# Looking Beyond The Gaussian Curve

Nicholas Mitchell

March 15, 2016

## Contents



## List of Tables



## List of Figures

# 1 Introduction

This project aims towards the development of a set of tools to aid in data simulation. Well-established tools already exist to simulate common distributions, however a goal of this paper is to explore novel methods of generating non-standard distribution, where the degree of abnormality can be introduced by the user via a set of hyper-parameters. All code is available, please contact the author[1].

Several approaches are investigated, each with its own advantages and disadvantages, depending on the desired outcome of the user. This report is focused on practical implementation of does therefore not delve deep into the theory of each individual method presented.

## 1.1 Motivation

This project with its tools aids the investigation of the common underlying assumption that financial returns are normally distributed. This assumption is seen almost ubiquitously within both theoretical and practical approaches of modelling financial markets and its agents. It also extends into the area of performance measurements for fund managers, which is our more specific area of interest.

The attraction of modelling stock returns as normal (or log-normal) distributions are numerous and consequently many theoretical and empirical works rely on such models, e.g. when pricing an option using the famous Black-Scholes formula. The assumption of normality is also seen in modelling and forecasting, where error measurements are assumed to be standard-normally distributed, i.e. with mean := 0. This is an example of where the symmetry of the normal distribution is important. Symmetry about zero as a principal represents the belief that an asset is equally likely to move upwards or downwards from its average value.

We are interested in stock returns on the basis of which the performance of a fund manager has been assessed. This assessment is as a standard performed under the assumption of standard normally distributed stock returns, however we simulate returns that serve as a means to re-assess a managers performance, should the underlying assumptions surrounding the distribution of the returns be violated.

The simulated returns to be specified by the user - expressed in terms of a normal distribution - should allow for common values of means and standard deviations to be achieved, while introducing differing, targeted values of skew and kurtosis. This will allow for performance assessment of managers on distributions the exhibit foe example fatter tails or levels of symmetry that are not present in the standard normal distribution, but traits that are detectable in actual market data.

## 1.2 Summary of Approaches

There are several possible approaches to simulating a realistic price-path i.e. returns for an asset. The three branches considered are:

1. linear combinations of well defined distributions, named 'naive approaches'
2. geometric Brownian motion and Fractional Brownian Motion (FBM)
3. distribution transformations - section 5

Each of the three approaches is explained more in their chapter introductions. It should be noted that, while this list does not exhaustively list known possibilities of distribution transformation, it does represent the most commonly used approaches.

[1]Via email: nickmitchell7 [-at-] gmail.com.

## 1.3 Trivial examples of assumption violation

### 1.3.1 Example one: the Sharpe ratio

When analysing the performance of a fund manager using any given of test or ratio, the values that we are accustomed to seeing are facets of the normal distribution. For example, let us look at the Sharpe ratio. This is a renowned measure of performance and conveys the average return a manager makes over the risk-free rate, per unit of volatility/risk. This is calculated using the mean and standard deviations of returns, or a market. If however there third and fourth moments of the market differ greatly from a standard normal distribution, we start to see some interesting results.

Let us consider two fictional portfolios, which both outperform the market. Portfolio A has a a 50% chance of beating the market by 5% and a 50% chance of beating the market by 15%. Portfolio B has a 50% chance of beating the market by 10% and a 50% chance of beating the market by 50%. When comparing the Sharpe ratios, the higher Sharpe ratio would be said to have performed better, or be more promising. When reading such probabilities of returns and asked to compare the two portfolios, any sane investor would say portfolio B is more promising and invest her money there, right? For the given portfolios, we have:

$$SR_A = \frac{0.5 \cdot 0.05 + 0.5 \cdot 0.15}{\sqrt{0.5 \cdot (0.05 - 0.10)^2 + 0.5 \cdot (0.15 - 0.10)^2}} = 2$$

and

$$SR_B = \frac{0.5 \cdot 0.10 + 0.5 \cdot 0.50}{\sqrt{0.5 \cdot (0.10 - 0.30)^2 + 0.5 \cdot (0.50 - 0.30)^2}} = 1.5$$

We see $\mathrm{SR}_1$ is the larger of the two and so is being shown to have better performance according to the Sharpe ratio, however we know that this is not the case, given that portfolio B clearly outperforms portfolio A. The Sharpe ratio can only be interpreted in the classical way when the returns and variance values lie more closely to those we know, e.g. 1% and 10% respectively. That is to say, when they follow a standard normal distribution! The large values of average returns above give the distribution a large negative skew, while the larger than usual variance values 'spread' the distribution, meaning more values lie far away from the mean.

