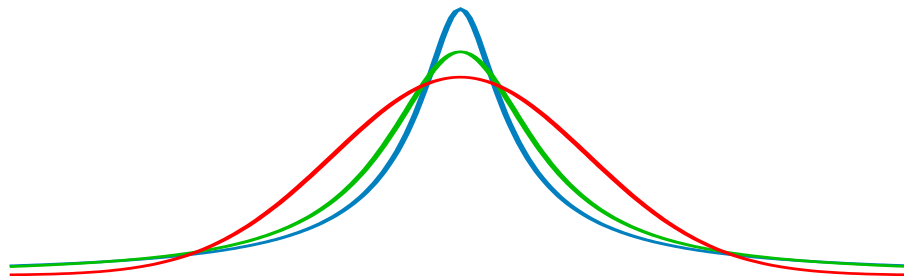


User Manual for STABLE 5.2

Excel Version



Abstract

This manual gives information about the STABLE library, which computes basic quantities for univariate stable distributions: densities, cumulative distribution functions, quantiles, and simulation. Statistical routines are given for fitting stable distributions to data and assessing the fit. Utility routines give information about the program and perform related calculations. Quick spline approximations of the basic functions are provided. Densities, cumulative distribution functions and simulation for discrete/quantized stable distributions are described.

The multivariate module gives functions to compute bivariate stable densities, simulate stable random vectors, and fit bivariate stable data. In the radially symmetric case, the amplitude densities, cumulative distribution functions, quantiles are computed for dimension up to 100.

©2002-2011 by Robust Analysis, Inc.

www.RobustAnalysis.com

info@robustanalysis.com

Revised December 13, 2010 (processed January 3, 2011)

Contents

1	Univariate Stable Introduction	4
2	Univariate Stable Functions	5
2.1	Basic functions	5
2.1.1	Stable densities	5
2.1.2	Stable distribution functions	5
2.1.3	Stable quantiles	5
2.1.4	Simulate stable random variates	6
2.1.5	Stable hazard function	6
2.1.6	Derivative of stable densities	6
2.1.7	Second derivative of stable densities	6
2.1.8	Stable score/nonlinear function	6
2.2	Statistical functions	6
2.2.1	Estimating stable parameters	6
2.2.2	Maximum likelihood estimation	7
2.2.3	Maximum likelihood estimation with restricted parameters	7
2.2.4	Maximum likelihood estimation with search control	7
2.2.5	Quantile based estimation	7
2.2.6	Empirical characteristic function estimation	7
2.2.7	Fractional moment estimation	7
2.2.8	Log absolute moment estimation	8
2.2.9	Quantile based estimation, version 2	8
2.2.10	U statistic based estimation	8
2.2.11	Confidence intervals for ML estimation	8
2.2.12	Information matrix for stable parameters	8
2.2.13	Log-likelihood computation	8
2.2.14	Chi-squared goodness-of-fit test	9
2.2.15	Kolmogorov-Smirnov goodness-of-fit test	9
2.2.16	Likelihood ratio test	9
2.2.17	Stable regression	10
2.3	Informational/utility functions	10
2.3.1	Version information	10
2.3.2	Modes of stable distributions	10
2.3.3	Set internal tolerance	10
2.3.4	Get internal tolerance	11
2.3.5	Convert between parameterizations	11
2.3.6	Omega function	11
2.4	Series approximations to basic distribution functions	11
2.4.1	Series approximation of stable pdf around the origin	11
2.4.2	Series approximation of stable cdf around the origin	12
2.4.3	Series approximation of stable pdf at the tail	12
2.4.4	Series approximation of stable cdf at the tail	12
2.5	Faster approximations to basic functions	12
2.5.1	Quick stable density computation	12
2.5.2	Quick stable cumulative computation	12
2.5.3	Quick stable log pdf computation	12
2.5.4	Quick stable quantile computation	13
2.5.5	Quick stable hazard function computation	13
2.5.6	Quick stable likelihood computation	13
2.5.7	Quick stable score/nonlinear function	13
2.6	Discrete stable distributions	13
2.6.1	Discrete stable density	13

2.6.2	Quick discrete stable density	13
2.6.3	Discrete stable cumulative distribution function	13
2.6.4	Quick discrete stable cumulative distribution function	14
2.6.5	Simulate discrete stable random variates	14
2.6.6	Simulate discrete stable random variates with specified saturation probability	14
2.6.7	Find scale γ to have a specified saturation probability for a discrete stable distribution	14
2.6.8	Discrete maximum likelihood estimation	14
3	Multivariate Stable Introduction	15
4	Multivariate Stable Functions	16
4.1	Define multivariate stable distribution	16
4.1.1	Independent components	17
4.1.2	Isotropic stable	17
4.1.3	Elliptical stable	17
4.1.4	Discrete spectral measure	17
4.1.5	Discrete spectral measure in 2 dimensions	17
4.1.6	Undefine a stable distribution	18
4.2	Basic functions	18
4.2.1	Density function	18
4.2.2	Cumulative function	18
4.2.3	Cumulative function (Monte Carlo)	18
4.2.4	Multivariate simulation	19
4.2.5	Find a 2-dimensional rectangle with probability at least p	19
4.3	Statistical functions	19
4.3.1	Estimate a discrete spectral measure - fit a stable distribution to bivariate data	19
4.3.2	Estimate parameter functions	19
4.3.3	Fit an elliptical stable distribution to multivariate data	19
4.4	Amplitude distribution	20
4.4.1	Amplitude cumulative distribution function	20
4.4.2	Amplitude density	20
4.4.3	Amplitude quantiles	20
4.4.4	Simulate amplitude distribution	20
4.4.5	Fit amplitude data	20
4.4.6	Amplitude score function	20
4.5	Faster approximations to multivariate routines	21
4.5.1	Quick log-likelihood for bivariate isotropic case	21
4.5.2	Quick amplitude density in bivariate case	21
4.6	Bivariate discrete stable distribution	21
4.6.1	Discrete bivariate density	21
4.7	Multivariate informational/utility functions	22
4.7.1	Information about a distribution	22
4.7.2	Compute projection parameter functions	22
4.7.3	Multivariate convert parameterization	22
5	Error/return codes	23
	References	24
	Index	25

1 Univariate Stable Introduction

Stable distributions are a class of probability distributions that generalize the normal distribution. Stable distributions are a four parameter family: α is the tail index, or index of stability, and is in the range $0 < \alpha \leq 2$, β is a skewness parameter and is in the range $-1 \leq \beta \leq 1$, γ is a scale parameter and must be positive, and δ is a location parameter, an arbitrary real number.

Since there are no formulas for the density and distribution function of a general stable law, they are described in terms of their characteristic function (see below). The main purpose of the STABLE program is to make these distributions accessible in practical problems. The package enables the calculation of stable densities, cumulative distribution functions, quantiles, etc. It also can fit data by several different estimation procedures.

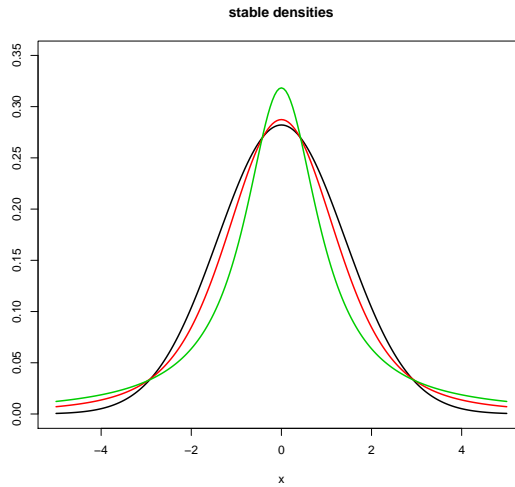


Figure 1: Symmetric stable densities ($\beta = 0$) with $\alpha = 2$ (Gaussian, in black), $\alpha = 1.5$ (red), and $\alpha = 1$ (Cauchy, green).

There are numerous meanings for these parameters. We will focus on two here, which we call the 0-parameterization and the 1-parameterization. The STABLE programs use a variable `param` to specify which of these parameterizations to use. If you are only concerned with symmetric stable distributions, the two parameterizations are identical. For non-symmetric stable distributions, we recommend using the 0-parameterization for most statistical problems, and only using the 1-parameterization in special cases, e. g. the one sided distributions when $\alpha < 1$ and $\beta = \pm 1$.

