Faculty: *Economics*   Subject of study: *Financial Mathematics*

OPTION PRICING WITH RESPECT TO THE FRACTIONAL BROWNIAN MOTION AND PORTFOLIO SELECTION ACCORDING TO α-STABLE DISTRIBUTIONS

SUPERVISOR
Prof. Gennaro Olivieri

STUDENT
Gerardo Manzo
matr. 613231

ASSISTANT SUPERVISOR
Prof. Fausto Gozzi

ACADEMIC YEAR 2008/2009

1st Ranked

Winner Award "Orazio Ruggeri"

sponsored by

ASSIOM FOREX

The Financial Markets Association of Italy

June 2011

I am grateful to the anonymous referees' committee for having believed in me and in my work.

Rome, June 2011 Gerardo Manzo

To Mum, Dad, Alfonso
and my grandparents
María and Alfonso

Contents

Acknowledgments

This dissertation is the outcome of my two years lasting master's degree at LUISS "Guido Carli" University in Rome.

First and foremost, I would like to express my deepest gratitude to Mom, Dad, Alfonso and my grandparents Maria and Alfonso for believing in me, for support and patient and for their enduring encouragement.

Thanks also to my friend Enrico who is more than a friend for me.

Above all, I want to thank my girlfriend Marta for giving moral support and for making my days more and more wonderful with her smile and love.

My thankfulness and appreciation are directed also to my academic supervisor and Prof. Gennaro Olivieri for his advice and guidance.

I gratefully acknowledge the technical support and patient of the following Professors (in alphabetical order): Prof. Fausto Gozzi (LUISS "Guido Carli" University), Prof. Giulia Rotundo ("La Sapienza" University), Prof. Luisa Beghin ("La Sapienza" University), Prof. Sergio Ortobelli (University of Bergamo), Prof. Sofia Buono (Professor of English), Prof. Tommaso Proietti (LUISS "Guido Carli" University).

Rome, January 2010 Gerardo Manzo

Introduction

"*Hey, one single observation, OK, can destroy thousands of years of confirmation*"

Nassim Nicholas Taleb

The above quotation summarizes the aim of this dissertation.

If we take a look at the past, many strange events occurred over a short period of time.

In 1711, after a war which left Britain with a 10 million pound debt, the Government proposed a deal to a financial institution, the South Sea Company, where Britain's debt would be financed in return for 6% interest. With this operation many illusions of richness came about increasing ten times the asset price of the Company. But in August 1720 the price fell by 80%.

In addition, in the USA the XIX century was characterized by many speculative bubbles which alternated every twenty or thirty years. When the Revolution took place, all the infrastructural projects were financed by issuing money without a concrete gold reserve and were coined by everyone. Everything collapsed in 1873.

Moreover, the XX century began with the worst crisis, the so-called *Great Depression*, which hit hard cities all around the world.

We can continue to report other crisis until today, mentioning the latest regarding the *subprime mortgage*.

Therefore, we could think that a crisis is a physiological aspect of the economy, and if so, why don't economists take into consideration this aspect when creating their models?

For centuries people have used an easy statistical tool, the Gaussian "bell curve", to model financial theories. But the Gaussian world is where extreme events could occur with a probability close to zero, so in case of a financial crash or not, like the daily fall in Dow Jones's stock prices by 22%[1], has a chance of occurrence of 10^{-107}, practically zero.

So, one of the two aims of this work is to present an alternative to the "bell curve", introducing the α-stable non-Gaussian distributions. These are concerned with a more general concept, where the *fractal geometry* is regarded the revolution of all the sciences existing in our world. It is a new *geometric language*, which studies different aspects of diverse subjects, either mathematical or natural, that are not smooth, but rough and fragmented to the same degree at all scales. In fact, scaling plays an important role, that is, *invariance under dilations and contractions*. As we can see, by considering daily, weekly and monthly financial data, we realize that their plots are really similar upon different scale. This is not the only characteristic of a financial series, but we can observe that also the continuity is no longer acceptable. A crash allows us to introduce *discontinuity* which is intimately connected with *concentration* and *cyclicity*.

A central role in this dissertation is played by the works of Benoit B. Mandelbrot (in figure), a brilliant mathematician, who has dedicated part of his life in studying nature in an empirical way, and applying his research to finance. He is considered "the father of fractals".

[1] The Black Thursday, October 19, 1987.

While Chapter 1 is dedicated to the presentation of the most taught and used classic financial models, Chapter 2 presents this *fractal geometry* in a fascinating manner, introducing some wonderful shapes which emerge from a simple iterative function with a useful power in several fields of research.

Chapter 3 introduces a detailed mathematical and statistical presentation of the α-stable non-Gaussian distribution, showing the Mandelbrot's three states of randomness and how to make the extreme events more probable, concluding by testing empirically the *scaling* property.

Moreover, we deal with another empirical observation, where asset returns show a *long-run dependence*, so every process of independent random variables, used to simulate the asset price movements, does not represent well the reality.

So, Chapter 4 introduces the alternative to the pure Wiener Brownian motion, that is, the *self-similar fractional Brownian motion* with respect to the *Hurst exponent* H. It is constrained between 0 and 1, and based on a different market hypothesis, the *Fractal Market Hypothesis* which differs from the Efficient one.

The last two chapters are an application of the previous theories.

Our purpose is concerned with two aspects: portfolio selection according to α-stable non-Gaussian distribution (Chapter 5) and option pricing with respect to the fractional Brownian motion (Chapter 6).

They both give us excellent results which are based on a more realistic hypothesis, where a stable efficient frontier is more risk preserving than the Gaussian one for a given value, expected return or scale parameter. And an option pricing taking into consideration the empirical long-run dependence included naturally in the data.

Acronyms

σ	volatility parameter of the stock
μ	drift parameter of the stock
E	expectation operator
Var	variance operator
Cov	covariance operator
H	Hurst parameter
α	exponent of an exponential distribution
α	parameter of an α-stable distribution (Peakedness index)
β	parameter of an α-stable distribution (Skewness index)
γ	parameter of an α-stable distribution (Scale parameter)
δ	parameter of an α-stable distribution (Location or shift parameter)
Γ	Gamma function or spectral measure
w	random event or path
B_t^H	process of fractional Brownian motion at time t
τ	time to maturity
t	current time
T	maturity time
$\diamond$	Wick multiplier
C_t	value of a European call option at time t
P_t	value of a European put option at time t
C_t^H	value of a fractional European call option at time t