### 1.3.2 Example two: The dangers of diversification

As a second short example, there is a comparable assumption made on a daily basis by fund managers regarding the diversification properties of a portfolio. It is believed that a large number of uncorrelated assets in one portfolio reduces the overall exposure of that portfolio to systematic risks in the market place - the large the portfolio the higher the level of diversification. While this holds well within certain limits, it isn't a water-tight assumption - as was seen in the 2007-2008 financial crisis. Huge news stories providing information to market participants can dramatically alter the correlation/co-variance matrix of any portfolio, meaning that the relationships between different asset classes change and diversification principles no longer hold. Such a phenomenon can also be measured by large changes in the skew and kurtosis of a market, which is another way of stating how the underpinning assumptions of normality are not always justified.

# 2 Initial approaches

## 2.1 Theory to validate naive approaches

A standard normal distribution (also know as the Gaussian distribution or bell curve) may be defined in many different ways, however we shall consider only the first four moments: the mean, the variance, the skew and the kurtosis. As we will manipulate them in this section in several ways, it is useful to know a little more of some relevant theory as to give confidence we are altering the parameters of normal distribution, without changing their definition as a normal distribution. The following theorems describe the characterisation of normal distributions and there combinations. As we claim that our perturbed normal distributions are indeed still normal distributions, two theorems with their proofs are provided by Quine[2].

**Bernstein's Theorem** states: *suppose $X_1$ and $X_2$ are independent random variables. Put $Y_1 = X_1 + X_2$, and $Y_2 = X_1 - X_2$. If $Y_1$ and $Y_2$ are independent, then $X_1$ and $X_2$ are normally distributed with the same variance*. If one imagines $X_1$ and $X_2$ are asset returns, with $Y_1$ and $Y_2$ being two portfolios. If both portfolios are long $X_1$, but are long versus short $X_2$, respectively, we would intuitively believe that the portfolios are independent of one another. This is true, but Bernstein's theorem proves this following the same logic in the reverse direction. It states if $Y_1$ and $Y_2$ are independent, $X_1$ and $X_2$ must be normally distributed. The second half of that statement, the *results* is what we assume to begin with: stock returns are normally distributed.

**Skitovitch's Theorem** is a generalisation of Bernstein's theorem. It states: *Suppose $X_1, ..., X_n$ are independent random variables, where $n \geq 2$ and that $Y_1 = a_1X_1 + a_2X_2 + ...a_nX_n$ and $Y_2 = b_1X_1 + b_2X_2 + ...b_nX_n$ are also independent, where $a_i, b_i > 0$. Then $X_i$ is normally distributed*. Stating the theorem in reverse: linear combinations of independent normal distributions return once again independent distributions. This is essential to know in the work below where we measure the moments of linear combinations of normal distributions using methodologies developed for normal distributions.

For more information on combining normal random variables, Statlect[3] provides a good reference to many ideas along with proofs.

## 2.2 Convergence methods, noise and auto-correlation

Using a simulated normal distribution for a set of base returns, it is possible to use Excel's Solver as a convergence methods over the returns targeting values of skew and kurtosis. Solver then changes the returns to reach these parameters. This doesn't not work very satisfactorily - it was not possible to converge to the exact desired moments, however some small levels of skew and kurtosis were introduced. Furthermore, although the moments recorded after Solver has finished seemed within an common limits, closer inspection of the actual data points shows that the returns generated - on an individual basis - were far from authentic. A large proportion of the periods showed zero returns, while there was a relatively high number with e.g. 60 % returns on single days. For this reason, the methodology is not recommended for further usage, as analysis based on these simulated returns would prove to be inconsistent with reality and so any conclusions drawn irrelevant.