A random variable X is $S(\alpha, \beta, \gamma, \delta; 0)$ if it has characteristic function (Fourier transform)

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

A random variable X is $S(\alpha, \beta, \gamma, \delta; 1)$ if it has characteristic function

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

Note that if $\beta = 0$, then these two parameterizations are identical, it is only when $\beta \neq 0$ that the asymmetry term (the imaginary factor involving $\tan \frac{\pi\alpha}{2}$ or $\frac{2}{\pi}$) becomes relevant. More information on parameterizations and about stable distributions in general can be found at <http://academic2.american.edu/~jpnolan>, which has a draft of the first chapter of Nolan (2010b).

The next section gives a description of the basic univariate functions in STABLE.

2 Univariate Stable Functions

Interfaced STABLE functions require input variables, and return the results of the computations. The interface computes the lengths of all arrays, specifies default values for some of the variables in some case, and handles return codes and results.

The parameters of the stable distribution must be specified. In Excel the parameter values `alpha`, `beta`, `gamma`, `delta`, `param` are passed individually.

The STABLE interface prints an error message when an error occurs. If an error occurs, execution is aborted; if a warning occurs, execution continues.

There is basic help information built into the interfaces. In Excel, you can see the parameters to each function call in the Function Wizard, which is started by clicking on the icon f_x near the top of the screen.

The Excel interface requires you to select a block of cells where the output will go, and then type the function you want to compute with its arguments, and then press CTRL-SHIFT-ENTER (all 3 keys pressed at the same time). If you just press ENTER, you will get the error message "You cannot change part of an array", and Excel will lock up until you correct the entry and press CTRL-SHIFT-ENTER. For example, to simulate stable random, you select a block of cells within a column where you want the output to go, then type the command `stablernd(1.7, 0, 1, 0, 0)` (to simulate random variables with $\alpha=1.7$, $\beta=0$, $\gamma=1$, $\delta=0$ in the 0-parameterization), then CTRL-SHIFT-ENTER. The program then fills the number of cells that were highlighted with stable random variates. In some cases, like the multivariate elliptical routines, the output will be a block of cells with different information in them.

The STABLE library is not reentrant; only one user should be using the library at once.

The user should be aware that these routines attempt to calculate quantities related to stable distributions with high accuracy. Nevertheless, there are times when the accuracy is limited. If α is small, the pdf and cdf have very abrupt changes and are hard to calculate. When some quantity is small, e.g. the cdf of the light tail of a totally skewed stable distribution, the routines may only be accurate to approximately ten decimal places. There are certain values of the parameters (α near 2, α near 1, β near ± 1 , etc.) where there are complicated numerical problems with calculations. In these cases, the STABLE program may approximate values by rounding parameters. For example, if you try to calculate a stable pdf or cdf for $\alpha = 1.01$ and $\beta = 0.01$, the STABLE program will round to $\alpha = 1$ and $\beta = 0$, and compute the value for these values of the parameters. Likewise, when x is near 0 in the 1-parameterization, STABLE will do a linear interpolation to compute the pdf or cdf at that point. The thresholds used in rounding and linear approximation are described on page 11. You can manually reset these values, but be careful: the algorithms may yield poor values in some cases.

The remainder of this section is a description of the functions in the STABLE library.

2.1 Basic functions

2.1.1 Stable densities

Excel function: `stablepdf(x, alpha, beta, gamma, delta, param)`

This function computes stable density functions (pdf): $y_i = f(x_i) = f(x_i|\alpha, \beta, \gamma, \delta; \text{param})$, $i = 1, \dots, n$. The algorithm is described in Nolan (1997).

2.1.2 Stable distribution functions

Excel function: `stablecdf(x, alpha, beta, gamma, delta, param)`

This function computes stable cumulative distribution functions (cdf): $y_i = F(x_i) = F(x_i|\alpha, \beta, \gamma, \delta; \text{param})$, $i = 1, \dots, n$. The algorithm is described in Nolan (1997).

2.1.3 Stable quantiles

Excel function: `stableinv(p, alpha, beta, gamma, delta, param)`

This function computes stable quantiles, the inverse of the cdf: $x_i = F^{-1}(p_i)$, $i = 1, \dots, n$. The quantiles are found by numerically inverting the cdf. Extreme tail quantiles may be hard to find because of

subtractive cancelation and the fact that cdf calculations may only be accurate to 10 decimal places; see the notes below.

Note that the accuracy of the inversion is determined by two internal tolerances. (See Section 2.3.3.) (1) tolerance 10 is used to limit how low a quantile can be searched for. The default value is $p = 10^{-10}$: quantiles below p will be set to the left endpoint of the support of the distribution, which may be $-\infty$. Likewise, quantiles above $1 - p$ will be set to the right endpoint of the support of the distribution, which may be $+\infty$. (2) tolerance 2 is the relative error used when searching for the quantile. The search tries to get full precision, but if it can't, it will stop when the relative error is less than tolerance 2.

2.1.4 Simulate stable random variates

Excel function: `stablernd(alpha, beta, gamma, delta, param)`

This function simulates n stable random variates: x_1, x_2, \dots, x_n with parameters $(\alpha, \beta, \gamma, \delta)$ in parameterization `param`. It is based on Chambers et al. (1976).

The algorithm that generates stable random variates requires uniform(0,1) random variates as input. In Excel, there is an function inside STABLE that does this. That method uses two integer values as “seeds” that determine the random values produced. When you start STABLE, these seeds are always set to the same values, so unless you change them using the function `stablernd_seed(seed1, seed2)`, you will always see the same random values. If you want to “randomize” the starting seeds based on clock time, you can call `stablernd_seed(0, 0)`. This will provide a different set of values for each run. If you want to repeat simulations: set the seeds manually to non-zero integers, then run your simulations. If you want to rerun the same simulations, you can reset the seeds and restart.

2.1.5 Stable hazard function

Excel function: not implemented in Excel

This function computes the hazard function for a stable distribution: $h_i = f(x_i)/(1 - F(x_i))$, $i = 1, \dots, n$.

2.1.6 Derivative of stable densities

Excel function: not implemented in Excel

This function computes the derivative of stable density functions: $y_i = f'(x_i) = f'(x_i|\alpha, \beta, \gamma, \delta; \text{param})$, $i = 1, \dots, n$.

2.1.7 Second derivative of stable densities

Excel function: not implemented in Excel

This function computes the second derivative of stable density functions: $y_i = f''(x_i) = f''(x_i|\alpha, \beta, \gamma, \delta; \text{param})$, $i = 1, \dots, n$.

2.1.8 Stable score/nonlinear function

Excel function: not implemented in Excel

This function computes the score or nonlinear function for a stable distribution: $g(x) = -f'(x)/f(x) = -(d/dx) \ln f(x)$. The routine uses `stablepdf` to evaluate $f(x)$ and numerically evaluates the derivative $f'(x)$. Warning: this routine will give unpredictable results when $\beta = \pm 1$. The problems occur where $f(x) = 0$ is small; in this region calculations of both $f(x)$ and $f'(x)$ are of limited accuracy and their ratio can be very unreliable.

2.2 Statistical functions

2.2.1 Estimating stable parameters

Excel function: `stablefit(x, method, param)`

Estimate stable parameters from the data in x_1, \dots, x_n , using `method` as described in the following table. This routine calls one of the functions described below to do the actual estimation.

method value	algorithm	notes
1	maximum likelihood	$\alpha \geq 0.4$
2	quantile	$\alpha \geq 0.1$
3	empirical characteristic function	$\alpha \geq 0.1$
4	fractional moment	$\alpha \geq 0.4, \beta = 0$, uses power $p = 0.2$
5	log absolute moment	$\beta = \delta = 0$
6	modified quantile	$\alpha \geq 0.4$
7	U statistic method	$\beta = \delta = 0$

Note that the fractional moment and log absolute moment methods do not work when there are zeros in the data set.

2.2.2 Maximum likelihood estimation

Excel function: no direct interface, use stablefit with method=1

Estimate the stable parameters for the data in x_1, \dots, x_n , in parameterization `param` using maximum likelihood estimation. The likelihood is numerically evaluated and maximized using an optimization routine. This program and the numerical computation of confidence intervals below are described in Nolan (2001). For speed reasons, the quick log likelihood routine is used to approximate the likelihood; this is where the restriction $\alpha \geq 0.4$ comes from.

2.2.3 Maximum likelihood estimation with restricted parameters

Excel function: not implemented in Excel

This is a modified version of maximum likelihood estimation, where some parameters can be estimated while the others are restricted to a fixed value. The function receives `theta = {alpha, beta, gamma, delta}` and if `restriction[i] = 1`, then `theta[i]` is fixed; to allow a parameter to vary, set `restriction[i] = 0`.

