The Importance of Being Normal

Prepared for the Institute and Faculty of Actuaries' Finance and Investment Conference, Edinburgh, June 2003.

Stephen Carlin and Andrew Smith

"No modelled tail can reproduce 1974, 1987, 1998 and September 2001"

Jeremy Goford (2002)

Abstract

Analysts have used a variety of statistical distributions to model losses due to operational risks [Embrechts, Cruz] or insurance claims [Hogg and Klugman]. In contrast, the normal distribution is the undisputed queen of capital market models.

This paper examines an alternative family of distributions, the conic moment family, for application to capital market models. We answer the following questions:

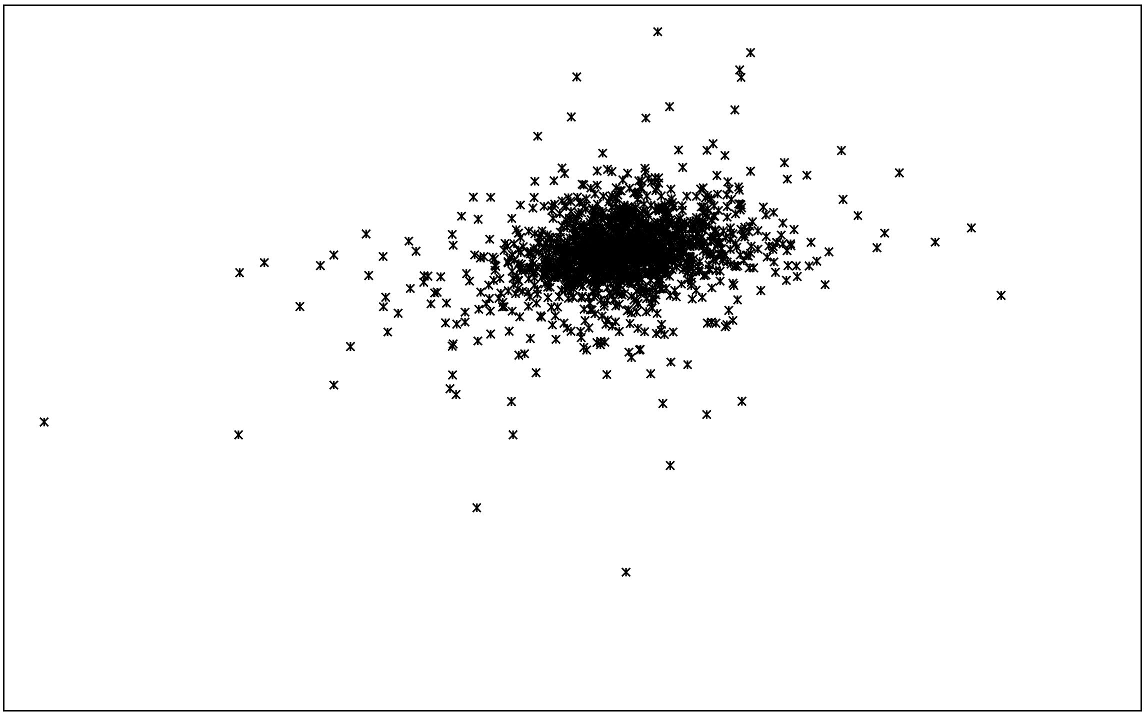
- What is the empirical evidence for normality or for other distributions in financial data?
- For what financial questions do non-normal distributions have a large impact on the answer?
- What tools and diagnostics are available for model calibration?
- How can I model many variables at once?
- How can I incorporate alternative distributions into financial time series models?
- What are the implications for prices of options and of corporate bonds?
- What is the extra effort in deviating from normal assumptions, and is it worth the trouble?

In our first section, we examine some empirical evidence, and fit some distributions. Our conclusions (section 7) summarise the business issues arising from choice of distributions. To get a brief overview of the paper, read just these two sections.

The interior covers the details. This is the meat that backs up our conclusions. We assume that you are already familiar with the normal theory, including multi dimensional problems, matrix manipulations and the logic behind Black and Scholes' formula. We continually make use of parallels and contrasts, between the familiar normal theory and the newer non-normal distributions.

1. Flypaper Models

A long and tall building has a west wall entirely covered with flypaper. A swarm of flies is released, and each fly buzzes according to a three dimensional Brownian motion. A breeze blows the flies gently in the direction of the building. Eventually every fly is stuck somewhere on the flypaper. The picture might look like this:



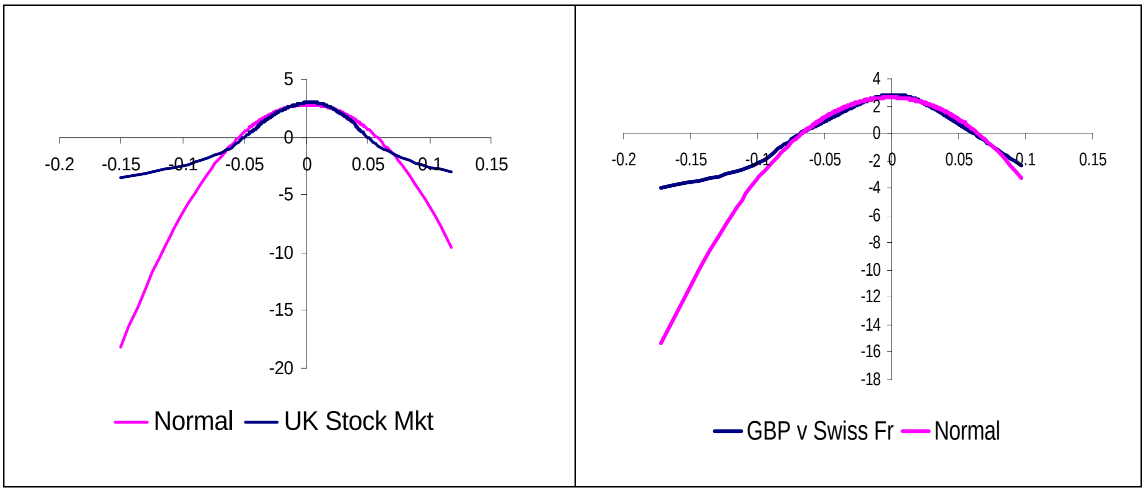
This is an example of a bivariate *conic moment* distribution. Let us suppose in 3-dimensional coordinates that the flies are released at the origin, and the flypaper corresponds to the plane $x = 1$. Let us suppose also that the flies' movements are multivariate normal with mean m and variance-covariance matrix V per unit time.

Arrivals on the flypaper are described by the bivariate density function:

$$f(y,z \mid m,V) = \left(\sqrt{m^T V^{-1} m} + \frac{1}{\sqrt{w^T V^{-1} w}} \right) \frac{\exp\left(m^T V^{-1} w - \sqrt{w^T V^{-1} w}\sqrt{m^T V^{-1} m} \right)}{2p \sqrt{|V|}\, w^T V^{-1} w}$$

where w is the vector $\begin{pmatrix} 1 \\ y \\ z \end{pmatrix}$.

Our plot of flies was in fact a plot of weekly changes in the Belgian and Singapore stock markets. This paper aims to demonstrate that the conic moment distribution is useful for describing these, and most other financial variables. To demonstrate this, we look first to some data.

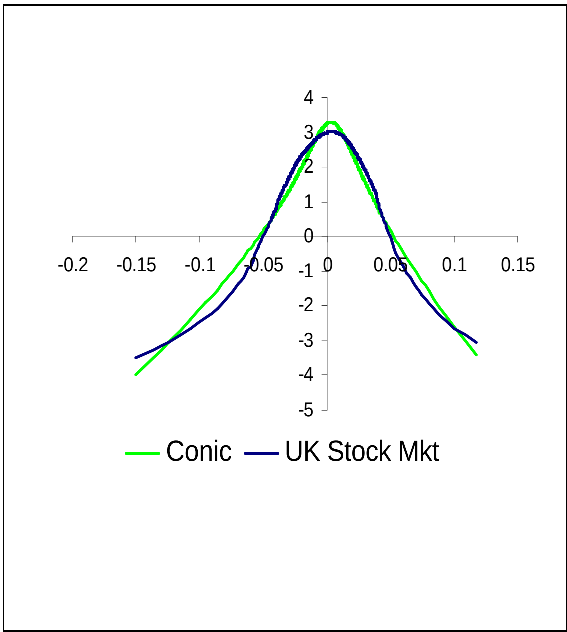


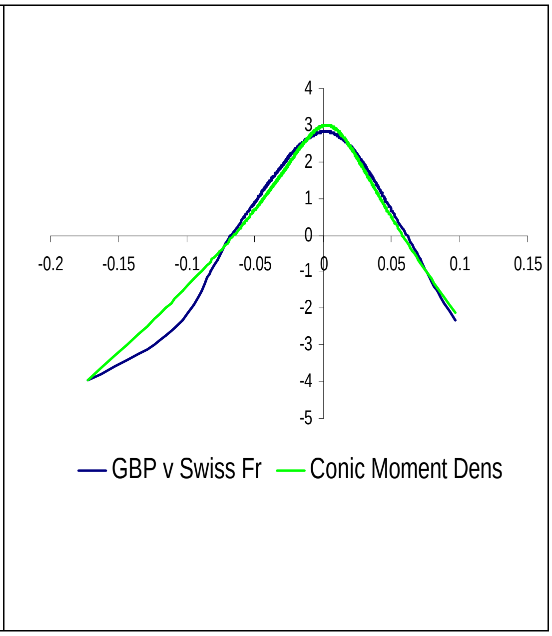
The charts show the distribution of weekly changes in the log of (i) the FTSE 100 share index, (ii) the UK pound / Swiss Franc exchange rate, each measured over the last 20 years. We have applied a kernel algorithm to estimate corresponding densities. On the same chart, we have shown the density of a normal distribution with the same mean and standard deviation. We have used a log scale on the y-axis for clarity.

The historical distributions are clearly bell-shaped, and look superficially like normal distributions. But a plot of empirical distributions next to normal densities, reveals a few systematic deviations from normality. There are more extreme values than we would expect from a normal distribution – the so-called *fat tails* effect. There are more small moves, and fewer intermediate values – the *thin peak* effect. Sometimes the historical density is asymmetric.

Until recently it was conventional to treat the large observations as *outliers*, that is, freak observations that are excluded from a model calibration process, or at least, are treated separately. This is a standard statistical approach, which would be justifiable if these large movements were typographical errors or inaccurate records. But for the financial market moves, this is clearly not the case. The large market moves are well attested; real investors have gained or lost large sums of money. As capital markets have become more skilled at using hedging tools to manage the risks of everyday movements, the remaining exposure to large moves assumes a greater relative importance.

In this paper, we will consider a wider class of distributions, the so-called *Conic Moment* or *Normal Inverse Gaussian* distributions. Ole Barndorff-Nielsen first described this class 1998. The one-dimensional Conic moment distribution has four parameters. As well as a mean and standard deviation we can also fit two other properties, for example skewness and kurtosis. Our idea is to use a single distribution to fit all the data, rather then proposing separate modes for normal movements and outliers. The charts below show the same historic densities, together with best-fit estimates from the conic moment family.





There is no magic to the improved fit – we can of course do better than with a normal distribution just by having two additional parameters to select. We will explain what the parameters mean in the next section.

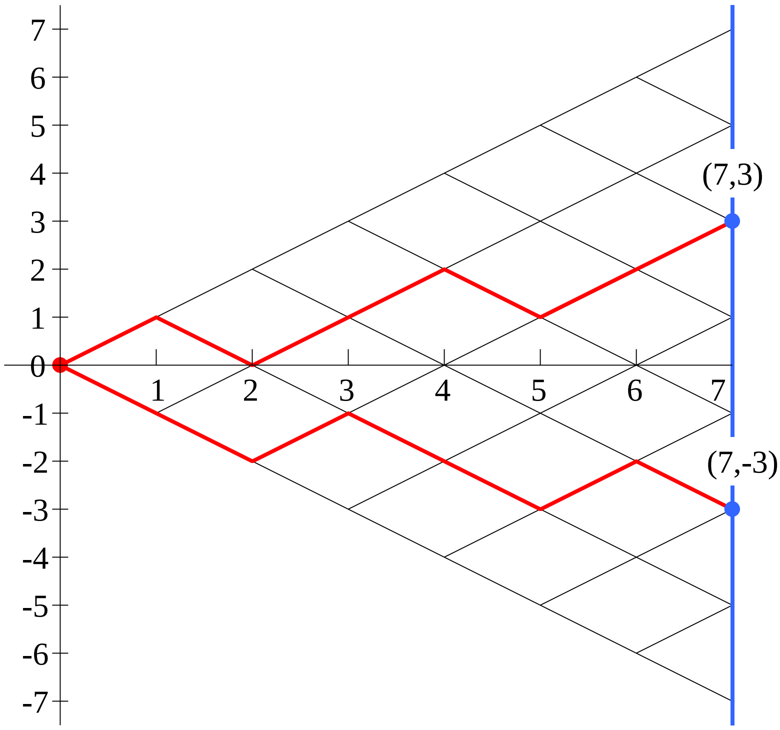
The literature on non-normal distributions in Finance starts with some observations from Mandelbrot in 1963. Since this time, many authors have proposed various families of distributions better to fit historic distributions. Mandelbrot (1963) suggested the Lévy stable distributions; Walter (1990) and Finkelstein (1997) have followed this up in the actuarial literature. Others have suggested log-*F* distributions (Bookstaber and MacDonald, 1987) or Student t-distributions (Blattberg & Gonedes, 1974). Madan (1998) and others have suggested so-called *variance gamma* distributions. Hardy (2001) advocates mixtures of normal distributions. Johnson and Kotz (1970) provide a compendium of other univariate distributions. The use of non-normal residuals has become a standard addition to the list of possible further work at the end of papers based on the normal distribution; see for example Wilkie (1995), Chan (2002).

2. Discrete Conic Moment Processes

This section describes a construction for the discrete conic moment distribution. We derive key results, and compare these to classical results from the binomial distribution. This builds important techniques and insights that are also relevant for the continuous time case.

Binomial Model

We start with a familiar model – the binomial random walk. A particle starts at the origin, and can move up or down with probabilities p_u or p_d respectively. At later dates, the location of the particle will have a binomial distribution. The chart shows two binomial paths, in red, stopped at time t=7 with values Z_7 = 3 and Z_7 = -3.

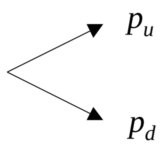


The structure of possible paths is called a *binomial tree*.