In addition to this approach, the same base data was taken and levels of noise and auto-correlation was introduced into the returns. Autocorrelation was created by adding a percentage (20% was used in the results below) of the return from the previous period to the current period. Noise was added by simply adding 5% of a second random normal distribution. This had the expected effect on the returns in terms of the means and variance, however the third and fourth moments were not greatly altered. What is more, this way offers no systematic or predictable way to specify the output moments. Table 1 summarises the input and output data for both approaches. For all simulations, 10,000 time periods were used.

[2] On the Characterizations of the normal distribution (1993).
[3] Available at `http://www.statlect.com/probability-distributions/normal-distribution-linear-combinations`.

| Parameters | Base distribution | Solver | Noise/auto-corr. |
|---|---|---|---|
| Mean (%) | 0.30 | 0.33 | 3.61 |
| Variance (%) | 15.00 | 15.12 | 15.28 |
| Skew | 0.00 | 0.11 | -0.0017 |
| Excess kurt. | 0.00 | 0.17 | 0.0756 |

Table 1: Results of both the Solver and noise/auto-correlation approaches

## 2.3 Linear combinations of Gaussian distributions

The next attempt was to combine a number of Gaussian that were simulated with different parameters in the hope that there may be a controlled method of combing them to attain desired moments, giving for example *fatter tails*. Given its intuitive nature and the fact that it is supported by the theorems described in Section 2.1, simulated returns by this method should more closely reflect reality.

The simulations were carried out using Excel once again. Twenty random distribution were created for a range of mean values ranging from -10 % to 10 %. The all took the same value of variance for the input: 10 %. Once the distributions were created, a final Gaussian was created by taking weighted averages of the 20 defined distribution on each time period (10,000 were used). The only remaining question was how best to weight the distributions. A trial and error approach, adding more weight to those distributions with higher means, was used in the hope that it would put more weight into the tails for the final distribution. Figure 1 displays a histogram of the returns generated by the final Gaussian.

