Modeling Risk with Unit-variance Leptokurtic Fratcal Normal Statistics

James Stacey and Alex Faseruk

Introduction:

One of the basic paradigms in the theory of finance is that in the long-run investors will be rewarded for risk-taking behavior. Stated more simply the greater the risk of an investment the greater the return that will be expected. Since the development of modern portfolio theory in the early 1950s, financial theorists have sought to appropriately define and measure risk and return in not only individual assets but also portfolios of assets. While Artzner, Delbaen, Eber and Heath examined various measures of risk, they concluded that the seminal work by Markowitz (1952) led researchers initially toward examining the first two moments of the statistical distribution becoming preoccupied with mean and variance (standard deviation). The underlying assumption was that asset prices were normally distributed. Later, researchers became concerned with the third moment of the statistical distribution, such as Friend and Westerfield (1980) and Kraus and Litzenberger (1976), while others laid the groundwork for moving towards higher moments as examined by Samuelson (1970) and Scott and Horvath (1980).

A study of cotton prices by Mandelbrot (1963) provided evidence that the stochastic process underlying financial time series deviated from the normal probability density function that had previously been the assumed paradigm in the finance literature. This led Mandelbrot to suggest that prices were more accurately described by stable Lévy distributions. In an extensive study of the statistical properties of the companies in the Dow Jones Industrial Average, Fama (1965) provided support for Mandelbrot, indicating that Mandelbrot's hypothesis seemed to be supported by the data.

Lévy probability density functions (pdfs) are characterized by much fatter tails in the distributions; that is, the probability of large fluctuations is much more likely than for the normal pdf. Analysis of Lévy distributions by Voit (2003) indicated that, for Lévy indices $q < 2$, the variance of this pdf is infinite but when for $q \leq 1$, the mean of this pdf is undefined. Stocks with returns distributed according to a Lévy distribution are inherently more risky than stocks with returns that are normally distributed.

This study examines the hypothesis that there exists a family of unit-variant leptokurtic probability density functions with the attractive properties of normal statistics.

James Stacey is a MBA graduate and was an instructor in the Faculty of Business Administration at Memorial University of Newfoundland. Alex Faseruk is Professor of Finance in the Faculty of Business Administration at Memorial University of Newfoundland, St. John's, Newfoundland.

The current study constructs numerically a family of symmetric Normal pdfs defined for fractal spaces with index q. These distributions have the attractive property of having unit variance for $1<q<2$ and approximately unit variance for $0.5<q<1$ (the limit of the range that was studied). This paper studies the conjecture that there exists a family of Normal distributions with unit variance that can be defined on metric spaces of dimension $m=q/2$ for $0<q<2$. Some analytical techniques that arise from this conjecture are explored, indicating how one could apply a broader theory of fractal statistics in practice. The power of these techniques follows from the property that the scale of the numerically generated probability distributions is independent of the fractal index q. This allows the non-linearity of fractal statistics to be absorbed into a universal non-linear function of a shape measure with the kurtosis chosen to illustrate the method: for what this study calls Fractal Normal, or Formal, pdfs. It is found that the application of Fractal Normal statistics retains the attractiveness of Normal statistics by allowing optimal solutions to be determined in a straightforward fashion.

Overview of Normal Statistics:

A time series of prices for stock j is given by

Eq. 1 $X_j = \{x_i; i = 1, T\}_j$

where $x_i > 0$ for all i, x_i is a shorthand notation for $x_i = x(t_i)$, and t_i is the time at which x is determined. As the index i increases by one unit, the value of time is presumed to increase by a constant Δt, such that $t_1 < t_2 < \cdots < t_T$. x is a random variable where the x_i are distributed according to the pdf $p(x)$.

The standard variable z is defined through the relation

Eq. 2 $z - z_0 = \dfrac{x-\mu}{A} = A^{-1}(x-\mu),$

where μ is the mean of the pdf for the x (in a manner that is yet to be defined) and A is the characteristic scale for the distribution of the x about the mean μ. Eq. 2 implies that $z = A^{-1}x$, and $z_0 = A^{-1}\mu$.

Conventionally, the scale A is identified with the standard deviation σ (the square root of the variance) but for Lévy distributions the second moment of the pdf (the variance) is infinite (or at least undefined) and, in some cases, even the first moment of the pdf (the mean) is undefined. This implies that, in general, identifying A as σ is incorrect. For the same reason, it is generally incorrect to identify μ as the expectation value $\langle x \rangle$ as this quantity may not be defined for some pdfs.

