Stable Distributions
Models for Heavy Tailed Data

John P. Nolan
jpnolan@american.edu
Math/Stat Department
American University
Copyright ©2004 John P. Nolan

Processed January 11, 2005
### 2.7.5 Multiple assets .................................. 34
### 2.7.6 Gumbel copula .................................. 35
### 2.8 Exponential power distributions ....................... 35
### 2.9 Signal processing .................................. 36
### 2.9.1 Independent component analysis ................... 36
### 2.9.2 Image processing ................................ 36
### 2.9.3 Queueing theory ................................ 36
### 2.9.4 Physics ................................ 37
### 2.9.5 Survival analysis and Reliability ................... 37
### 2.9.6 Wrapped stable distribution ..................... 37
### 2.9.7 Network traffic ................................ 38
### 2.9.8 Computer Science ................................ 38
### 2.9.9 Biology and medicine ........................... 39
### 2.9.10 Discrepancies ................................ 39
### 2.9.11 Punctuated change ............................ 39
### 2.9.12 Central Pre-Limit Theorem ...................... 39
### 2.10 Miscellaneous applications/Historical notes ................. 39
### 2.11 Miscellaneous .................................. 39
### 2.12 Appropriateness of infinite variance models ................. 40
### 2.13 Problems .................................. 41

### 3 Technical Results on Univariate Stable Distributions 45
### 3.1 Proofs of Basic Theorems of Chapter 1 ................... 45
### 3.2 Densities and distribution functions .................... 55
### 3.2.1 Series expansions ................................ 63
### 3.2.2 Modes ................................ 64
### 3.2.3 Duality ................................ 69
### 3.3 Numerical algorithms ................................ 70
### 3.3.1 Computation of distribution functions and densities ....... 70
### 3.3.2 Spline approximation of densities .................. 73
### 3.3.3 Simulation ................................ 73
### 3.4 More on parameterizations ........................... 75
### 3.5 Tail behavior .................................. 79
### 3.6 Moments and other transforms ........................ 91
### 3.7 Convergence of stable laws in terms of \((\alpha, \beta, \gamma, \delta)\) .............. 96
### 3.8 Combinations of stable random variables ................ 98
### 3.9 Exponential-stable, log-stable and discrete stable distributions .... 105
### 3.10 Stable distributions arising as functions of other distributions ........ 106
### 3.11 Stochastic series representations .................... 108
### 3.12 Generalized Central Limit Theorem and Domains of Attraction .......... 109
### 3.13 Central Pre-Limit Theorem .......................... 117
### 3.14 Differential equations and stable semi-groups ............... 117
### 3.15 Problems .................................. 119

### 4 Univariate Estimation 125
### 4.1 Order statistics .................................. 125
Contents

4.2 Tail based estimation ........................................... 126
   4.2.1 Hill estimator ........................................... 126
4.3 Extreme value theory estimate of $\alpha$ ..................... 129
4.4 Quantile based estimation ..................................... 130
4.5 Characteristic function based estimation ...................... 133
4.6 Moment based methods of estimation ......................... 135
4.7 Maximum likelihood estimation ................................ 136
   4.7.1 Asymptotic normality and Fisher information matrix ...... 137
   4.7.2 The score function ....................................... 141
4.8 Other methods of estimation .................................. 141
4.9 Comparisons of estimators ................................... 143
4.10 Assessing a stable fit ....................................... 143
   4.10.1 Likelihood ratio tests and goodness-of-fit tests .......... 144
   4.10.2 Testing the stability hypothesis ........................ 144
   4.10.3 Diagnostics ............................................ 144
4.11 Applications ................................................. 147
4.12 Fitting stable distributions to concentration data .......... 153
4.13 Estimation for discrete stable distributions .................. 153
4.14 Discussion .................................................. 153
4.15 Problems ................................................... 153

II Multivariate Stable ............................................. 155

5 Basic Properties of Multivariate Stable Distributions ......... 157
   5.1 Definition of jointly stable ................................ 157
   5.2 Parameterizations .......................................... 160
      5.2.1 Projection based description .......................... 161
      5.2.2 Spectral measures ...................................... 162
      5.2.3 Stable stochastic integrals ............................ 164
      5.2.4 Stochastic series representation ........................ 164
   5.3 Multivariate stable densities and probabilities .......... 165
      5.3.1 Multivariate tail probabilities ........................ 166
   5.4 Sums of stable random vectors - independent and dependent 167
   5.5 Classes of multivariate stable distributions ............... 168
      5.5.1 Independent components ............................... 168
      5.5.2 Discrete spectral measures ............................. 169
      5.5.3 Radially and elliptically contoured stable laws ....... 170
      5.5.4 Sub-stable laws ....................................... 171
      5.5.5 Linear combinations ................................. 171
   5.6 Multivariate generalized central limit theorem ............ 171
   5.7 Simulation .................................................. 171
   5.8 misc .................................................. 172
   5.9 Problems .................................................. 172