2.2.4 Maximum likelihood estimation with search control

Excel function: not implemented in Excel

This is maximum likelihood estimation with greater control over the search and ranges for the parameters. It is used internally.

2.2.5 Quantile based estimation

Excel function: no direct interface, use stablefit with method=2

Estimate stable parameters for the data in x , using the quantile based on the method of McCulloch (1986). It sometimes has problems when α is small, say $\alpha < 1/2$, and the data is highly skewed. Try the modified version below in such cases.

2.2.6 Empirical characteristic function estimation

Excel function: no direct interface, use stablefit with method=3

Estimate stable parameters for the data in x using the empirical characteristic function based method of Koutrovelis-Kogon-Williams, described in Kogon and Williams (1998). An initial estimate of the scale `gamma0` and the location `delta0` are needed to get accurate results. We recommend using the quantile based estimates of these parameters as input to this routine.

2.2.7 Fractional moment estimation

Excel function: no direct interface, use stablefit with method=4

Estimate stable parameters for the data in x , using the fractional moment estimator as in Nikias and Shao (1995). This routine only works in the symmetric case, it will always return $\beta = 0$ and $\delta = 0$. In this case the 0-parameterization coincides with the 1-parameterization, so there is no need to specify parameterization. `p`

is the fractional moment power used. A reasonable default value is $p = 0.2$; take $p < \alpha/2$ to get reasonable results.

This method does not work if there are zeros in the data set - negative sample moments do not exist. Remove zero values (and possibly values close to 0) from the data set if you want to use this method.

2.2.8 Log absolute moment estimation

Excel function: no direct interface, use stablefit with method=5

Estimate stable parameters for the data in x , using the log absolute moment method as in Nikias and Shao (1995). This routine only works in the symmetric case, it will always return $\beta = 0$ and $\delta = 0$. In this case the 0-parameterization coincides with the 1-parameterization, so there is no need to specify parameterization.

The log absolute moment method does not work when there are zeros in the data set, because $\log|x|$ is undefined when x is 0. Remove zero values (and possibly values close to 0) from the data set if you want to use this method.

2.2.9 Quantile based estimation, version 2

Excel function: no direct interface, use stablefit with method=6

Estimate stable parameters for the data in x , using a modified quantile method of Nolan (2010b). It should work for any values of the parameters, but some extreme values are inaccurate.

2.2.10 U statistic based estimation

Excel function: no direct interface, use stablefit with method=7

Estimate stable parameters for the data in x , using the method of Fan (2006). It only works for the symmetric case.

2.2.11 Confidence intervals for ML estimation

Excel function: stablefitmleci(theta, z, n, sigtheta)

This routine finds confidence intervals for maximum likelihood estimators of all four stable parameters. The routine returns a vector `sigtheta` of half widths of the confidence interval for each parameter in `theta=(alpha,beta,gamma,delta)`. These values depend on the confidence level you are seeking, specified by z , and the size of the sample n . The z value is the standard critical value from a normal distribution, i.e. use $z = 1.96$ for a 95% confidence interval. For example, the point estimate of α is `theta[1]`, and the confidence interval is `theta[1]±sigtheta[1]`. For β , the confidence interval is `theta[2]±sigtheta[2]`, for γ , the confidence interval is `theta[3]±sigtheta[3]`, For δ , the confidence interval is `theta[4] ± sigtheta[4]`. These values do not make sense when a parameter is at the boundary of the parameter space, e.g. $\alpha = 2$ or $\beta = \pm 1$.

These values are numerically approximated using a grid of numerically computed values in Nolan (2001). The values have limited accuracy, especially when $\alpha \leq 1$.

2.2.12 Information matrix for stable parameters

Excel function: not implemented in Excel

Returns the 4×4 information matrix for maximum likelihood estimation of the stable parameters for parameter values `theta`. This is done in the continuous 0-parameterization. These are approximate values, interpolated from a grid of numerically computed values in Nolan (2001) for $\alpha \geq 0.5$. The values have limited accuracy, especially when $\alpha \leq 1$.

2.2.13 Log-likelihood computation

Excel function: not implemented in Excel

Compute the log-likelihood of the data, assuming an underlying stable distribution with the specified

parameters.

2.2.14 Chi-squared goodness-of-fit test

Excel function: not implemented in Excel

Compute chi-squared goodness-of-fit statistic for the data in x_1, \dots, x_n using `nclass` equally probable classes/bins.

2.2.15 Kolmogorov-Smirnov goodness-of-fit test

Excel function: not implemented in Excel

This function computes the Kolmogorov-Smirnov two-sided test statistic:

$$D = \sup_{-\infty < x < \infty} |F(x) - \hat{F}(x)|,$$

where $F(\cdot)$ is the stable cdf with parameters $\alpha = \text{theta}[1]$, $\beta = \text{theta}[2]$, $\gamma = \text{theta}[3]$, $\delta = \text{theta}[4]$ and $\hat{F}(\cdot)$ is the sample cdf of the data in x . Use `method=0` for quick computations (`stableqkcdf` is used to compute cdf), use `method=1` for slower computations (`stablecdf` is used to compute cdf). The routine returns the observed value of D and an estimate of the tail probability $P(D > d)$, i.e. the significance level of the test. This tail probability is calculated using Stephen's approximation to the limiting distribution, e.g. $(n^{1/2} + 0.12 + 0.11n^{-1/2})D$ is close to the limiting Smirnov distribution. This is close to $n^{1/2}D$ for large n , and a better approximation on the tails for small n . Note, this calculation is not very accurate if the tail probability is large, but these cases aren't of much interest in a goodness-of-fit test. (If you don't like this approximation, the function returns D , and you can compute your own tail probability.) **WARNING:** the computation of the significance level is based on the assumption that the parameter values `theta=($\alpha, \beta, \gamma, \delta$)` were chosen independently of the data. If the parameters were estimated from the data, then this tail probability will be an overestimate of the significance level.

2.2.16 Likelihood ratio test

Excel function: not implemented in Excel

This function computes the likelihood ratio L_0/L_1 , where L_0 is the maximum likelihood of the data x under the assumption that x is an i.i.d. sample from a stable distribution with α and β restricted to the range `abnd[1] ≤ α ≤ abnd[2]` and `bbnd[1] ≤ β ≤ bbnd[2]`, and L_1 is the maximum likelihood of the data under an unrestricted stable model. The function computes the maximum likelihood using the quick approximation to stable likelihoods, so is limited to α in the range [0.4,2].

The vector `results` will contain the results of the computations:

```
results[1] = ratio of the likelihoods
results[2] = -2*log(ratio of likelihoods)
results[3] = log likelihood of the data for the restricted H0
results[4] = log likelihood of the data for the unrestricted H1
results[5] = estimated value of alpha under H0
results[6] = estimated value of beta under H0
results[7] = estimated value of gamma under H0
results[8] = estimated value of delta under H0
results[9] = estimated value of alpha without assuming H0
results[10] = estimated value of beta without assuming H0
results[11] = estimated value of gamma without assuming H0
results[12] = estimated value of delta without assuming H0
```

Note that under the standard assumptions, `results[2]` converges to a chi-squared distribution with d.f. = (# free parameters in H1 parameter space - # free parameters in H0 parameter space) as the sample size tends to ∞ .

For example, to compute the likelihood ratio test for the null hypothesis H0: data comes from a normal distribution vs H1: data comes from stable distribution, use `abnd=(2,2)` and `bbnd=(0,0)`, in which case

results[2] will have 2 d.f. To test H0: data comes from a symmetric stable distribution vs H1: data comes from a general stable distribution, use abnd=(0.4,2) and bbnd=(0,0), in which case results[2] will have 3 d.f.

2.2.17 Stable regression

Excel function: stableregression(x, trimprob, theta_init, theta_trim, theta, symmetric)

Computes linear regression coefficients $\theta_1, \theta_2, \dots, \theta_k$ for the problem

$$y_i = \theta_1 x_{i,1} + \theta_2 x_{i,2} + \dots + \theta_k x_{i,k} + e_i, \quad i = 1, \dots, n$$

where the error term e_i has a stable distribution. In matrix form, the equation is $\mathbf{y} = \mathbf{X}\theta + \mathbf{e}$. The algorithm uses maximum likelihood and is described in Nolan and Ojeda (2010).

y is a vector of length n of observed responses. x is a $n \times k$ matrix, with the columns of x representing the variables and the rows representing the different observations. NOTE: if you want an intercept term, you must include a column of ones in the x matrix. Typically one sets the first column of x to ones, and then θ_1 is the intercept.

trimprob is a vector of length 2, e.g. (0.1,0.9), which gives the lower and upper quantiles for the trimmed regression. (Trimmed regression trims off extreme values and then performs ordinary least squares regression. The resulting coefficients are used to get an initial estimate of the stable regression coefficients.) symmetric can be used to force the fitting program to assume symmetry in the error terms e_i .