The standard variable z is a dimensionless quantity. For following discussion, z_0 can be taken to be zero, without loss of generality. However, as noted below, all distributions in z should be considered proxies for distributions expressed in terms of $z - z_0$. Eq. 2 can be rearranged as

Eq. 3 $x = \mu + Az$.

Probability distributions for which it is possible to introduce a standard variable z in this way constitute a location-scale family of pdfs. Both Normal and Lévy pdfs are examples of distributions belonging to location-scale families.

The time series of price data x for a stock can be modeled as a random walk

Eq. 4 $$x_i = x_{i-1} + Az,$$

where z is a random variable drawn according to some pdf $p(z)$, as will be defined below. In practice, studying the distribution of stock price data is not as illuminating as the distribution of stock return data. Returns may be defined in several ways. In this paper stock returns are defined as

Eq. 5 $$r_i = \frac{x_i - x_{i-1}}{x_{i-1}}.$$

Eq. 4 and Eq. 5 can be combined to yield

Eq. 6 $$r_i = \frac{A}{x_{i-1}} \cdot z \cong A'z.$$

For stocks where $x >> A$ the distribution of returns r_i are well approximated by the pdf $p(z)$ scaled by $A' = A/\langle x \rangle$.

Towards Fractal Normal Statistics:

With normal random variables, the additivity of variances is similar to the Pythagorean relation in multiple dimensions. In this work, the authors are motivated to retain an analogue of Pythagoras' fundamental relationship with fractional dimensional spaces.

A metric space can be defined on the positive real axis with a single-valued, non-negative real distance function, given by

Eq. 7 $$\delta(x, x_0) = |x - x_0|^{q/2}.$$

The exponent $q/2$ is the dimensionality m of the metric space, although the theory of fractals identifies q as the dimension of the fractal space as identified in Peters (1991). Following the derivation from Hein as recounted in Gardner (1977), the equation of a circle with this definition of metric is written as

Eq. 8 $$\delta((x,y),(x_0,y_0)) = \left[|x - x_0|^q + |y - y_0|^q\right]^{1/q}.$$

The circumference of a unit circle in this metric space can be calculated (by performing a path integral, integrating $ds = (dx^q + dy^q)^{1/q}$ around the circumference of the unit circle) and thereby determine that the value for π is no longer constant, but a unique function of q. For example, for $q = 1$, the value of π is 4. In general $\pi(q) > \pi(2)$ for all $q \neq 2$.

In this paper, a Normal distribution is then a 1-dimensional pdf, and fractal pdfs have dimensionality defined on the interval (0,1) which implies pdfs with index q will be considered as fractal distributions with dimension $m = q/2$, although throughout the study it will more commonly refer to the fractal *index q* rather than the fractal *dimension m* to conform with standard practice.

The algebra of the underlying metric space with fractal dimension $m = q/2$ affects the statistics of the Normal pdf. This analysis cannot start with the definition of a characteristic function since the value of π is not constant and an appropriate fractal calculus and corresponding fractal Fourier transform are not available. Instead, the analysis begins with a Normal random number generator based on the Polar Marsaglia variant of the Box-Muller algorithm from Gardner (1977). This algorithm does not require knowledge of π, nor does it require knowledge of the trigonometry of the underlying *q/2*-dimensional metric space. It relies on the equation of a unit circle in this metric space in Eq. 8.

A Normal pdf with unit variance is obtained by generating a uniform distribution of points on a disk circumscribed by a unit circle. This algorithm is extended by generating a uniform distribution of points on a unit disk in a fractal space of dimension *q/2*. The resulting distributions share the attractive features of normal pdfs and shall be referred to as Fractal-Normal distributions (or Formal pdfs).

The method is defined by the following algorithm, which generates a pair of random variates F_1 and F_2 distributed according to a Formal pdf, as follows. First, a pair of uniformly distributed temporary random variables t_1 and t_2 are drawn on the interval $[-1,+1]$ by

Eq. 9
$$t_1 \leftarrow U(-1,1) = 2 \cdot U(0,1) - 1$$
$$t_2 \leftarrow U(-1,1) = 2 \cdot U(0,1) - 1$$

One then defines a point in the unit square in the space $\mathrm{R}^m \times \mathrm{R}^m$ using

Eq. 10
$$v_1 = \operatorname{sign}(\mathrm{t}_1) \cdot |t_1|^{q/2}$$
$$v_2 = \operatorname{sign}(t_2) \cdot |t_2|^{q/2}$$

To generate a point within the unit disk in this fractal space, the radius r is calculated as