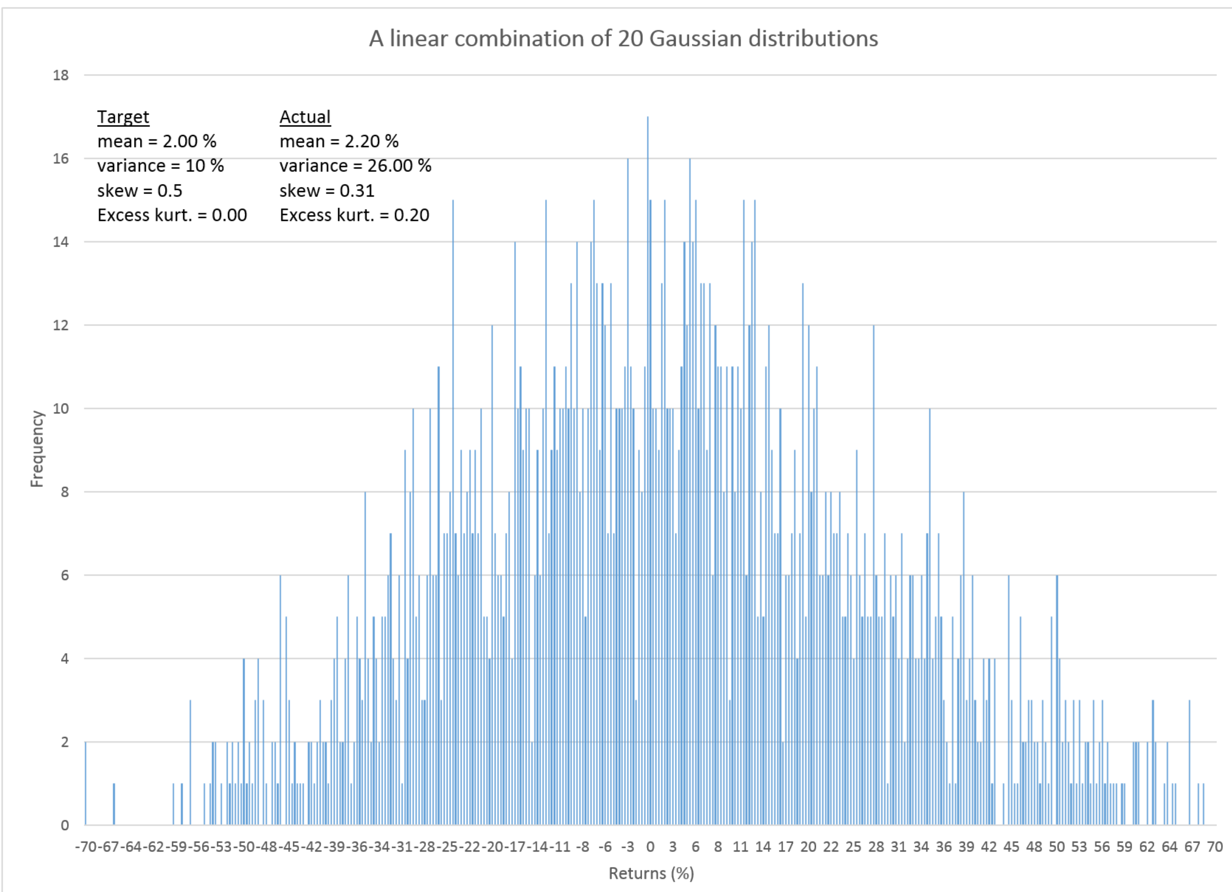



Figure 1: A Gaussian distribution - target and actual moments displayed in frame

While this method did manage to alter the skew and kurtosis of the final distribution, through amending the weights, it did also alter the mean and especially the variance of the distribution quite drastically. The results may still be representative for several markets or perhaps single stocks, but do not make for easy comparison to a larger group of assets, such as a fund or index.

### 2.3.1 Extension on linear combination theme

This methodology could be further explored with the inclusion of further distributions, such as the Gamma and Beta distributions. A selection of distributions are given with their properties and in Table 2.

| Dist | Parameters |
|---|---|
| Gaussian | mean, variance, skew, kurtosis |
| Gamma | alpha (shape), beta (rate) |
| Beta | alpha (shape), beta (shape) |

Table 2: Descriptive parameters for three distributions

As can be seen from the example distributions in Table 2, any of these could be coerced into fitting our expectations of a somewhat perturbed normal distribution. The question is: which of these provides the flexibility to be altered in a way that best aids our investigation? Perhaps more than one would be required.

A still simpler approach would be to perturb the normal distribution by adding some randomly generated noise with different mean and variance values in a Monte Carlo fashion until desired parameters are achieved. The noise could also be drawn from a normal distribution, as to comply with Bernstein and Skitovich in section 2.1. This would of course alter the values of skew and kurtosis. If we would like to compare two distributions with varying skew and kurtosis, but with very similar mean and variance we will have to find another method, as re-normalising the new distribution back to the desired mean and variance would also re-adjust the skew and kurtosis values away from the desired values.

# 3 Stochastic Monte Carlo methods

## 3.1 Definition of algorithm

This second approach to linear combinations involves combining well-defined Gaussian distributions in a more methodical way, that is without paying attention to the price-path and stochastic properties of the price-path, and without looking too much into intuitive weighting systems, but rather using a sweeping Monte Carlo method. The method is constructed as follows:

**Step 1:** simulate $\mathbf{G} : \{g_1, \ldots, g_n\}$ Gaussian distributions of length $t$ with unique mean and variance

**Step 2:** define $\mathbf{W_1} : \{w_1, \ldots, w_n\}$ to be a vector of weights, where $0 \leq w_i \leq 1$ and $\sum_{i=1}^{n} w_i = 1$

**Step 3:** create a new Gaussian distribution using $\mathbf{W_1}$:

$$g_1^* := \sum_{i=1}^{n} w_i \cdot g_i$$

**Step 4:** randomly re-order $w_i$ to create a new vector $\mathbf{W_2}$

**Step 5:** repeat Step 3 using $\mathbf{W_2}$ to define a new linear combination:

$$g_2^* := \sum_{i=1}^{n} w_i \cdot g_i$$

**Step 6:** repeat Steps 4 and 5 $m$ times, resulting in $\mathbf{G}^*$

The output of this algorithm, $\mathbf{G}^*$, is an array of $m$ Gaussian distribution each of length $t$.