6 Technical Results on Multivariate Stable Distributions ....... 175
6.1 Proofs of basic properties of multivariate stable distributions ............................... 175
6.2 Parameterizations .................................................................................................................. 177
6.3 Stable Stochastic Integrals and Series .................................................................................. 179
6.4 Multivariate stable densities and probabilities ..................................................................... 179
   6.4.1 Duality ......................................................................................................................... 185
   6.4.2 Tail probabilities ........................................................................................................... 186
6.5 Radially symmetric and elliptically contoured stable distributions ...................................... 188
   6.5.1 The amplitude distribution .......................................................................................... 189
   6.5.2 Densities of isotropic and elliptically contoured distributions ...................................... 192
6.6 Complex stable random variables ....................................................................................... 194
6.7 Order statistics with dependent components ....................................................................... 195
6.8 Multivariate domains of attraction ..................................................................................... 196
6.9 Problems .............................................................................................................................. 196

7 Multivariate Estimation ........................................................................................................... 197
   7.1 Tail estimators .................................................................................................................... 197
      7.1.1 Rachev-Xin-Cheng method ....................................................................................... 197
      7.1.2 Davydov-Nagaev-Paulauskas method ....................................................................... 198
   7.2 Sample characteristic function method .............................................................................. 198
   7.3 Projection method .............................................................................................................. 200
   7.4 Estimation for radially symmetric stable distributions ...................................................... 202
      7.4.1 Maximum likelihood estimation .................................................................................. 202
      7.4.2 Moment based estimation .......................................................................................... 202
      7.4.3 Estimation based on log $R$ ......................................................................................... 203
      7.4.4 Quantile based estimation .......................................................................................... 203
      7.4.5 Characteristic function based estimation ...................................................................... 203
   7.5 Estimation for elliptically contoured distributions ............................................................. 204
   7.6 Diagnostics ........................................................................................................................ 204
   7.7 Examples ............................................................................................................................ 204
   7.8 Examples ............................................................................................................................ 204
   7.9 Discussion .......................................................................................................................... 206

8 Dependence and Metrics ......................................................................................................... 207
   8.1 Conditional distributions and moments ............................................................................... 207
   8.2 Dependence measures ....................................................................................................... 208
      8.2.1 Covariation ............................................................................................................... 208
      8.2.2 Codifference .............................................................................................................. 208
      8.2.3 Other measures ......................................................................................................... 208
   8.3 Preliminaries ....................................................................................................................... 209
   8.4 Measures of dependence .................................................................................................... 212
      8.4.1 Comparisons to roundness ....................................................................................... 213
      8.4.2 Comparisons to independence ................................................................................. 213
   8.5 Metrics ............................................................................................................................... 215
   8.6 Symmetric case ................................................................................................................... 217
   8.7 Non-symmetric case ......................................................................................................... 219
   8.8 Closeness in terms of spectral measures ............................................................................ 222
8.9 Proofs of technical lemmas ........................................... 225
8.10 Problems .................................................................. 230