Eq. 11 $$r = \left(v_1^2 + v_2^2\right)^{1/q}$$

which is equivalent to

Eq. 12 $$r^{q/2} = \left(v_1^2 + v_2^2\right)^{1/2}.$$

When $r \leq 1$, a pair of Formally distributed random variates is given by

Eq. 13

$$F_1 = \frac{v_1}{r^{q/2}} \cdot \left(-2\ln\left(r^q\right)\right)^{1/q}$$

$$F_2 = \frac{v_2}{r^{q/2}} \cdot \left(-2\ln\left(r^q\right)\right)^{1/q}$$

This algorithm is similar in form to that for the exponential power distribution (also known as the generalized error distribution). For $q=1$, the above algorithm generates a distribution similar to the Laplace distribution while, for $q=2$, a Normal pdf is generated. The essential difference is that the Formal distribution generated according to the above algorithm has a unit variance where, for example, the Laplace distribution has a variance equal to 8 (which, notably, is equal to two π for $q=1$).

The radial distribution of points arising from this algorithm are presented in Figure 1 as a function of q = 2, 1.5, 1.0, and 0.5. The linear dependence of the distributions on the radius is consistent with the points being uniformly distributed on the unit disk for each value of q. In Figure 2, Formal distributions are compared to the corresponding Lévy distributions at q=2, 1.5, 1.0 and 0.5. In order for Lévy distributions with scale $A = 1$ to be directly compared with Formal distributions with scale $A=1$, the following form for the characteristic function is used:

Eq. 14 $$\hat{L}_q(k) = \exp\left(-\frac{1}{2}|Ak|^q\right).$$

Lévy pdfs are then generated by performing an inverse Fourier transform on this characteristic function. For $q<1$, the Formal distributions become increasingly cusp-shaped at $z=0$. Lévy distributions have less pronounced, but similar, behaviour.

Normal probability density functions have zero kurtosis, or $\kappa = 0$. Pdfs with $\kappa > 0$ are leptokurtic and have peaks that are narrower and have tails that are fatter than that of a Normal pdf. Figure 3 plots (a) the standard deviation σ versus the fractal index q, (b) the kurtosis κ versus the fractal index q and (c) the fractal index q of Formal pdfs as a function of kurtosis κ for $0.5 \le q \le 2$.

The standard calculation of σ (available in spreadsheets and most analysis programs) was used to determine the scale A for the Formal pdf. Surprisingly, Formal pdfs have unit standard deviations on the interval [1,2] with only a weak drop-off for $q < 1$, dropping by a factor of less than 15% at $q = 0.5$. This drop-off may be due to the increased sensitivity of this algorithm to the performance of the uniform random number generator at the extremes of the range [0, 1]. If so, it means that the scale $A = \sigma$ of the Formal pdf is independent of q, the index of the underlying metric space. These numerical results also suggest that Formal pdfs may well have unit variance in general. This would be a defining property of Formal distributions.

This independence of sigma from q also implies that q uniquely determines the shape of the Formal distribution. Since the kurtosis κ is one measure of the shape of a pdf, it should be possible to uniquely determine q from the kurtosis, i.e. $q = q(\kappa)$. A sum of exponentials was fitted to numerically generated Formal data of index q versus kurtosis κ (See Figure 3c). Each point in the fit represented a Formal pdf with 200,000 generated values. The function $q(\kappa)$ was then approximated by the following functional form:

Eq. 15 $q(\kappa) \approx c_1 \exp\left(-b_1 \kappa^{a_1}\right) + c_2 \exp\left(-b_2 \kappa\right).$

Formal pdfs are leptokurtic. The peak is generally sharper than Lévy pdfs and the tails are less broad, particularly as q approaches 1, but for intermediate values of q (say 1.2 to 2.0) there is little difference between Lévy and Formal distributions, apart from the faster drop-off in the tails which renders the variance unity for Formal pdfs and infinite for Lévy pdfs.

Figure 4 and Figure 5 compares Lévy and Formal distributions against S&P 500 daily data (9-July-2002 through 26-June-2006) with identical values of the fractal index q, scale σ and mean μ. (All stock data used in this paper were obtained online from http://finance.yahoo.com and were corrected for dividends and stock splits.) For the index q, the value 1.4, as quoted by Voit (2003) from studies of S&P 500 data by Mantegna, was used.

Heuristics for Fractal-Normal Statistical Analysis:

The preceding numerical analysis indicates that Formal statistics can be characterized by the fractal index q (with the kurtosis κ acting as a proxy for q), scale σ and mean μ. In this section, an outline of how Fractal Normal statistics might be applied to the general case of stock portfolios with correlations is given.

First, the means μ_i for each stock return are calculated. In general, once a column vector of weights for the stocks is given, the weighted means of the stock returns are added to determine the mean of the portfolio return. The means are fundamental to calculating the higher central moments.