P_t^H	value of a fractional European put option at time t
$\hat{B}_t^H$	conditional expectation of B_t^H at time t
$B_t^{H(n)}$	discrete n-step approximation of B_t^H
r	interest rate
ε	normal random variable with zero mean and unit variance
S_t	value of the basic risky asset at time t
K	strike price of an option
A_t	value of a deterministic bond
$\hat{\sigma}_{T,t}^2$	conditional variance of B_t^H at time t
$N(x)$	value of the standard normal distribution function
$\mathbb{R}$	set of real number
ρ_H	narrowing factor of the conditional distribution of fractional Brownian motion
WBm	Wiener Brownian motion
fBm	fractional Brownian motion
P	probability measure
R_t	value of a dynamic portfolio at time t
$\Im_t$	information set at time t
ξ	binomial random variable with zero mean and unit variance
S_n	dimension of a stochastic integral
$S\alpha S$	multivariate symmetric α-stable distribution
μ^0	mean vector
r_i	return of the asset i
r_p	portfolio return
ω_i	amount of money invested in the asset i
ω	n-vector of the portfolio weights
$\bar{e}$	n-vector of ones
$\Phi_{\bar{R}}(t)$	characteristic function of α-stable distribution
Q	var-covariance matrix (for $\alpha = 2$) or dispersion matrix (for $1 < \alpha < 2$)

$[\tilde{r}_i, \tilde{r}_j]_\alpha$ covariation between two jointly symmetric α-stable random variables

σ_j another way to indicate the scale parameter of the asset j

p rate of convergence of the empirical matrix $\hat{Q}$ to the unknown matrix Q

$\bar{q}_{ij}$ elements of the dispersion matrix

α_i stability index of the i-th asset

$\beta_{i,m}$ CAPM coefficient

$EBIT$ Earning Before Interest and Tax

$EBITDA$ Earning Before Interest, Tax, Depreciation and Amortization

$NOPAT$ Net Operating Profit After Tax

UCF Operating Cash Flow or Unlevered Cash Flow

LCF Available Cash Flow or Levered Cash Flow

Chapter 1

THE EDIFICE OF MODERN FINANCE

Taking a look at the past, many "rare" events occurred during the last century in the economic world. So a spontaneous question comes to mind: are these events really rare? Or rather, what makes us define an event so rare? The economic and financial literature is full of models that were born to forecast the future values of some fundamental economic variables and to price financial assets. We can give a list of those models but they all have in common the Gaussian *bell curve*, a statistical distribution which puts us in a symmetric world where the "rare" events may occur every 100000 years, defining it as a "peaceful world".

How many crisis have occurred in the last century? We could mention many drastic events, that is, the *Great Depression* considered the longest, most widespread, and deepest depression of the 20th century. It originated in the United States, starting with the stock market crash on October 29, 1929 (known as *Black Tuesday*), and quickly spreading to almost every country in the world. And the *Black Monday*? On 19 October 1987 stock markets around the world crashed, shedding a huge value in a very short time; indeed, the Dow Jones Industrial Average (DJIA) dropped by 508 points to 1738.74 (22.61%) in one day. Moreover, the *Russian financial crisis* (also called *Ruble crisis*) hit Russia on 17

August 1998 and triggered by the Asian financial crisis. A crisis which started in July 1997, where petroleum, natural gas, metals, and timber accounted for more than 80% of Russian exports, leaving the country vulnerable to swings in world prices. Oil was also a major source of government tax revenue. Furthermore, the burst of the *dot-com bubble* (or sometimes the *I.T. bubble*), was a speculative bubble covering roughly the period from 1998 to 2001 (with a climax on March 10, 2000), where stock markets in Western nations saw their equity value rise rapidly latest in the more recent Internet sector and related fields. Besides, the latest crisis is the deepest of the 21st century, still lasting and involving a significant loss of jobs, money and investments.

So these crisis prove that our economy is not stable and as symmetric as we wish, because a revolution could occur in economics and finance, so a careful regulation in several fields is necessary but it's not our last hope; we have to review the most important economic and financial models taught and used in the whole world by economists and analysts. After reading the Mandelbrot's studies, one could wonder why, in the universities around the world, lectures continue to teach these *incorrect* models. Instead of answering this question, this chapter tries to show the most used models in economics and finance and their weaknesses, while their alternatives are presented in the next chapters. Therefore, before repairing the edifice, its foundations must be reconstructed.

1.1 The ancient vision

The risk can be studied in several ways and one of the easiest is the "fundamental analysis" which consists in researches the cause of a price variation to forecast its successive trend. For example, if the dollar falls down because of an imminent war, it drives up the oil price. It's only a subjective problem, indeed, reading a newspaper, we can fall into what N.N. Taleb defines *narrative fallacy* to indicate the unbridled and fruitless search of the causes made by journalists, only to publish an article. But it's not so easy. In the real world the causes are not so clear and sometimes the relevant information is ignored, unknown or even hidden. Also their interpretations can be wrong. So to avoid these weaknesses, financial engineering developed other instruments. The latest is the "technical analysis" which deals with searching structures, configurations and schemes through a careful empirical analysis of charts of several historical price series. Many *graphic* models are used, like double-minimum, double-max, head-and-shoulders model, and others[2]. Afterwards, "Modern Finance" was born and it is based on mathematics of uncertainty and statistics. The main concept is that the prices are not forecasted but their variations can be described by certain mathematic laws. Therefore, we can manage and measure the risk. But this is a true orthodoxy. The research in this field began in 1900, when a young French mathematician, Louis Bachelier, studied financial market in a period when no "true" mathematicians dealt with money. Bachelier created the foundations for the next big wave in the field of probability theory, invented by Pascal and Fermat in the XVII century (whose "last theorem" required 350 years to be demonstrated) to help some aristocrats with the game of chance. His main model is called *Random Walk* and is similar to the one by Pascal and Fermat. The concept is that prices have the same probability of variation up or down, as a fair coin may show one or the other side. In fact, for him, the negotiations on a trading floor is equal to a static discharge, a *White Noise*. The majority of price variations, 68 percent, is due to little positive or negative gap from the average, for an amount lower than one standard deviation; in 95 percent of the cases, the gap is lower than two times the standard deviation, while in 98 percent of the cases, it is lower than three times the standard deviation; finally, the conspicuous variations are little probable. If we line up all these variations on a graph paper, we may see a *bell*

[2] For a review of these models, see the first part of *Statistical Analysis of Mondadori Risk*, G. Manzo, 2007.

curve which is called Gaussian distribution deriving from the name of the German mathematician Carl Friedrich Gauss.

Therefore, why does N.N. Taleb talk about the "fraud of the bell curve"? We can show this through an example[3]. The average height of the US adult male population is about 178 cm, with a standard deviation of about 5 cm. This means that 68 per cent of the US men have a height between 173 and 183 cm, 98 per cent with a height between 168 and 188 cm and, 99 per cent between 163 and 193 cm. The bell curve does not exclude either a 3 meters tall man or a man with a under average height, but their respective probabilities are so rare that nobody expects to encounter them in real life. *The same* is true of gains and losses, so when we take into consideration the average of many data, we may expect to find an average height and gains equal to zero. This does not mean that the fundamental or exogenous factors are unimportant as they are very significant, however, we aren't able to forecast all these events, conseguently, we can rely on probability theory.

The generalization of Bachelier is thought is due to Eugene F. Fama of the University of Chicago, a PhD student of Mandelbrot, who asserted in his thesis the Efficient Market Hypothesis, according to which, in an ideal market all information is already included in the current price of assets. So, in his opinion, every price variation is independent, and it isn't influenced by the past data and does not influence the future ones; therefore, no one can forecast anything.