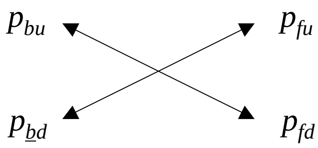
Generalisation - Conic Moments

Our conic moment construction generalises this by allowing our particle to go backwards as well as forwards in time, subject to a net forward drift. We retain the property that future movements in the process are independent of any movements in the past. We can contrast the possible movements under the two models:

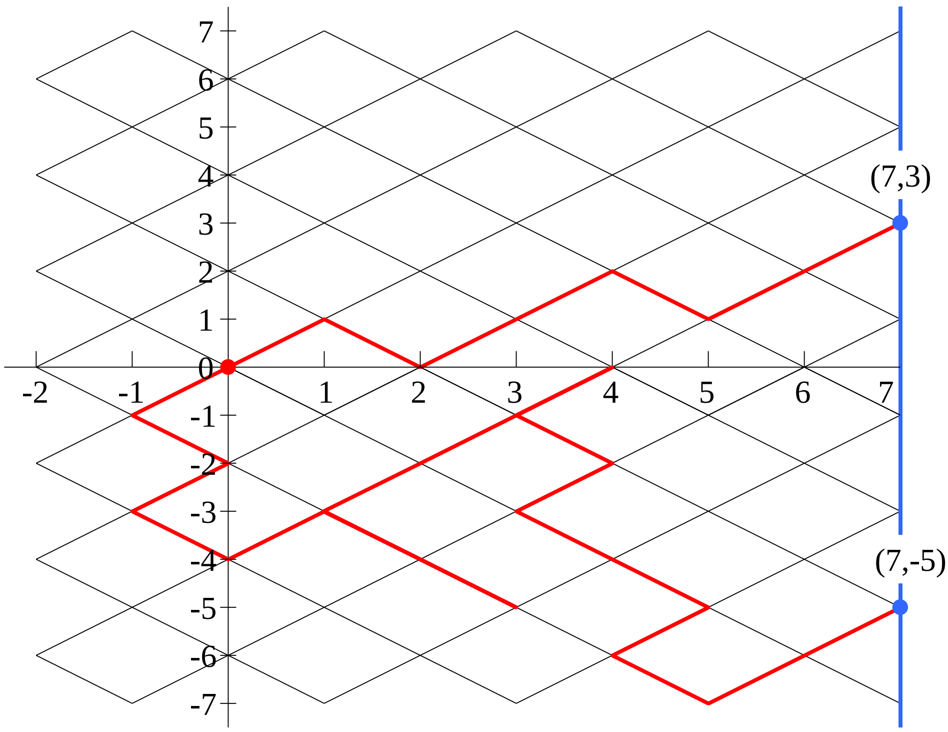
Binomial Tree
$p_u + p_d = 1$



Conic Moment Tree
$p_{bu} + p_{bd} + p_{fu} + p_{fd} = 1$
$p_{fu} + p_{fd} > ½$



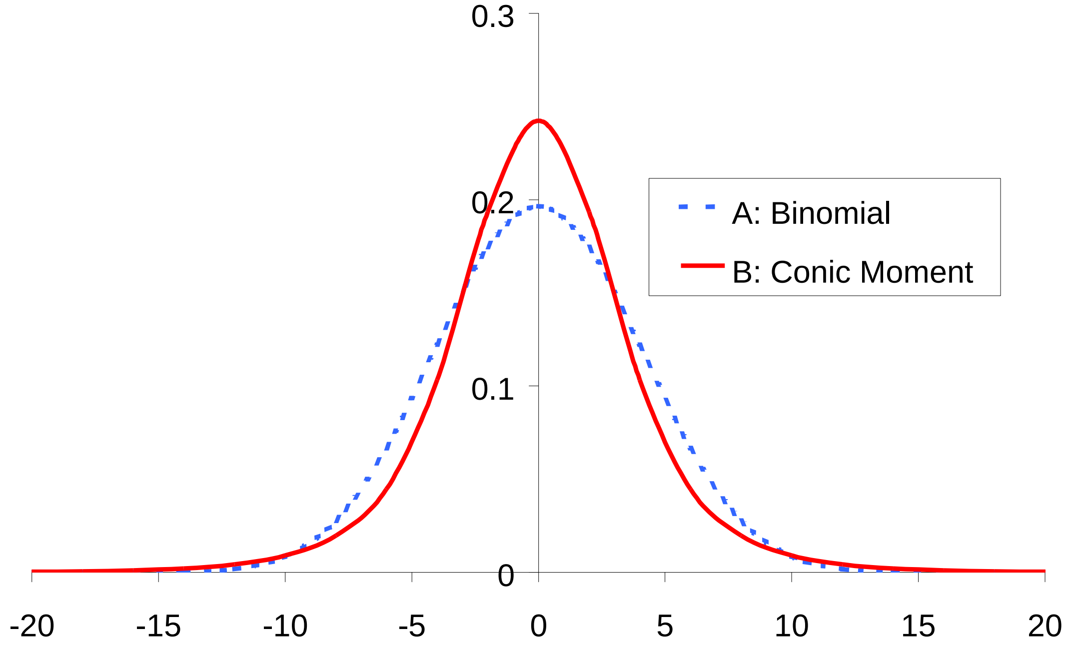
In addition to the forward paths possible in the binomial tree, the conic moment tree also allows paths to loop back on themselves. We observe the conic moment variable Z_t the first time that the particle hits a given barrier t. Two possible paths are shown in the chart below, with the barrier set at t = 7. We observe Z_7 = 3 on the upper path and Z_7 = -5 on the lower path.



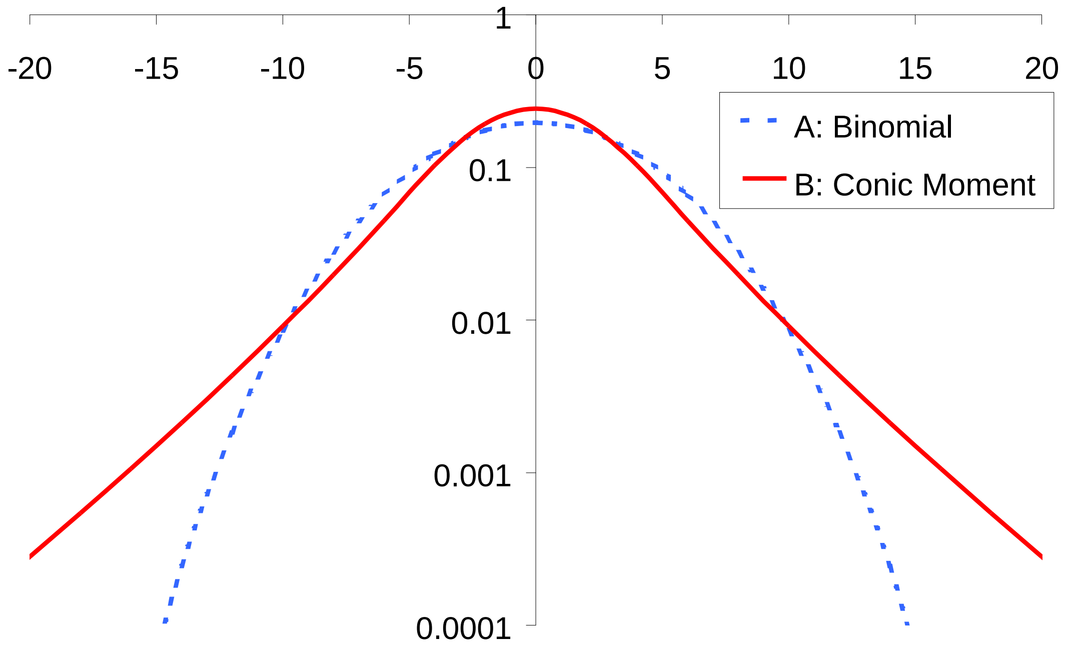
In the binominal case, we could guarantee that $|Z_t| = t$. In contrast, our conic moment construction allows a range of Z from plus to minus infinity, although Z_t must still be an integer of the same parity as t. In this case, we can trivially observe that the conic moment distribution is fatter tailed than the binomial, as the binomial tails vanish for $|Z_t| > t$ but the conic moment tails do not.

Comparison: Conic Moment vs Binomial

The chart below shows a comparison of the binomial and conic moment distributions. The horizontal axis is the random outcome, and the vertical axis is the probability. Curve A is binomial while curve B is a conic moment distribution.



Strictly speaking, we should not have joined the points on this chart, as *Z* can take only even integer values. We chose two distributions with the same mean and standard deviation. The conic moment distribution shows classic fat-tailed behaviour, with greater likelihood attached to very small and very large values. This contrast is clearer if we use a log scale for the probability:

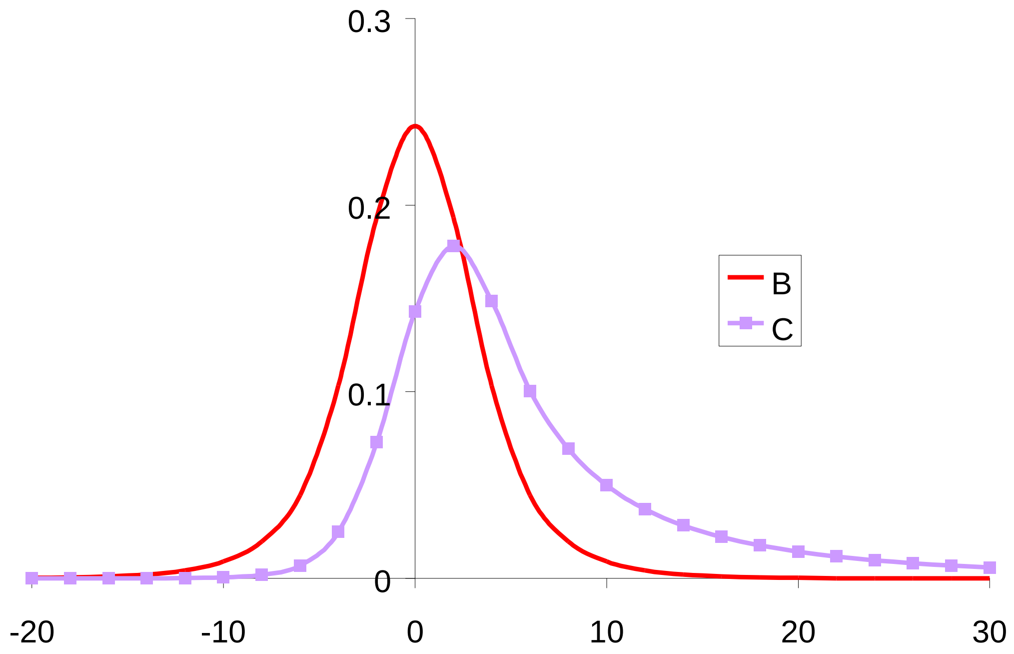


In passing, we also observe what seem to be linear asymptotes of the probability function, which implies exponential tails on both the left and right.

Skew Distributions

We now move on to consider asymmetric distributions. One way of generating these is to use *Esscher Transforms*. To effect an Esscher transform, we must multiply the probability function $p(k)$ by some exponential function u^k and then re-scale so that the probabilities add to 1. If we had started with a binomial distribution, then the transform would produce another binomial distribution with different parameters. The same is true for the conic moment distribution family. The Esscher transform, applied to a conic moment distribution, produces another conic moment distribution. Bühlmann et al (1998) discuss the use of Esscher transforms in option pricing – a theme which we develop further in section 6 of this paper.

Starting from our base case B, we have applied an Esscher transform with $u = 1.25$ to derive the curve C. Curve C is conspicuously right-skewed.

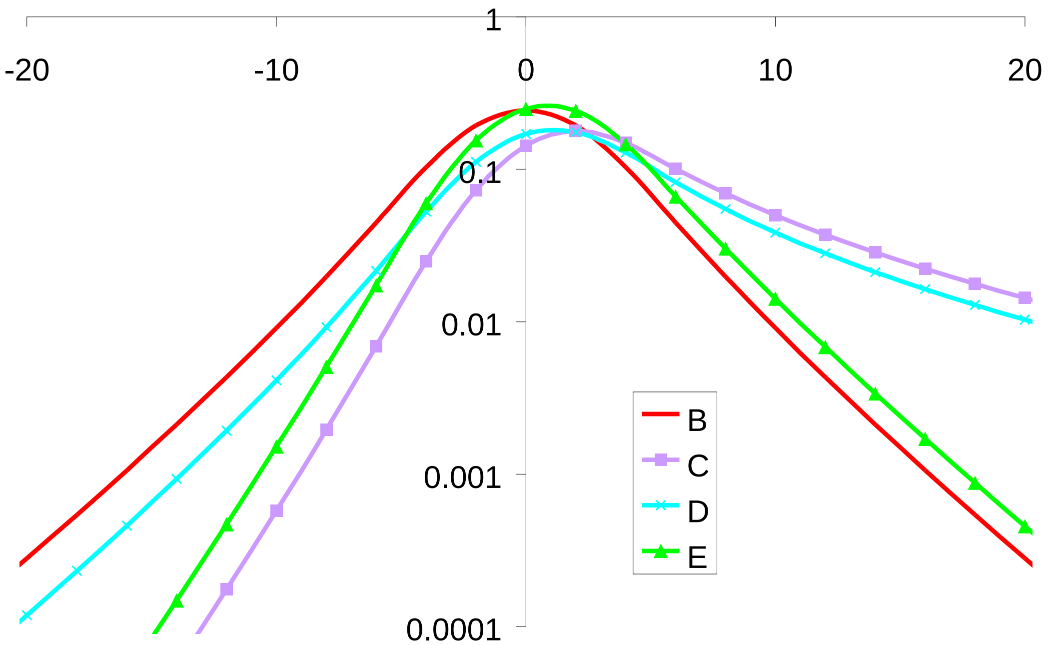


On a log scale, it appears we have derived curve C by adding a linear function to curve B. This is a consequence of the Esscher transform.

To show the flexibility of the conic moment family, we have added two further distributions.

- Curve D has the fattest of all tails, taking its left tail behaviour from curve B and its right tail behaviour from curve C
- Curve E has the thinnest of all tails, taking the left tail of C and right tail of B.

The chart shows densities for all four of these conic moment distributions, this time on a log scale. Once again the asymptotic exponential behaviour is clear.



We generated these curves from the parameters in the table below:

	steps	Probabilities				Moments			
Curve	T	p_{bu}	p_{bd}	p_{fu}	p_{fd}	mean	stdev	skew	kurt
A	16	0.0000	0.0000	0.5000	0.5000	0.00	4.00	0.00	-0.13
B	4	0.1875	0.1875	0.3125	0.3125	0.00	4.00	0.00	2.69
C	4	0.2675	0.1712	0.3422	0.2191	7.16	11.58	3.75	24.88
D	4	0.2527	0.2021	0.3029	0.2423	4.92	10.49	3.86	28.25
E	4	0.1846	0.1477	0.3709	0.2968	1.33	3.60	0.68	2.60

The skewness and kurtosis in this table are both normalised to unit standard deviation.

Probability Generating Function

This section can be skipped on a first reading. We derive the probability generating function (pgf), as defined by the following formula:

$$m(w,t) = \mathbf{E}\left[w^{Z_t}\right] = \sum_{k=-\infty}^{\infty} \mathbf{Prob}\{Z_t = k\} w^k \ .$$

We will see that the pgf is tractable for both the binomial and conic moment distributions. In the binomial case, we can read off the probabilities by repeated differentiation of the pgf at ? =0. Unfortunately, we cannot follow the same procedure for the conic moment distribution as the mgf is a doubly infinite sum and so is not differentiable at ? =0. Nevertheless, the pgf is still useful, for example in calculating moments.