Second, the covariance matrix Σ^2 of the portfolio is calculated. The elements of the covariance matrix are given by

Eq. 16 $\left[\Sigma^2\right]_{ij} = \mathrm{E}\left\langle \left(x_i - \mu_i\right)\left(x_j - \mu_j\right)\right\rangle.$

The variance of the portfolio σ_p^2 is given by

Eq. 17 $\sigma_p^2 = w^T \cdot \Sigma^2 \cdot w$

where w is a column vector of weights for each stock. Determining the optimal weights of a portfolio (e.g. Markowitz mean-variance portfolio theory), calculating the efficient frontier, calculations of betas, constructing the CAPM and like models based on the variance and standard deviation of the portfolio return as a measure of the *scale* of risk (i.e. the volatility), are included within Formal statistics without modification.

Where Formal statistics departs from Normal statistics is that while the scale of risk may be unchanged, and the algebra involving the scale of risk is unchanged, the probability of any given return depends on the shape of the Formal distribution. This requires analysis of higher moments to determine the fractal index of the underlying Formal distribution.

Therefore, thirdly, determining the fractal index q of the Formal distribution depends on determining a shape measure (for example, this paper uses the kurtosis) for the distribution of portfolio returns. For a general portfolio of N stocks, with correlations between stocks, one can proceed in two ways. If the weights are known (optimized according to Markowitz portfolio analysis, for example) then the kurtosis of the distribution of portfolio returns can be calculated directly. In this case, the advantages of Formal statistics adds little value beyond giving an estimate of risk measures like value-at-risk or VaR.

The more interesting case is presented when the weights are still to be determined, say by optimizing the portfolio's σ_p and VaR simultaneously. In this case, further analysis is complicated by the necessity of calculating a 4-th rank tensor analogous to the covariance matrix calculated above (which can be thought of as a second rank tensor). If there are N stocks in the portfolio, then there are N^4 elements in this tensor, given by

Eq. 18 $$\mathrm{K}_{rstu} = \frac{E\langle (x_r - \mu_r)(x_s - \mu_s)(x_t - \mu_t)(x_u - \mu_u)\rangle}{\sigma_p^4}.$$

This tensor is interesting because K admits the calculation of correlations between 2, 3 and 4 stocks simultaneously. Given K, the kurtosis of the portfolio is then given by

Eq. 19 $$\kappa = \sum_{r,s,t,u=1}^{N} w_r w_s w_t w_u \mathrm{K}_{rstu} - 3.$$

Note that one subtracts 3 after the summation is performed.

Finally, once the kurtosis of the portfolio is determined, one can predict the value of the fractal index q for the portfolio using the function $q(\kappa)$. One can then generate an appropriate Formal distribution with the predicted index and compare with the distribution of portfolio returns. The non-linearity of Formal statistics is contained in $q(\kappa)$. This is a significant enhancement over the current non-linear theory employing Lévy distributions, where the fractal index and scale of the pdf must be obtained with non-linear fitting of data to Lévy distributions, or other non-trivial analytical methods.

Discussion:

Probability density functions that can be cast into a universal form as a function of the dimensionless variable z with unit scale constitute a location-scale family of pdfs. All of the distributions discussed in this paper (Normal, Lévy and Formal pdfs) satisfy this criterion. For such pdfs, Eq. 2 provides a means of determining finite values for both the mean μ and the scale A where the expectation value $\langle x \rangle$ and standard deviation σ might not formally exist. Eq. 2 is a fundamental relation. The existence of the scale A can be established independently of the standard deviation of the pdf. For Normal and Formal pdfs, algebraic relations for the expectation values of the first moment and second central moment provide unbiased estimators for the mean μ and the scale *A*. For Lévy pdfs, in principle, a non-linear fit to a distribution of observed data provides a means of estimating these same quantities.

Lévy pdfs have the desirable statistical property of being leptokurtic and include the Normal pdf as a limiting case. On the other hand, Lévy pdfs are not closed; i.e. the product of two Lévy distributions with different indices is not itself a Lévy pdf which may be a limiting problem. Financial data indicates that returns for different stocks are generally distributed with different values of the index q, when Lévy pdfs are chosen as a model.

In Figure 6, the shape of the sum of two Formal distributions with different values of the index q is shown. For this case, the two distributions are uncorrelated. The agreement between the measured and predicted distributions is excellent.