Based on these theories, many economists have built several models to analyze markets, to measure volatility and beta-coefficient of different assets and to classify investment portfolio on their risk probability. So we can say that the ancient vision of economics and finance is based on two of Bachelier's *pillars*: the statistical independence of price variation and their normal distribution.

[3] This example is taken from *The (Mis)Behavior of Markets*, Benoit B. Mandelbrot, 2005

What we want to show in the next chapters is the empirical studies conducted by Mandelbrot and others from the 1960s to now, showing the long-run dependence of price variations and their non-Gaussian distribution.

1.2 Focus on Bachelier

Let a man get drunk and put him on his way home. We can defy anyone to predict the trajectory of his steps. It seems to be very difficult, almost impossible, and this is more or less the idea behind the model of Bachelier. Past data are insignificant to forecast the future: two steps on the right, one on the left, other four on the right. Moreover we don't know if he reaches his house. This is the "Random Walk on the street".

The first reference to the concept of *random walk* appeared in 1905, in an open letter published by the British scientific journal *Nature*, entitled *The Problem of Random Walk*. Professor Karl Pearson, a member of the Royal Society, asked if any readers can provide a solution to the following problem:

> a man begins his walk from a certain point O for l meters and walks straight, then turns a corner and walks for another l meters, again in a straight line. He repeats the process n times. I would like to know the probability that after n strokes is at a distance of between r and δr from the point of departure O.

The answer came from a distinguished scientist, Lord Rayleigh, to whom Pearson said: "Lord Rayleigh's solution shows that in an open place, where it is more likely to find a drunken man, barely able to stand on his feet, is close to the point where he started".

Randomness is an intrinsically difficult idea that seems to clash in finance with instances of clear casualty, economic rationality and perhaps even with free-will[4]. What does *random* mean? In everyday language, a fair coin is called

[4] Mandelbrot's idea exposed in *Fractals and Scaling in Finance*, 1997, with whom who is writing is agreed.

random, but not a coin that shows *head* more often than *tail*. A coin that keeps a memory of its own record of *heads* and *tails* is viewed as even less random.

1.2.1 Wiener Brownian motion

Some models assume that prices change at random and each price change is statistically independent of all past ones. Random walk variations proceed in equal steps, up or down, equally spaced in time. Other models assume that asset returns follow the Gaussian distribution. Here we present the graphic version of random walk, also called *Brownian motion*, the name of the Scottish botanist Robert Brown[5], where the word *walk* denotes a motion that proceeds in *steps*, while the alternative presented afterwards, proceeds in *jumps*.

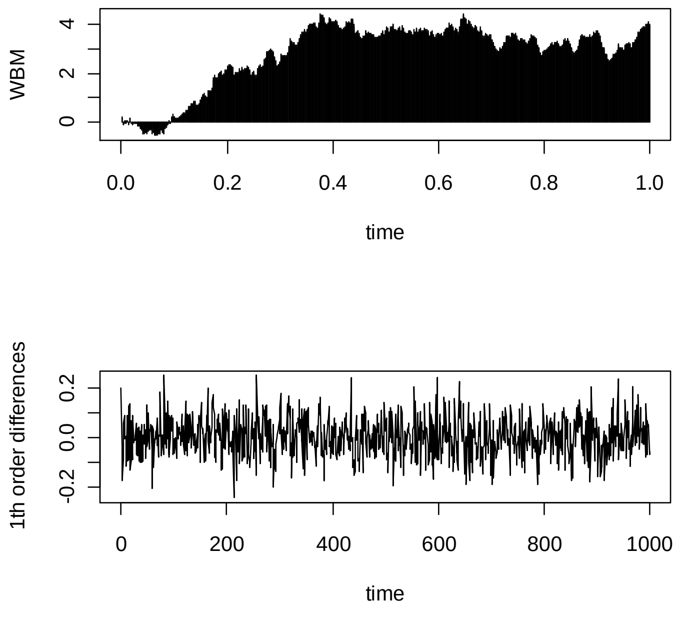


FIGURE 1-C1. Graph of a sample of Brownian motion (top), and its white noise increments in unit of 1 standard deviation (bottom).

[5] R. Brown studied the seemingly random movement of particles suspended in a fluid (i.e. a liquid or gas) closed in a tube on both sides.

Figure 1-C1 shows the simulation of Weiner Brownian motion (WBm), $B(t)$, with a sample of 1000 data and its increments of the first order. This motion is also called Wiener because Bachelier's discovery of Brownian motion in financial speculation occurred years before physicists discovered it in the motion of small particles, and decades before a mathematical theory of $B(t)$ was provided by Norbert Wiener.

The main property of WBm are best listed in two categories, as follow. It's important to define its invariance, indeed. A line, place or space, and the homogeneous distribution on them are invariant under both *displacement* and *change of scale*. Therefore, they are both *stationary* and *scaling*[6].
Both properties are extended to WBm:

a) *Statistical stationarity of increments of price*. Series of a stationary time are those whose statistical properties such as mean, variance, autocorrelation, etc. are all constant over time. Samples of increments of WBm taken over equal time can be superimposed in a statistical sense. However, equal parts of a straight line can be precisely superimposed in each other, but this is not possible for the parts of a random process.
b) *Price scaling*. Parts of a sample of WBM corresponding to the increments of non-overlapping time of different duration can be suitably rescaled so they can also superimposed in a statistical sense. This key property implements the principle of scaling, a concept which will be dealt with in details afterwards.

But *stationary* and *scaling* aren't sufficient to determine the Brownian motion, a motion which also has the following properties:

a) *Independence of price increments*. The past data are insignificant to forecast the future.

[6] That of Mandelbrot is a different concept of *scaling* we will explain in successive chapters.

b) *Continuity of price variation*. A sample of Brownian motion is a continuous curve, even though it has no derivation from anywhere. This property can be associated with liquid and depth markets where the price changes between transactions are relatively slight; so each trade results in a minimal price changes, as if the proceeding price continued through the next transaction.

c) *Rough evenness of asset returns.* A record of Wiener Brownian asset returns, in increments over equal time, Δt, is a sequence of independent Gaussian variables, called *White Noise* process. Mandelbrot shows us an empirical analysis of processes, claiming that "*the eye and ear are more sensitive to records of changes than of actual values; indeed, the ear hears it* [white noise] *like the hum on a low-fidelity radio not tuned to any station. The eye sees it as a kind of evenly spread "grass" that sticks out nowhere*"… as is showed by Figure 1-C1.

d) *Absence of clustering in the time locations of the large changes.*

e) *Absence of cyclic behavior*.

Now it's time to show the analytic view of WBM.

Wiener processes represent a particular type of Markov processes, with zero mean and variance equal to 1.