III Stable Regression, Time Series and Processes ............. 231

9 Stable Regression .......................................................... 233

10 Stable Time Series ......................................................... 235

11 Stable Processes ............................................................ 237
  11.1 Definitions .................................................................. 237
  11.2 Classes of stable processes ........................................... 238
    11.2.1 Stable sequences .................................................. 238
    11.2.2 Stable Lévy processes and stable noise ....................... 238
    11.2.3 Stochastic integral representation of stable processes ......... 239
    11.2.4 Self-similar processes .............................................. 239
    11.2.5 Harmonizable processes .......................................... 239
    11.2.6 Moving averages ................................................... 239
    11.2.7 Sub-Gaussian processes .......................................... 239
    11.2.8 Fractional stable motions ......................................... 239
  11.3 Dependence .............................................................. 239
  11.4 Stable stochastic processes ........................................... 239
  11.5 Path properties .......................................................... 240
    11.5.1 Regularity ........................................................... 240
    11.5.2 Erraticism ........................................................... 240
  11.6 Model identification ..................................................... 240
  11.7 Constructing the spectral measure from a stochastic integral 242
    11.7.1 Discrete Spectral Measures ....................................... 243
    11.7.2 Independent increment processes .............................. 244
    11.7.3 Moving average processes ....................................... 244
    11.7.4 Harmonizable processes .......................................... 245
    11.7.5 Substable processes .............................................. 246
  11.8 Constructing a stochastic integral from the spectral measure 248
  11.9 Ergodicity and mixing .................................................. 250
  11.10 Simulating stable processes ......................................... 250
  11.11 Miscellaneous .......................................................... 250

IV Related Topics .............................................................. 251

12 Related Distributions ...................................................... 253
  12.1 Pareto distributions ..................................................... 253
  12.2 t distributions .......................................................... 257
  12.3 Laplace distributions .................................................... 257
  12.4 Geometric stable distributions ....................................... 258
12.5 Max-stable, min-stable, multiplication-stable, and s-stable distributions, median-stable .....................................258
12.6 Discrete stable and Linnik distributions ..........................................................259
12.7 Semi-stable ..................................................................................................259
12.8 Marginal stable and operator stable distributions ........................................259
12.9 Mixtures and convolutions of normals .........................................................260
12.10 Univariate mixtures and convolutions of stable distributions ............260
12.11 Infinitely divisible distributions .................................................................261
12.12 Misc. ...........................................................................................................261
12.13 Problems .....................................................................................................261

V Appendices ..................................................................................................263

A Mathematical Facts 265
A.1 Sums of random variables .................................................................265
A.2 Symmetric random variables .............................................................265
A.3 Moments .................................................................................................266
A.4 Characteristic functions ........................................................................266
A.5 Laplace transforms ................................................................................267
A.6 Mellin transforms ....................................................................................268
A.7 Spheres, balls and polar coordinates in $\mathbb{R}^d$ ....................................268
A.8 Gamma function .......................................................................................269

B Stable Quantiles 273

C Stable Modes 275

D Asymptotic standard deviations 279

Index 297

Author Index 300

Symbol Index 305
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 ?? 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 7 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 Theo-
rem 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 modelling 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 modelling 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 $b$, $aX_1 + bX_2 \overset{d}{=} cX + d$, (1.1)

for some positive $c$ and some $d \in \mathbb{R}$. (The symbol $\overset{d}{=} \text{ means equality in distribution, i.e. both expressions have the same probability law.}$) 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. $X \overset{d}{=} -X$.

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 $X \sim \mathcal{N}(\mu, \sigma^2)$, then the terms on the left hand side above are $\mathcal{N}(a\mu, (a\sigma)^2)$ and $\mathcal{N}(b\mu, (b\sigma)^2)$ respectively, while the right hand side is $\mathcal{N}(c\mu + d, (c\sigma)^2)$. 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
1.1 Definition of stable and geometric stable distributions (see Chapter 12). 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 exists constants \(A > 0\) and \(B \in \mathbb{R}\) with \(X \overset{d}{=} AY + B\). 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.

**Example 1.2** Normal or Gaussian distributions. \(X \sim \mathcal{N}(\mu, \sigma^2)\) 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 is \(F(x) = P(X \leq x) = \Phi((x-\mu)/\sigma)\), where \(\Phi(z)\) = probability that a standard normal r.v. is less than or equal \(z\).

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

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

For the d.f. of a Cauchy distribution, see Problem 1.4.

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

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

For the d.f. of a Lévy distribution, see Problem 1.17.

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) more than 100 times as many 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. 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
1. Basic Properties of Univariate Stable Distributions

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

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

<table>
<thead>
<tr>
<th>( c )</th>
<th>( P(X &gt; c) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Cauchy</td>
</tr>
<tr>
<td>0</td>
<td>0.5000</td>
</tr>
<tr>
<td>1</td>
<td>0.1587</td>
</tr>
<tr>
<td>2</td>
<td>0.0228</td>
</tr>
<tr>
<td>3</td>
<td>0.001347</td>
</tr>
<tr>
<td>4</td>
<td>0.00003167</td>
</tr>
<tr>
<td>5</td>
<td>0.0000002866</td>
</tr>
</tbody>
</table>

There are other equivalent definitions of stable random variables. Two are stated here, the proof 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

\[
X_1 + \cdots + X_n \overset{d}{=} c_n X + d_n,
\]

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.
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 way of parameterizing stable distributions. The most concrete way to describe all possible Limit Theorem, Theorem 1.20. While useful, these conditions do not give a concrete properties of $X$.)