The interfaced versions of this function returns a structure with different fields.

- theta is the vector of coefficients
- theta_init is the initial vector of coefficients from the OLS regression
- theta_trim is the initial vector of coefficients from the trimmed regression
- alpha, beta, gamma are the stable parameters estimated from the residuals. They can be regarded as nuisance parameters if you only care about the coefficients. Note that all parameters are in the 2-parameterization. You can convert to another representation using the convert parameterization function in Section 2.3.5. It is assumed that the location parameter δ is zero.

Note that in the non-Gaussian stable case, some of the traditional assumptions in regression are no longer true. In particular, it is NOT always the case that $Ee_i = 0$. First, if $\alpha \leq 1$, the heavy tails will mean that Ee_i is undefined. Second, in the non-symmetric case, for $\beta \neq 0$, we do not require $Ee_i = 0$. Instead, the regression uses the 2-parameterization so that the mode of e_i is zero. The reason for this is to make the regression line go through the center of the data points.

2.3 Informational/utility functions

2.3.1 Version information

Excel function: stableversion

stableversion() returns a string with version information.

2.3.2 Modes of stable distributions

Excel function: not implemented in Excel

Returns the mode of a $S(\alpha = \text{theta}[1], \beta = \text{theta}[2], \gamma = \text{theta}[3], \delta = \text{theta}[4]; \text{param})$ distribution. If $\beta \neq 0$, the mode is determined by a numerical search of the pdf.

2.3.3 Set internal tolerance

Excel function: stablesettolerance(inum, value)

Sets the value of internal tolerances that are used during computations. You change these values at your own risk: computation times can become very long, inaccuracy can accumulate, and some choices of the parameters can cause infinite loops.

inum	meaning
0	relative error for pdf numerical integration
1	relative error for cdf numerical integration
2	relative error for quantile search
3	alpha and beta rounding
4	x tolerance near zeta
5	exponential cutoff
6	peak/strim location tolerance
7	stabletrim tolerance
8	minimum alpha
9	minimum xtol
10	threshold for quantile search
11	x tolerance

2.3.4 Get internal tolerance

Excel function: `stablegettolerance (inum, value)`

Returns the value of the internal tolerances, see the preceding function for the meanings of each variable.

2.3.5 Convert between parameterizations

Excel function: `stableconvert (param, alpha, beta, gamma, delta, newparam)`

Convert from the parameters given in `theta=(alpha, beta, gamma, delta)` given in the `param`-parameterization to the parameters `thetaneu` given in the `newparam`-parameterization. Currently `param` and `newparam` are restricted to the values 0,1,2 and 3.

2.3.6 Omega function

Excel function: not implemented in Excel

Compute the function $\omega(u_i|\alpha, \beta; k), i = 1, \dots, n$ where

$$\omega(u|\alpha, \beta; 0) = \begin{cases} |u|^\alpha [1 + i\beta(\tan \frac{\pi\alpha}{2})(\text{sign } u)(|u|^{1-\alpha} - 1)] & \alpha \neq 1 \\ |u| [1 + i\beta\frac{2}{\pi}(\text{sign } u) \ln |u|] & \alpha = 1, \end{cases} \quad (3)$$

$$\omega(u|\alpha, \beta; 1) = \begin{cases} |u|^\alpha [1 - i\beta(\tan \frac{\pi\alpha}{2})(\text{sign } u)] & \alpha \neq 1 \\ |u| [1 + i\beta\frac{2}{\pi}(\text{sign } u) \ln |u|] & \alpha = 1. \end{cases}$$

These functions are from the characteristic functions of standardized univariate stable distributions: if $Z \sim \mathbf{S}(\alpha, \beta, 1, 0; k)$, then $E \exp(iuZ) = \exp(-\omega(u|\alpha, \beta; k))$. As before, $k = 0$ or $k = 1$ correspond to two different parameterization. The function returns two vectors containing the real and imaginary parts of $\omega(u|\alpha, \beta; k)$

2.4 Series approximations to basic distribution functions

These functions use the Bergstrom series for stable densities and cdfs, which are only defined for $\alpha \neq 1$.

2.4.1 Series approximation of stable pdf around the origin

Excel function: not implemented in Excel

Computes the stable probability distribution function using a series approximation with `nterms` in it. This function is best used to calculate the density near the origin in the 1-parameterization. The series is not

defined for $\alpha = 1$. Note that `nterms=1` corresponds to a constant term, `nterms=2` corresponds to a linear term, etc.

2.4.2 Series approximation of stable cdf around the origin

Excel function: not implemented in Excel

Computes the stable cumulative distribution function using a series expansion with `nterms` in it. This function is best used to calculate the cdf near the origin in the 1-parameterization. The series is not defined for $\alpha = 1$. Note that `nterms=1` corresponds to a constant term, `nterms=2` corresponds to a linear term, etc.

2.4.3 Series approximation of stable pdf at the tail

Excel function: not implemented in Excel

Computes the stable probability distribution function using a series approximation with `nterms` in it. This function is best used to calculate points on the tail of a distribution. The series is defined only for $x > 0$. (For $x < 0$, replace x by $-x$ and β by $-\beta$. The series is not defined for $\alpha = 1$.)

2.4.4 Series approximation of stable cdf at the tail

Excel function: not implemented in Excel

Computes the stable cumulative distribution function using a series approximation with `nterms` in it. This function is best used to calculate points on the tail of a distribution. The series is defined only for $x > 0$. (For $x < 0$, replace x by $-x$ and β by $-\beta$. The series is not defined for $\alpha = 1$.)

2.5 Faster approximations to basic functions

The functions described in preceding sections are accurate, but can take a long time to compute. For evaluating a single pdf or cdf at a single set of parameter values, they are fine. However, when the functions must be evaluated many times, the previous routines are slow. For example, when estimating stable parameters by maximum likelihood estimation, the likelihood is evaluated at each data point for a large number of parameter values during the numerical search for the point where the likelihood is maximized. In these cases, speed is more desirable than great accuracy.

The functions described below are approximations to the functions above, and are based on pre-computed values using those basic functions. They are designed to evaluate the quantity of interest at many x values for fixed values of α and β . Each routine has a setup time, and if you change α or β , that setup code must be rerun. It can be slower to run these routines than the basic routines above if you only want to calculate the quantity at a few x values. These routines work for $0.2 \leq \alpha \leq 2$ and all $-1 \leq \beta \leq 1$.

2.5.1 Quick stable density computation

Excel function: `stableqkpdf(x, alpha, beta, gamma, delta, param)`

Call is identical to Section 2.1.1, results are approximately the same.

2.5.2 Quick stable cumulative computation

Excel function: `stableqkcdf(x, alpha, beta, gamma, delta, param)`

Call is identical to Section 2.1.2, results are approximately the same.

2.5.3 Quick stable log pdf computation

Excel function: not implemented in Excel

Approximates $\log(\text{pdf})$ for stable distributions. Results are approximately the same as $\log(f(x))$.

2.5.4 Quick stable quantile computation

Excel function: `stableqkinv(p, alpha, beta, gamma, delta, param)`

Call is identical to Section 2.1.3, but much faster. Note the comments in that section about extreme upper quantiles.

2.5.5 Quick stable hazard function computation

Excel function: Not implemented in Excel

Call is identical to Section 2.1.5.

2.5.6 Quick stable likelihood computation

Excel function: Not implemented in Excel

Call is identical to Section 2.2.13.

2.5.7 Quick stable score/nonlinear function

Excel function: Not implemented in Excel

This function approximates the score or nonlinear function for a stable distribution: $g(x) = -f'(x)/f(x) = -(d/dx) \ln f(x)$. The algorithm used depends on the value of `method`. When `method=1`, `stableqkpdf` is used to compute $f(x)$ and in the numerical evaluation of $f'(x)$. When `method=2`, `stablescorefn` is used to compute $g(x)$ on a grid, then a spline is fit to those values. The resulting spline is used to approximate $g(x)$. If n is large, this is noticeably faster than either `stablescorefn` or `method=1` above. When `method=3`, a rational function approximation is used to approximate $g(x)$. This is the fastest method, but the accuracy depends on the values of `alpha` and `beta`. If `alpha` is between 1 and 1.9 and `beta` is near 0, the approximation is good.