Formally, we can say that a variable Z follows the Wiener process if it satisfies these two properties:

PROPERTY 1. The variation dz in a small interval Δt is

$$dz = \varepsilon\sqrt{dt} \quad \text{where } \varepsilon \sim N(0,1) \tag{1.1}$$

so also $dz \sim N(0, \Delta t)$

PROPERTY 2. dz values in each intervals are independent.

So z follows the Markov process.

Summarizing this process, we can write it with a *drift rate* and a *variance rate*, which measure respectively, the average rate at which a value increases in a stochastic process and the variance of changes, for example a change equals to 1, the variance of changes of z in an interval of length T is $1 \times T$.

Therefore, the "Generalized Wiener Process" is

$$dX = adt + bdz \tag{1.2}$$

where a and b are constant, and through simple steps, we can show that the *drift rate* is a and the *variance rate* is b^2.

Defining $B(t)$ as being a random process with Gaussian increments, or *Brownian motion*, we can show this "Fickian"[7] diffusion rule:

$$\text{For all } t \text{ and } T,\ E\{B(t+T) - B(t)\} = 0 \ \text{ and } \ E\{B(t+T) - B(t)\}^2 = T \tag{1.3}$$

A Fickian variance is an automatic consequence if the increments are considered independent, while Fickian variance guarantees the orthogonality of the increments. In a Gaussian world, orthogonality is independence and successively on these hypothesis we will present the *Fractional Brownian motion*.

Another stochastic process is *Itô's process*, where the previous parameters a and b depends on the underlining variable, X, and on the time,t, so that we can write

$$dX = a(X,t)dt + b(X,t)dz \tag{1.4}$$

Therefore, the stochastic process for the prices of the assets can be written as follows:

[7] Only for intellectual curiosity. Fick's law relates the diffusive flux to the concentration field, by postulating that the flux goes from regions of high concentration to regions of low concentration, with a magnitude that is proportional to the concentration gradient. They were derived by Adolf Fick in the year 1855.

$$dS = \mu S dt + \sigma S dz \quad (1.5)$$

Where S is the spot price of an asset, μ is the average, σ is the standard deviation and dz follows the Wiener process $\varepsilon\sqrt{\Delta t}$.

1.2.2 Bachelier's *Theory of Speculation* and its discrepancies

Louis Bachelier's PhD thesis in 1900 *Théorie de la Spéculation* introduced mathematical finance to the world and also provided a kind of agenda for probability theory and stochastic analysis for the next 65 years or so. The agenda was carried out by the best mathematician and physicists of the 20th century, but the economic side of Bachelier's work was completely ignored until it was taken into consideration by Paul Samuelson, who introduced the quoted geometric Brownian motion in the 1960s.

Based on these hypothesis, Bachelier built his *Theory of Speculation*, which is presented here.

Consider a series of price in time, $S(t)$, and also consider its logarithmic relative by $L(t,T)$

$$L(t,T) = \ln S(t,T) - \ln S(t)$$

This model, the simplest and most important assumes that successive differences of the form $S(t+T) - S(t)$ are independent Gaussian random variables, with zero mean and with variance proportional to the several intervals T. Indeed, it implicitly assumes that the variance of the differences $S(t+T) - S(t)$ is independent of the level of $S(t)$. So we may expect that the standard deviation of $dS(t)$ will be proportional to the price level, which is why many authors suggest that the original assumption of independent increments of $S(t)$

will be replaced by the assumption of the independent and Gaussian increments of $\ln S(t)$.

Summing up, the basic model of price variation assumes that successive increments $L(t, T)$ are (a) random, (b) statistically independent, (c) identically distributed and (d) Gaussian with zero mean, so that is why it is called "stationary Gaussian random walk" or "Brownian motion".
But this model contradicts the evidence in at least two ways: Firstly, the sample variance of $L(t, T)$ varies in time and the histogram is fatter than in the Gaussian case. Secondly, no reasonable mixture of Gaussian distributions can account for the largest price changes, so owning to this, they are treated as "outliers". According to Mandelbrot, it was Bachelier who noted this discrepancies.
Other important evidences also show that considering the assumption of statistical independence of successive $L(t, T)$ is only a simplification of reality. But simplification may mean the loss of money in finance, especially if we move too much from reality. In fact, independence implies that no investor can use his knowledge of past data to increase his expected profit. But there are processes where the expected profit vanishes, but dependence is extremely long term. In this case, knowledge of the past may be profitable to these investors whose utility function differs from the market's. This type of process is called *Martingale* and will be treated in the next section.
Moreover, also the stationarity of the series of prices series is clearly contradicted. Figure 2-C1 shows a sample second moment of the daily change of $\ln S(t)$, where $S(t)$ is the spot price of The Coca-Cola Co.

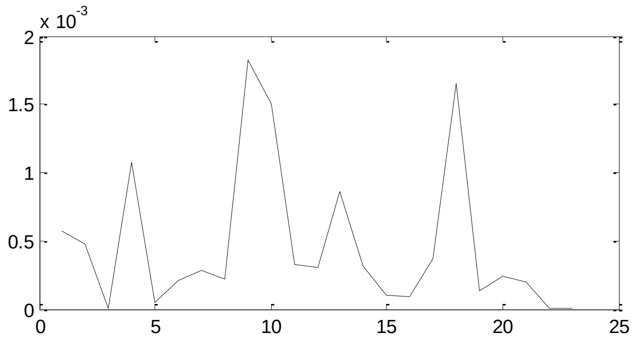


Figure 2-C1. The graph represents the variation of a sample second moment of Coca-Cola price changes. The horizontal scale represents time in weeks, with origins T_0 on January 12, 2009.

In fact, the tails of the distributions of asset returns are so extraordinary long that the sample second moment typically varies in an erratic trend. For example, the second moment reproduced in Figure 2-C1 does not seem to tend to any limit even though the sample size is enormous by economic standards.

Therefore, we can make a list of discrepancies between Brownian motion and the facts. See Figure 3-C1. It comes spontaneous to ask anyeone what is the "true" graph among those shown. It may be easier if we also show their price changes respectively as follows in Figure 4-C1.

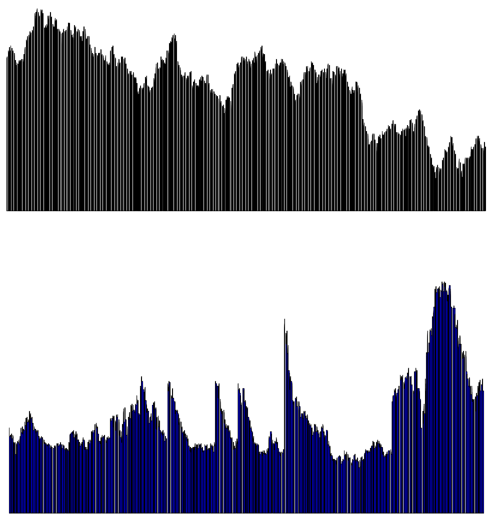


Figure 3-C1. Which is the true graph?

See Figure 3-C1, it is difficult to establish what is true and false, but if we add respectively the increments of the first order, it will become easier to differentiate them.