Problems 1.1, 1.2 and 1.3 show that a $\mathcal{N}(\mu, \sigma^2)$ 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$(\gamma, \delta)$ 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 $\gamma > 0$ and a location parameter $\delta \in \mathbb{R}$. We will use $\gamma$ for the scale parameter and $\delta$ for the location parameter to avoid confusion with the symbols $\sigma$ and $\mu$, 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 $\delta$ when it simplifies the statement of a result. Since $\alpha$ and $\beta$ 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 $S_d(\sigma, \beta, \mu)$ is used for the class of stable laws. We will use a modified notation of the form $S(\alpha, \beta, \gamma; \delta; k)$ for three reasons. First, the usual notation singles out $\alpha$ 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 $\gamma$ for the scale (not $\sigma$) and $\delta$ for the location (not $\mu$). 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 $S(\alpha, \beta, \gamma; \delta; 0)$ if

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

(1.3)

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

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

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

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

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

(1.5)

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

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

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

Note that if $\beta = 0$, then these two parameterizations are identical, it is only when $\beta \neq 0$ that the asymmetry factor (the imaginary term) becomes an issue. The symbol $S_{0S}$ is used as an abbreviation for symmetric $\alpha$-stable. When a scale parameter is used, $S_{0S}(\gamma) = S(\alpha, 0, \gamma, 0; 0) = S(\alpha, 0, \gamma, 0; 1)$.

When $\alpha = 2$, a $S(2, 0, \gamma, \delta; 0) = S(2, 0, \gamma, \delta; 1)$ distribution is normal with mean $\delta$, but the variance is not $\gamma$. Because of the way the characteristic function is defined above, it is actually $S(2, 0, \gamma, \delta; 0) = \mathcal{N}(\delta, 2\gamma^2)$. This fact is a frequent source of confusion when
one tries to compare, e.g. stable quantiles to normal quantiles. It is an unfortunate fact; the 2-parameterization and 3-parameterizations defined in Section 3.4 rescale to avoid this problem, but the above scaling is standard in the literature.

The different parameterizations have caused considerable confusion. Hall (1981a) describes a “comedy of errors” caused by parameterization choices. The most common mistake concerns the sign of the skewness parameter when $\alpha = 1$. The present author was confused by the integral formulas for the density given in Theorem 2.2.3 of Zolotarev (1986) which are stated in yet a different parameterization (Zolotarev’s (B)). Another example is the stable random number generator of Chambers et al. (1976) which has two arguments: $\alpha$ and $\beta$. Most users expect to get a $\mathcal{S}(\alpha, \beta; 1)$ result, however, the routine actually returns random variates with a $\mathcal{S}(\alpha, \beta; 0)$ 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 $\mathcal{S}(\alpha, \beta, \gamma, \delta; 0)$ 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 $\alpha$ and $\beta$ determine the shape of the distribution, while $\gamma$ and $\delta$ determine scale and location in the standard way: if $X \sim \mathcal{S}(\alpha, \beta, \gamma, \delta; 0)$, then $(X - \delta)/\gamma \sim \mathcal{S}(\alpha, \beta, 1, 0; 0)$. This is not true for the $\mathcal{S}(\alpha, \beta, \gamma, \delta; 1)$ 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 $\mathcal{S}(\alpha, \beta, \gamma, \delta; 1)$ 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 $\mathcal{S}(\alpha, \beta, \gamma, \delta; 1)$ parameterization is that the location of the mode is un-
bounded in any neighborhood of $\alpha = 1$: if $X \sim S(\alpha, \beta, \gamma, \delta; 1)$ 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 $S(\alpha, \beta, \gamma, \delta; 1)$ 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.

