γ, δ).

For the Lévy case, Problem 1.17 shows

$$X = \gamma \frac{1}{Z^2} + \delta \tag{1.12}$$

is Lévy(γ, δ) if $Z \sim \mathbf{N}(0,1)$.

In the general case, the following result of Chambers et al. (1976) gives a method for simulating any stable random variate.

Theorem 1.19 ***Simulating stable random variables*** *Let* Θ *and* W *be independent with* Θ *uniformly distributed on* $(-\frac{\pi}{2}, \frac{\pi}{2})$, *W exponentially distributed with mean 1,* $0 < \alpha \le 2$.
(a) The symmetric random variable

$$Z = \begin{cases} \dfrac{\sin \alpha\Theta}{(\cos\Theta)^{1/\alpha}} \left[\dfrac{\cos((\alpha-1)\Theta)}{W}\right]^{(1-\alpha)/\alpha} & \alpha \ne 1 \\ \tan\Theta & \alpha = 1. \end{cases}$$

has a $\mathbf{S}(\alpha, 0; 0) = \mathbf{S}(\alpha, 0; 1)$ *distribution.*
(b) In the nonsymmetric case, for any $-1 \le \beta \le 1$, *define* $\theta_0 = \arctan(\beta \tan(\pi\alpha/2))/\alpha$ *when* $\alpha \ne 1$. *Then*

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

has a $\mathbf{S}(\alpha, \beta; 1)$ *distribution.*

It is easy to get Θ and W from independent Uniform(0,1) random variables U_1 and U_2: set $\Theta = \pi(U_1 - \frac{1}{2})$ and $W = -\log U_2$. To simulate stable random variables with arbitrary shift and scale, (1.3) is used for the 0-parameterization and (1.5) is used for the 1-parameterization. Since there are numerical problems evaluating the expressions involved when α is near 1, the STABLE program uses an algebraic rearrangement of the formula. Section 3.3.3 gives a proof of this formula and a discussion of the numerical implementation of Chambers et al. (1976).

1.8 Generalized Central Limit Theorem and Domains of Attraction

The classical Central Limit Theorem says that the normalized sum of independent, identical terms with a finite variance converges to a normal distribution. To be more precise, let $X_1, X_2, X_3, \ldots$ be independent identically distributed random variables with mean μ and variance σ^2. The classical Central Limit Theorem states that the sample mean $\overline{X}_n = (X_1 + \cdots + X_n)/n$ will have

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} Z \sim \mathrm{N}(0,1) \text{ as } n \to \infty.$$

To match the notation in what follows, this can be rewritten as

$$a_n(X_1 + \cdots + X_n) - b_n \xrightarrow{d} Z \sim \mathrm{N}(0,1) \text{ as } n \to \infty, \tag{1.13}$$

where $a_n = 1/(\sigma\sqrt{n})$ and $b_n = \sqrt{n}\mu/\sigma$.

The Generalized Central Limit Theorem shows that if the finite variance assumption is dropped, the only possible resulting limits are stable.

Theorem 1.20 ***Generalized Central Limit Theorem*** *A nondegenerate random variable Z is α-stable for some $0 < \alpha \leq 2$ if and only if there is an independent, identically distributed sequence of random variables $X_1, X_2, X_3, \ldots$ and constants $a_n > 0$, $b_n \in \mathbb{R}$ with*

$$a_n(X_1 + \cdots + X_n) - b_n \xrightarrow{d} Z.$$

The following definition is useful in discussing convergence of normalized sums.

Definition 1.21 A random variable X is in the *domain of attraction* of Z if there exists constants $a_n > 0$, $b_n \in \mathbb{R}$ with

$$a_n(X_1 + \cdots + X_n) - b_n \xrightarrow{d} Z,$$

where $X_1, X_2, X_3, \ldots$ are independent identically distributed copies of X. $DA(Z)$ is the set of all random variables that are in the domain of attraction of Z.

Theorem 1.20 says that the only possible nondegenerate distributions with a domain of attraction are stable. Section 3.13 proves the Generalized Central Limit Theorem, characterizes the distributions in $DA(Z)$ in terms of their tail probabilities, and gives information about the norming constants a_n and b_n. For example, suppose X is a random variable with tail probabilities that satisfy $x^\alpha P(X > x) \to c^+$ and $x^\alpha P(X < -x) \to c^-$ as $x \to \infty$, with $c^+ + c^- > 0$ and $1 < \alpha < 2$. Then $\mu = EX$ must be finite and Theorem 3.58 shows that the analog of (1.13) is

$$a_n(X_1 + \cdots + X_n) - b_n \xrightarrow{d} Z \sim \mathbf{S}(\alpha, \beta, 1, 0; 1) \text{ as } n \to \infty,$$

when $a_n = \left((2\Gamma(\alpha)\sin(\frac{\pi\alpha}{2}))/(\pi(c^+ + c^-))\right)^{1/\alpha} n^{-1/\alpha}$, $b_n = na_n\mu$ and $\beta = (c^+ - c^-)/(c^+ + c^-)$. In this case, the rate at which of the tail probabilities of X decay determines the index α and the relative weights of the right and left tail determine the skewness β.