I claim the following formula for the pgf of the conic moment variable Z at time t:

$$\mathbf{E}\left(w^{Z_t}\right) = \left[\frac{2(p_{fu}w + p_{fd}w^{-1})}{1 + \sqrt{1 - 4(p_{bu}w + p_{bd}w^{-1})(p_{fu}w + p_{fd}w^{-1})}}\right]^t$$

This is valid provided ? is in an annulus around the origin in the complex plane where the appropriate expectation converges. For real ? , this is puts ?^{2} in a range where the radicand is positive. We know a band of such ? exists as the radicand is negative at ? = 0 and 8 but positive at ? = 1. If there are no leftward jumps in the construction, then $p_{bu} = p_{bd} = 0$ and we recover the probability generating function for a binomial distribution.

A heuristic derivation of the moment generating function for Z_t is as follows. We start at the point (0,0). There are four possible movements from here. If particle jumps forward in time, then the terminal value of Z will consist of a known upward or downward movement, plus however the path moves in future over the remaining t-1 time periods. In the same way, if the particle jumps backward in time, Z will be the first movement plus the future movements over the remaining t+1 periods. Therefore, we have the following four cases:

? t	? k	Probability	Conditional distribution of Z_t
-1	-1	p_{bd}	-1 + Z_{t+1}
-1	+1	p_{bu}	+1 + Z_{t+1}
+1	-1	p_{fd}	-1 + Z_{t-1}
+1	+1	p_{fu}	+1 + Z_{t-1}

These equations imply the following difference equation for the probability generating function:

$$\begin{aligned}\mathbf{E}\left(w^{Z_t}\right) = m(w,t) = {} & p_{bd}w^{-1}m(w,t+1) \\ & + p_{bu}wm(w,t+1) \\ & + p_{fd}w^{-1}m(w,t-1) \\ & + p_{fu}wm(w,t-1)\end{aligned}$$

It is easy to verify our claimed pgf does indeed satisfy this difference equation. Unfortunately, the equation also has other solutions; for example, take our proposed formula and take the negative root on the denominator. The proof that we have in fact picked the correct solution is difficult, and not given here.

Mean, Variance and Higher Moments

We can use the pgf to derive moments of the conic moment distribution. We list the first three here.

$$m = \mathbf{E}(Z_t) = \frac{B}{A}t$$

$$\mathbf{E}(Z_t - m)^2 = \frac{A^2 - 2ABC + B^2}{A^3}t$$

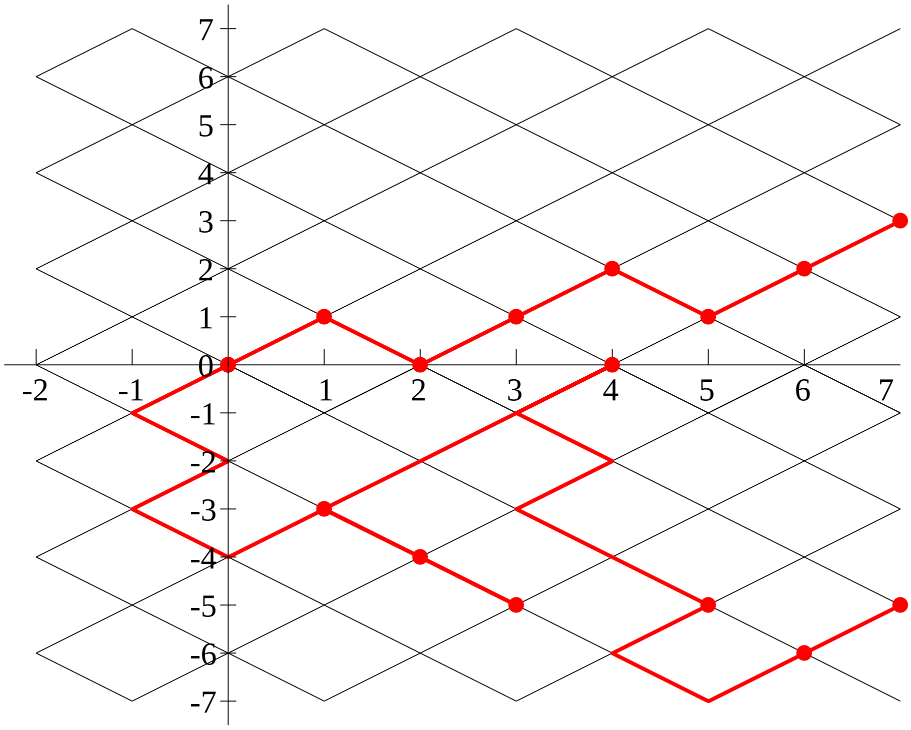
$$\mathbf{E}(Z_t - m)^3 = \frac{-2A^4B - 3A^3C + 3A^2B + 2A^2B^3 + 6A^2BC^2 - 9AB^2C + 3B^3}{A^5}t$$

where

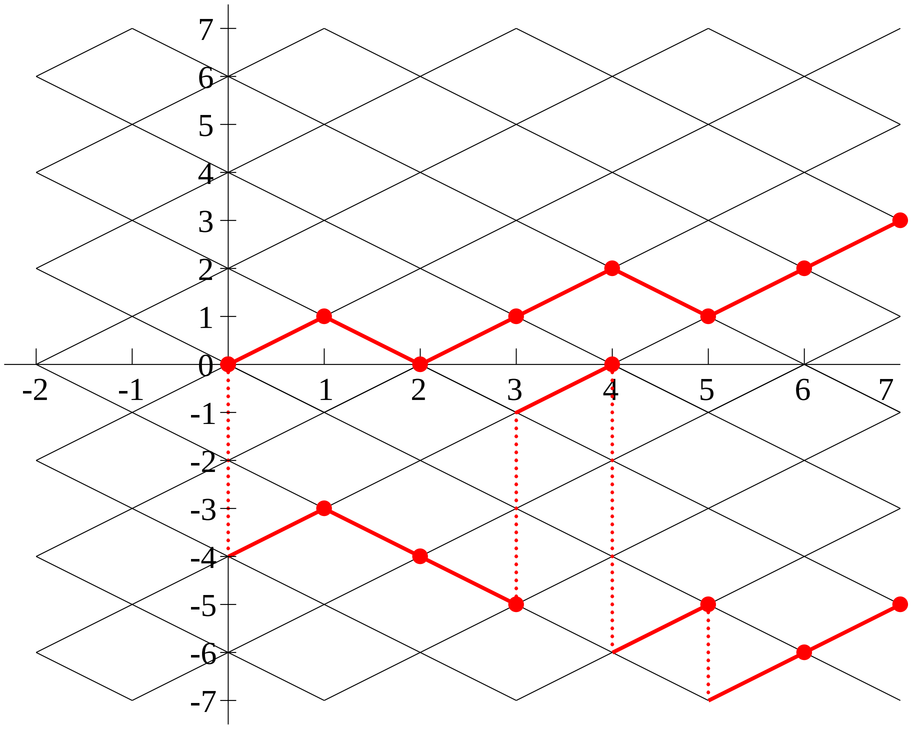
$$A = p_{fd} + p_{fu} - p_{bd} - p_{bu}$$
$$B = p_{bu} + p_{fu} - p_{bd} - p_{fd}$$
$$C = p_{bd} + p_{fu} - p_{bu} - p_{fd}$$

Building a Process

We now extend out construction to build a process. Rather than looking at Z_t for a particular terminal value of t, we can consider the progression of the Z_t collectively over time t. We might use this for example as a model of log stock prices. We have used red circles to outline possible paths of Z_t in the charts below. While the particle itself can visit the same t more than once, conventional stochastic processes do not permit this. Therefore, we drop a red dot only the *first* occasion that the particle hits a new time point. Subsequent visits do not count, until the path re-emerges at a new maximum time.



This path may appear to branch, but actually the path retraced its own steps. On the lower path, we can see how the process Z_t can jump by more than 1 over a single time period. This happens when the particle goes backwards in time and re-emerges somewhere completely different. From the particle's internal perspective, it moves continuously over the lattice. However, time travel is invisible to the outside observer, who sees discontinuities at the point where the particle tunnels back in time. Thus, while the mathematician can imagine the forward and backward paths, the trader can act only on the subset of nodes that create new records on the horizontal axis. The trader's observed paths are shown in the chart below, dotted lines representing the discontinuous jumps.



The new path has the important property of *independent increments*. Let us suppose that the path has just reached a new maximum t=4. Then, from here, the path follows the same probability law as it did at (0,0), but commencing from a new starting point. The future path has no memory of how it got to its starting point. This implies that the increment $Z_7 - Z_4$ is independent of Z_4, and has the same distribution as Z_3. In particular, Z_7 has the same distributions as a sum of independent samples of Z_3 and Z_4. We could also have deduced this fact from the form of probability generating function, namely that $m(?\,, t) = m(?\,, 1)^t$.

Lessons from Discrete Time

Here are some of the key lessons we learn from the discrete time model, which, we will see, translate readily to the continuous case:

- There is a wide range of useful distributions showing fat tailed and skewed behaviour, while retaining smooth, unimodal distribution functions.
- The chosen family includes the familiar binomial distribution as a special case. When we move to continuous limits, we will see that the limiting normal distribution retains a special place, but is supplemented by fat tailed and skew distributions.
- The conic moment distribution has exponential tails, consistent with the empirical literature on investment returns. The thickness of the left and right tails can be adjusted by changing the model parameters.
- The greater generality comes at a price, that is, the loss of an analytical formula for the probability function. However, we can still summarise the probability generating function in closed form, and so derive expressions for mean, variance and higher moments.
- The conic moment family of distributions can be used to describe the marginal distributions of a random walk process with independent increments – and so may be suitable for modelling the movements of log prices over time.

- The binomial lattice allows movements of ± 1 at each step. When used as a model of prices, this has important implications for risk management, as it limits the extent to which prices can move before management action is taken. In contrast, the conic moment processes impose no limit on price moves in a single time period, and so expose risks arising from market jumps.
- In particular, the binomial model of option pricing relies crucially on price movements being small. The classical hedge construction fails in the general conic moment case. Additional assumptions are required in order to derive option prices.

3. The Continuous Conic Moment Distribution

We now generalise the discrete conic moment distribution to continuous space and time.

Our construction is as follows. Let us consider a bivariate random walk $W_a = \begin{pmatrix} W_a^{(0)} \\ W_a^{(1)} \end{pmatrix}$.

This is a bivariate normal random walk in continuous time $a = 0$. We denote its drift per unit time by a matrix $m = \begin{pmatrix} m_0 \\ m_1 \end{pmatrix}$ at its variance–covariance matrix per unit time by $V = \begin{pmatrix} V_{00} & V_{01} \\ V_{01} & V_{11} \end{pmatrix}$, which must be positive semi-definite. We define A as the first time a when the first coordinate $W_a^{(0)}$ hits 1, that is:

$$A = \min\{ a = 0;\ W_a^{(0)} = 1 \}$$

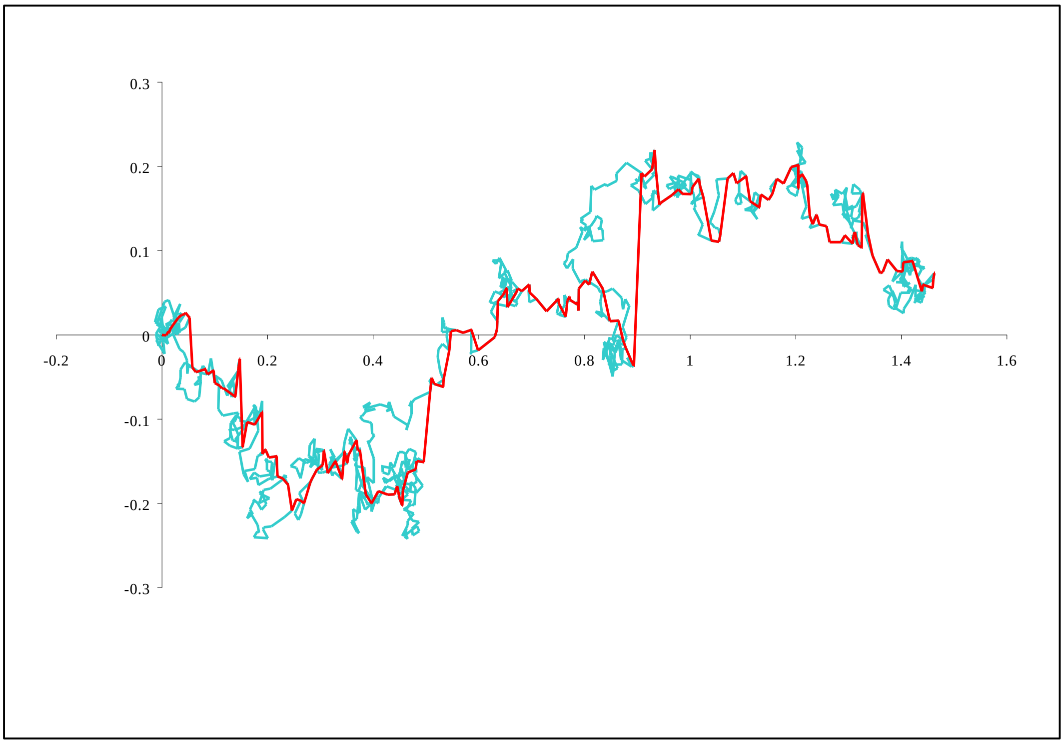
We need to be sure that A is finite. This means one of the following must apply.

- either $W_a^{(0)}$ has positive drift $m_0 > 0$

- or $W_a^{(0)}$ has zero drift $m_0 = 0$ and positive volatility $V_{00} > 0$.

Let us consider the value of W_A at time A. As the random walk has continuous paths, we can be sure that $W_A^{(0)} = 1$. The other coordinate $W_A^{(1)}$ defines the conic moment distribution.