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 $S(\alpha, \beta, \gamma, \delta; 2)$ 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. 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. We will generally use the $S(\alpha, \beta, \gamma, \delta; 0)$ and $S(\alpha, \beta, \gamma, \delta; 1)$ parameterizations in what follows to avoid (or at least limit) confusion. In these two parameterizations, $\alpha$, $\beta$ and the scale $\gamma$ are always the same, but the location parameters will have different values. The notation $X \sim S(\alpha, \beta, \gamma, \delta_0; k)$ for $k = 0, 1$ will be shorthand for $X \sim S(\alpha, \beta, \gamma, \delta_0; 0)$ and $X \sim S(\alpha, \beta, \gamma, \delta_1; 1)$ simultaneously. In this case, the parameters are related by (see Problem 1.9)

\[
\delta_0 = \begin{cases} 
\delta_1 + \beta \gamma \tan \frac{\pi}{2} & \alpha \neq 1 \\
\delta_1 + \frac{\beta \gamma}{\log \gamma} & \alpha = 1
\end{cases} \\
\delta_1 = \begin{cases} 
\delta_0 - \beta \gamma \tan \frac{\pi}{2} & \alpha \neq 1 \\
\delta_0 - \frac{\beta \gamma}{\log \gamma} & \alpha = 1
\end{cases}
\] (1.7)
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 \( S(\alpha, \beta, \gamma, \delta; k) \) 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 \( Z \sim S(\alpha, \beta; 0) \), 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.

![Stable densities graph](image-url)
Lemma 1.10 The support of a stable distribution in the different parameterizations is

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

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

The constant \(\tan \frac{\pi \alpha}{2}\) appears frequently when working with stable distributions, so it is worth recording its behavior. Figure 1.5 shows that 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 \(\alpha\) and \(\beta\), \(Z \sim S(\alpha, \beta; k)\), \(k = 0, 1, 2\)

\[Z(\alpha, -\beta) \overset{d}{=} -Z(\alpha, \beta).
\]

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

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 1.6 shows the bell-shaped density of symmetric stable distributions. As \(\alpha\) 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 $\alpha$ 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 $\beta$. In symbols, $Z(2, \beta) \overset{d}{=} Z(2, 0)$. In general, as $\alpha \uparrow 2$, all stable distributions get closer and closer to being symmetric and $\beta$ becomes less meaningful in applications (and harder to estimate accurately).

Figure 1.7 shows the density and d.f. when $\alpha = 1.9$, with varying $\beta$, and there is little visible difference as $\beta$ varies. As $\alpha$ decreases, the effect of $\beta$ becomes more pronounced: the left tail gets lighter and lighter for $\beta > 0$, see Figure 1.8 ($\alpha = 1.3$), Figure 1.9 ($\alpha = 0.7$), and Figure 1.10 ($\alpha = 0.1$). The last figure shows that when $\alpha$ 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 $|\beta| = 1$.

Finally, all stable densities are unimodal, but there is no known formula for the location of the mode. However, the mode of a $Z \sim S(\alpha, \beta; 0)$ distribution, denoted by $m(\alpha, \beta)$, has been numerically computed. The values of $m(\alpha, \beta)$ are shown for $\beta \geq 0$ in Figure 1.11 and a table of modes is given in Appendix C. By the reflection property, $m(\alpha, -\beta) = -m(\alpha, \beta)$. 
1. Basic Properties of Univariate Stable Distributions

Figure 1.7: Stable densities and cumulative distribution functions for $Z \sim S(1.9, \beta; 0)$, $\beta = 0, 0.5, 1$.

Numerically, it is also observed that $P(Z > m(\alpha, \beta)) > P(Z < m(\alpha, \beta))$ 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 $\beta < 0$. 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. The tail probabilities in the non-Gaussian cases are known asymptotically. The statement $h(x) \sim g(x)$ as $x \to a$ will mean $\lim_{x \to a} h(x)/g(x) = 1$.

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

\[
P(X > x) \sim \gamma^\beta c_\alpha (1 + \beta) x^{-\alpha} \\
f(x|\alpha, \beta, \gamma, \delta; 0) \sim \alpha \gamma^\beta c_\alpha (1 + \beta) x^{-(\alpha + 1)}
\]

where $c_\alpha = \sin(\pi/\alpha) \Gamma(\alpha)/\pi$.

Using the reflection property, the lower tail properties are similar: for $-1 \leq \beta < 1$, $P(X < -x) \sim \gamma^\beta c_\alpha (1 - \beta) x^{-\alpha}$ and $f(-x|\alpha, \beta, \gamma, \delta; 0) \sim \alpha \gamma^\beta c_\alpha (1 - \beta) x^{-(\alpha + 1)}$. For all $\alpha < 2$ and $\beta > -1$, the upper tail probabilities and densities are asymptotically power laws. (When $\beta = -1$, the right hand sides in Theorem 1.12 are 0. In these cases, the tail probabilities are not asymptotically power laws, see Section 3.5 for more information.) Pareto distributions
Figure 1.8: Stable densities and cumulative distribution functions for $Z \sim S(1.3, \beta; 0)$, $\beta = 0, 0.5, 1$.