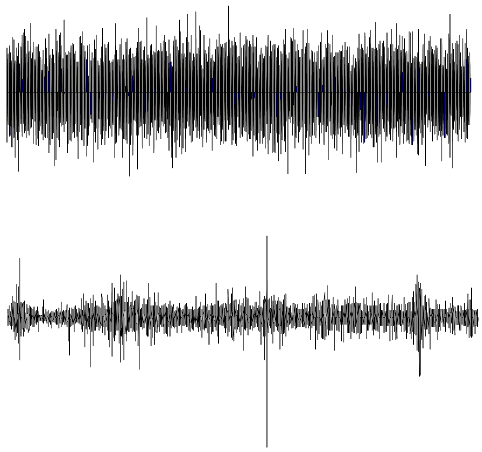


Figure 4-C1. Both graphs represent the increments of the first order of the series of the price showed in Figure 3-C1 respectively.

What we want to put in evidence is that a careful observer should perceive the rift between the two graphs. While the upper graph represents the continous process of the *white noise*, without any shadow of discontinuity, the real graph concerning the change of price of Coca-Cola Co., presents some discontinuities due to several events which occurred during the last century, that is, what N.N. Taleb defines the "black swans".

Now we are able to make the list of discrepancies previously quoted as follow:

- *Apparent non-stationarity of the underlying rules*. The diagram in Figure 3-C1 is an actual record of prices. Different pieces look dissimilar to such an extend that one is tempted not to credit them to a generating process that remains constant in time. While a record of Brownian motion changes look like a kind of "grass"[8], a record of actual price

[8] This is a metaphor used by Mandelbrot in *Fractals and Scaling in Finance*, 1997.

changes looks like an irregular alteration of quiet periods and bursts of volatility that stand out from the grass;

- Repeated instances of *discontinuous change*. In the graph of Figure 3-C1, the discontinuities appear as sharp “peaks” rising from the “grass”;
- Clear-cut *concentration*. The “peaks” rising from the “grass” are not isolated, but bunched together;
- Conspicuously *cyclic, but non periodic, behavior*;
- *The long-tailed character of the distribution of price changes* (leptokurtic);
- The existence of *long-term dependence*.

Some of these concepts have been previously explained, but the last ones, leptokurtic and long-run dependence, will be treated in details afterwards.

1.2.3 Martingale

Bachelier introduced $B(t)$ as the easiest example he had known among a broader class of processes now called *martingales*, which embodies the notion of “efficient market” and successful arbitraging. This term was first used to describe a type of wagering in which the bet is doubled or halved after a loss or win, respectively. The concept of martingales is due to Lévy, and it was developed extensively by Doob.

Prices follow a martingale process if they somehow have the following desirable property: whether the past is known in full, in part, or not at all, price changes over all future time spans have zero mean as expectation.

The analytical view.

In the probability theory, *martingale* is a stochastic process (i.e., a sequence of variables at random) where we have the conditional expected value of an observation at some time t, given all the observations up to some earlier time s, and equal to the observation at that earlier time s. Precise definitions are given below.

A sequence of variables at random $X_0, X_1, \ldots$ with finite means about the conditional expectation of X_{n+1} given $X_0, X_1, X_2 \ldots, X_T$, is equal to X_T, i.e., $E\{X_{T+1}|X_0, X_1, X_2 \ldots, X_T\} = X_T$.

A one-dimensional random walk with steps equally likely in either direction $(p = q = 1/2)$ is an example of a martingale.

This concept remains attractive, but raises serious difficulties.

The first one is this. A positive martingale always converges; that is, it eventually settles down and ceases to vary randomly[9]. Conversely, a martingale that continues to vary randomly must eventually became negative. For example, a random walk eventually becomes negative. In fact, it can be postulated that it is the *logarithm of price* that is Brownian, or at least a martingale. So, with this transformation, the price itself cannot become negative, but ceases to be a martingale. Moreover, the justification of the "efficient market" for martingales disappears.

A second difficulty is more serious. While a martingale implements an efficient market as ideal, it is not possible to implement it by arbitraging. Mandelbrot postulated that non-arbitraged price follows *Fractional Brownian motion*, a generalization of $B(t)$ will be discussed in Chapter 4.

1.2.4 The Efficient Market Hypothesis

The name of Bachelier was known, in economics only in 1965, after being ignored for many years, in the thesis of a student of Paul Samuelson, a MIT economist[10], and then in that of Fama's one.

[9] Samuelson, P. A. 1965. *Proof that properly anticipated prices fluctuate randomly*, Industrial Management Review.

[10] Doctoral student of Samuelson was Richard J. Kruizenga; his thesis is entitled *Put and Call Options: A Theoretical and Market Analysis*.

The *Efficient Market Hypothesis* (EMH) is an academic concept of study developed in the early 1960s by Eugene F. Fama, Mandelbrot's PhD student, and now professor at the University of Chicago Booth School of Business. EMH asserts that financial markets are efficient *in the sense of information*, or that the prices on traded assets (*e.g.,* stocks, bonds, or property) already reflect all known information, and instantly change to reflect new information. Therefore, according to this theory, it is impossible to outperform consistently the market by using any information already known by the market, except for a stroke of luck. Information or *news* in the EMH is defined as anything that may affect prices unknown in the present and appearing randomly in the future.

So the reader can understand the link between the concepts explained in the previous sections. In fact, after having assimilated all possible information, a price fluctuates until reaching the "new equilibrium" between buyers and sellers, so that the next variation has the same probability to go up or down. Hence, one we relax without trying to search new information because it is only a waste of time and money. For this reason Samuelson was in favour of the hypothesis that "[…] *the major part of the fund manager should abandon the business and repair plumping or teach ancient Greek or contribute to GDP as corporate managers*"[11]. This is a nihilist message, but as Mandelbrot wrote, the main characteristic of Wall Street is the flexibility and so what could be its epitaph became a war cry, that is, the development of Bachelier's thesis which takes a character of orthodoxy. But, what we are demonstrating is based on a structure built on sand.

1.3 Focus on Black, Scholes and Merton

At the beginning of the 1970s, Fisher Black, Myron Scholes and Robert Merton gave a fundamental contribution to *Option Pricing Theory*, developing the *Black & Scholes-Merton model* (B&S-m). This model is already used by many traders to

[11] These words were publicized on *The Journal of Portfolio Management* in 1974.

price derivatives and to make hedging. In 1997, three authors received the Nobel prize for economics for having created this model.
The model adopted by Black & Scholes and Merton to describe price variations of an asset is the one we have developed in Section 1.3, defining it as *Itô's process.* Before introducing B&S-m, we have to present *Itô's Lemma.*

1.3.1 Itô's Lemma

According to Lemma, the price of any derivatives is a function of the stochastic underlying variables and time.