Figure 3.1: Bivariate random walk and conic moment variable
We denote the cumulative distribution function by ? , where:



$$\textbf{Prob}\left\{W_A^{(1)} \le k\right\} = \Psi\left[k \left| \begin{pmatrix} m_0 \\ m_1 \end{pmatrix}, \begin{pmatrix} V_{00} & V_{01} \\ V_{01} & V_{11} \end{pmatrix} \right.\right]$$

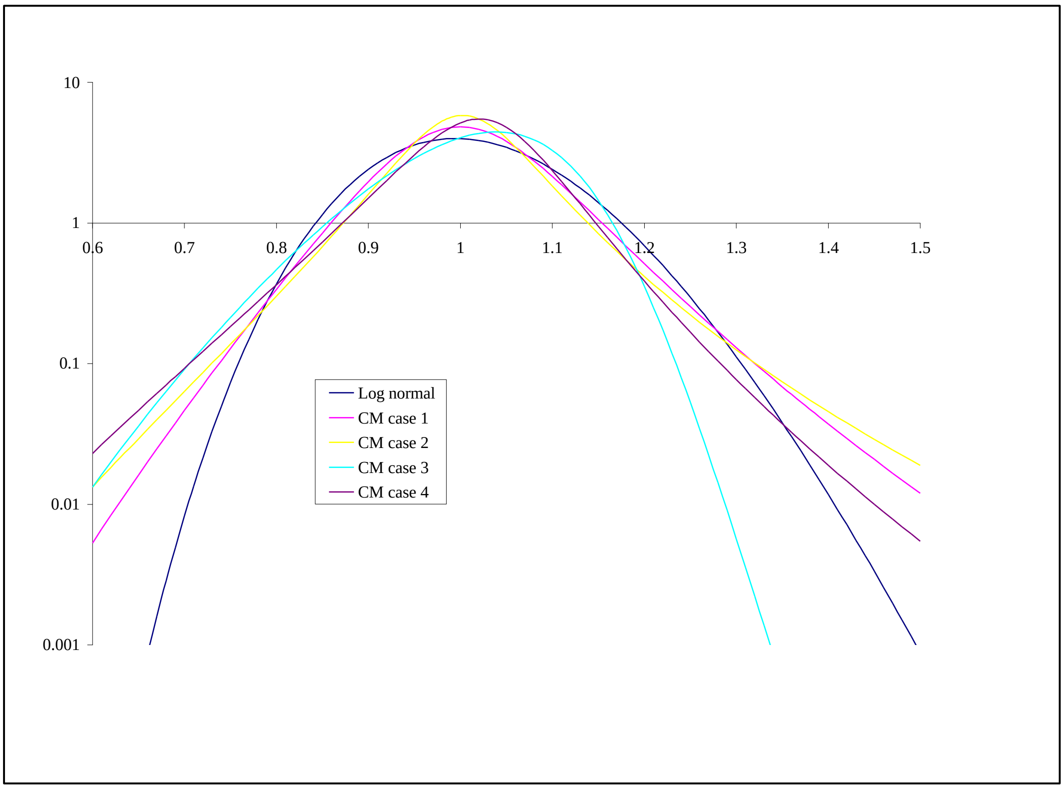
This is a flexible distribution family. It is closed under translating and scaling, but we can also set skewness and kurtosis subject to some constraints.

The conic moment distribution takes values on the whole real line, while asset prices are positive. We therefore propose a log transformation to link our asset price model to the conic moment distribution. We ensure asset prices are positive by writing:

$$S = F \exp(X)$$

where F is the forward price and X has a conic moment distribution.

Figure 3.2: Density plots (log scale) for example distributions:



Bivariate Distribution

We also define a bivariate conic moment distribution. This will allow us to build joint models of a share price and deflator. We derive the distribution from a three-dimensional random walk W_a, with drift m and variance–covariance V per unit time. As before, we wait for the first time A that the 0th component $W_a^{(0)}$ hits 1, and read off the other two components as the conic moment distribution.

Happily, the two dimensional conic moment distribution has a tractable density function, as follows:

$$f(x_1, x_2 \mid m, V) = \left(\sqrt{m^T V^{-1} m} + \frac{1}{\sqrt{w^T V^{-1} w}}\right) \frac{\exp\left(m^T V^{-1} w - \sqrt{w^T V^{-1} w}\sqrt{m^T V^{-1} m}\right)}{2p\sqrt{|V|}\, w^T V^{-1} w}$$

Here:

- m is a 3 dimensional vector $m = \begin{pmatrix} m_0 \\ m_1 \\ m_2 \end{pmatrix}$ with $m_0 = 0$

- V is a 3 × 3 positive definite symmetric matrix $V = \begin{pmatrix} V_{00} & V_{01} & V_{02} \\ V_{01} & V_{11} & V_{12} \\ V_{02} & V_{12} & V_{22} \end{pmatrix}$

- w is the vector $w = \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix}$.

Barndorff-Nielsen (1998) gives evidence that this density integrates to 1, and that it does indeed correspond to our random walk construction. It is also possible to build limiting conic moment distributions where V is positive semi-definite. This limiting cases include the normal distribution when $V_{00} = V_{01} = V_{02} = 0$ and m_0=1. The remaining elements of m and V then signal the mean and variance covariance matrix of the bivariate normal distribution.

Under this bivariate distribution, the marginal distribution of X_1 depends only on the vector $\begin{pmatrix} m_0 \\ m_1 \end{pmatrix}$ and matrix $\begin{pmatrix} V_{00} & V_{01} \\ V_{01} & V_{11} \end{pmatrix}$ – a fact we will soon prove using moment generating functions. We can therefore define the univariate conic moment distribution as the marginal distribution of X_1, given the parameters $\begin{pmatrix} m_0 \\ m_1 \end{pmatrix}$ and $\begin{pmatrix} V_{00} & V_{01} \\ V_{01} & V_{11} \end{pmatrix}$. We denote the univariate density by ? ?, using the standard integral definition of marginal density:

$$\Psi'\left[k \middle| \begin{pmatrix} m_0 \\ m_1 \end{pmatrix}, \begin{pmatrix} V_{00} & V_{01} \\ V_{01} & V_{11} \end{pmatrix} \right] = \int_{x_2=-\infty}^{\infty} f(k, x_2 \mid m, V)\, dx_2$$

The corresponding cumulative distribution function is as follows.

$$\mathbf{Prob}\{X_1 \le k\} = \Psi\left[k \middle| \begin{pmatrix} m_0 \\ m_1 \end{pmatrix}, \begin{pmatrix} V_{00} & V_{01} \\ V_{01} & V_{11} \end{pmatrix} \right] = \int_{x_1=-\infty}^{k} \int_{x_2=-\infty}^{\infty} f(x_1, x_2 \mid m, V)\, dx_2\, dx_1$$

Table 3.1 Properties of Univariate Normal and Conic moment Distributions

Property	Normal distribution	Conic moment distribution
Parameters	μ, σ > 0	$m=\begin{pmatrix} m_0 \\ m_1 \end{pmatrix}, V=\begin{pmatrix} V_{00} & V_{01} \\ V_{01} & V_{11} \end{pmatrix}; m_0 \geq 0; V_{00} \geq 0; V_{00}V_{11} \geq V_{01}^2$
Shifting	$Z \sim N(\mu, \sigma^2) \Rightarrow Z+a \sim N(\mu+a, \sigma^2)$	If Z ~ *Conic moment*(m,V) then $Z+a$ ~ *Conic moment*(m',V') where $m'=\begin{pmatrix} m_0 \\ m_1+am_0 \end{pmatrix}, V=\begin{pmatrix} V_{00} & V_{01}+aV_{00} \\ V_{01}+aV_{00} & V_{11}+2aV_{01}+a^2V_{00} \end{pmatrix}$
Scaling	$Z \sim N(\mu, \sigma^2) \Rightarrow cZ \sim N(c\mu, c^2\sigma^2)$	If Z ~ *Conic moment*(m,V) then $cZ+a$ ~ *Conic moment*(m',V') where $m'=\begin{pmatrix} m_0 \\ cm_1 \end{pmatrix}, V'=\begin{pmatrix} V_{00} & cV_{01} \\ cV_{01} & c^2V_{11} \end{pmatrix}$
Independent Addition	If Z and Z' are independent normal, then $Z+Z' \sim N\left(m+m', s^2+(s')^2\right)$	If Z and Z' are independent *Conic moment* and satisfy some additional constraints, then $Z+Z' = Z''$ ~ *Conic moment*(m'', V'')
Bivariate	For any normal distributions Z, Z' and correlation ρ in (-1,1) we can construct a bivariate normal distribution with marginals Z and Z'.	For *conic moment* distributions Z, Z' with $(m_0')^{-1}V_{00}' = m_0^{-1}V_{00}$ we can construct a bivariate *conic moment* distribution with marginals Z and Z', but only for a restricted range of correlations ρ.
Option pricing	Option prices are completely determined by the parameter σ; the mean μ is irrelevant.	Option prices are determined by parameters [x, y] and a third parameter determining investor preferences. Given [x and y], the other two parameters are irrelevant for pricing options

Moment Generating Functions

As in the discrete case, it is not simple to write down a formula for the density of a univariate conic moment distribution. It is often more convenient to work instead with *moment generating functions*. For a univariate random variable Z, the moment generating function is defined as follows:

$$\text{Moment generating function} = \mathbf{E}\exp(pZ) = \int_{-\infty}^{\infty} e^{pz} f(z)dz$$

Often, the integral diverges, for large positive or negative p. So there is a range of p for which this moment generating function is defined. We can define the function not only for real p but for complex numbers as well. At the very least, we know that the moment generating function always exists when p is pure imaginary.

We can characterise a conic moment distribution as one whose log moment generating function is a conic section (that is, a parabola, hyperbola, ellipse or two straight lines). It turns out that the hyperbola can never happen, and the ellipse is a general case. But before launching into the general case, let’s see some better known examples of parabolas and straight lines. In this table, Φ is the cumulative standard normal distribution function.

Table 3.2 Distribution Functions and Moment Generating Functions for Univariate Conic moment Distributions

Distribution	Cumulative distribution function $\mathbf{P}\{Z \le z\}$	Moment generating function $\exp\{\psi(p)\} = \mathbf{E}\exp(pZ)$	Conic section
Normal(μ, σ^2) $\sigma > 0$	$\Phi\left(\frac{z-m}{s}\right)$	$\exp\left(mp + \frac{s^2p^2}{2}\right)$	$\begin{pmatrix}-y\\p\end{pmatrix}^T\begin{pmatrix}0&0\\0&s^2\end{pmatrix}\begin{pmatrix}-y\\p\end{pmatrix}+2\begin{pmatrix}1\\m\end{pmatrix}^T\begin{pmatrix}-y\\p\end{pmatrix}=0$
Inverse Gaussian(a, b) $a > 0$, $b \ge 0$	$\Phi\left(-\frac{a}{\sqrt{2z}}+\sqrt{2bz}\right)$ $+\exp\left(2a\sqrt{b}\right)\Phi\left(-\frac{a}{\sqrt{2z}}-\sqrt{2bz}\right)$ $(z > 0)$	$\exp\left(a\sqrt{b}-a\sqrt{b-p}\right)$ $(p \le b)$	$\begin{pmatrix}-y\\p\end{pmatrix}^T\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}-y\\p\end{pmatrix}+2\begin{pmatrix}a\sqrt{b}\\\frac{1}{2}a^2\end{pmatrix}^T\begin{pmatrix}-y\\p\end{pmatrix}=0$
Cauchy (a, c) [density picture]	$\frac{1}{2}+\frac{1}{p}\tan^{-1}\left(\frac{z-c}{a}\right)$	$\begin{cases}\exp[p(c+ia)] & \mathbf{Re}(p)=0;\mathbf{Im}(p)\ge 0\\ \exp[p(c-ia)] & \mathbf{Re}(p)=0;\mathbf{Im}(p)\le 0\end{cases}$	$\begin{pmatrix}-y\\p\end{pmatrix}^T\begin{pmatrix}1&c\\c&a^2+c^2\end{pmatrix}\begin{pmatrix}-y\\p\end{pmatrix}+2\begin{pmatrix}0\\0\end{pmatrix}^T\begin{pmatrix}-y\\p\end{pmatrix}=0$
Conic moment(m, V) $m_0 \ge 0$ V positive semi-definite	$\Psi\left(z\middle\vert\begin{pmatrix}m_0\\m_1\end{pmatrix},\begin{pmatrix}V_{00}&V_{01}\\V_{01}&V_{11}\end{pmatrix}\right)$	$\exp\left[\frac{m_0+V_{01}p}{V_{00}} - V_{00}^{-1}\sqrt{m_0^2+2(V_{01}m_0-V_{00}m_1)p - (V_{00}V_{11}-V_{01}^2)p^2}\right]$	$\begin{pmatrix}-y\\p\end{pmatrix}^T\begin{pmatrix}V_{00}&V_{01}\\V_{01}&V_{11}\end{pmatrix}\begin{pmatrix}-y\\p\end{pmatrix}+2\begin{pmatrix}m_0\\m_1\end{pmatrix}^T\begin{pmatrix}-y\\p\end{pmatrix}=0$

The formula for the log moment generating function of the conic moment distribution looks like the most general conic section in two variables. We have the quadratic terms, we have the linear terms – but where is the constant? It turns out that the constant has to be zero. This is because, putting $p = 0$, we can see that $\exp(\psi) = \mathbf{E}\exp(0.Z) = \mathbf{E}(1) = 1$. This means that the conic section must pass through the origin – and therefore the constant term can only be zero. From this representation it is simple to verify our tabulated transformations for shifting, scaling, symmetry and independent addition.

Tail Thickness

We lack closed formulas for the cumulative distribution function of a conic moment random variable. What we can do is build some approximations, which are best for large (positive or negative) values of Z.

For large values of Z, we have the following formulas:

$$F(z \mid m, V) \sim \begin{cases} \left|z^{3/2}\right| \exp\left[-(\ldots)\left|z\right|\right] & z \downarrow -\infty \\ 1 - z^{3/2} \exp\left[-(\ldots) z\right] & z \uparrow \infty \end{cases}$$

The corresponding inverse functions are as follows:

$$F^{-1}(p \mid m, V) = \begin{cases} \ldots & p \downarrow 0 \\ \ldots & p \uparrow 1 \end{cases}$$

The Multivariate Case

We have looked at a univariate case, but we often need to build multivariate models too. The conic moment distribution has a natural multivariate version. In the n-dimensional case, we say that a vector Z has a conic moment distribution with parameters m and V if:

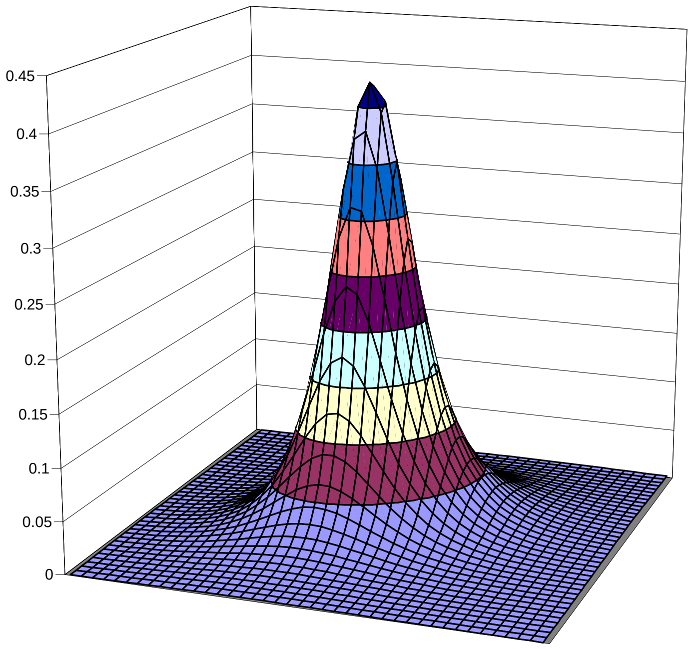
$$\begin{pmatrix} -\log \mathbf{E}\exp(pZ) \\ p \end{pmatrix}^T V \begin{pmatrix} -\log \mathbf{E}\exp(pZ) \\ p \end{pmatrix} + 2m^T \begin{pmatrix} -\log \mathbf{E}\exp(pZ) \\ p \end{pmatrix} = 0$$

This equation must hold for all vectors p (real or complex) where the expectation is finite. This is a multivariate extension of the moment generating function. The matrix V should be an $(n+1)\times(n+1)$ symmetric matrix and the vector m should have $n+1$ elements. We number the elements from 0 to n. Not all combinations of V and m will give rise to valid distributions. Specifically, V should be positive semi definite, and $m_0 \geq 0$, with strict inequality if $V_{00} = 0$.