(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.

A general distribution is said to be heavy tailed if its tails are heavier than exponential. 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.

One consequence of heavy tails is that not all moments exist. In most statistical problems, the first moment $E[X]$ and variance $\text{Var}(X) = E(X^2) - (E[X])^2$ 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 $p \geq \alpha$. Explicit formulas for moments of strictly stable laws are given in Section 3.6.

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 2.12 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 $X \sim S(\alpha, \beta, \gamma, \delta; k)$ for $k = 0, 1$ is

$$\mu = EX = \delta_1 = \delta_0 - \beta \gamma \tan \frac{\pi}{\alpha}.$$
Consider what happens to the mean of $X \sim S(\alpha, \beta; 0)$ as $\alpha \downarrow 1$. Even though the mode of the distribution stays close to 0, it has a mean of $\mu = \beta \tan \frac{\pi}{\alpha}$. 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 $\alpha$ reaches 1, the tails are too heavy for the integral $EX = \int_{-\infty}^{\infty} xf(x)\,dx$ to converge. In contrast, a $S(\alpha, \beta; 1)$ 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 3.13 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 $S(\alpha, \beta; 2)$ distribution keeps the mode exactly at 0, and the mean as a function of $(\alpha, \beta)$ is continuous, like the mean of a $S(\alpha, \beta; 0)$ distribution.

Note that the skewness parameter $\beta$ is not the same thing as the classical skewness parameter. The latter is undefined for every non-Gaussian stable distribution because neither the third moment or the variance exist. Likewise, the kurtosis is undefined, because the fourth moment is undefined for every non-Gaussian stable distribution.

Tables of standard normal quantiles or percentiles are given in most basic probability and statistic books. Let $z_\lambda$ be the $\lambda^{th}$ quantile, i.e. the $z$ value for which the standard normal distribution has lower tail probability $\lambda$, i.e. $P(Z < z_\lambda) = \lambda$. The value $z_{0.975} = 1.96$ is commonly used: for $X \sim N(\mu, \sigma^2)$, 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.
1.5 Tail probabilities, moments and quantiles

Quantiles of the standard stable distributions are used in the same way. The difficulty is that there are different quantiles for every value of \( \alpha \) and \( \beta \). The symbol \( z_\lambda(\alpha, \beta) \) will be used for the \( \lambda \)th quantile of a \( S(\alpha, \beta; 0) \) 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 \( \alpha \) and \( \beta \) values. Appendix B shows selected quantiles for \( \alpha = 0.1, 0.2, \ldots, 1.9, 1.95, 1.99, 2.0 \) and \( \beta = 0, 0.1, 0.2, \ldots, 0.9, 1 \). (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 symmetric. Second, the way the quantiles scale depend on what parameterization is being used. In the \( S(\alpha, \beta, \gamma, \delta; 0) \) parameterization, it is straightforward; in other parameterizations one has to either convert to the \( S(\alpha, \beta, \gamma, \delta; 0) \) 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 5th and 95th quantiles for \( X \sim S(1.3, 0.5, 2, 7; 0) \). From Appendix B, the 5th quantile is \( z_{0.05}(1.3, 0.5) = -2.355 \) and the 95th quantile is \( z_{0.95}(1.3, 0.5) = +5.333 \) for a standardized \( S(1.3, 0.5, 1, 0; 0) \) 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 \( X \sim S(1.3, 0.5, 2, 7; 1) \), then using the previous example, the 5th and 95th quantiles are \( \gamma(-2.355) + (\delta + \beta \tan \varphi) = 0.327 \) and \( \gamma(5.333) + (\delta + \beta \tan \varphi) = 15.704 \). Alternatively, \( S(1.3, 0.5, 2, 7; 1) = S(1.3, 0.5, 2, 5.037; 0) \) by (1.7), so the 5th and 95th 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 $\alpha$-stable random variables are $\alpha$-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 $\alpha$, as Problem 1.12 shows that otherwise the sum will not be stable. Section 12.10 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 Chapter 5.