This algorithm was performed with $n = 20$ simulated Gaussian distributions, again using mean values ranging from -10 % to 10 %, however with a constant variance of 20 %. With $m = 1000$ unique linearly weighted Gaussian distributions created from the original $n = 20$.

## 3.2 Interpreting results

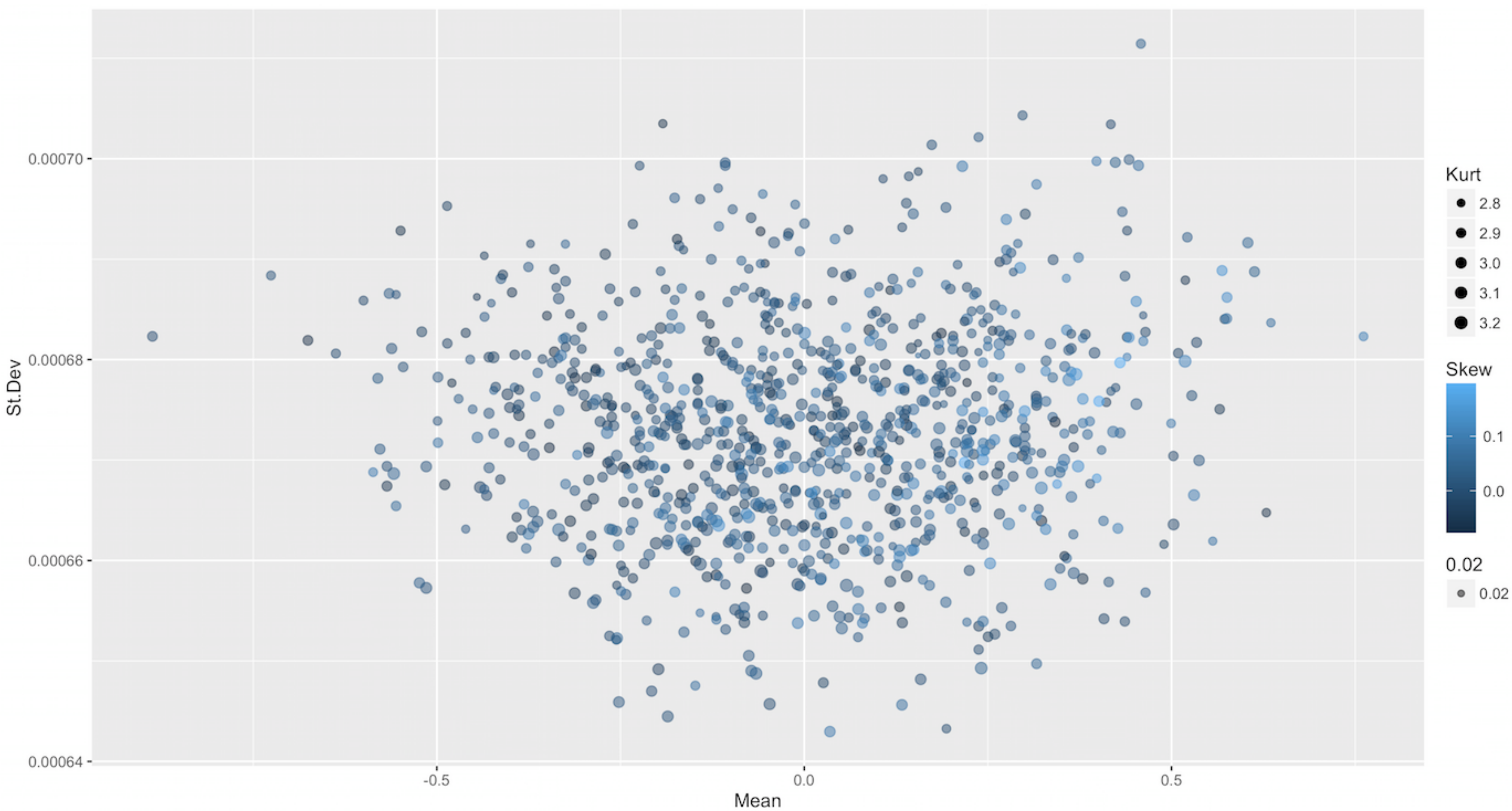
Rather than looking at the density function of each element of $\mathbf{G}^* : \{g_1^*, \ldots, g_i^*\}$ or even at a table of their moments, a more practical first step is to visualise the output in a way that quickly allows the user to narrow down the selection of outcome to those with moments that are desired. Once the most attractive set of moments is discovered, the user is able to look into the results to locate the unique configuration of weightings $w^*$ that were used to create that final Gaussian of interest $g_j^*$.