2.6 Discrete stable distributions

Given a stable distribution $X \sim S(\alpha, \beta, \gamma, \delta; \text{param})$ and a pair of cutoff values $a < b$, the random variable

$$Y = \text{integer part of } \max(a, \min(X, b))$$

is a discrete stable distribution. These distribution arise in signal processing where a continuous quantity is quantized/digitized and limited accuracy is kept. It is assumed that the cutoff values are integers. The saturation probability is $P(X < a - 1/2) + P(X > b + 1/2)$, and is a measure of how much of the distribution is lost by truncating at the cutoff values. In the routines below, the cutoff is specified by a vector of length 2: `cutoff=(a,b)`. In this section X will always refer to the continuous stable distribution, while Y will always refer to a discrete/quantized/integer valued distribution.

In the internal routines, the x values are integers. The matlab/R/Mathematica interfaces use double precision values.

2.6.1 Discrete stable density

Excel function: Not implemented in Excel

Calculates $f_i = P(Y = x_i), i = 1, \dots, n$.

2.6.2 Quick discrete stable density

Excel function: Not implemented in Excel

Calculates $f_i = P(Y = x_i), i = 1, \dots, n$. Faster than above, less accurate.

2.6.3 Discrete stable cumulative distribution function

Excel function: Not implemented in Excel

Calculates $F_i = P(Y \leq x_i), i = 1, \dots, n$.

2.6.4 Quick discrete stable cumulative distribution function

Excel function: Not implemented in Excel

Calculates $F_i = P(Y \leq x_i), i = 1, \dots, n$. Faster than above, less accurate.

2.6.5 Simulate discrete stable random variates

Excel function: Not implemented in Excel

Simulates discrete stable random variates with the specified parameters and cutoffs.

2.6.6 Simulate discrete stable random variates with specified saturation probability

Excel function: Not implemented in Excel

Simulates discrete stable random variates, where the scale is computed internally to make the saturation probability=`psaturation`. Note that in cases where the stable parameters are passed individually, `gamma` is NOT used. In the cases where the vector `theta` is used, the value of $\gamma = \text{theta}[3]$ is ignored. The following function is used to compute γ , then the previous function is called to generate the values.

2.6.7 Find scale γ to have a specified saturation probability for a discrete stable distribution

Excel function: Not implemented in Excel

Given α, β, δ and `cutoff= (a,b)`, the scale γ is computed to get the requested saturation probability, e.g. `psaturation=P(X < a - 1/2) + P(X > b + 1/2)`.

2.6.8 Discrete maximum likelihood estimation

Excel function: Not implemented in Excel

Estimate the stable parameters for the discrete stable data in x_1, \dots, x_n , in parameterization `param` using maximum likelihood estimation. The likelihood is numerically evaluated and maximized using an optimization routine. When `method=1`, `stablepdfdiscrete` is used to calculate likelihood, when `method=2`, symmetry is assumed ($\beta = 0$) and a faster method is used to compute the likelihood.

3 Multivariate Stable Introduction

WARNING: The Excel interface for multivariate stable distributions currently only implements two functions: `mvstablern-delliptical()` and `mvstablefitelliptical()`.

To specify a multivariate stable distribution $\mathbf{X} = (X_1, X_2, \dots, X_d)^T$ in d dimensions requires an index of stability $\alpha \in (0, 2]$, a finite Borel measure Λ on the unit sphere $\mathbb{S} = \{\mathbf{s} \in \mathbb{R}^d : |\mathbf{s}| = 1\}$ and a shift vector $\boldsymbol{\delta} \in \mathbb{R}^d$. The measure Λ is called the spectral measure of the distribution. The joint characteristic function of $\mathbf{X} \sim \mathbf{S}(\alpha, \Lambda, \boldsymbol{\delta}; k)$ is given by:

$$E \exp(i \langle \mathbf{u}, \mathbf{X} \rangle) = \exp \left(- \int_{\mathbb{S}} \omega_k(\langle \mathbf{u}, \mathbf{s} \rangle | \alpha, 1; k) \Lambda(ds) + i \langle \mathbf{u}, \boldsymbol{\delta} \rangle \right),$$

where $\omega(u|\alpha, \beta; k)$ is defined in (3). As in one dimension, the 1-parameterization is more common in theoretical research, while the 0-parameterization is better suited to computation and statistical problems. Here and below, $\langle \mathbf{u}, \mathbf{X} \rangle = \mathbf{u} \mathbf{X}^T = u_1 X_1 + \dots + u_d X_d$ is the inner product. Symmetric stable distributions are defined by the condition $\mathbf{X} \stackrel{d}{=} -\mathbf{X}$, which is equivalent to Λ being a symmetric measure on \mathbb{S} , i.e. $\Lambda(A) = \Lambda(-A)$ for any Borel subset $A \subset \mathbb{S}$. As in the univariate case, in the symmetric case the 0-parameterization and the 1-parameterization coincide.

The general case is beyond current computational capabilities, but several special cases: isotropic (radially symmetric), elliptical, independent components and discrete spectral measure are computationally accessible.

isotropic The spectral measure is continuous and uniform, leading to isotropic/radial symmetry for the distribution. The characteristic function is

$$E \exp(i \langle \mathbf{u}, \mathbf{X} \rangle) = \exp(-\gamma_0^\alpha |\mathbf{u}|^\alpha + i \langle \mathbf{u}, \boldsymbol{\delta} \rangle). \tag{4}$$

elliptical The characteristic function is

$$E \exp(i \langle \mathbf{u}, \mathbf{X} \rangle) = \exp \left(-(\mathbf{u}^T R \mathbf{u})^{(\alpha/2)} + i \langle \mathbf{u}, \boldsymbol{\delta} \rangle \right) \tag{5}$$

where R is a positive definite matrix. ($R = \gamma_0^2 I$ is equivalent to the isotropic case above.) More information on this accessible class of distributions is given in Nolan (2010a).

independent components If components are independent with $X_j \sim \mathbf{S}(\alpha, \beta_j, \gamma_j, \delta_j; k)$, then the characteristic function is

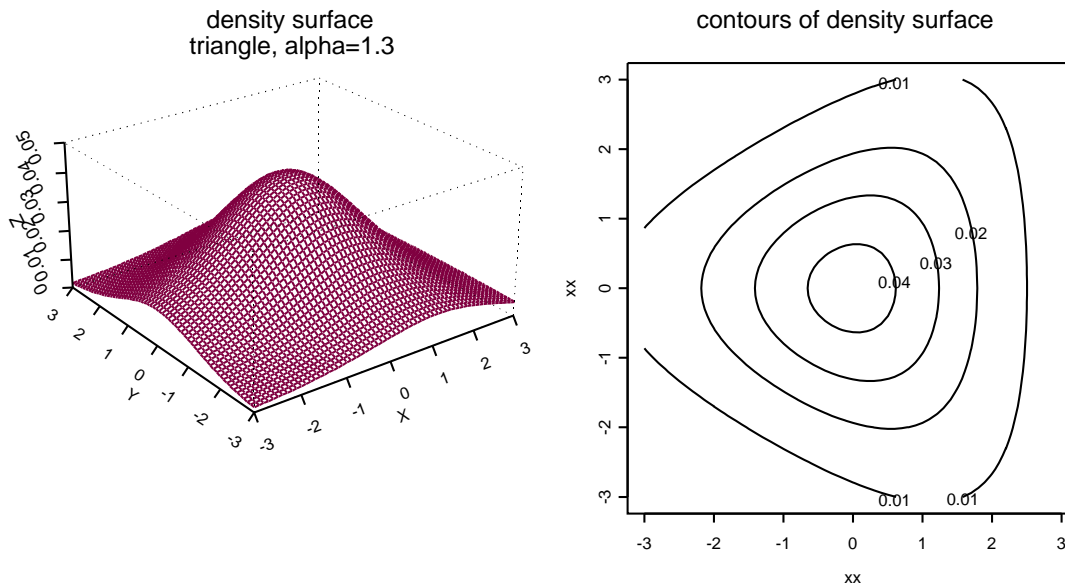
$$E \exp(i \langle \mathbf{u}, \mathbf{X} \rangle) = \exp \left(- \sum_{j=1}^d \omega(u_j | \alpha, \beta_j; k) \gamma_j^\alpha + i \langle \mathbf{u}, \boldsymbol{\delta} \rangle \right) \tag{6}$$