**Proposition 1.16** The $S(\alpha, \beta, \gamma, \delta; 0)$ parameterization has the following properties.

(a) If $X \sim S(\alpha, \beta, \gamma, \delta; 0)$, then for any $a \neq 0$, $b \in \mathbb{R}$,

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

(b) The characteristic functions, densities and distribution functions are jointly continuous in all four parameters $(\alpha, \beta, \gamma, \delta)$.

(c) If $X_1 \sim S(\alpha, \beta_1, \gamma_1, \delta_1; 0)$ and $X_2 \sim S(\alpha, \beta_2, \gamma_2, \delta_2; 0)$ are independent, then $X_1 + X_2 \sim S(\alpha, \beta, \gamma, \delta; 0)$ where

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

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

Figure 1.11: The location of the mode of a $S(\alpha, \beta; 0)$ density.
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 $\alpha^{th}$ power of the scale parameters, not the scale parameters themselves.

**Proposition 1.17** The $S(\alpha, \beta, \gamma, \delta; 1)$ parameterization has the following properties.

(a) If $X \sim S(\alpha, \beta, \gamma, \delta; 1)$, then for any $a \neq 0$, $b \in \mathbb{R}$,

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

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

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

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

The corresponding results for the $S(\alpha, \beta, \gamma, \delta; 2)$ parameterization are given in Proposition 3.38.

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

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

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

where

$$\gamma^\alpha = \sum_{j=1}^n |w_j\gamma_j|^\alpha$$

$$\beta = \frac{\sum_{j=1}^n \beta_j \text{sign} w_j |w_j\gamma_j|^\alpha}{\gamma^\alpha}$$

$$\delta = \begin{cases} \sum_j w_j \delta_j + \tan \frac{\pi}{2} (\beta \gamma - \sum_j \beta_j w_j \gamma_j) & k = 0, \alpha \neq 1 \\ \sum_j w_j \delta_j + \frac{\alpha}{\epsilon} (\beta \gamma \log \gamma - \sum_j \beta_j w_j \gamma_j \log |w_j\gamma_j|) & k = 0, \alpha = 1 \\ \sum_j w_j \delta_j - \frac{\alpha}{\epsilon} \sum_j \beta_j w_j \gamma_j \log |w_j| & k = 1, \alpha \neq 1 \\ \sum_j w_j \delta_j - \frac{\alpha}{\epsilon} \sum_j \beta_j w_j \gamma_j \log |w_j| & k = 1, \alpha = 1. \end{cases}$$

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 $X_j \sim S(\alpha, \beta, \gamma, \delta; k)$, then

$$X_1 + \cdots + X_n \sim S\left(\alpha, \beta, n^{1/\alpha} \gamma, \delta_n; k\right)$$ (1.9)
where
\[
\delta_n = \begin{cases} 
n\delta + \gamma \beta \tan \frac{\pi}{2} (n^{1/\alpha} - n) & k = 0, \alpha \neq 1 
n\delta + \gamma \beta n \log n & k = 0, \alpha = 1 
n\delta & k = 1.
\end{cases}
\]

This is a restatement of Definition 1.5: the shape of the sum 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 \( X \sim S(\alpha, \beta, \gamma, \delta; k) \) for \( k = 0, 1 \).

(a) If \( \alpha \neq 1 \), then \( X \) is strictly stable if and only if \( \delta_1 = \delta_0 - \beta \gamma \tan \frac{\pi}{2} = 0 \).

(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 \( \delta \) can be anything. In the \( \alpha \neq 1 \) cases, any \( \beta \) 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
\[
\begin{align*}
X_1 &= \mu + \sigma \sqrt{-2 \log U_1} \cos 2\pi U_2 \\
X_2 &= \mu + \sigma \sqrt{-2 \log U_1} \sin 2\pi U_2
\end{align*}
\] (1.10)
give two independent \( \mathcal{N}(\mu, \sigma^2) \) random variables. This is known as the Box-Muller algorithm.

For the Cauchy case, Problem 1.16 shows
\[
X = \gamma \tan(\pi(U - 1/2)) + \delta
\] (1.11)
is Cauchy(\( \gamma, \delta \)).

For the Lévy case, Problem 1.17 shows
\[
X = \gamma \frac{1}{Z^2} + \delta
\] (1.12)
is Lévy(\( \gamma, \delta \)) if \( Z \sim \mathcal{N}(0,1) \).

