Stable Distributions

Abstract

In the pages set out below we explore *stable* distributions. A special case is the normal distribution. Stable distributions are also called *Levy stable* or (if not normal) *stable Paretian* distributions. Stable distributions other than the normal distribution are fat-tailed and may be skewed. As with the normal distribution there are theoretical justifications for using them to model financial data, if the returns on the exposures being modelled can be expected to arise from a large number of smaller innovations with suitable characteristics.

Contents

1. Introduction

[StableDistributions1]

1.1 Stable distributions are a class of probability distributions that have interesting theoretical and practical properties that make them potentially useful for modelling financial data. In a sense that we will explore further below, they generalise the Normal distribution. They also allow fat tails and skewness, characteristics that are also frequently observed in financial data. Traditionally they have been perceived to be subject to the practical disadvantage that they have infinite variances (apart from the special case of the Normal distribution) and thus are not particularly easy to manipulate mathematically. However, more recently, mathematical tools and programs have been developed that simplify such manipulations.

1.2 Whether stable distributions are actually good at modelling financial data is not something that we explore in depth in these pages. Longuin (1993), when analysing the distribution of U.S. equity returns, concluded that their distribution was not sufficiently fat-tailed to be adequately modelled by Levy stable distributions, even if it was fatter tailed than implied by the normal distribution. Moreover, implicit within the theoretical justification for (non-Normal) stable distributions in such a context is an assumption that aggregate returns arise from the combined impact of a large number of smaller independent innovations, so that a generalisation of the Central Limit Theorem applies, see Section 4. Fat-tailed behaviour in the distribution of aggregate returns in line with stable laws can then be expected to arise if it is assumed that each of these smaller innovations is also (suitably) fat-tailed. The challenge is that this is not necessarily how fat tails arise in aggregate return data. Fat tails may instead arise partly or wholly due to distributional mixtures, e.g. regime shifts or time-varying volatility, or from one-off (systemic) ‘shocks’ that cannot be conceptually decomposed in to lots of smaller independent elements, see e.g. Kemp (2009). The latter might include the impact of an aggregate loss of risk appetite (and feedback effects that might then arise because of changed perceptions amongst market participants regarding the views of others).

1.3 The implicit assumption underlying stable distributions referred to in the previous paragraph is revealed by their defining characteristic, and the reason for the term *stable*, which is that they retain their shape (suitably scaled and shifted) under addition. The definition of a stable distribution is that if $X, X_1, X_2, \dots, X_n$ are independent, identically distributed random variables coming from such a distribution, then for every n we have the following relationship for some constants $c_n > 0$ and d_n:

$$X_1 + X_2 + \cdots + X_n \triangleq c_n X + d_n$$

Here $\triangleq$ means equality in distributional form, i.e. the left and right hand sides have the same probability distribution. The distribution is called *strictly stable* if $d_n = 0$ for all n. Some authors use the term *sum stable* to differentiate from other types of stability that might apply.

1.4 Normal distributions satisfy this property, indeed they are the only distributions with finite variance that do so. Other probability distributions that exhibit the stability property described above include the *Cauchy* distribution and the *Levy* distribution.

1.5 The class of all distributions that satisfy the above property is described by four parameters, $(\alpha, \beta, \gamma, \delta)$. In general there are no simple closed form formulae for the probability densities, f, and cumulative distribution functions, F, applicable to these distributional forms (exceptions are the normal, Cauchy and Levy distributions), but there are now reliable computer algorithms for working with them.

2. Parameterisation of stable distributions

[StableDistributions2]

2.1 As noted in Section 1 any specific stable distributional form is characterised by four parameters $(\alpha, \beta, \gamma, \delta)$. Nolan (2005) notes that there are multiple definitions used in the literature regarding what these parameters mean. He focuses there on two, which he denotes by $\mathbf{S}(\alpha, \beta, \gamma, \delta_0; 0)$ and $\mathbf{S}(\alpha, \beta, \gamma, \delta_1; 1)$, that are differentiated according to the meaning given to δ. The first is the one that he concentrates on, because it has better numerical behaviour and intuitive meaning, but the second is more commonly used in the literature. We call the former the '0-parameterisation' and the latter the '1-parameterisation' in these pages.

2.2 In either of these descriptions:

(a) α is the *index* of the distribution, also known as the *index of stability* or *characteristic exponent*, and must be in the range $0 < \alpha \leq 2$. The constant c_n in the formula in Section 1.3 must be of the form $n^{1/\alpha}$;

(b) β is the *skewness* of the distribution and must be in the range $-1 \leq \beta \leq 1$. If $\beta = 0$ then the distribution is symmetric, if $\beta > 0$ then it is skewed to the right and if $\beta < 0$ then it is skewed to the left;

(c) γ is a scale parameter and can be any positive number; and

(d) δ is a location parameter, shifting the distribution right if $\delta < 0$ and left if $\delta > 0$.