This is a special case of the discrete spectral measure below: the spectral mass is concentrated on the points where the coordinates axes intersect the unit sphere.

discrete When the spectral measure is discrete with mass λ_j at $\mathbf{s}_j \in \mathbb{S}$, $j = 1, \dots, m$ the characteristic function is

$$E \exp(i \langle \mathbf{u}, \mathbf{X} \rangle) = \exp \left(- \sum_{j=1}^m \omega(\langle \mathbf{u}, \mathbf{s}_j \rangle | \alpha, 1; k) \lambda_j + i \langle \mathbf{u}, \boldsymbol{\delta} \rangle \right) \tag{7}$$

This discrete class is dense in the class of all stable distributions: any finite spectral measure Λ can be approximated by a discrete measure, see Byczkowski et al. (1993). Below is a plot of the density surface of a bivariate stable density with three point masses, each of weight 1 at locations $(\cos(\pi/3), \sin(\pi/3))$, $(-1, 0)$, and $(\cos(5\pi/3), \sin(5\pi/3))$



4 Multivariate Stable Functions

Since the specification of a multivariate stable distribution is somewhat cumbersome, a different approach from the univariate case is taken in these routines. Two steps are needed to work with a multivariate stable distribution. First, the distribution is specified by calling a function to define the distribution. Second, call a separate functions to compute densities, cumulatives, simulate, etc.

The programs for working with multivariate stable distributions are less well developed and generally limited to 2 dimensions. At the current time, when dimension $2 < d \leq 100$, you can:

- simulate using `mvstablern`
- fit multivariate data with an elliptical model
- calculate the pdf using `mvstablepdf` if the components are independent OR the spectral measure has exactly d point masses OR the distribution is isotropic or elliptical
- calculate the cdf using `mvstablecdf` if the components are independent
- calculate the cdf using `mvstablecdfMC` by Monte Carlo estimation for any type of distribution

The accuracy of the multivariate pdf and cdf calculations are limited. In all cases, X is a column vector, this is important to remember when you specify x for calculating the multivariate pdf or cdf.

4.1 Define multivariate stable distribution

STABLE has the ability to work with multiple distributions. When a multivariate stable distribution is defined, a ‘distribution descriptor’ is returned. That descriptor must be used when computing quantities for that

distribution. Note: The descriptor should not be changed by a user. The descriptor may change between calls, and contents may vary in future versions of STABLE.

There are different functions used to define each of the different types of distributions that STABLE can work with. They are described below.

4.1.1 Independent components

Excel function: Not implemented in Excel

Define a multivariate stable distribution with dimension $d \leq 100$ and independent components with characteristic function (6). `beta`, `gamma` and `delta` should be vectors of length d = the dimension of the distribution.

4.1.2 Isotropic stable

Excel function: Not implemented in Excel

Define a multivariate isotropic stable distribution with with dimension $d \leq 100$ and characteristic function (4). d is the dimension of the distribution, `gamma0` is the scale parameter, `delta` is the location vector.

4.1.3 Elliptical stable

Excel function: Not implemented in Excel

Define a multivariate elliptically contoured/sub-Gaussian stable distribution with dimension $d \leq 100$ and with characteristic function (5). The dimension of the distribution is determined from the size of R , a positive definite $d \times d$ shape matrix, and `delta` is the location vector.

4.1.4 Discrete spectral measure

Excel function: Not implemented in Excel

Define a multivariate stable distribution with discrete spectral measure with dimension $d \leq 100$ and characteristic function (7). `s` should be a $d \times nlambda$ matrix specifying the location of the point masses as columns, `lambda` should be a row vector of length `nlambda` containing the weights. `beta` should be a row vector of length `nlambda` specifying the skewness at each point mass. `delta` is the shift as a column vector. `param` is the parameterization, must be 0 or 1.

The spectral measure is defined by putting mass $lambda[j]*(1+beta[j])/2$ at s_j and mass $lambda[j]*(1-beta[j])/2$ at $-s_j$. Setting all `beta` equal to 1 gives the standard definition of a spectral measure, with mass $lambda[j]$ at $s[j]$. Setting all `beta` equal to 0 guarantees that the distribution is symmetric, putting weight $lambda[j]/2$ at $\pm s_j$. If any element of `beta` is not 0, the distribution is assumed to be nonsymmetric. (It is possible to manually make a spectral measure symmetric with nonzero `beta` by defining antipodal points and weights and values of `beta` that balance correctly. However, STABLE does not detect this.) Some parts of STABLE are significantly faster and more accurate in the symmetric case, e.g. density calculations and simulations.

4.1.5 Discrete spectral measure in 2 dimensions

Excel function: Not implemented in Excel

Define a bivariate stable distribution with discrete spectral measure. This is a special case of the previous function. In two dimensions the locations of the point masses can be specified by angles: `angle[j]` gives the angle (in radians) of the location of $s_j = (\cos(angle[j]), \sin(angle[j]))$.

There are several special cases that are handled differently internally:

- When `angle` and `lambda` are of length 2, densities can be calculated in terms of univariate densities.
- The special case of the previous one is when `angle`=(0, $\pi/2$). This corresponds to a distribution with independent components. Both density and cdf are calculated in terms of products of univariate density and cdf respectively.

- If all elements of `beta` are 0, the distribution is symmetric. Cumulative distribution function calculations only work in the symmetric case (though Monte Carlo based cdf estimation works for any case you can simulate, including skewed.)

4.1.6 Undefine a stable distribution

Excel function: Not implemented in Excel

Clears the definition of the stable distribution `dist`. When a multivariate stable distribution is defined by one of the above functions, the STABLE library allocates memory for a distribution descriptor. The amount of memory used is not large, and this should not normally be a problem. But if your program loops and defines many multivariate stable distributions, you could have a memory problem as the descriptors accumulate. Calling this function when you are done with a stable distribution will deallocate the memory.

A technical detail: distribution descriptors are allocated internally in the STABLE library, not in the workspace of the user session. If you exit, these distribution descriptors are lost, EVEN IF YOUR SESSION INFORMATION IS SAVED. You must redefine all multivariate distributions when you restart the STABLE library.

4.2 Basic functions

4.2.1 Density function

Excel function: Not implemented in Excel

Computes the density $f(\mathbf{x})$ for stable distribution `dist` at each value in `x`. Note: this routine assumes that the density exists. The density will not exist if the spectral measure is supported on a proper linear subspace of the domain.

When dimension $d > 2$, the pdf can be calculated in special cases:

- the components are independent components
- the spectral measure is discrete AND the number of point masses is equal the dimension of the problem
- the isotropic case
- the elliptical case

Otherwise, only 2-dimensional computations can be done. The symmetric case uses the method in Abdul-Hamid and Nolan (1998), the nonsymmetric case uses the method in Nolan and Rajput (1995). The symmetric case is faster and more accurate than the nonsymmetric case. Both routines are accurate near the center of the distribution, and have limited accuracy near the tails.

4.2.2 Cumulative function

Excel function: Not implemented in Excel

This function approximates $P(\mathbf{a} \leq \mathbf{X} \leq \mathbf{b})$. If the components are independent, this works in dimensions up to 100.

In the symmetric two-dimensional case, the probability is evaluated by numerically integrating the (numerically computed) 2-dimensional density $f(\mathbf{x})$. Due to the limited precision in the numerical calculation of the density, and the approximate nature of the integration of this density, this routine gives only a few digits of accuracy. To find the probability of an unbounded regions, it is best to truncate the region using the routine in Section 4.2.5 to find a bounded rectangle containing most of the probability.

Use the function in Section 4.2.3 to approximate in 2-dimensional nonsymmetric case or in higher dimensions.

4.2.3 Cumulative function (Monte Carlo)

Excel function: Not implemented in Excel

This function approximates $P(\mathbf{a} \leq \mathbf{X} \leq \mathbf{b})$ by simulating n independent random vectors with the same distribution as \mathbf{X} and counting how many are in the interval $[\mathbf{a}, \mathbf{b}]$. It works for any distribution and dimension that can be simulated.

4.2.4 Multivariate simulation

Excel function: `mvstablerndeelliptical(alpha, delta, R)`

In Excel, one can only simulate from a multivariate elliptical distribution. To do this, specify the stable index α , a location vector δ (length d), and a shape matrix R (a $d \times d$ matrix). The output should be a block of $n \times d$ columns, where n is the number of values to simulate.