1.9 Problems

Problem 1.1 Show directly using the convolution formula (A.1) that the normal distributions are stable. Show that $a^2+b^2=c^2$ in (1.1), so $\alpha=2$. Conclude that $\mathbf{N}(\mu,\sigma^2) = \mathbf{S}\left(2,0,\sigma/\sqrt{2},\mu;0\right) = \mathbf{S}\left(2,0,\sigma/\sqrt{2},\mu;1\right)$ and $\mathbf{S}(2,0;\gamma,\delta;0) = \mathbf{S}(2,0;\gamma,\delta;1) = \mathbf{N}(\delta,2\gamma^2)$.

Problem 1.2 Show directly using the convolution formula that the Cauchy distributions are stable. Show that $a+b=c$ in (1.1), so $\alpha=1$ and conclude that Cauchy$(\gamma,\delta) = \mathbf{S}(1,0,\gamma,\delta;0) = \mathbf{S}(1,0,\gamma,\delta;1)$.

Problem 1.3 Show that the cumulative distribution function of a Cauchy distribution is $F(x|1,0,\gamma,\delta;0) = F(x|1,0,\gamma,\delta;1) = (1/2)+\arctan((x-\delta)/\gamma)/\pi$.

Problem 1.4 Show directly using the convolution formula that Lévy distributions are stable. Show that $a^{1/2}+b^{1/2}=c^{1/2}$ in (1.1), so $\alpha=1/2$ and conclude that Lévy$(\gamma,\delta) = \mathbf{S}(1/2,1,\gamma,\delta;1) = \mathbf{S}(1/2,1,\gamma,\delta+\gamma;0)$.

Problem 1.5 Show that the cumulative distribution function of a Lévy distribution $X \sim \mathbf{S}(1/2,1,\gamma,\delta;1)$ is, for $x>\delta$

$$F(x|1/2,1,\gamma,\delta;1) = 2\left(1-\Phi\left(\sqrt{\gamma/(x-\delta)}\right)\right),$$

where $\Phi(x)$ is the d.f. of a standard normal distribution.

Problem 1.6 What is wrong with the following argument? If $X_1,\ldots,X_n \sim$Gamma(α,β) are independent, then $X_1+\cdots+X_n \sim$Gamma$(n\alpha,\beta)$, so gamma distributions must be stable distributions.

Problem 1.7 Use the characteristic function (1.2) to show that $Z(\alpha,-\beta)\stackrel{d}{=}-Z(\alpha,\beta)$. This proves Proposition 1.11.

Problem 1.8 Use the definitions of the different parameterizations and the characteristic function (1.2) to show that the characteristic functions in (1.4) and (1.6) are correct.

Problem 1.9 Show that the conversions between the parameterizations in (1.7) are correct. (Use either the characteristic functions in (1.4) and (1.6) or the definitions of the parameterizations in terms of $Z(\alpha,\beta)$.)

Problem 1.10 A Pareto(α,c) ($\alpha>0$ and $c>0$) distribution has density $f(x)=\alpha c^{\alpha}x^{-(1+\alpha)}$, $x>c$. Show that if $p<\alpha$, then EX^p exists and find its value, but if $p\geq\alpha$, then $EX^p=\infty$.

Problem 1.11 Extend the previous problem to show that if X is any random variable with a bounded density for which both left and right tail densities are asymptotically equivalent to Pareto(α,c), then $E|X|^p$ is finite if $p<\alpha$ and infinite if $p\geq\alpha$. (The left tail is defined to be asymptotically Pareto if $f(x)\sim\alpha c^{\alpha}|x|^{-(1+\alpha)}$ as $x\to-\infty$.)

Problem 1.12 Show that the sum of two independent stable random variables with different αs is not stable. Section 13.9 gives a brief discussion of what happens when you combine different indices of stability.

Problem 1.13 Derive (1.8) and (1.9) for the sums of independent α-stable r.v.

Problem 1.14 Simulate $n = 1,000$ uniform random variables and let s_k^2 be the sample variance of the first k values. A "running sample variance" plot is a graph of (k, s_k^2), $k = 2, 3, 4, \ldots, n$. Repeat the process with normal random variables, Cauchy random variables and Pareto random variables (see Section 13.1 for a method of simulating Pareto distributions) with $\alpha = 0.5, 1, 1.5$. Contrast the behavior of s_k^2.

Problem 1.15 Show directly that (1.10) gives independent $\mathbf{N}(\mu, \sigma^2)$. Theorem 1.19 also works when $\alpha = 2$ to generate normal random variates, but it requires two uniforms to generate one normal, whereas (1.10) generates two normals from two uniforms.

Problem 1.16 Use the cumulative distribution function for a Cauchy(γ, δ) distribution from Problem 1.3 to prove (1.11).

Problem 1.17 Use Problem 1.5 to prove (1.12).