Similarly, in Figure 7, a portfolio of an equally weighted sum of two stocks, Microsoft and IBM, with non-zero correlation is constructed. Five thousand daily returns (03-September-1986 through 26-June-2006) were used for each stock. On the basis of the kurtosis for each stock, 11.0 and 9.4 for IBM and Microsoft respectively, the fractal indexes were 0.79 and 0.83 respectively. The predicted (measured) fractal index for the two stock portfolio was 0.80 (0.77). From the Figure it is evident that the predicted Formal pdf does a poor job of fitting the peak but does a much better job fitting the tails of the distribution. Two conclusions can be drawn; first, that the kurtosis is not the best measure of the shape of a pdf as it is strongly biased towards the tails of the distribution and, second, that Formal pdfs that provide a better fit for the central peak of the distribution may still underestimate the tails of the distribution (relative to a Lévy pdf, for example).

A better illustration of the promise of the approach is illustrated in Figure 8 with a portfolio of twenty stocks with random weights from the top 25 stocks in the S&P500, with 1716 daily returns over six years (22-November-1999 through 19-September-2006).

In this case, the agreement between the predicted and actual distributions of returns is much improved. It should also be noted that the fractal index of the portfolio was observed to be greater than the fractal indices of the constituent stocks, indicating the portfolio is trending towards a normal distribution as stocks are added, an indication that the central limit theorem may be supported by Formal statistics.

It is evident that portfolio optimization using Fractal Normal statistics promises to be significantly more challenging than that given by Markowitz. Correlations between 2, 3 and 4 stocks are introduced through the kurtosis tensor defined above. As the size of the portfolio increases, the number of terms in the kurtosis tensor scales by the fourth power of the number of stocks in the portfolio. Fortunately, there is significant duplication between entries which renders the calculation tractable in practice.

Conclusions:

This study has explicitly constructed a family of unit variance Formal pdfs which includes the Normal pdf as a limiting case. This family of pdfs is generated using the geometry of the underlying metric space as the guiding principle, rather than using a characteristic function and Fourier transforms. Formal pdfs are leptokurtic with a shape that is uniquely characterized by the fractal index q. Formal statistics has been shown in this paper to have an algebra that is consistent with the algebra of Normal statistics.

Fractal Normal statistics provides a means of significantly enhancing the application of familiar Normal statistics. Most of the attractive features of a non-linear theory are included (leptokurtic pdfs, underlying fractal theory) while the undesirable features of the non-linear theory are excluded (infinite-variance Lévy pdf, lack of closure, difficult analytical techniques).

An intuitive leap has been made to arrive at the fundamental conjecture of this work – that there exists a family of unit-variance leptokurtic probability density functions with the attractive properties of normal statistics. It would have been more desirable if this conjecture had been based more on fundamental theorems in the analysis of fractal metric spaces, fractal trigonometry, fractal calculus and fractal Fourier transforms. However, the existent literature in these topical areas is not readily applicable to the current problem.

In summarizing the implications of these findings for finance, several are readily identifiable. The first finding is that the standard deviation of a distribution of portfolio returns is independent of q, that the foundation for CAPM and Markowitz portfolio theory will be validated, even for distributions of data described by pdfs with $q<2$.

A generalizable result arising from this study is the importance of drawing a distinction between the scale of a probability distribution and the shape of a probability distribution. The standard deviation is a measure of scale. One can associate measures of risk with the scale of fluctuations (as in CAPM) or one can associate measures of risk with the shape distribution of fluctuations (essential for calculating VaR measures, for example).

Finally, this work indicates that the theory of fractals does not necessarily demand that fractal statistics is necessarily non-linear.

References

Artzner, P., F. Delbaen, J. Eber and D. Heath, "Coherent Measures of Risk, " *Mathematical Finance,* 9, 1999, 203-228.

Fama, E.F., "The Behavior of Stock Prices," *Journal of Business*, January 1965, 34-105.

Friend, I. and R. Westerfield, "Co-skewness and Capital Asset Pricing," *Journal of Finance,* 35, 1980, 897-914.

Gardner, M., "Piet Hein's Super-ellipse," in Martin Gardner, *Mathematical Carnival: A New Round-Up from Scientific American,* New York: Vintage, 1977, 240-254.

Gielis, J., *Inventing the Circle: The Geometry of Nature*, Antwerp: Geniaal Press, 2003.

Kloeden, P.E. ,E. Platen, and H. Schurz, *Numerical Solution of SDE through Computer Experiments*, Springer-Verlag: Berlin, 1994.

Kraus, A. and R. Litzenberger, "Skewness Preference and the Valuation of Risky Assets," *Journal of Finance,* 1976, 1085-1100.

Mandelbrot, B., "The Variation of Certain Speculative Prices," Journal *of Business*, October 1963, 394-419.