Further work must be done to optimise the distribution of the weights that are specified in $\mathbf{W}$ to be stochastically sampled from, as this will have a large effect on how far the final Gaussian distributions $\mathbf{G}^*$ can differ from the inputs, or more accurately from the average of the inputs. Inspecting Figure 2 the slight drawback that is notices is that the range of kurtosis is rather limited. This is something that may be addresses by implementing the aforementioned optimisation steps to extend this study.

Of course, the number of starting simulated Gaussian distributions and the total number of linear combinations created can be made arbitrarily large, thereby increasing the likelihood of identifying an optimal set of weights, and so the desired moments. This is of course a facet of Monte Carlo methods.

Figure 2:

Visualisation of the four moments of 1000 Gaussian distributions. Each point represents a unique stochastic linear combination of twenty simulated Gaussian distributions. The colour scale represent skew whereas the size of the points signifies the level of kurtosis



## 4 Fractional Brownian Motion

The idea of random behaviour in stock markets, meaning stochastic returns, is often modelled with Brownian motion. An extension of this is fractional Brownian motion (FBM), which assesses momentum and breaks (i.e. mini-trends) along a an stochastic price-path. The Multifractal Model of Asset Returns (MMAR) suggested by Mandelbrot *et al.* addresses the question of *how realistic is simulated data?*, as it contains many properties that one associates with market returns such as 'memory', which describes market cycles e.g. bear/bull markets. This is appealing as it will not only simulate returns around a given mean and variance, but will also factor in likelihoods of uncommon market phenomenon, bubbles or busts. In the MMAR model, this behaviour is facilitated through the implementation of a log-dependence that returns have on previous returns. The concept used is that of *local Hölder Exponents.* MMAR bridges the gap between the the common notion of geometric Brownian motion and other stochastic processes. Examples of the latter are jump-diffusion models[4] that allow for a slow diffusion process while also incorporating jumps via a stochastic Poisson process. The stochastic nature in that field requires heavy usage of Itô calculus, which is somewhat circumnavigated within MMAR, as the model permits a multiplicity of Hölder Exponents.

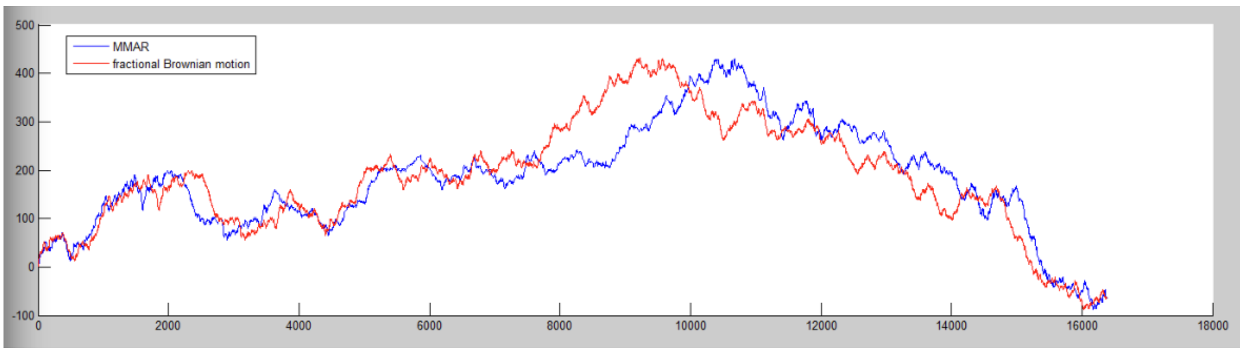



Figure 3: Stochastic price-path of an asset, clearly displaying both *bull* and *bear* phases