4.2.5 Find a 2-dimensional rectangle with probability at least p

Excel function: Not implemented in Excel

Find a number r so that the 2-dimensional rectangle $A = A(r) = [-r, r] \times [-r, r]$ has $P(\mathbf{X} \in A) \geq p$, where \mathbf{X} is a bivariate stable distribution defined by `dist`. This is used for technical calculations, e.g. in approximating the probability of unbounded regions. The method uses univariate projections and will generally give an overestimate of r . The method is less accurate for small p or if the distribution is not centered or highly skewed, it gets more accurate if p is close to 1 and the distribution is centered and symmetric.

If p is not too close to 1, one can get a better value of r by making repeated calls to the multivariate cdf function with rectangles of the form $A(r)$ and search for a value of r that makes $P(\mathbf{X} \in A(r))$ close to p . That procedure involves bivariate numerical integration will take much longer than this function.

4.3 Statistical functions

4.3.1 Estimate a discrete spectral measure - fit a stable distribution to bivariate data

Excel function: Not implemented in Excel

`x` contains the data values, `nspectral` is the number of points in the estimated spectral measure (must be divisible by 4), `method1d` is the method to use for estimating univariate stable parameters internally (see Section 2.2.1 for codes; only used if `method2d=1`), `method2d` is the method to use in estimating bivariate distribution. Use `method2d=1` for Rachev-Xin-Cheng method, `method2d=2` for projection method, `method2d=3` for empirical characteristic function method. The methods are described in Nolan et al. (2001), see Nolan and Panorska (1997) for some discussion of suggested values and diagnostics. Suggest using `nspectral=40`, `method1d=3`, `method2d=2`, `param=1`.

The function returns a list/structure that contains information about the fit, which is always done as a discrete spectral measure. The fields in the fit are: the estimated value of α , the estimated shift/location vector δ , `angle` which is a uniform grid from 0 to 2π of length `nspectral`, and `lambda` for the estimated weights at each position.

4.3.2 Estimate parameter functions

Excel function: Not implemented in Excel

Estimate the parameter functions for the bivariate data in `x`. The data is projected in each direction given by `angle` and the parameters are estimated in the `param` parameterization. `method1d` is the univariate method used to estimate the parameters (see Section 2.2.1 for codes).

The result is a matrix of dimension `length(x)×5`. The columns of the result are (1) for the angle, (2) for the estimate of α , (3) for the estimate of β , (4) for the estimate of γ , (5) for the estimate of δ at that angle.

4.3.3 Fit an elliptical stable distribution to multivariate data

Excel function: `mvstablefitelliptical(x, method1d)`

`x` contains the data values (d -vectors), `method1d` is the method to use for estimating univariate stable parameters internally (see Section 2.2.1 for codes).

The output should be a block of cells with d rows and $d + 2$ columns. The first column will have the estimated alpha in the first row (and no other entries). The second column will have the estimated shift vector delta. The remaining columns will contain an estimate of the $d \times d$ shape matrix R .

4.4 Amplitude distribution

For d -dimensional random vector \mathbf{X} , the univariate quantity $R = |\mathbf{X}|$ is called the amplitude of \mathbf{X} . When \mathbf{X} is isotropic, the radial symmetry allows one to reduce the dimension of the problem to a univariate problem. The following routines compute the cdf, pdf, quantiles, simulate and estimate for amplitudes of isotropic stable random vectors. Since these are univariate quantities, and it is required that the distribution is isotropic, one does NOT have to define the isotropic distribution separately. Because of computational difficulties, these routines are limited to dimension $d \leq 100$.

4.4.1 Amplitude cumulative distribution function

Excel function: Not implemented in Excel

Compute the cdf of the amplitude distribution: $F_R(r) = P(R \leq r)$ for $R = |\mathbf{X}|$, where \mathbf{X} is an d dimensional isotropic stable random vector with characteristic function $E \exp(i \langle \mathbf{u}, \mathbf{X} \rangle) = \exp(-\gamma_0^\alpha |u|^\alpha)$. Current implementation works for $\alpha \in [0.8, 2]$. There seems to be a relative error of approximately 3% for large r .

4.4.2 Amplitude density

Excel function: Not implemented in Excel

Compute the density $f_R(r)$ where R is described above. Current implementation works for $\alpha \in [0.8, 2]$. There seems to be a relative error of approximately 3% for large r .

4.4.3 Amplitude quantiles

Excel function: Not implemented in Excel

Compute the quantiles of the amplitude R described above. Current implementation works for $\alpha \in [0.8, 2]$.

4.4.4 Simulate amplitude distribution

Excel function: Not implemented in Excel

Simulate n i.i.d. values of the amplitude distribution R as described above. Current implementation works for $\alpha \in (0.2, 2]$.

4.4.5 Fit amplitude data

Excel function: Not implemented in Excel

Estimate the parameters α and γ_0 for amplitude data. \mathbf{r} contains the (univariate) amplitude data values, d is the dimension of the underlying distribution that the amplitude data comes from, `method` is the method to use for estimating. This initial implementation allows only `method=5`, which uses method of moments on the log of the amplitude data. Other methods are planned for the future. The function returns the estimated value of α and γ_0 .

4.4.6 Amplitude score function

Excel function: Not implemented in Excel

Compute the score function of the amplitude: $g(r) = -f'(r)/f(r)$, where $f(r) = f_R(r|\alpha, \gamma_0, d)$ is the amplitude pdf defined above. Current implementation works for $\alpha \in [0.8, 2]$, $d \leq 98$. There seems to be a

relative error of approximately 3% for large r .

4.5 Faster approximations to multivariate routines

There are a limited number of functions for quickly calculating multivariate functions in the 2-dimensional isotropic case. Such a distribution is specified by the index of stability α , the scale γ_0 , and the location $\boldsymbol{\delta} = (\delta_1, \delta_2)$. Because the description is simple, these functions use those arguments directly and do not use a distribution descriptor.

4.5.1 Quick log-likelihood for bivariate isotropic case

Excel function: Not implemented in Excel

Compute the log likelihood of the bivariate isotropic stable data in \mathbf{x} with stable index `alpha`, scale `gamma0`, and location vector `delta`. An internal approximation is used to compute the single value

$$\ell(\alpha, \gamma_0, \boldsymbol{\delta} | \mathbf{x}_1, \dots, \mathbf{x}_n) = \log \prod_{i=1}^n f_{\mathbf{X}}(\mathbf{x}_i | \alpha, \gamma_0, \boldsymbol{\delta}).$$

This function is designed to compute the log likelihood for a fixed α many times. In this case, it is much faster than trying to compute the right hand side above using the bivariate pdf routine in Section 4.2.1. It is also more accurate than that routine, especially on the tails. The program initially computes an approximation that depends on α ; if α changes, the approximation must be recomputed and it will be slower.

4.5.2 Quick amplitude density in bivariate case

Excel function: Not implemented in Excel

Compute the amplitude function $f_R(r | \alpha, \gamma_0, d = 2)$ for a 2-dimensional isotropic stable vector \mathbf{r} . For large n , it is much faster than the function in Section 4.4.2.

4.6 Bivariate discrete stable distribution

A bivariate discrete stable distribution is defined by digitizing and truncating a continuous bivariate stable distribution $\mathbf{X} = (X_1, X_2)^T$: discrete $\mathbf{Y} = (Y_1, Y_2)^T$ has components

$$Y_i = \text{integer part of } \max(a, \min(X_i, b)),$$

where `cutoff=(a,b)` are the upper and lower cutoff values. Note that the same cutoff is used for both components of \mathbf{X} . These distributions arise in signal processing where a bivariate continuous quantity is quantized/digitized and limited accuracy is kept. It is assumed that the cutoff values are integers. The saturation probability is $p_{sat} = P(X_1 < a - 1/2) + P(X_1 > b + 1/2) + P(X_2 < a - 1/2) + P(X_2 > b + 1/2)$, and is a measure of how much of the distribution is lost by truncating at the cutoff values.

In the internal routines, the x values are integers. The R, Mathematica and matlab interfaces store these integer values in double precision numbers.

4.6.1 Discrete bivariate density

Excel function: Not implemented in Excel

Compute the pdf of a discrete bivariate stable distribution. \mathbf{x} should be a $2 \times n$ matrix of integer values, `cutoff` is a vector of length 2 with upper and lower cutoff values for the truncation. The typical value for `cutoff` is (-128,127); both components of (X_1, X_2) are truncated at the same value. The function returns a vector `p` of length n with

$$p_i = P(\mathbf{Y} = \mathbf{x}_i) = P(Y_1 = x_{1i}, Y_2 = x_{2i}).$$

Note that `eps` is the attempted accuracy for each probability p_i , not for the total error. The probabilities are computed using the bivariate cdf function above and thus only works for symmetric stable two dimensional distributions. It's accuracy is limited: it is likely that when all possible values of \mathbf{x} are used, $\sum_i p_i$

will be slightly different from 1. The current implementation is slow. The `method` variable is unused at the current time; it will be used for faster approximations in future implementations.