Density Function

Another way of specifying distributions is by a probability density function. The density for a conic moment variable takes a different algebraic form according to how many dimensions we want to consider. This is inconvenient, particularly when it comes to parameter estimation. Let us focus for a moment on the bivariate distribution of a pair (Y, Z) of random variables – which illustrates the main issues arising. Figure 3.3 shows a plot of this bivariate density.

Figure 3.3: Plot of bivariate density:



The two dimensional case (other than the limiting normal, inverse Gaussian and Cauchy) distributions is as follows:

$$f(y,z\,|\,m,V)=\left(\sqrt{m^TV^{-1}m}+\frac{1}{\sqrt{w^TV^{-1}w}}\right)\frac{\exp\left(m^TV^{-1}w-\sqrt{w^TV^{-1}w}\sqrt{m^TV^{-1}m}\right)}{2p\sqrt{|V|}w^TV^{-1}w}$$

In this representation, $w=\begin{pmatrix}1\\ y\\ z\end{pmatrix}, m=\begin{pmatrix}m_0\\ m_1\\ m_2\end{pmatrix}$ are vectors, with $m_0 > 0$. The matrix $V=\begin{pmatrix}V_{00} & V_{01} & V_{02}\\ V_{01} & V_{11} & V_{12}\\ V_{02} & V_{12} & V_{22}\end{pmatrix}$ is a positive definite symmetric matrix. Wherever possible, we will use vector and matrix notation to spare us lengthy algebra. We can see that this density is unchanged if we scale both m and V by a common scalar. Without loss of generality therefore we might take $m_0 = 1$. This leaves 8 parameters to this bivariate distribution- twice as many as for the univariate case. The one thing we won't prove here is that this truly is a probability density function – you have to turn to Barndorff-Nielsen (1998) for that, or integrate it yourself.

Moment Generating Function

Reconciling the Density

Here's a chance for the eager reader to test their new knowledge. Try to verify that our stated probability density function really does imply the moment generating function we claim. Then you have to show that

$$\begin{pmatrix} -y \\ p_x \\ p_y \end{pmatrix}^T V \begin{pmatrix} -y \\ p_x \\ p_y \end{pmatrix} + 2m.\begin{pmatrix} -y \\ p_x \\ p_y \end{pmatrix} = 0 \text{ where } \exp\lfloor y\,(p_y, p_z)\rfloor = \mathbf{E}\exp\left(p_y Y + p_z Z\right)$$

This quadratic will have 2 real roots, or none, for general choices of (p_y, p_z). I claim we have the right root if $m_0 - V_{00}\psi + V_{01}p_x + V_{02}p_y > 0$. To prove these results, we need to show that for such a ψ:

$$\exp(y\ ) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(p_y y + p_z z)\, f(y, z \mid m, V)\, dydz$$

We won't write out the full proof here, but here's a hint for the keen reader. Define: $\tilde{m} = m + V\begin{pmatrix} -y \\ p_x \\ p_y \end{pmatrix}$ and show that $\tilde{m}^T V^{-1}\tilde{m} = m^T V^{-1} m$. Then demonstrate that $f(y, z \mid \tilde{m}, V) = \exp(-y\ + p_y y + p_z z)\, f(y, z \mid m, V)$. The result follows on integrating both sides over all y and z.

4. Random Walks

The normal distribution is interesting because of the way we can build up normally distributed processes in continuous time. The increments of a normal random walk (or Brownian motion) are normally distributed. The mean is zero, and the variance of an increment is proportional to the length of time over which the increment is measured. Increments over non-overlapping periods are independent.

This construction works because of the reproductive property of the normal distribution. The sum of two independent normal distributions is still normal. The variance of the sum is the sum of the variances.

Conic moment distributions allow a similar construction. The increments of a conic moment random walk have conic moment distributions. Increments over disjoint intervals are independent. Intervals of a given length all share a common increment distribution. A random walk like this, in continuous time, is called a Lévy process. The standard reference on Lévy processes is the book by Sato (1999).

The conic moment random walk works because the sum of two or more independent samples from a conic moment distribution still has a conic moment distribution, although with different parameters. Conversely, for any conic moment distribution, we can find some other conic moment distribution which, when added to itself (say) 61 times, gives the distribution we first thought of.

5. Calibrating a Conic moment Distribution

How can we calibrate a suitable conic moment distribution, given a sample of data?

We offer two approaches. The method of moments is easy to execute, but has poor sampling properties. Maximum likelihood has best sampling properties but requires messy numerical methods to implement.

Let us start with the method of moments. Let Z be a random variable where $\mathbf{E}(Z^4)$ is finite. The first four *semi-invariants* are then defined as follows:

$$
\begin{aligned}
k_1 &= \mathbf{E}(Z) \\
k_2 &= \mathbf{E}\left[(Z-k_1)^2\right] \\
k_3 &= \mathbf{E}\left[(Z-k_1)^3\right] \\
k_4 &= \mathbf{E}\left[(Z-k_1)^4\right] - 3k_2^2
\end{aligned}
$$

For a univariate conic moment process, we can relate these to the matrix parameters m and V. We can deduce the expressions by differentiating the moment generating functions repeatedly.

Univariate Conic moment Distribution: Moments

To calculate m and V:

$$
m = l\begin{pmatrix} 9k_2^3 \\ 9k_1k_2^3 \end{pmatrix}
$$

$$
V = l\begin{pmatrix} 3k_2k_4 - 4k_3^2 & 3k_2(k_1k_4 - k_2k_3) - 4k_1k_3^2 \\ 3k_2(k_1k_4 - k_2k_3) - 4k_1k_3^2 & 9k_2^4 + 3k_1^2k_2k_4 - 6k_1k_2^2k_3 - 4k_1^2k_3^2 \end{pmatrix}
$$

Check that: $\kappa_2 > 0$ (this happens automatically)
$\kappa_4 \geq 0$

$$
|k_3| \leq \sqrt{\frac{3}{5}k_2k_4}
$$

The parameter λ is an arbitrary positive scalar.

Let us now consider the bivariate conic moment distribution. We now need to estimate a three dimensional vector parameter m and a three-by-three matrix V. As far as possible, we wish to populate these from the corresponding parameters from the marginal distributions. But we encounter two problems. The first problem is that each marginal distribution gives values for m_0 and V_{00}. The second problem is that we still need to determine the cross term V_{12}.

Let us tackle first the problem of common estimates for m_0 and V_{00}. We have freedom to scale m and V by some constant λ. So our real problem is two estimates of the ratio V_{00}/m_0. From our calibration equations, we know that each marginal distribution must satisfy the following equation:

$$\frac{k_4}{k_2^2} = 3\frac{V_{00}}{m_0} + \frac{4k_3^2}{3k_2^3}$$

This gives us a minimum value for the fourth semi-invariant, and also implies a quadratic relationship between the third and fourth semi invariants. This same quadratic curve must apply for all the marginals of a multivariate conic moment distribution. We do not therefore have the flexibility to set marginals independently of each other. In practice, for example, we might try to estimate the semi-invariants for each component, and then re-estimate the fourth semi-invariant by choosing a best-fit quadratic. In this case, we must bear in mind that the third semi-invariant should not get too large; we must have $\left|\frac{k_3}{k_2^{3/2}}\right| \le 3\sqrt{\frac{V_{00}}{m_0}}$. This is required in order to satisfy the univariate constraint $|k_3| \le \sqrt{\frac{3}{5}k_2 k_4}$.

Let us turn now to the calibration of the cross terms in the *V* matrix. Denoting our conic moment vector by *Z*, we can show that (for *i*, *j* ≥ 1):

$$\mathbf{Cov}(Z_i, Z_j) = \frac{V_{ij}}{m_0} - \frac{m_i}{m_0^2}V_{oj} - \frac{m_j}{m_0^2}V_{oi} + \frac{m_i m_j}{m_0^3}V_{00}.$$

This then gives us an expression for the whole matrix *V*, given the values in the row and column zero.

One hurdle remains. Unfortunately, we do not have complete freedom to specify our desired covariance structure. For example, if *Y* has a symmetric distribution and *Z* is 100% correlated with *Y*, then *Z*, too, must have a symmetric distribution. There is a limit on how closely two variables can be correlated, depending on their relative skewness. If we try do go beyond these limits, we face problems in the form of a matrix *V* which is not positive definite.

To determine the limit, for each component *i* we define θ_i in (-π/2, π/2) by

$$\frac{k_3}{k_2^{3/2}} = 3\sqrt{\frac{V_{00}}{m_0}}\sin(q_i)$$

The correlation must now satisfy the following bounds:

$$-\cos(q_i + q_j) \le \mathbf{Corr}(Z_i, Z_j) = \frac{\mathbf{Cov}(Z_i, Z_j)}{\sqrt{\mathbf{Var}(Z_i)\mathbf{Var}(Z_j)}} \le \cos(q_i - q_j)$$

Robust Calibration

Sampling error is the perennial challenge for any estimation technique. Estimating fat tails is particularly error-prone, relying as they do on estimation of rare events. There are more robust tools for model estimation, which seek instead to capture the thin peak effect, and so calibrate the shape of a distribution.

The method of moments calibrates unambiguously to the expectations of the first few powers of a variable *Z*. In the case of a conic moment distribution we fit the first four powers – the so-called *test functions*. The sampling problem arises because powers, in particular the fourth power, grow so quickly with Z, making any outlying observations even more extreme.

In contrast, the robust estimation tools use test functions, which are better behaved for large values of z. The complication is that we need to know the distribution before we know which test function to choose. This leads to an iterative process – we start by guessing a distribution, and calculate the test functions. We then use those test functions to calibrate a new distribution, from which we recalculate the test functions. And so the process continues until we find a distribution, which correctly replicates the expected value of its own test functions.

6. Option Pricing

Classical option pricing theory assumes that log share prices perform a normal random walk. The theory gives rise to elegant solutions; most notably Black & Scholes' famous option pricing formula.

How does this theory stand up if we replace the normal random walk with the conic moment random walk? At first blush, we lose everything. The Black-Scholes approach uses a hedging methodology. The hedge only works if share prices move continuously. If the share price can jump, then the hedge breaks down. This is called the "incomplete market" problem. We cannot price options in a conic moment world by appeal to arbitrage arguments alone.

But all is not lost. In this section we will investigate pricing methods for conic moment distributed variables. We use the idea of "no good deal", developed by Cochrane and Saa Requejo (2000). The approach gives us bounds on option prices which are narrow enough to be useful, and collapse to the Black-Scholes formula in the limit of normal distributions. Our findings challenge a popular actuarial prejudice, exemplified by Wise (1999, surreally misquoting Cochrane in support), that pricing in incomplete markets is necessarily arbitrary.

So how can we proceed? One approach is to assume a specific functional form for the deflator, that is, for the stochastic discount factor. In a conic moment context, it is natural to suppose that the log share price and the log deflator come from a bivariate conic moment distribution.

Reasons why hedging may fail

There are many cashflows that we may wish to price, but for which exact hedging instruments are unavailable. Particular issues that may cause difficulty in hedging include:

- cashflows on insurance or savings products containing embedded options in favour of the policyholder;
- cashflows subject to the discretion of a financial institution, for example participation in insurance or pension fund profit sharing arrangements;
- cashflows due after the longest dated comparable bond has expired;
- cashflows depending on non-financial risks, including mortality, sickness, natural catastrophes or operational risks;
- cashflows dependent on behaviour of customers, for example to cash in an investment or pay down a mortgage;
- options or other derivatives on an underlying asset whose price exhibits jump behaviour, thus invalidating hedging algorithms that depend on diffusion processes; and
- cashflows contingent on non-traded macroeconomic quantities such as inflation in wages or retail prices.

For all these cashflows, it is possible to estimate their statistical distribution and correlations with other assets, to some degree of accuracy. What we lack is a market price. The idea shown here is to identify the relationship between price and distribution for traded cashflows, and then to extrapolate this to the non-traded situation.

Consequences of hedge failure

If we remove the ability to trade continuously, or we relax the assumption of normality, then the classical framework falls apart. Dynamic hedging is no longer capable of replicating options. Well-constructed pricing models can still avoid arbitrage, but this constraint is now insufficient to force a unique option price. Arbitrage arguments are still valid, but inconclusive. These points have generated some criticism of the wider applications of option-pricing theory, particularly in applications where the theoretical ideals of normal distributions and continuous frictionless trading are hard to support.

This chapter introduces deflators as a tool for cashflow pricing, focusing specifically on the pricing of options. Our option prices are consistent with market prices where these are observable. Classical option prices are recovered as a limiting case, but our tools require neither normal distributions nor continuous trading. The techniques are useful in many situations where no exact replicating hedge exists, but real world distributional assumptions can be estimated. This extends not only to options, but also to the valuation of illiquid liabilities such as can arise in insurance and pension work.

Deflators

We are now in a position to define a deflator. We focus on a single period model, starting now at time 0 and terminating at time 1. A deflator is a positive random variable D such that, for any future cashflow X at time 1, the value today is $\mathbf{E}(DX)$. This is an established technique for analysing market prices and their relationship to cashflow distributions. The existence of a deflator is equivalent to the absence of arbitrage, see Harrison and Kreps (1979) or the opening chapter of Duffie (2001).

In situations where a cashflow has an exact replicating hedge, the deflator will reproduce the price deduced from hedging arguments. See Cochrane (2001) for more on the basics of deflators.

The beauty of the deflator technique is that we can use the same formal expression $\mathbf{E}(DX)$ whether X is traded or not. This is the key template, which we can use to extrapolate market prices to other cashflows.

Fundamentals of Option Pricing

Notation

All our analysis considers options of a one-year maturity, on a stock S, which may pay dividends. The time now is t=0. The price at time zero of a one-year zero-coupon bond is v.

The forward price of the stock is F. In other words, there are counterparties who will agree now to buy or sell the stock for a price F on one year’s time. While nobody knows the value of the stock S in one year’s time, the forward price F is fixed at time zero. In what follows, we shall assume that v, F and the distribution of S are known inputs to the modelling process.

We will use the following notation throughout the chapter:

$$m_S = \mathbf{E}\ln(S) - F$$

$$s_S^2 = \mathbf{Var}(\ln S_1) = \mathbf{E}[\ln(S) - F - m_S]^2$$

$$g_S = s_S^{-3}\mathbf{E}[\ln(S) - F - m_S]^3$$

$$k_S = s_S^{-4}\mathbf{E}[\ln(S) - F - m_S]^4 - 3$$

Based on our earlier observations on the non-normality of returns, and with the objective of seeing the effect of skewness and kurtosis, we will follow these four examples in the rest of this chapter:

	Normal	Case I	Case 2	Case 3	Case 4
μ_S	0.01				
s_S	0.1				
$?_S$	0	0	0	-1	-1
$?_S$	0	2	5	2	5

Table 6.1: Example parameters for quarterly return distributions: Normal distributions and four alternative cases.