2.3In either description, the distributional form is normally defined via the distribution's characteristic function, i.e. the (complex) function $\varphi(u) = \mathrm{E}\left(e^{iuX}\right)$, where $\mathrm{E}(.)$ is the expectation operator. Nolan (2005) uses the following definitions:

(a) A random variable X is $\mathbf{S}(\alpha,\beta,\gamma,\delta_0;0)$ if it has characteristic function

$$\varphi(u) = \begin{cases} -\exp\left(-\gamma^\alpha|u|^\alpha\left(1+i\beta\left(\tan\frac{\pi\alpha}{2}\right)(\mathrm{sign}\, u)(|\gamma u|^{1-\alpha}-1)\right)+i\delta_0 u\right), & \alpha\neq 1 \\ \exp\left(-\gamma|u|\left(1+i\beta\frac{2}{\pi}(\mathrm{sign}\, u)\log(\gamma|u|)\right)+i\delta_0 u\right), & \alpha = 1 \end{cases}$$

(b) A random variable X is $\mathbf{S}(\alpha,\beta,\gamma,\delta_1;1)$ if it has characteristic function

$$\varphi(u) = \begin{cases} -\exp\left(-\gamma^\alpha|u|^\alpha\left(1-i\beta\left(\tan\frac{\pi\alpha}{2}\right)(\mathrm{sign}\, u)\right)+i\delta_1 u\right), & \alpha\neq 1 \\ \exp\left(-\gamma|u|\left(1+i\beta\frac{2}{\pi}(\mathrm{sign}\, u)\log(|u|)\right)+i\delta_1 u\right), & \alpha = 1 \end{cases}$$

2.4 The location parameters are related by:

$$\delta_0 = \begin{cases} \delta_1+\beta\gamma\tan\frac{\pi\alpha}{2}, & \alpha\neq 1 \\ \delta_1+\beta\frac{2}{\pi}\gamma\log(\gamma), & \alpha = 1 \end{cases}$$

or

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

2.5 Nolan (2005) notes that if $\beta = 0$ then the 0-parameterisation and the 1-parameterisation coincide. When $\alpha \neq 1$ and $\beta \neq 0$ then the parameterisations differ by a shift $\beta\gamma\tan\frac{\pi\alpha}{2}$ which gets infinitely large as $\alpha \to 1$. Nolan argues that the 0-parameterisation is a better approach because it is jointly continuous in all four parameters, but accepts that the 1-parameterisation is simpler algebraically, so is unlikely to disappear from the literature.

3. Other features

[StableDistributions3]

3.1 Three special cases of stable laws, which have closed form expressions for their probability densities are:

(a) Normal (i.e. Gaussian). $X\sim N(\mu,\sigma^2) = \mathbf{S}\left(2,0,\sigma/\sqrt{2},\mu;0\right) = \mathbf{S}\left(2,0,\sigma/\sqrt{2},\mu;1\right)$ if X has density

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

(b) Cauchy. $X \sim Cauchy(\gamma, \delta) = \mathbf{S}(1,0,\gamma,\delta;0) = \mathbf{S}(1,0,\gamma,\delta;1)$ if X has density

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

(c) Levy. $X \sim Levy(\gamma, \delta) = \mathbf{S}(1/2, 1, \gamma, \gamma + \delta; 0) = \mathbf{S}(1/2, 1, \gamma, \delta; 1)$ if X has density

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

3.2 Generic features of stable distributions noted by Nolan (2005) include:

(a) They are unimodal

(b) When α is small then the skewness parameter is significant, but when α is close to 2 then it matters less and less.

(c) When $\alpha = 2$ (i.e. the Normal distribution), the distribution has 'light' tails and all moments exist. In all other cases (i.e. $0 < \alpha < 2$), stable distributions have heavy tails and an asymptotic power law (i.e. Pareto) decay. The term *stable Paretian* is thus used to distinguish the $\alpha < 2$ case from the Normal case. A consequence of these heavy tails is that not all *population* moments exist. If $\alpha < 2$ then the population variance does not exist, and if $\alpha \leq 1$ then the population mean does not exist either. Fractional moments, e.g. the p'th absolute moment, defined as $E(|X|^p)$, exist if and only if $p < \alpha$ (if $\alpha < 2$). Of course all *sample* moments exist, if there are sufficient observations in the sample, but these may exhibit unstable behaviour as the sample size increases if the corresponding population moment does not exist.