Supposing that a variable x follows the *Itô's process*:

$$dx = a(x,t)dt + b(x,t)dz \tag{1.6}$$

where dz is the Wiener's process and a and b are functions of x and t. The variable x has a drift rate a and a variance rate b^2. On the base of *Itô's Lemma*, the function G of x and t follows the process:

$$dG = \left(\frac{\partial G}{\partial x}a + \frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2\right)dt + \frac{\partial G}{\partial x}b\,dz \tag{1.7}$$

where dz is the same Wiener's process of the equation 1.6. Therefore, also G follows *Itô's process*. Here, the rigorous demonstration of *Itô's Lemma* is beyond the goal of this work.
In Section 1.3 we have already shown that

$$dS = \mu S dt + \sigma S dz \tag{1.8}$$

with μ and σ constants, is a model of price variations. Based on *Itô's Lemma,* the process followed by a function G of S and t is

$$dG = \left(\frac{\partial G}{\partial S} \mu S + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial G}{\partial S} \sigma S \, dz \tag{1.9}$$

Note that both G and S are influenced by the same uncertain source, dz.

1.3.2 Black & Scholes-Merton model

The model adopted by Black, Scholes and Merton to describe the price trend of assets is the one we have presented in Section 1.3.1.

The rate of changes in stock price at times dt has mean μ and standard deviation $\sigma\sqrt{dt}$. So

$$\frac{dS}{S} \sim \varphi\left(\mu \, dt, \sigma\sqrt{dt}\right) \tag{1.10}$$

where dS is the stock price change at times dt and $\varphi(m, s)$ indicates a Gaussian distribution with mean m and standard deviation s. Considering the log-normal price, that is, the function $G = \ln(S)$ and developing it with Itô's Lemma, we find that the process follows by G is

$$dG = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma \, dz \tag{1.11}$$

that is, The Wiener's process with the drift rate $\mu - \sigma^2/2$ and the constant variance rate σ^2. Therefore, we can write

$$\ln(S_T) - \ln(S_0) \sim \varphi\left[\left(\mu - \frac{\sigma^2}{2} \right) T, \sigma\sqrt{T} \right] \tag{1.12}$$

and

$$\ln(S_T) \sim \varphi\left[\ln(S_0) + \left(\mu - \frac{\sigma^2}{2}\right)T, \sigma\sqrt{T}\right] \qquad (1.13)$$

that shows that $ln(S_T)$ in normal, so S_T is log-normal.

Leaving out the analytic part for a moment, now we present a list of assumptions, underlying the easier version of the B&S-m:

- it is possible to borrow and lend cash at a known constant *risk-free* interest rate;
- the price follows a geometric Brownian motion with constant drift and volatility;
- there are no transaction costs;
- the stock does not pay a dividend (but there are some extensions to handle dividend payments);
- all securities are perfectly divisible (*i.e.* it is possible to buy any fraction of a share);
- there are no restrictions on short selling;
- there is no arbitrage opportunity.

So we can assume that the underlying spot price of a share follows the process of the equation (1.8).

The payoff of an option $f(S,T)$ at maturity is known. To find out its value at an earlier time we need to know how f evolves as a function of *S* and *T*. By Itô's Lemma with two variables we have

$$df = \left(\frac{\partial f}{\partial S}\mu S + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2\right)dt + \frac{\partial f}{\partial S}\sigma S\, dz \qquad (1.14)$$

where both dz of the equations (1.8) and (1.14) are the same, so the Wiener's process can be eliminated.

Now we have to consider a trading strategy under which the appropriate portfolio is composed by

$$\begin{cases} -1: derivative \\ +\dfrac{\partial f}{\partial S}: shares \end{cases}$$

or a short-position on one derivative and long-position on $\partial f/\partial S$ shares. At time t, the value Π of these holdings will be

$$\Pi = -f + \frac{\partial f}{\partial S}S \tag{1.15}$$

The composition of this portfolio, called the *delta-hedge* portfolio, will vary from step to step. Denoting $d\Pi$ as the accumulated profit or loss by following this strategy, over a period $[t, t + dt]$, the instantaneous profit or loss is

$$d\Pi = -df + \frac{\partial f}{\partial S}dS \tag{1.16}$$

By substituting the equations above we have

$$d\Pi = \left(-\frac{\partial f}{\partial t} - \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2\right)dt \tag{1.17}$$

This equation contains no dz term, and it is completely riskless (*delta neutral*). Black and Scholes' claim that in an ideal condition, the rate of return on this portfolio must be always equal to the rate of return on any other riskless situation; otherwise, there would be opportunities for arbitrage:

$$d\Pi = r\Pi dt \tag{1.18}$$

where r is *risk-free interest rate*. By substituting the values Π and dΠ of the equations (1.15) and (1.17) we obtain

$$\left(-\frac{\partial f}{\partial t} - \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2\right)dt = r\left(f - \frac{\partial f}{\partial S}S\right)dt \tag{1.19}$$

and so

$$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf \tag{1.20}$$

The equation (1.20) is the differential equation of Black-Scholes-Merton. It has many solutions, one for each derivatives depending on S. The particular solution obtained is based upon the *boundary conditions* defining the derivative's value on a range of extreme values of S and t.

In the case of a *European Call*, the main boundary condition is

$$f = max(S - K, 0)$$

while for a *European Put* is

$$f = max(K - S, 0)$$

when $t = T$ and where S and K represent the *spot price* and the *strike price*, respectively.

But the current value of the portfolio Π is not permanently risk-free, that is, only at a small infinitesimal time, so the hedging should be reviewed after each price change.

Finally, we can present the assessment of Black and Scholes formula. The values of European call c and put p of a share at time zero which does not pay dividends are

$$c = S_0 N(d_1) - K e^{-rT} N(d_2) \tag{1.21}$$

$$p = K e^{-rT} N(-d_2) - S_0 N(-d_1) \tag{1.22}$$

where

$$d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}$$

$$d_2 = \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

$N(x)$ is a Gaussian distribution with zero mean and standard deviation equal to 1, that is, the probability that a normal standardized variable assumed a value less than x. Then, S_0 is the spot price at zero time, K the strike price, r risk-free interest rate, σ the volatility and T the maturity of the derivative.

After this analytic part, a careful reader could be aware that everything is based on a large pillar, the Gaussian distribution. But asset returns do not follow the “bell curve”, and this is demonstrated by graph. In fact if we look at Figure 5-C1, the daily price changes of the Dow Jones Industrial Average from 1916 to 2003 do not have a bell curve. The extreme values are too many. Based on this theory, the days when the variations are more than 3.49% should be 58; but they are effectively 1,001. The theory forecasts 6 days with a variation more than 4.5%, but in reality, they are 366. Besides, increments of more than 7% occur once every 300,000 years; but they occurred 48 times in the XX century.

According to this theory, the probability to see a financial collapse like that of 19 October 1987 is about 10^{-107}, or 0.000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 00001%. Hence, it's almost unlikely! We could think that there's something wrong in those used instruments. Pricing an option assuming that price changes follow a Gaussian distribution is only a simplification of the reality, so we agree with Taleb when he talks about the double-ball, the derivatives' mispricing of probabilities and profits.