In later parts of this paper, we will use F to denote the cumulative standard normal distribution, and ? to denote the cumulative function for the conic moment distribution, which we shall later define.

Option Prices

In this chapter, we consider only European-style options, which cannot be exercised early. A call option is an option to buy the stock for a price *K* in three month's time. The value of the option in one year is then max{ 0, *S* – *K* }. The price of the call option today is some function *Call*(*K*). *K* is called the "strike price".
A put option is an option to sell the stock for *K* >0 in three month's time. The future value is max{ 0, *K* – *S* }, and today's market value of the put option is some function *Put*(*K*).

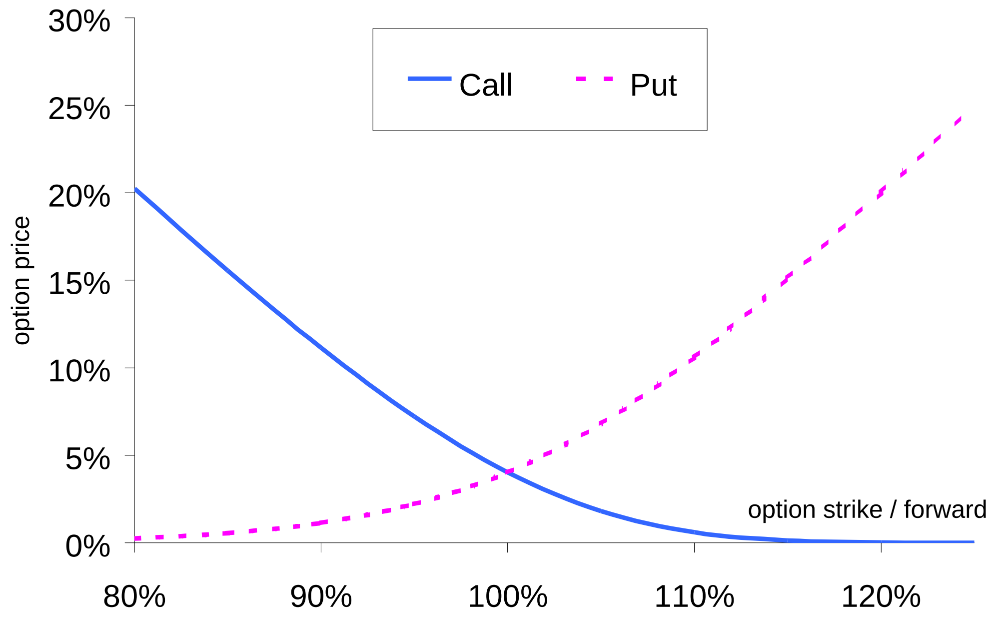


Figure 6.1: Call and put prices for distribution Case 3.

In some markets, the functions *Call*(K) and *Put*(K) are observable or partially observable. In such a situation, analysts may seek to calibrate their pricing models to known option prices. In this chapter, however, we treat the situation where no option prices are observable and we have to propose formulas on the basis of a known distribution of S. An important test of the theory is whether our formulas describe the option prices we really observe.

Arbitrage Bounds on Option Prices

In the Black–Scholes model, hedging arguments lead to option prices. For other distributions, particularly those arising from jump processes, such dynamic hedging arguments do not apply. Nevertheless, there are some weaker bounds that arise from static arbitrage arguments, and which apply irrespective of the distribution of S. Merton (1973) first proposed the arguments in this section.

We will argue one of these bounds; we have tabulated the others, which arise from similar logic. We assume there is no bid–ask spread, so that option prices are the same for buyers and for sellers. For our example, we explain why, always:

$$Put(K) = v \max\{0, K\text{-}F\}$$

Consider two investors. The first investor has a put option. He can decide at time 1 whether to sell a stock S for a price K. He will exercise this option to sell, provided that S ends up below the strike K. Otherwise the option expires worthless. We know that this option is worth *Put*(K).

Our second investor also has an option to sell the share for K at time 1. This investor, however, must make his decision at time 0 not at time 1. He can make this decision not on the basis of the unknown future price S but on the current forward price F. Our second investor has a worthless option if $K < F$, as he would be better off selling the share forward for F. On the other hand, if $K > F$ then the option is worth the discounted different between strike and forward. Thus, taking all the cases, our second investor has an asset worth $v \times \max\{0, K\text{-}F\}$.

It is plain that the first investor has an advantage relative to the second. He has the advantage of time, and he waits another year before deciding to exercise the option. Our claimed inequality reflects the fact that the first investor has a stronger position. The difference between the two sides is sometimes called the "time value" of the option.

These and other familiar option-pricing results are summarised in Table 6.1.

	Call Option	Put Option
Payoff	$\max\{ 0, S\text{-}K\}$	$\max\{ 0, K\text{-}S\}$
Price Bounds	$v \max\{0, F\text{-}K\} = Call(K) = vF$	$v \max\{0, K\text{-}F\} = Put(K) = vK$
Put-Call Parity	$Call(K) = v(F\text{-}K) + Put(K)$	$Put(K) = v(K\text{-}F) + Call(K)$
Black Formula	$Call(K) = vF\Phi\left(\frac{\ln(F/K)}{a} + \frac{a}{2}\right) - vK\Phi\left(\frac{\ln(F/K)}{a} - \frac{a}{2}\right)$	$Put(K) = vK\Phi\left(\frac{\ln(K/F)}{a} + \frac{a}{2}\right) - vF\Phi\left(\frac{\ln(K/F)}{a} - \frac{a}{2}\right)$

Table 6.2: Properties of Option Prices

Implied Volatility

Finally, we define the implied volatility a(K) for a strike price K. Let ? be the cumulative normal distribution function. The implied volatility a(K) is then defined as the value of a to solve the following implicit equation:

$$Call(K) = vF\Phi\left(\frac{\ln(F/K)}{a} + \frac{a}{2}\right) - vK\Phi\left(\frac{\ln(F/K)}{a} - \frac{a}{2}\right)$$

The right-hand side, called the "Black Formula", is a strictly increasing function of a, moving from $v \max\{0, F\text{-}K\}$ to vF as a moves from 0 to infinity. As the right-hand side covers the feasible range for $Call(K)$, so there must be a unique positive implied volatility a(K). Conversely, any positive implied volatility gives a feasible option price.

This implied volatility a(K) is also the unique solution to

$$Put(K) = vK\Phi\left(\frac{\ln(K/F)}{a} + \frac{a}{2}\right) - vF\Phi\left(\frac{\ln(K/F)}{a} - \frac{a}{2}\right)$$

We can deduce this result from put–call parity and the fact that ? (-z) = 1- ? (z).
Classical option-pricing theory seeks to *prove* the Black Formula, or rather, seeks to prove that a(K) = s_S for all K, based on restrictive assumptions about the stochastic process governing S. But in this chapter, we allow more general process for S, and treat the Black Formula merely as a convention for quoting prices. Instead of quoting a dollar price for a call or put option, a trader will quote the implied volatility a(K). Actual prices reveal that a(K) does, in fact, depend on K, contrary to the classical predictions.

Using Deflators to Price Options

We can use deflators to price options. The relevant formulas are as follows:

$$Call(K) = \mathbf{E}[\ D \max\{0, S\text{-}K\}\ ]$$
$$Put(K) = \mathbf{E}[\ D \max\{0, K\text{-}S\}\ ]$$

Put-call parity requires that

$$\mathbf{E}(D) = v$$
$$\mathbf{E}(DS) = vF$$

In general, there may be very many possible deflators. The challenge then is to find a deflator, which is, in some suitable sense, the best for pricing a particular option.

Finding A Deflator

Having developed a distribution for *S*, we now look for a deflator *D*. Our strategy will be as follows:

1. assume a functional form for *S* and *D*;
2. solve for parameters ensuring put–call parity;
3. eliminate redundant parameters; and
4. apply equilibrium constructions.

Each of these steps is discussed in turn.

Functional Form for Deflators

We now generalise our model for the share price, to include the rest of the economy and, in particular a deflator. Let us suppose that the log prices of all assets in the economy are drawn from some huge multivariate conic moment distribution. In this situation, it is natural to suppose also the log deflator would also be a component of this multivariate family. While the parameters for observable assets would be empirically calibrated, we must apply more economic theory to suggest parameters for the deflator.

We assume therefore that the log price and log deflator are drawn from a bivariate conic moment distribution (x_S, x_D). To simplify subsequent algebra, we apply the distribution to ln(S/F) and ln(D/v), so that

$$S = F\exp(X_S)$$
$$D = v\exp(X_D)$$

The variables X_S and X_D are assumed to have a joint conic moment distribution with parameters:

$$m = \begin{pmatrix} m_0 \\ m_S \\ m_D \end{pmatrix}; V = \begin{pmatrix} V_{00} & V_{0S} & V_{0D} \\ V_{0S} & V_{SS} & V_{SD} \\ V_{0D} & V_{SD} & V_{DD} \end{pmatrix}$$

Notice that the parameters m_0, m_D, V_{00}, V_{0D} and V_{DD} relate to the distribution of the deflator, which is a function of the economy as a whole. We should not expect to choose the deflator distribution based only on the distribution of *S*.
To ensure put–call parity, we need to verify that,

$$\mathbf{E}(D) = v$$
$$\mathbf{E}(DS) = vF$$

This we can ensure using the moment generating function for a bivariate conic moment distribution. The conditions on m and V are that:

$$\begin{pmatrix}0\\0\\1\end{pmatrix}^T V \begin{pmatrix}0\\0\\1\end{pmatrix} + 2m.\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}0\\1\\1\end{pmatrix}^T V \begin{pmatrix}0\\1\\1\end{pmatrix} + 2m.\begin{pmatrix}0\\1\\1\end{pmatrix} = 0$$

Written out long-hand, these are:

$$V_{DD} + 2m_D = 0$$
$$V_{SS} + 2V_{SD} + 2m_S = 0$$

with the constraints that

$$m_0 + V_{0D} \geq 0$$
$$m_0 + V_{0S} + V_{0D} \geq 0$$

and V must be positive definite. We now need to evaluate when these conditions together have a solution.

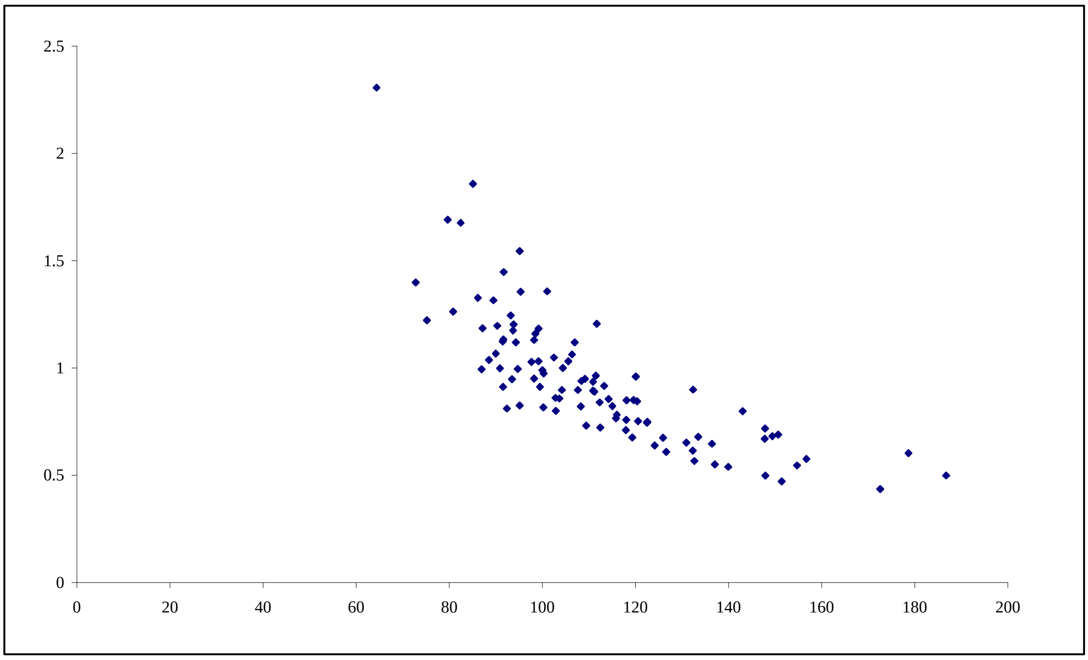
We dismiss the positive definite constraint first. Given that $\begin{pmatrix}V_{00} & V_{0S}\\ V_{0S} & V_{SS}\end{pmatrix}$ is positive definite by construction, we can achieve positive definiteness of the larger matrix simply by ensuring V_{DD} is sufficiently large. This would affect m_D, but as both m_D and V_{DD} occur only in the first equation, we can adjust them without affecting other variables.

Thus, we have boiled down the put–call parity equations to the following constraints:

$$V_{SS} + 2V_{SD} + 2m_S = 0$$
$$m_0 + V_{0D} \geq 0$$
$$m_0 + V_{0S} + V_{0D} \geq 0$$

This plainly has infinitely many solutions, provided we choose a V_{0D} that is big enough.

Figure 6.2: Share prices (horizontal axis) and deflators (vertical axis) from a typical model with 1-year horizon



Option-Pricing Formulas

We are now in a position to derive option-pricing formulas for our deflator model, expressed in terms of the univariate conic moment distribution. Let us focus first on put options. We know that

$$\begin{aligned}
Put(K) &= \mathbf{E}\left[D\max\{0, K-S\}\right] \\
&= v\mathbf{E}\left[\exp(X_D)\max\{0, K - F\exp(X_S)\}\right] \\
&= vK\mathbf{E}\left[\exp(X_D); X_S \le \ln(K/F)\right] - vF\mathbf{E}\left[\exp(X_S + X_D); X_S \le \ln(K/F)\right] \\
&= vK\int_{-\infty}^{\ln(K/F)}\int_{-\infty}^{\infty}\exp(x_D)\, f(x_S, x_D \mid m, V)\,dx_2\,dx_1 \\
&\quad - vF\int_{-\infty}^{\ln(K/F)}\int_{-\infty}^{\infty}\exp(x_S + x_D)\, f(x_S, x_D \mid m, V)\,dx_2\,dx_1 \\
&= vK\int_{-\infty}^{\ln(K/F)}\int_{-\infty}^{\infty} f\left(x_1, x_2 \,\middle|\, \begin{pmatrix} m_0 + V_{0D} \\ m_S + V_{SD} \\ m_D + V_{DD} \end{pmatrix}, V\right) dx_2\,dx_1 \\
&\quad - vF\int_{-\infty}^{\ln(K/F)}\int_{-\infty}^{\infty} f\left(x_1, x_2 \,\middle|\, \begin{pmatrix} m_0 + V_{0S} + V_{0D} \\ m_S + V_{SS} + V_{SD} \\ m_D + V_{SD} + V_{DD} \end{pmatrix}, V\right) dx_2\,dx_1
\end{aligned}$$

Now finally we can evaluate the integrals in terms of the univariate distribution:

$$\begin{aligned}
Put(K) &= vK\Psi\left(\ln\left(\frac{K}{F}\right) \,\middle|\, \begin{pmatrix} m_0 + V_{0D} \\ -\tfrac{1}{2}V_{SS} \end{pmatrix}, \begin{pmatrix} V_{00} & V_{0S} \\ V_{0S} & V_{SS} \end{pmatrix}\right) \\
&\quad - vF\Psi\left(\ln\left(\frac{K}{F}\right) \,\middle|\, \begin{pmatrix} m_0 + V_{0S} + V_{0D} \\ \tfrac{1}{2}V_{SS} \end{pmatrix}, \begin{pmatrix} V_{00} & V_{0S} \\ V_{0S} & V_{SS} \end{pmatrix}\right)
\end{aligned}$$

We notice that m_S, m_D and V_{DD} have disappeared from the option-pricing formula. We have substituted for V_{12} from the moment generating function. Our free choice then relates to V_{0D}, which must merely be chosen large enough.