4.7 Multivariate informational/utility functions

4.7.1 Information about a distribution

Excel function: Not implemented in Excel

Returns information about distribution `dist`. Useful for checking that definition.

4.7.2 Compute projection parameter functions

Excel function: Not implemented in Excel

Compute the exact parameter functions for a bivariate stable distribution. For direction $\mathbf{t} \in \mathbb{R}^2$, $(\mathbf{X}, \mathbf{t}_j)$ is univariate stable with parameters $(\alpha, \beta(\mathbf{t}), \gamma(\mathbf{t}), \delta(\mathbf{t}))$. This function computes the parameter functions $\beta(\cdot)$, $\gamma(\cdot)$ and $\delta(\cdot)$ at the values $\mathbf{t} = (\cos \text{angle}[j], \sin \text{angle}[j])$. Angles in `angle` are given in radians.

4.7.3 Multivariate convert parameterization

Excel function: Not implemented in Excel

Converts between multivariate stable parameterizations. `newparam` must be 0 or 1. In the interfaced versions of STABLE, the input distribution `dist` is converted to the new parameterization, and a new distribution descriptor is returned.

5 Error/return codes

An error is unrecoverable and stops execution. For example, if you ask to compute the density of a stable parameter with $\alpha = 3$, you will get a return code of 1 and your function will stop. In contrast, a warning is informational and is usually not serious. It alerts you to the fact that the results of a calculation may have some inaccuracy. For example, stable densities have radical changes of the tail behavior when $\alpha = 2$ or $\beta = \pm 1$, and the computations have small inaccuracies in them. In practical terms this usually means little, as the difference between an $\alpha = 1.99$ stable distribution and an $\alpha = 2$ stable distribution in an statistical problem is likely to be unobservable in practice.

Return codes for STABLE program are given in the tables below. Univariate routines return error codes in the range 1-99, multivariate routines return error codes in the range 100-199.

code	type	meaning
0		No error
1	error	Invalid input parameter
2	error	alpha parameter outside of tabulated values in QKSTABLE
3	error	Too many data points for internal array
4	error	Error computing the likelihood, e.g. pdf=0
5	warning	Possible approx. error while using QKSTABLE for alpha or beta near boundary
6	warning	Possible error in confidence intervals because parameter is near boundary
7	warning	alpha and/or beta rounded to a special value, adjust tol(4)
8	warning	alpha is at lower bound for search, may not have found best value for alpha
9	error	Too many bins (distinct possible values) in sdiscretemle
10	error	beta must be 0 to use this function
11	error	beta near +1 or -1 does not work in this function
12	error	sinc error in sfitfracmoment
13	error	Internal error in sfitlogabs
14	error	Data value near zero in sfitfracmoment or sfitlogabs
15	error	Error in subroutine
16	error	Internal error while computing derivatives
17	error	f(a) and f(b) have the same signs
18	error	Too many function evaluations
19	error	Not enough memory
20	error	X zero value
21	error	Internal error in quickstable
30	error	Two parameterization is required in skewed case

Table 1: Univariate error codes

Warning code 7 can arise in several ways. The purpose of this warning is to avoid numerical problems in internal calculations that can occur near the boundary in the parameter space or to use special cases to increase speed, but to let the user know that something nonstandard is being done. In the following discussion, let $\epsilon =$ the value of tolerance(4). The default value is $\epsilon = 0.01$. (You can change the value of tolerance(4) by using the function `stablestolerance` above, and query it's value by using function `stablegettolerance`. The default value was picked in an ad hoc way; you can make it smaller, even 0, if you wish to calculate certain quantities in one of the cases below. But be aware that numerical errors may arise.) Special cases where warning code 7 occur are:

1. α near 2: if $\alpha \in (2 - \epsilon, 2)$, then α is set to 2 and β is set to 0.
2. Near $\alpha = 1$ but not Cauchy: if $|\alpha - 1| < \epsilon$ and $|\beta| \geq \epsilon$, then α is set to 1 and β is left unchanged. This is to avoid computations involving $\beta \tan(\pi\alpha/2)$, which blows up as $\alpha \rightarrow 1$ if $\beta \neq 0$.
3. Near Cauchy case: if $|\alpha - 1| < \epsilon$ and $|\beta| < \epsilon$, then α is set to 1 and β is set to 0.
4. Near Lévy case: if $|\alpha - 1/2| < \epsilon$ and $|\beta - 1| < \epsilon$, then α is set to 1/2 and β is set to 1; if $|\alpha - 1/2| < \epsilon$ and $|\beta + 1| < \epsilon$, then α is set to 1/2 and β is set to -1.

code	type	meaning
101	error	Invalid input parameter
102	warning	Accuracy warning, $\alpha < 1$
103	warning	vmax exceeded in mvstablepdf
104	error	Too many points in spectral measure
105	error	nspectral must be divisible by 4
106	error	This parameterization is not allowed in this function
107	error	Too few uniform(0,1) input values for simulation
108	error	Distribution not defined
109	error	mvstablecdf not implemented for nonsymmetric case
110	error	Matrix is not positive definite
111	error	alpha must be at least 0.8
112	error	Definition error
113	error	Dimension is greater than the max allowed
115	error	Spline error
150	error	Not enough memory
151	error	Error in a subroutine

Table 2: Multivariate error codes

References

- Abdul-Hamid, H. and J. P. Nolan (1998). Multivariate stable densities as functions of one dimensional projections. *J. Multivar. Anal.* 67, 80–89.
- Byczkowski, T., J. P. Nolan, and B. Rajput (1993). Approximation of multidimensional stable densities. *J. Multivar. Anal.* 46, 13–31.
- Chambers, J., C. Mallows, and B. Stuck (1976). A method for simulating stable random variables. *Journal of the American Statistical Association* 71(354), 340–344.
- Fan, Z. (2006). Parameter estimation of stable distributions. *Communications in Statistics. Theory and Methods* 35, 245–256.
- Kogon, S. and D. Williams (1998). *Characteristic function based estimation of stable parameters*. In R. Adler, R. Feldman, and M. Taqqu (Eds.), *A Practical Guide to Heavy Tailed Data* pp. 311–338. Boston, MA: Birkhäuser.
- McCulloch, J. H. (1986). Simple consistent estimators of stable distribution parameters. *Communications in Statistics. Simulation and Computation* 15, 1109–1136.
- Nikias, C. L. and M. Shao (1995). *Signal Processing with Alpha-Stable Distributions and Applications*. New York: Wiley.
- Nolan, J. P. (1997). Numerical calculation of stable densities and distribution functions. *Commun. Statist. -Stochastic Models* 13, 759–774.
- Nolan, J. P. (2001). Maximum likelihood estimation of stable parameters. In O. E. Barndorff-Nielsen, T. Mikosch, and S. I. Resnick (Eds.), *Lévy Processes: Theory and Applications*. Boston: Birkhäuser.
- Nolan, J. P. (2010a). Multivariate elliptically contoured stable distributions: theory and estimation. Submitted.
- Nolan, J. P. (2010b). *Stable Distributions - Models for Heavy Tailed Data*. Boston: Birkhäuser. In progress, Chapter 1 online at academic2.american.edu/~jpnolan.
- Nolan, J. P. and D. Ojeda (2010). Linear and nonlinear regression with stable errors. Preprint.
- Nolan, J. P., A. Panorska, and J. H. McCulloch (2001). Estimation of stable spectral measures. *Mathematical and Computer Modelling* 34, 1113–1122.

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

Nolan, J. P. and B. Rajput (1995). Calculation of multidimensional stable densities. *Commun. Statist. - Simula.* 24, 551–556.

Index

Excel functions, 6–14, 17–22
 mvstablefitelliptical, 19
 mvstablerndeelliptical, 19
 stablecdf, 5
 stableconvert, 11
 stablefit, 6
 stablefitmleci, 8
 stablegettolerance, 11
 stableinv, 5
 stablepdf, 5
 stableqcdf, 12
 stableqkinv, 13
 stableqkpdf, 12
 stableregression, 10
 stablernnd, 6
 stablesettolerance, 10
 stableversion, 10