1.4 Focus on Markowitz and Sharpe

Strong applications of Bachelier's ideas were made by Harry Markowitz, who achieved a PhD at the University of Chicago with a thesis on the *Modern Portfolio Theory*, and by Sharpe, an American economist who idealized an important model of asset pricing, *CAPM*, at the beginning of the 1970s, and by Black & Scholes-Merton, who we have dealt with in the previous Section.

1.4.1 Markowitz's *Modern Portfolio Theory*

In 1950 when Markowitz presented his work to obtain a doctorate, in that period, the main opinion about shares, was to choose their best asset allocation. But nothing was written on that argument in details until Markowitz wrote his work. He thought that investors do not consider only their potential profits, because if so, they would buy only one share waiting for the arrival of gains. Instead, they also think of diversifying their investments, choosing carefully different risk-based assets, creating a portfolio.

Showing this theory, we can focus our attention on the usage of some concepts typical of Gaussian distribution, mean and variance.

Having a first idea of Markowitz’s *Modern Portfolio Theory* (MPT), we present what he wrote to introduce his work, then a Nobel lecture, as follows:

The process of selecting a portfolio may be divided into two stages. The first stage starts with observation and experience, ending with beliefs about the future performances of available securities. The second stage starts with the relevant beliefs about future performances and ends with the choice of portfolio. This part is concerned with the second stage. We first consider the rule that the investor does (or should) maximize discounted, expected or anticipated, returns. This rule is rejected both as a hypothesis (to explain), and as a maximum to guide investment behavior. Next we have to take into consideration the rule that the investor does (or should) consider expected return a desirable thing *and* variance of return an undesirable thing. This rule has many sound points, both as a maxim and hypothesis on investment behavior. We want to illustrate geometrically relations between beliefs and choice of portfolio according to the "expected returns-variance of returns" rule.

One type of rule concerning the choice of portfolio is that the investor does (or should) maximize the discounted (or capitalized) value of future returns. As the future is not known with certainty, so the future must be “expected” or “anticipated” returns which we discount. Variations of this type of rule can be suggested. Following Hicks, we must considered that "anticipated" returns should include an allowance for risk, or, we could let the rate at which we capitalize the returns from particular securities vary with risk.

The hypothesis (or maxim) that the investor does (or should) maximize discounted return must be rejected. If we ignore market imperfections the foregoing rule never implies that there is a diversified portfolio which is preferable to all non-diversified portfolios. Diversification is both observed and sensible; a rule of behavior which does not imply the superiority of diversification must be rejected both as a hypothesis and as a maxim.

The fundamental concept behind MPT is that the assets in an investment portfolio cannot be selected individually, each on their own merits. Rather, it is important to consider how each asset changes in price relative to how every other assets in the portfolio changes in price.

MPT assumes that investors are risk averse, meaning that given two assets that offer the same expected return, investors will prefer the less risky one. So, an investor will take on increased risk only if compensated by higher expected returns. Conversely, an investor who wants higher returns must accept more risk[12]. The exact trade-off will differ by investors based on individual risk aversion characteristics. The implication is that a rational investor will not invest in a portfolio if a second portfolio exists with a more favorable risk-return profile – i.e., if for that level of risk an alternative portfolio exists which has better expected returns.

This theory uses a parameter, volatility, as a proxy for risk, while return is an expectation on the future. This is in line with the EMH.

Therefore, under the model, *portfolio return* R_P is the proportion-weighted combination of the constituent assets' returns, while *portfolio volatility* σ_P is a function of the correlation ρ of the component assets. The change in volatility is non-linear as the weighting of the component assets changes. In formulas:

$expected\ return$:

$$E[R_P] = \sum_{i=1}^{N} w_i E[R_i] \tag{1.24}$$

where R_i is return of asset i, and $w_i$is the weighting of component asset i.

$portfolio\ variance$:

$$\sigma_P^2 = \sum_{i=1}^{N} w_i^2\, \sigma_i^2 + \sum_{i=1}^{N} w_i w_j \sigma_i\, \sigma_j \rho_{ij} = \sum\nolimits_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \sigma_i\, \sigma_j \rho_{ij} \tag{1.25}$$

[12] The trade-off between return and risk will be formalized some years after by Tobin in his *Separation Theorem*.

where $i \neq j$ and for $i = j$, $\rho_{ij} = 1$

$$portfolio\ volatility: \qquad \sigma_P = \sqrt{\sigma_P^2} \tag{1.25}$$

1.4.1.1 Diversification strategy

Diversification is used to reduce risk and it is based on the correlation parameter constrained between -1 and 1. See the several cases when we have only two assets i and j.

When $\rho_{ij} = -1$, we get

$$\sigma_P = \sqrt{\sigma_P^2} = \sqrt{w_i^2\sigma_i^2 + (1-w_i)^2\sigma_j^2 - 2w_i(1-w_i)\sigma_i\sigma_j\rho_{ij}}$$
$$= \sqrt{\left[w_i\sigma_i - (1-w_i)\sigma_j\right]^2}$$

end $$\sigma_P = w_i\sigma_i - (1-w_i)\sigma_j$$

the portfolio variance and hence volatility is less than the sum of the individual asset volatilities, so there's a perfect diversification.

When $\rho_{ij} = 1$, we get

$$\sigma_P = \sqrt{\sigma_P^2} = \sqrt{w_i^2\sigma_i^2 + (1-w_i)^2\sigma_j^2 + 2w_i(1-w_i)\sigma_i\sigma_j\rho_{ij}}$$
$$= \sqrt{\left[w_i\sigma_i + (1-w_i)\sigma_j\right]^2}$$

and $$\sigma_P = w_i\sigma_i + (1-w_i)\sigma_j$$

that is, there's no volatility combinations able to reduce portfolio risk.

In the case $\rho_{ij} = 0$, all the assets are completely uncorrelated, hence the portfolio variance is the sum of the individual asset weights squared times the individual asset variance (and the standard deviation is the square root of this sum):

$$\sigma_P = \sqrt{\sigma_P^2} = \sqrt{w_i^2\sigma_i^2 + (1 - w_i)^2\sigma_j^2}$$

1.4.1.2 n asset case

Consider n assets, we have a matrix form as follows:

Let $\bar{x}$ be the vector of n assets and let $\bar{E}$ and S be respectively vector yields and variance-covariance matrix:

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \bar{E} = \begin{bmatrix} E(R_1) \\ E(R_2) \\ \vdots \\ E(R_n) \end{bmatrix} \qquad S = \begin{bmatrix} \sigma_{R_1} & \sigma_{R_1R_2} & \cdots & \sigma_{R_1R_n} \\ \sigma_{R_2R_1} & \sigma_{R_2} & \cdots & \sigma_{R_2R_n} \\ \vdots & \vdots & \cdots & \vdots \\ \sigma_{R_nR_1} & \sigma_{R_nR_2} & \cdots & \sigma_{R_n} \end{bmatrix}$$

So that, portfolio return and portfolio variance are severally:

$$E[R_P] = \bar{x}^t\bar{E} \tag{1.26}$$

$$\sigma_P = \bar{x}^t S\bar{x} \tag{1.27}$$

Moreover, let $\bar{y}$ be the vector of a different combination of the n assets:

$$\bar{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

So that portfolio covariance is

$$\sigma_{P_1P_2} = \bar{x}^t S \bar{y} \tag{1.28}$$

After this analytic explanation we will show the *Efficient Frontier*.