We can use put–call parity to determine the price of call options

$$\begin{aligned}
Call(K) &= vF - vF\Psi\left(\ln\left(\frac{K}{F}\right) \,\middle|\, \begin{pmatrix} m_0 + V_{0S} + V_{0D} \\ \tfrac{1}{2}V_{SS} \end{pmatrix}, \begin{pmatrix} V_{00} & V_{0S} \\ V_{0S} & V_{SS} \end{pmatrix}\right) \\
&\quad + vK - vK\Psi\left(\ln\left(\frac{K}{F}\right) \,\middle|\, \begin{pmatrix} m_0 + V_{0D} \\ -\tfrac{1}{2}V_{SS} \end{pmatrix}, \begin{pmatrix} V_{00} & V_{0S} \\ V_{0S} & V_{SS} \end{pmatrix}\right)
\end{aligned}$$

We notice that this has two expressions of the form (1-?). We can turn these into expressions of the form ? by changing the sign of the implied underlying distribution.

$$\begin{aligned}
Call(K) &= vF\Psi\left(\ln\left(\frac{F}{K}\right) \,\middle|\, \begin{pmatrix} m_0 + V_{0S} + V_{0D} \\ -\tfrac{1}{2}V_{SS} \end{pmatrix}, \begin{pmatrix} V_{00} & -V_{0S} \\ -V_{0S} & V_{SS} \end{pmatrix}\right) \\
&\quad - vK\Psi\left(\ln\left(\frac{F}{K}\right) \,\middle|\, \begin{pmatrix} m_0 + V_{0D} \\ \tfrac{1}{2}V_{SS} \end{pmatrix}, \begin{pmatrix} V_{00} & -V_{0S} \\ -V_{0S} & V_{SS} \end{pmatrix}\right)
\end{aligned}$$

We can now see some symmetry in the analysis; the call option formula looks like the put option but with K and F interchanged. Traders tend to quote option prices using implied volatilities. The corresponding implied volatility charts are shown in Figure 8.

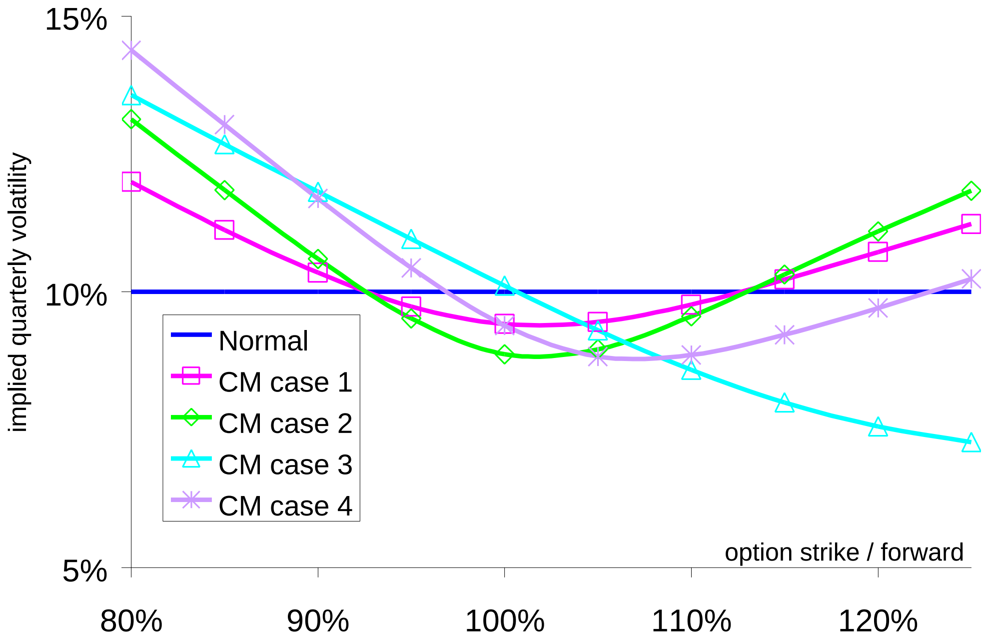


Figure 6.3: Implied quarterly volatilities for the normal distribution and four conic moment distributions.

The Final Optimisation

We have developed formulas for options on a stock, given a conic moment distribution for that stock. These option-pricing formulas introduced a new parameter V_{0D}. Any value of V_{0D} could give consistent option prices. There are a wide range of possible prices. This choice widens further if we had used a wider class of functional forms for D. In the limit, we get back to Merton's arbitrage bounds. So, reducing the class of D's under consideration is crucial.

Portfolio Selection and Optimal Deflators

The Incomplete Market Problem

There is an extensive literature on how to pick a pricing law out of many alternatives. The most popular technique uses the "minimal martingale measure", defined by Follmer and Schweizer (1986). Picking a pricing law is equivalent to picking a utility function. This makes the process seem arbitrary.

However, this arbitrariness is not necessarily a problem, because many utility functions could imply the same prices. For example, Cochrane (2001) shows how all quadratic utilities would lead to Ross's APT model. However, the literature does present a number of practical problems, including the following:

1. Prices based on quadratic hedging error may give rise to negative option prices.

2. Optimisation for jump processes often results in singular solutions. For example, the optimal portfolio under a log utility often results in zero holdings for some asset classes; these asset classes are not then priced to market by the utility gradient.
3. With exponential utility and jump process, the prices of options often become infinite.
4. Most risk measures require specification of an accounting currency, which is necessarily arbitrary. Log utility is an exception.
5. Most of the literature is concerned with one risky asset. The proposed deflator is then a function of that asset only. But in a multi-asset world, the optimal deflator would be a function of a market portfolio that is not necessarily the underlying for the option you wish to price.
6. There are practical difficulties in solving the equations numerically for many realistic models, especially those involving jumps.

The other extreme is to insist on analytical tractability – this leads for example to Esscher transform methods (Buhlmann *et al*., 1998). In a single asset case this could be equivalent to power law utilities, but in a multiple asset case this breaks down because real portfolios are arithmetic means but Esscher transform gives a deflator that is a product of powers of prices.

In this chapter we have tried to steer a middle course between tractability and economic coherence. This is why we have combined an analytical formula with free parameters subject to constraints on maximum utility.

Good Deal Bounds

Cochrane and Saá-Requejo (2000) develop the crucial concept we need to complete our option-pricing example. They develop the concept of *good deal* bounds. The idea is to place a limit on investors' ability to earn high rewards with low risk. In this note we achieve this by placing bounds on the utility achievable by the logarithmic investor. This takes us into the realms of portfolio optimisation.

Logarithmic utility

Let us step back a little and consider an investment with initial wealth L_0. The investor wishes to invest for one year, after which time the terminal wealth will be L_1. We assume that the investor's objective is to maximise the utility function $\mathbf{E}\log(L_1\text{-}c)$, where c is the minimal acceptable wealth. For this to make sense, the investor must at least be capable of ensuring the minimal wealth, so that $vc < L_0$. To model the portfolio process, we need to allow an economy with more than one stock. We therefore give our investor complete flexibility, to invest in any asset in the economy, to hold cash at the risk-free rate, to buy or sell the assets forward, or even to trade in options on other assets, assuming these are priced according to some specified deflator.

One thing we can say for sure is that $\mathbf{E}(DL_1) = L_0$. This applies for all investments in the economy, from zero coupon bonds and forwards through to call and put options. As this is a linear constraint, it follows that it also applies to portfolios.

Constraints on Optimised Utility

The focus is now shifted to proving the following simple inequality:

$$\mathbf{E}\ln(L_1\text{-}c) = \ln(L_0\text{-}vc) - \mathbf{E}\ln(D)$$

which claims that the expected log deflator puts a constraint on investors' utility. If $\mathbf{E}\ln(D)$ is large and negative, then a high utility may be possible. On the other hand, if $\mathbf{E}\ln(D)$ is closer to zero then the opportunities for enhancing utility are more limited. The inequality is deceptively easy to prove, for we know that

$$\mathbf{E}[\ D(L_1\text{-}c)\ ] = L_0\text{-}vc$$

As the natural logarithm is a concave function, Jensen's inequality (see Williams, 1991) applies and we can see that

$$\mathbf{E}\ln(L_1\text{-}c) + \mathbf{E}\ln(D)\ \mathbf{E}\ \ln[D(L_1\text{-}c)] = \ln\ \mathbf{E}[D(L_1\text{-}c)] = \ln(L_0\text{-}vc)$$

Incidentally, similar inequalities would apply had we originally chosen a different utility. The logarithm is the easiest to deal with, but we could also have chosen quadratic or power law utilities.

Market Efficiency

We now revert to our problem of choosing between different models that share a distribution of S but have different deflators. Our idea is to choose a deflator that provides a meaningful market efficiency constraint.

Market efficiency is difficult to define, but one if its characteristics is the limitation of the return opportunities from a given level of risk. If markets permit arbitrage, then free lunches allow unbounded expected utility. Inefficient markets provide the possibility of high returns with low risk, which lead to high utility. Thus, we could seek to define market efficiency in terms of the bounds on achievable utility. This would ensure that the derived option prices were not out of line with returns available on other investments. Equivalently, given our previous inequalities, if we wish to model efficient markets we need to ensure $\mathbf{E}\log D$ is sufficiently high.

We recall that $v = \mathbf{E}(D)$, so the mean of D is fixed. By Jensen's inequality, $\mathbf{E}\ln(D) = \ln v$. The gap depends on the variability of D. If we want to place a lower bound on $\mathbf{E}\log D$, we must place an upper bound on the variability of D. If we had selected some other utility function, we would still need to control the variability of D but we would be using a different definition of variability.

Controlling the variability of D has another incidental benefit, for applications where Monte Carlo simulation is used.. A large deflator variance would be inconvenient as it implies large sampling errors, and so large numbers of simulations are needed to achieve a desired accuracy. The constraints we will place on the variability of D also control sampling errors so underpin the efficiency of Monte Carlo pricing tools.

We now return to the option-pricing example. The mean log deflator corresponds to the mean of our conic moment distribution, and is given by

$$\mathbf{E}\ln(D) = \ln v + m_D/m_0 = \ln v - V_{DD} / (2m_0)$$

Now, as V is positive semi-definite, we know that

$$V_{DD} \geq \begin{pmatrix} V_{0D} \\ V_{SD} \end{pmatrix}^T \begin{pmatrix} V_{00} & V_{0S} \\ V_{0S} & V_{SS} \end{pmatrix}^{-1} \begin{pmatrix} V_{0D} \\ V_{SD} \end{pmatrix}$$
$$= \frac{V_{SS}V_{0D}^2 - 2V_{0S}V_{0D}V_{SD} + V_{00}V_{SD}^2}{V_{00}V_{SS} - V_{0S}^2}$$
$$= \frac{V_{SS}\left(V_{0D} - \frac{V_{0S}V_{SD}}{V_{SS}}\right)^2}{V_{00}V_{SS} - V_{0S}^2} + \frac{V_{SD}^2}{V_{SS}}$$

We can also substitute for V_{SD}, and apply back to the log deflator:

$$\ln v - \mathbf{E}\ln(D) = \frac{V_{DD}}{2m_0} \geq \frac{V_{SS}\left(V_{0D} - \frac{V_{0S}V_{SD}}{V_{SS}}\right)^2}{2m_0\left(V_{00}V_{SS} - V_{SS}^2\right)} + \frac{V_{SD}^2}{2m_0V_{SS}}$$
$$= \frac{V_{SS}\left(V_{0D} + \frac{V_{0S}}{V_{SS}}\left[\frac{1}{2}V_{SS} + m_S\right]\right)^2}{2m_0\left(V_{00}V_{SS} - V_{SS}^2\right)} + \frac{\left[\frac{1}{2}V_{SS} + m_S\right]^2}{2m_0V_{SS}}$$

Given existing constraints on V_{0D}, the right-hand side is minimised when

$$V_{0D} = -\min\left\{m_0, m_0 + V_{0S}, \frac{V_{0S}}{V_{SS}}\left[\frac{1}{2}V_{SS} + m_S\right]\right\}$$

Three cases can now occur:
In Case I, the usual case, there is a range of possible V_{0D} that give acceptable values for **E**log$D$. This range would lie either side of the value in the equation above. The option price based on distributions alone is indeterminate; further information is required.

In Case II, the range of permissible V_{0D} is so narrow that it collapses to a point. This happens in the limit of a lognormal distribution, in which the model collapses to Balck and Scholes' formula. In many examples, the range of V_{0D} is so narrow that the possible option prices lie within a narrow band.

In Case III, it is possible that there is no value of V_{0D} giving an acceptably high value of **E**log$D$. This can occur if the distribution has a very high return in relation to risk, or if the distribution is very fat-tailed. While this is rare in the univariate case, it is much more frequently a problem for multivariate models. Special calibration tools are required in order to derive consistent risk and return assumptions to obviate the problem.

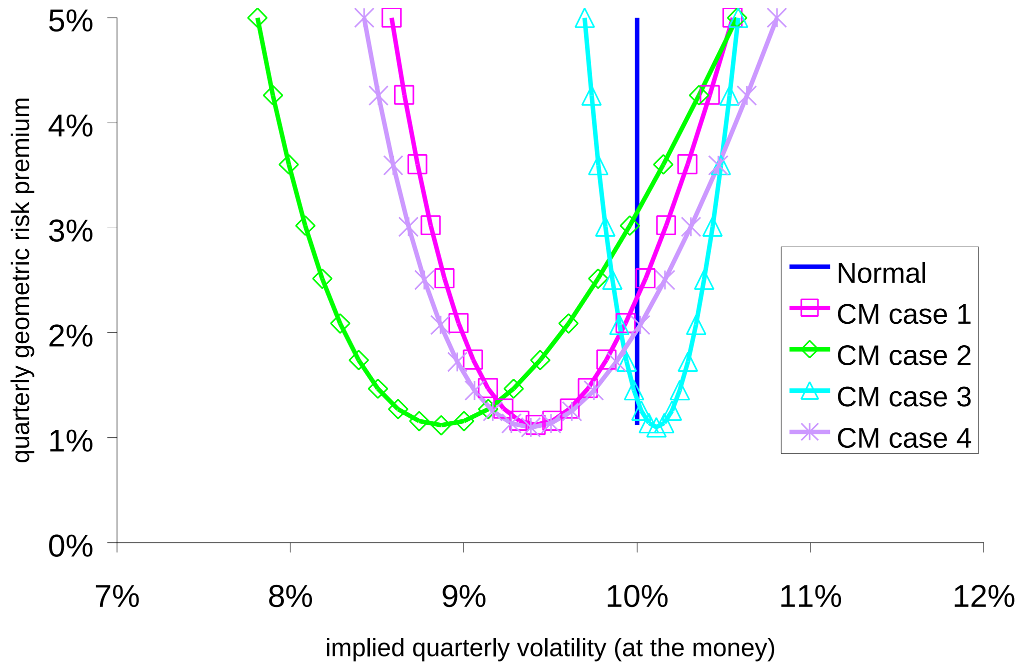


Figure 6.4. Market Efficiency and Implied Volatilities.