Markowitz, H. "Portfolio Selection," *Journal of Finance*, March 1952, 77-91.

Peters, E., *Chaos and Order in the Capital Markets: A New View of Cycles, Prices and Market Volatility*, John Wiley & Sons: New York, 1991.

Samuelson, P., "The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moments," *Review of Economic Studies*, 37, 1970, 537-542.

Scott, R. and P. Horvath, "On the Direction and Preference for Moments of Higher Order than the Variance," *Journal of Finance*, 1980, 915-919.

Voit, J., *The Statistical Mechanics of Financial Markets 2nd Edition*, Springer: Berlin, 2003.

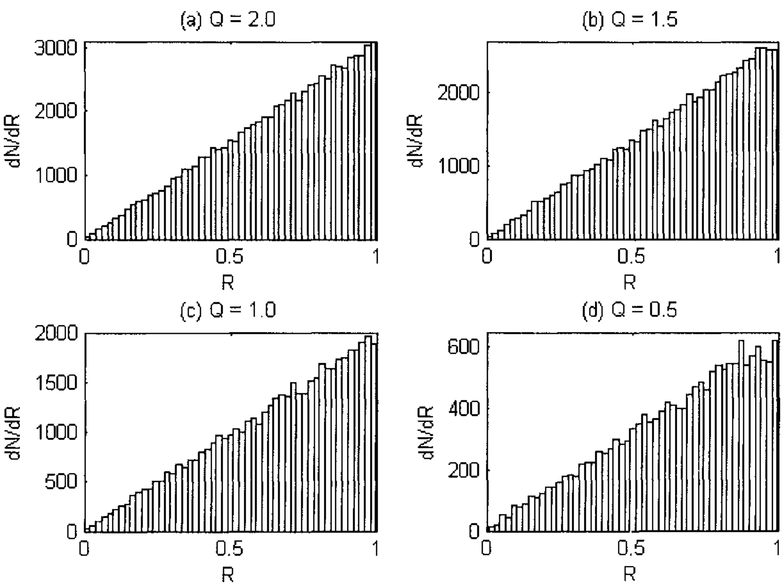

Figure 1: Radial distribution of points distributed uniformly on unit disk with fractal index q.

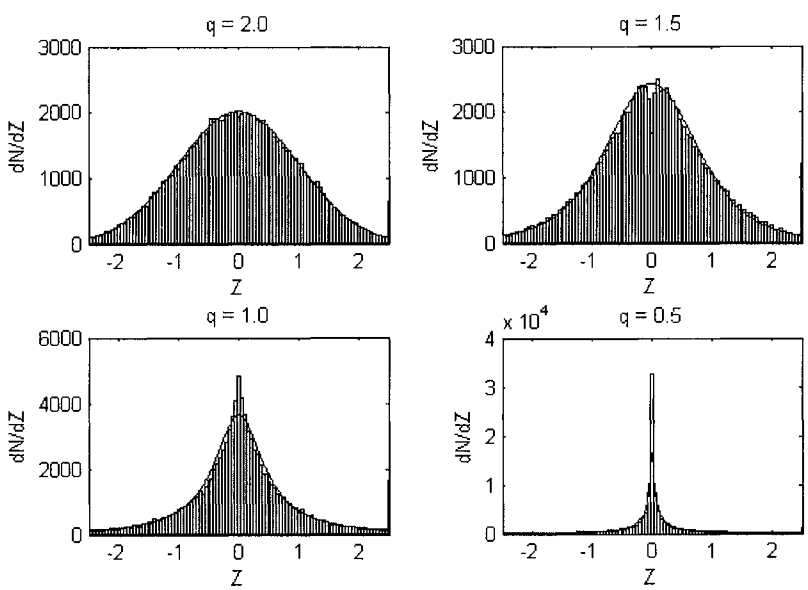

Figure 2: Comparison of generated Formal pdf versus calculated Lévy pdf with fractal index q and unit scale.