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 $\Theta$ and $W$ be independent with $\Theta$ uniformly distributed on $(-\frac{\pi}{2}, \frac{\pi}{2})$, $W$ exponentially distributed with mean 1, $0 < \alpha \leq 2$.
(a) The symmetric random variable
\[ Z = \begin{cases} \frac{\sin \alpha \Theta}{(\cos \Theta)^{1/\alpha}} \left[ \frac{\cos(\alpha - 1) \Theta}{W} \right]^{(1-\alpha)/\alpha} & \alpha \neq 1 \\ \tan \Theta & \alpha = 1. \end{cases} \]
has a $S(\alpha, 0; 0) = S(\alpha, 0; 1)$ distribution.
(b) In the nonsymmetric case, for any $-1 \leq \beta \leq 1$, define $\theta_0 = \arctan(\beta \tan(\pi \alpha / 2)) / \alpha$ when $\alpha \neq 1$. Then
\[ Z = \begin{cases} \frac{\sin \alpha(\theta_0 + \Theta)}{(\cos \alpha \theta_0 \cos \Theta)^{1/\alpha}} \left[ \frac{\cos(\alpha \theta_0 + (\alpha - 1) \Theta)}{W} \right]^{(1-\alpha)/\alpha} & \alpha \neq 1 \\ \frac{\pi}{\alpha} \frac{(\beta + \Theta) \tan \Theta - \beta \log \left( \frac{\pi}{2} W \cos \frac{\pi}{2} + \beta \Theta \right)}{\alpha = 1}. \end{cases} \]
has a $S(\alpha, \beta; 1)$ distribution.

It is easy to get $\Theta$ 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 $\alpha$ 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, Mallows and Stuck.

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 $\mu$ and variance $\sigma^2$. The classical Central Limit Theorem states that the sample mean $\overline{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$ will have
\[ \frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}} \xrightarrow{d} Z \sim N(0, 1) \text{ as } n \to \infty. \]

To match the notation in what follows, this can be rewritten as
\[ a_n(X_1 + \cdots + X_n) - b_n d \xrightarrow{d} Z \sim N(0, 1) \text{ as } n \to \infty, \quad (1.13) \]
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.
1. Basic Properties of Univariate Stable Distributions

**Theorem 1.20** Generalized Central Limit Theorem

Let $X_1, X_2, X_3, \ldots$ be an independent, identically distributed sequence of random variables. There exist constants $a_n > 0, b_n \in \mathbb{R}$ and a nondegenerate random variable $Z$ with

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

if and only if $Z$ is $\alpha$-stable for some $0 < \alpha \leq 2$.

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 and only 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 distributions with a domain of attraction are stable. Section 3.12 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 3.50 shows that the analog of (1.13) is

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

as $n \to \infty$,

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

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 $\alpha = 2$ and conclude that $N(\mu, \sigma^2) = S\left(2, 0, \sigma/\sqrt{2}, \mu; 0\right)$.

**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($\gamma, \delta$) = $S\left(1, 0, \gamma, \delta; 0\right)$.

**Problem 1.3** 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($\gamma, \delta$) = $S\left(1/2, 1, \gamma, \delta; 1\right)$.

**Problem 1.4** Show that the cumulative distribution function of a Cauchy distribution is $F(x|1, 0, \gamma, \delta; 0) = F(x|1, 0, \gamma, \delta; 1) = (1/2) + \arctan((x - \delta)/\gamma)/\pi$. (Hint: differentiate.)
Problem 1.5 Show that the cumulative distribution function of a Lévy distribution $X \sim S(1/2, 1, \gamma, \delta; 1)$ 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. (Hint: differentiate.)

Problem 1.6 What is wrong with the following argument? If $X_1, \ldots, X_n \sim \text{Gamma}(\alpha, \beta)$ are independent, then $X_1 + \cdots + X_n \sim \text{Gamma}(n\alpha, \beta)$, 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($\alpha, c$) 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($\alpha, c$), 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 $\alpha$s is not stable. Section 12.10 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 $\alpha$-stable r.v.

Problem 1.14 Simulate $n = 1,000$ uniform random variables and plot a “running sample variance” $(k, s_k^2)$, $k = 2, 3, 4, \ldots, n$. Repeat the process with normal random variables, Cauchy random variables and Pareto random variables (see Section 12.1 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 $N(\mu, \sigma^2)$. 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($\gamma, \delta$) distribution from Problem 1.4 to prove (1.11).

Problem 1.17 Use Problem 1.5 to prove (1.12).