For our example distributions, Figure 9 summarises the relationship between option prices and marekt efficiency. Towards the top of the chart, we have the less efficient marekts. The risk premium of 5% per quarter means that an investor with a logarithmic utility function could achieve 5% per quarter abovce the risk free rate – an assumption most economists would regard as on the high side. These are achievable either when implied volatilities are low (in which case our investor puys options) or when they are high (when the investor sells).

So efficiency constraints limit option implied volatilities to a central range. Financial theory has not delivered a unqiue price for options, but has at least excluded a whole range of unreasonable values.

7. Conclusions

There is no perfect financial forecasting model. Real markets show a huge range of phenomena, and a model will only capture some of them. Traditional normal-based tools capture all sources of variability into a single parameter: σ. Unpicking different effects is messy – we often need to use different values of σ for different problems, depending what part of the distribution we are interested in. Why not capture the variability by using conic moment distributions instead? This better describes the shape of many historic return distributions. This solves some, but not all, of the calibration problems faced by normal distributions. Devising models that work adequately at all time scales remains challenging.

There are many applications where normal distributions may continue to be used confidently. The effect of fat tails is most noticeable over short time horizons, of less than 5 years. Fat tailed distributions shed little light on the distributions of returns over many years, nor on prices of long dated options and guarantees. In these cases, there is often room for debate regarding the correct volatility assumption; adding further complexity to the shape of the distribution does not overcome this basic calibration problem.

The commercial area where normality is most crucial is also the area where it is least likely to be challenged – that is, economic capital. Using conic moment distributions instead of normal ones, could increase the economic capital by (10x) as measured weekly with an acceptable failure frequency of 1 in 2000 years. This vast difference generates a debate about the right way to extrapolate observed distributions, and indeed over whether anyone could reliably estimate such extremes of a distribution. In the meantime, management are likely to prefer the lower economic capital numbers, which come from the assumption of normal distributions.

The use of fat tailed distributions does hold out a hope of reconciling the credit spreads in corporate bonds to their probability of default and to their likely recovery rates. The difficulty arises because deflators must be very large in the event of default. Indeed, one of the longstanding theoretical puzzles in corporate bond defaults is to understand why defaults happen at all. Would shareholders not optimally reduce the risk in a company as it approached insolvency, thus postponing indefinitely the day when control passes to bondholders? This strategy would fail only if there are jumps in the net assets of a firm – an effect that conic moment distributions capture well.

Many statistical tests rely on normality. In particular, evidence for time-varying volatility and for auto-regressive effects, is typically assessed using tests based on normal distributions. One or two outliers in a data sample may force complicated additional terms into a model. If we switch to fat tailed distributions, then more outliers are allowed before we have to complicate the model. We should not expect fat tailed distributions to result in more complex models overall. We may have more complex error distributions, but this is usually offset by simplifications in a model's structure.

For the time being, the normal distribution retains her crown as the queen of capital market models. But as her frailty becomes more apparent, attention is likely to focus increasing on a suitable heir – the conic moment distribution family.

Acknowledgements

We are grateful to many friends and colleagues who have shared insights with or have provided helpful comments on earlier drafts of this paper. Particular thanks are due to Alessandra Gambini-Weigel, Frances Southall, Hui Ming Ng, Tim Sheldon and Elliot Varnell. Any remaining errors are the responsibility of the authors alone.

Bibliography

Aase, K (2002): “Representative Agent Pricing of Financial Assets Based on Lévy Processes with Normal Inverse Gaussian Marginals”, Annals of Operations Research 114 pp 15-31.

Albanese, C, Jaimungal, S and Rubisov, D (2001): "Jumping in line", Risk Magazine, February, pp 65-68.

Barndorff-Nielsen, O E & Shephard, N (1998): "Incorporation of a Leverage Effect in a Stochastic Volatility Model", MaPhySto Research Report No. 18, University of Aarhus.

Barndorff-Nielsen, O E & Shephard, N (2000): "Integrated Ornstein Uhlenbeck Processes", Research Report, Oxford University.

Barndorff-Nielsen, O E & Shephard, N (2**000):** "Modeling with Lévy processes for financial econometrics", MaPhySto Research Report No. 16, University of Aarhus.

Barndorff-Nielsen, O E (1977): "Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions", Zeitschrift fur Wahrscheinlichkeitstheorie une verwandte Gebiete, 38, 309-312.

Barndorff-Nielsen, O E (1977): “Exponentially decreasing distributions for the logarithm of particle size”, Prc Royal Soc London A 353, pp 401-419

Barndorff-Nielsen, O E (1997): "Normal inverse gaussian distributions and stochastic volatility modeling", Scandinavian Journal of Statistics, 24, 1-13.

Barndorff-Nielsen, O E (1998): “Processes of normal inverse Gaussian type”, Finance and Stochastics 2(1) pp 41-68.

Barndorff-Nielsen, O E 1995. “Normal inverse Gaussian processes and the modelling of stock returns”, Research report #300, Aarhus University.

Barndorff-Nielsen, O E, Mikosch T. & Resnick S. (Eds.) (2001): "Lévy processes-theory and applications", Boston: Birkhauser.

Bauer, C (2000): “Value at Risk using Hyperbolic Distributions”, Journal of Economics and Business 52 pp 455-467.

Bertoin, J (1996): "Lévy processes", Cambridge University Press.

Blattberg, R C & Gonedes, N J (1974): “A comparison of the stable and student distributions as statistical models for stock prices”, Journal of Business 47(2) pp 244 – 280.

Bookstaber, R M & Mcdonald, J B (1987): “A general distribution for describing security price returns”, Journal of Business 60 401-424.

Boyarchenko, S I and Levendorskii, S Z (2002): “Non-Gaussian Merton-Black-Scholes Theory”, World Scientific Publishing.

Bretagnolle, J (1973): "Processus à incréments indépendants", Ecole d'Eté de Probabilités, Lecture Notes in Mathematics, Vol. 237, pp 1-26. Berlin: Springer.

Bühlmann, H, Delbaen, F, Embrechts, P and Shiryaev, A N (1998): “On Esscher Transforms in Discrete Finance Models”, ASTIN Bulletin 28(2) pp171-186.

Carr, P, Geman, H, Madan, D & Yor, M (2000): "The fine structure of asset returns: an empirical investigation", Working Paper.

Carr, P, Geman, H, Madan, D & Yor, M (2002): “Self Decomposability and Option Pricing”, Working Paper, University of Maryland.

Carr, P, Madan, D (1998): "Option valuation using the fast Fourier transform", Journal of Computational Finance, 2, 61-73.

Cerny, A (2003): "Mathematical Techniques in Finance: Tools for Incomplete Markets", (Princeton University Press), Forthcoming.

Chan, K (1999): "Pricing contingent claims on stocks driven by Lévy processes", Annals of Applied Probability, 9, 504-528.

Chan, W S (2002): “Stochastic Investment Modelling : a Multiple Time Series Approach” British Actuarial Journal 8(3) pp545-592.

Cochrane, J H (2001): "Asset Pricing", (Princeton University Press).

Cochrane, J H and J Saá Requejo (2000): “Beyond Arbitrage: Good Deal Asset Price Bounds in Incomplete Markets”, Journal of Political Economy, 108, pp. 79–119.

Cont, R (2001): "Empirical properties of asset returns: stylized facts and statistical issues", Quantitative Finance, 1, No. 2.

Cont, R, Bouchaud, J P & Potters, M (1997): "Scaling in financial data: stable laws and beyond", in: B Dubrulle, F Graner & D Sornette (Eds.): Scale invariance and beyond, Berlin: Springer.

Dixit, A K, and R S Pindyck (1994): "Investment Under Uncertainty", (Princeton University Press).

Duffie, D (2001): "Dynamic Asset Pricing Theory", Third Edition, (Princeton University Press).

Eberlein, E & Jacod, J (1997): "On the range of options prices", Finance & Stochastics, 1, No. 1, 131-140.

Eberlein, E and von Hammerstein, E A (2002): “Generalised Hyperbolic and Inverse Gaussian Distributrions: Limiting Cases and Approximation of Processes".

Eberlein, E and Keller, U 1995. “Hyperbolic distributions in Finance”, Bernoulli 1, pp 281-299.

Eberlein, E and Prause, K 1999 “The generalized hyperbolic model: financial derivatives and risk measures”. Working paper, Freiburg University.

Eberlein, E, Keller, U & Prause, K (1998): "New insights into smile, mispricing and Value at Risk: the hyperbolic model", Journal of Business, 71, No. 3, 371-405.

Eberlein, E, Raible, S (1999): "Term structure models driven by general Levy processes", Mathematical Finance, 9, 31-53.

Finkelstein, G S (1997): “Maturity guarantees revisited: allowing for extreme stochastic fluctations using stable distributions”. British Actuarial Journal 3, pp 411-482

Follmer, H, and M Schweizer (1986): “Hedging of Non-Redundant Contingent Claims”, Contributions to Mathematical Economics, pp. 205–23.

Frittelli, M (2000): "The Minimal Entropy Martingale Measures and the Valuation Problem in Incomplete Markets", Mathematical Finance, 10, No. 1, 39-52.

Fujiwara, T and Miyahara, Y (2001): "On the Minimal Entropy Martingale Measure for Geometric Levy Processes", (In preparation).

Geman, H, Madan, D & Yor, M (1999): "Time changes for Levy processes", Preprint.

Geman, H, Madan, D & Yor, M (2000): "Time changes in subordinated Brownian motion", Preprint.

Gerber, H U and Shiu, E S W (1994): “Option Pricing by Esscher Transforms”. Transactions of the Society of Actuaries 46, pp99-192

Goford, J (2002): Remarks in the discussion of “Fair Value of Liabilities” British Actuarial Journal 8(2) p326.

Hardy, M R (2001): “A Regime-Switching Model of Long-Term Stock Returns”, North American Actuarial Journal, Volume 5, Number 2.

Harrison, J M and D M Kreps (1979): “Martingales and Arbitrage In Multiperiod Securities Markets”, Journal of Economic Theory, 20, pp. 381–408.

Harrison, J M, and S R Pliska (1981): “Martingales and Stochastic Integrals in the Theory of Continuous Trading”. Stochastic Processes and Applications, 11, pp. 215–60.

Hogg, R V and Klugman, S A (1984): "Loss Distributions", Wiley.

Howison, S and Lamper (2000): “Trading volume in models of financial derivatives”, OCIAM working paper

Hull, J C (1993): "Options, Futures and Other Derivative Securities", (Prentice Hall).

Hurst, S R, Platen, E and Rachev, S T (1999): "Option pricing for a logstable asset price model", Math. Comput. Modelling, 29 , no. 10-12, 105--119.

Jacod, J & Shiryaev, A N (1987): "Limit theorems for stochastic processes", Berlin, Springer.

Johansen, A and Sornette, D (2000): “Large Stockmarket Price Drawdowns are Outliers”, UCLA working paper.

Johnson, M E and Kotz, S (1970): "Continuous univariate distributions", Houghton Miflin, Boston

Kallsen, J (2001): “Utility-based derivative pricing in incomplete markets.” In H. Geman, D.
Madan, S. Pliska, and T. Vorst (Eds.), Selected Proceedings of the First Bachelier Congress, held in Paris, 2000, Berlin. Springer.

Klüppelberg, C (2002): “Risk Management with Extreme Value Theory”. University of München working paper.

Küchler, U, Neumann, M, Sørensen, M and Streller, A (1994): “Stock returns and hyperbolic distributions”, Preprint, Humbolt University, Berlin.

Kühn, C (2002): “Shocks and Choices – an Analysis of Incomplete Market Models”, PhD thesis, München

Lévy, P (1937): "Théorie de l'addition des variables aléatoires", Paris: Gauthier Villars.

Madan, D and E Seneta (1990): "The variance gamma model for share market returns", Journal of Business, 63, 511-524.

Madan, D, Carr, P and Chang, E (1998): "The variance gamma process and option pricing", European Finance Review, v.2, 79-105.

Madan, D & Milne, F (1991): “Option pricing with variance gamma martingale components”, Mathematical finance, 1, 39-55.

MandelBrot B B (1963): “The variation of certain speculative prices”, Journal of Business 36(4) pp 394-419

Mandelbrot, B B (1997): "Fractals and scaling in finance", Berlin: Springer.

Merton, R C (1973): “Theory of Rational Option Pricing”, Bell Journal of Economic Management Science, 4, pp. 141–83.

Merton, R C (1974): “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates”, Journal of Finance, 29, pp. 449–70.

Musiela, M and M Rutkowski (1977): "Martingale Methods in Financial Modelling", (Berlin: Springer).

Musiela, M and T Zariphopoulou (2001): “Pricing and Risk Managment of Derivatives Written on Non-Traded Assets”, BNP Paribas Working Paper

Plerou, V, Parameswaran, G, Nunes-Amaral, L A, Meye, M and Stanley, H E (1999): “Scaling of the distribution of price fluctuations of individual companies”, Physical Review E 60(6) pp 6519-6529.

Protter, P (1990): "Stochastic integration and differential equations: a new approach", Berlin: Springer.

Raible, S (2000): “Lévy processes in Finance: Theory, Numerics and Empirical Facts”. PhD thesis, Freiburg University.

Rejman, A, Weron, A, and Weron, R (1997): “Option pricing proposals under the generalised hyperbolic model”. Cummunications in Statistics – Stochastic Models 13, 867-885.

Ribereiro, C and Webber, N (2003): "A Monte Carlo Method for the Normal Incverse Gaussian Option Valuation Model using an Inverse Gaussian Bridge".

Rydberg, T H (1997): "The normal inverse Gaussian Lévy process: Simulation and approximation", Commun. Stat., Stochastic Models 13(4), 887-910.

Samorodnitsky, G & Taqqu, M (1994): "Stable non-gaussian random processes", New York: Chapman and Hall.

Sato, K (1999): "Levy Processes and Infinitely Divisible Distributions", Cambridge University Press.

Walter, C (1990): “Mise en évidence de distributions Lévy-stables et d’une structure fractable sur le marché de Paris”, Transactions of the 1st AFIR International Colloquium, 3, pp241-259. Paris.

Wilkie, A D (1987): “An Option Pricing Approach to Bonus Policy”. Journal of the Institute of Actuaries 114, pp 21-77

Wilkie, A D (1995): “More on a stochastic asset model for actuarial use” British Actuarial Journal 1, pp777-964.

Williams, D (1991): "Probability with Martingales", (Cambridge University Press).