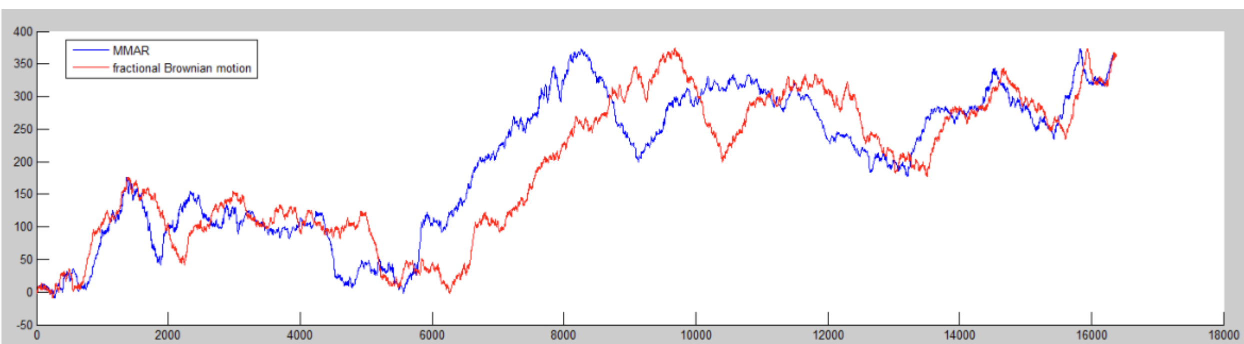



Figure 4: Stochastic price-path of an asset, showing the capability of MMAR to include phases of high volatility

Some simulations of market returns using this method were created using a Matlab implementation, with two examples of the stochastic price-paths shown in Figure 3 and Figure 4. In each figure $2^{14}$ time period were simulated, with the initial value set to zero. At first glance the price-paths in both figure do appear completely stochastic, but one does notice clear cycles and spikes. This is the strength of MMAR, that returns are realistic on a daily basis as well as over longer time periods, the log-dependency within the timeline translating to market memory. The pictures additionally compare MMAR to the underlying idea FBM.

Detailed information regarding MMAR can be found in the series of three papers released by Mandelbrot *et al.* in 1997[5]. In those papers it the theory is described and a practical implementation is provided that

[4]First described by Merton (1976): Option Pricing When Underlying Stock Returns Are Discontinuous.
[5]Included electronically with this report.

analysis the Dollar-Deutschmark exchange rates.

Further work into this direction could yield some interesting results, especially if combined with the (somewhat limited) effects of linear combinations of distributions.

# 5 Distribution transformations

The main three methodologies that were investigated were the:

1. Pearson distribution,

2. Johnson distribution, and

3. Generalised Lambda Distribution

While the area was researched in detail, no results are to be provided here - the code is however included in the tool-set.

Distribution transformations do offer a very flexible way to mutate data and could perhaps be combined with the Monte Carlo method outlined in Section 3. Please see the provided tool-set and Literature folder for more information on these methods.

# 6 Conclusions

Having evaluated many different approaches, we see there is mixed success. The goal of the study was to provide the user with as much flexibility as possible when introducing a desired level of perturbation to a base distribution. It is the author's opinion the the Monte Carlo method holds the most promise for further improvements, but that it may be enhanced by through a combination with further methods named in this study. One may choose to extend this study, in the recommended vain, by implementing some ensemble methods, where the input distributions are passed through a number of steps, including both numerical (Monte Carlo) and analytical (distribution transformation) phases.

However, as was mentioned in the introduction, the best approach is also likely to depend on the use case, which is why such a wide spectrum of tool has been assessed and provided as a result of this studied.

## 6.1 Future work

With strong methods now set in place to generate stock returns, future work utilising these results could lead to answering the question proposed in the introduction within the context of fund management: Would a fund manager's performance assessment show different results if a more accurate description of the underlying distributions were to be provided?

# 7 Summary of Tools

All the computer code and preliminary results of this study are provided electronically with this report.

1. Preliminary analysis of the distributions listed in table1 was carried out in Microsoft Excel

2. The work investigating MMAR was completed using Matlab

3. Looking at distribution transformations, including the Pearson, Johnson and Lambda distributions, was completed in a mixture of Matlab and R.