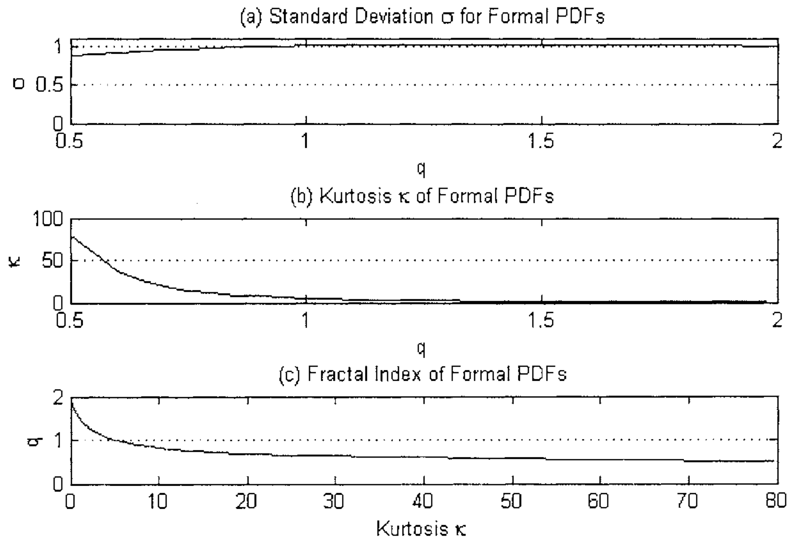

Figure 3: Properties of Formal pdfs as a function of fractal index q.

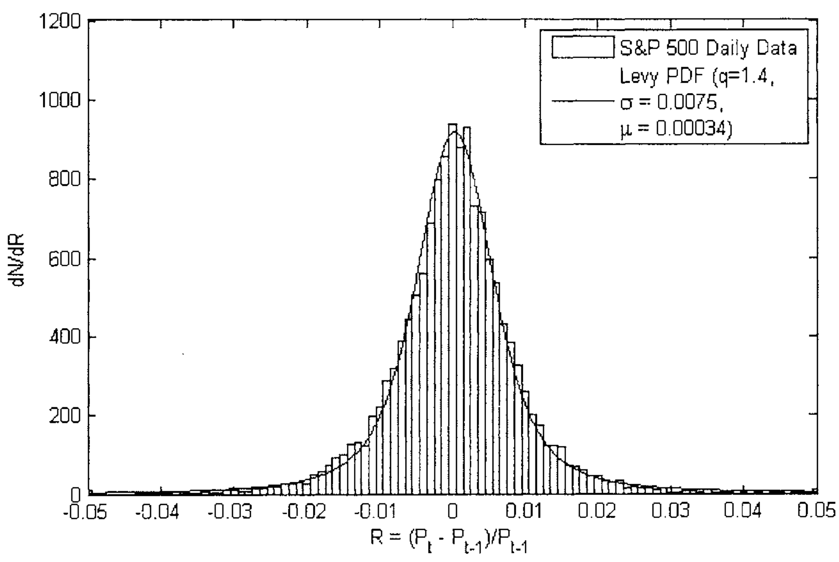

Figure 4: Comparison of Lévy pdf (curve) with S&P 500 daily data (histogram) for the period 9-July-2002 through 26-June-2006.

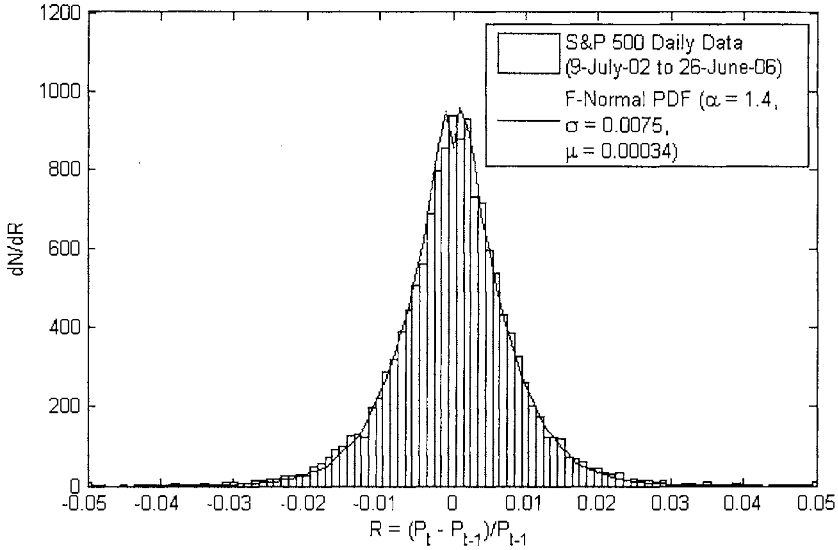

Figure 5: Comparison of Formal pdf (curve) with S&P500 daily data (histogram) for the period 9-July-2002 through 26-June-2006.