3.3 Linear combinations of independent stable distributions with the same index, α, are stable. If $X_j \sim \mathbf{S}(\alpha, \beta_j, \gamma_j, \delta_j; k)$ for $j = 1, \dots, n$ then

$$\sum_{j=1}^{n} a_j X_j \sim \mathbf{S}(\alpha, \beta, \gamma, \delta; k)$$

where:

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

$$\beta = \frac{\sum_{j=1}^{n} \beta_j (\operatorname{sign} a_j) |a_j \gamma_j|^\alpha}{\gamma^\alpha}$$

$$\delta = \begin{cases} \sum_{j=1}^{n} \delta_j + \gamma\beta \tan\frac{\pi\alpha}{2} & k = 0, \alpha \neq 1 \\ \sum_{j=1}^{n} \delta_j + \beta\frac{2}{\pi}\gamma \log\gamma & k = 0, \alpha = 1 \\ \sum_{j=1}^{n} \delta_j & k = 1 \end{cases}$$

In this generalisation of the definition of stable distributions given in Section 1.3 it is essential for the α's to be the same. Adding stable random variables with different α's does not result in a stable law.

4. The Generalised Central Limit Theorem

[StableDistributions4]

4.1 The two main reasons why stable laws are commonly proposed for modelling return series are:

(a) The *Generalised Central Limit Theorem*. This states that the only possible non-trivial limit of normalised sums of independent identically distributed terms is stable; and

(b) *Empirical*. Many large data sets exhibit fat tails (and skewness), and stable distributions form a convenient family of distributions that can cater for such features (with choice of α and β allowing different levels of fat-tailed-ness or skewness to be accommodated).

We focus below on the former, since there are other families of distributions that can be parameterised in ways that can fit different levels of fat-tailed-ness or skewness, including ones simpler to handle analytically such as ones with quantile-quantile plots versus the Normal distribution that are polynomials rather than straight lines, see e.g. Kemp (2009).

4.2 The classical Central Limit Theorem states that the normalised sum of independent, identically distributed random variables converges to a Normal distribution. The Generalised Central Limit Theorem shows that if the finite variance assumption is dropped then the only possible resulting limiting distribution is a stable one as defined above. Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random variables. Then there exist constants $a_n > 0$ and b_n and a non-degenerate random variable Z with

$$a_n(X_1 + \cdots + X_n) - b_n \overset{\Delta}{\rightarrow} Z$$

if and only if Z is stable (here $\overset{\Delta}{\rightarrow}$ means tends as $n \to \infty$ to the given distributional form).

4.3 A random variable X is said to be in the *domain of attraction* of Z if there exist constants $a_n > 0$ and b_n such that the equation in Section 4.2 holds when $X_1, X_2, \ldots$ are independent identically distributed copies of X. The Generalised Central Limit Theorem thus shows that the only possible distributions with a domain of attraction are stable distributions as described above. Distributions within a given domain of attraction are characterised in terms of tail probabilities. If X is a random variable with $x^\alpha P(X > x) \to c^+ \geq 0$ and $x^\alpha P(X < x) \to c^- \geq 0$ with $c^+ + c^- > 0$ for some $0 < \alpha < 2$ as $x \to \infty$ then X is in the domain of attraction of an α-stable law. a_n must then be of the form $a_n = an^{-1/\alpha}$.

Nomenclature

[StableDistributionsNomenclature]

α = *index* parameter
β = *skewness* parameter
γ = *scale* parameter

δ = *location* parameter
δ_0 = location parameter if distributional family is parameterised as per the '0-parameterisation'
δ_1 = location parameter if distributional family is parameterised as per the '1-parameterisation'
$\mathrm{E}(.)$ = expectation operator
$\mathbf{S}(\alpha, \beta, \gamma, \delta_0; 0)$ = stable distributional form if parameterised as per the '0-parameterisation'
$\mathbf{S}(\alpha, \beta, \gamma, \delta_1; 1)$ = stable distributional form if parameterised as per the '1-parameterisation'
$\triangleq$ = equality in distributional form
$\overset{\Delta}{\rightarrow}$ = tends to the given distributional form

References

[StableDistributionsRefs]

Kemp, M.H.D. (2009). *Market consistency: Model calibration in imperfect markets*. John Wiley & Sons [for further information on this book please MarketConsistency]

Longuin, F. (1993). Booms and crashes: application of extreme value theory to the U.S. stock market. *Institute of Finance and Accounting, London Business School*, Working Paper No 179

Nolan, J.P. (2005). Modelling financial data with stable distributions.