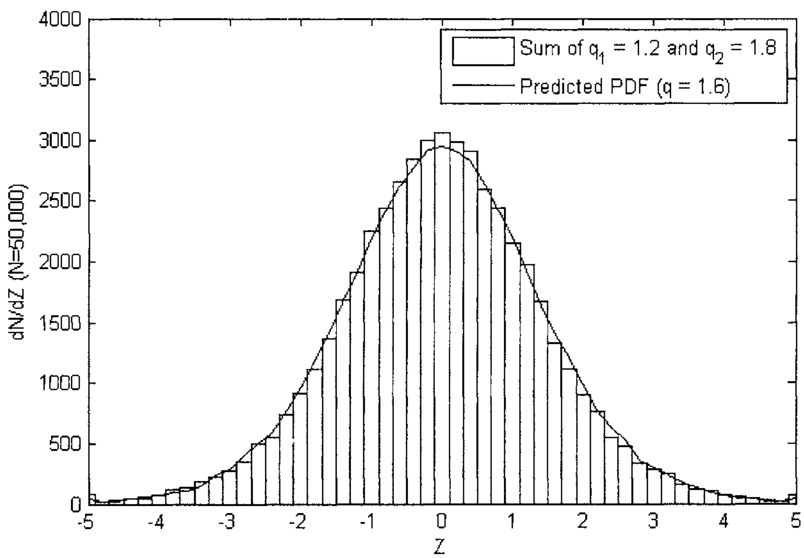

Figure 6: Sum of two uncorrelated Formal pdfs with different fractal indices q versus predicted Formal pdf.

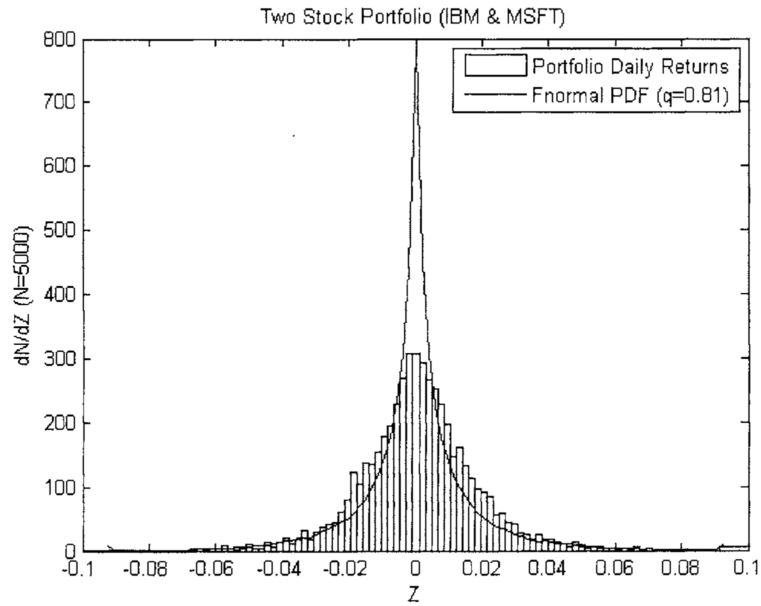

Figure 7: Two stock portfolio of equally weighted IBM and MSFT stocks (histogram) versus predicted Formal pdf (curve). (Daily data 03-September-1986 through 26-June-2006.)

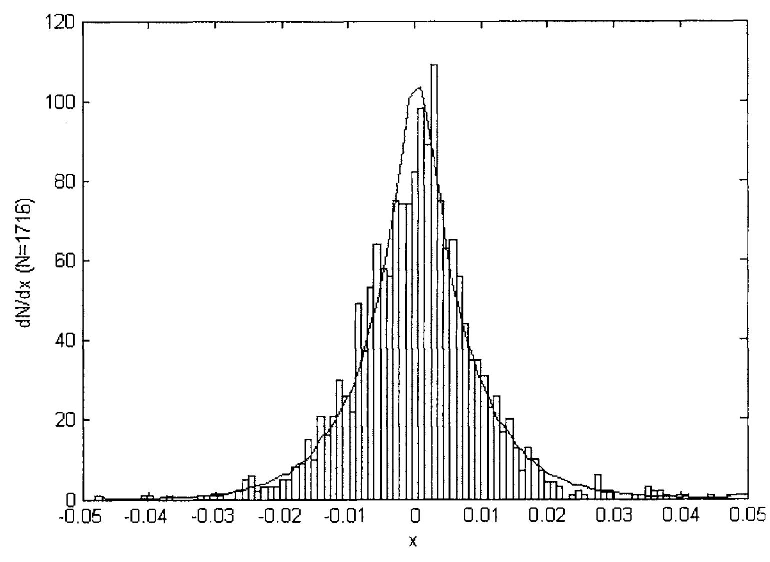

Figure 8: Comparison of randomly weighted portfolio of 20 of the Top 25 stocks of the S&P500 (histogram of 1716 daily returns) versus predicted Formal distribution (curve). (Daily data (22-November-1999 through 19-September-2006.)