1.4.1.3 The *Efficient Frontier*

Every possible asset combination can be plotted in risk-return space, and the collection of all such possible portfolios defines a region in this space.

A *Markowitz Efficient Portfolio* is one where no added diversification can lower the portfolio's risk for a given return expectation, alternately, no additional expected return can be gained without increasing the risk of the portfolio. The Markowitz *Efficient Frontier* is the set of all portfolios that will give the lowest expected risk for each given level of expected return.

Therefore, the bonded optimization problem can assume two alternative ways:

$$\begin{aligned} &\min_{\bar{x}} \sigma_P = \bar{x}^t S \bar{x} \\ &u.c. \\ &E[R_P] = \bar{x}^t \bar{E} = given \\ &\bar{x}^t \bar{I} = 1 \end{aligned} \tag{1.29}$$

or

$$\begin{aligned} &\max_{\bar{x}} E[R_P] = \bar{x}^t \bar{E} \\ &u.c. \\ &\sigma_P = \bar{x}^t S \bar{x} = given \\ &\bar{x}^t \bar{I} = 1 \end{aligned} \tag{1.30}$$

where $\bar{I}^t = [1 \quad 1 \quad \dots \quad 1]$ n times.

Considering a portfolio composed of ten assets, and their log-returns from 5/1/2004 to 30/12/2008, a sample of 1274 data and putting in practice the equations' set (1.29), we obtain the whole frontier (on the left of Figure 5-C1) while the *efficient one* (on the right of Figure 5-C1) begins from the point of minimum variance.

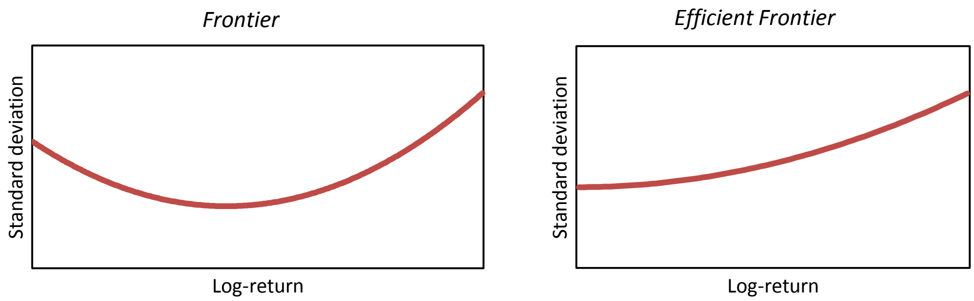


Figure 5-C1. Both graphs represent the whole frontier (left side) and the efficient one (right side). Optimal portfolio is composed by the shares of ten companies: Mondadori Inc., Fastweb, Eni, Enel, A2A, MPS, Saipem, Pirelli, Erg and Tenaris.

The *efficient frontier* is convex because the risk-return characteristics of a portfolio change in a non-linear trend as its component weightings are changed. Indeed, portfolio risk is a function of the correlation of the component assets showed by the equation (1.27), and the changes in a non-linear trend as the weighting of component assets changes. The *efficient frontier* is a parabola (hyperbola) when expected returns are plotted against standard deviations.

The region above the frontier is unachievable by holding risky assets alone. No portfolios can be constructed corresponding to the points in this region. The points below the frontier are suboptimal. A rational investor will hold a portfolio only on the frontier.

1.4.2 Sharpe's CAPM

Other contributes to Markowitz's theory was due to William F. Sharpe. He was one of the founders of the Capital Asset Pricing Model (CAPM) who created the *Sharpe ratio* for risk-adjusted investment performance analysis and contributed to the development of the binomial method for the valuation of options. It is the gradient method for asset allocation optimization, and returns-based style analysis for evaluating the style and performance of investment funds.

In his pricing model, CAPM, he found out that one would buy a share only if he forecasts profits more than Treasury bills' ones. This "more" is proportional to the accuracy with which the share reflects the performance of the market on the whole. This concept is summarized in the so-called *expected β-return* relationship, according to which the expected rate of return on an investment is directly proportional to its risk premium, as signified by its β coefficient.
In formulas, let $E(R_i)$ be the expected return of the asset i, let $E(R_M)$ be the market performance and let R_f be free-risk interest rate, we have:

$$E(R_i) = R_f + \beta\left[E(R_M) - R_f\right] \tag{1.31}$$

where $\beta = Cov(R_i, R_M)/Var\,(R_M)$.
Even in this case, we exploit the properties of the Gaussian bell curve, mean, variance and covariance, so this construction holds well only if the returns follow the Gaussian distribution.
Moreover, these approximations are not the only ones. In fact, according to two works by Fama and French[13], there are evident inconsistencies with Sharpe's model. Their works have received a great deal of attention, both in academic circles and in the popular press, with newspaper articles displaying headlines

[13] Eugene F. Fama and Kenneth R. French, *The Cross-Section of Expected Stock Returns, Journal of Finance* 47 (1992), pp. 427–66, and E. F. Fama and K. R. French, *Common Risk Factors in the Returns on Stocks and Bonds*, *Journal of Financial Economics* 17 (1993), pp. 3–56

such as “*Beta Is Dead!*”. The two main points of these papers are: firstly after, analyzing two sets of historical data from 1941 to 1990 and from 1963 to 1990, they concluded that, over the first period, the relationship between expected return and beta is weak, while, over the second one, it is virtually nonexistent; Secondly, they argue that the expected returns are negatively related to both the firm’s price-earning (P/E) ratio and the firm’s market-to-book (M/B) ratio. These debates damage CAPM which states that the average return on stocks should be related *only* to β, and not to other factors such as P/E and M/B [14].

1.5 The monstrous logical link

The edifice of modern finance stands – on condition that you can prove the correctness of Bachelier and his successors. According to Markowitz, variance and the standard deviations are the best substitute for risk – on condition that price changes follow exactly the Gaussian *bell curve*. β coefficient and CAPM make sense – on condition that Markowitz is right and Bachelier too. The Black & Scholes-Merton’s formula is correct – on the condition that price changes are described by the Gaussian distribution and are continuous.

Therefore, as Mandelbrot asserts, *this intellectual edifice is* an *extraordinary testimony of human genius, but the entire set is not stronger than its weakest*[15].

1.5.1 And now?

In 2004, Mandelbrot, talking about his model of Fractional Brownian motion based on his fractal theory, which we will deal with in the next chapters, asserted that his results are not yet used to choose shares, to negotiate derivatives or to price them; *only time, and other researches, could be decisive*.

[14] But also these critique are questioning, for example, P/E and M/B are merely two of an infinite number of possible factors.

[15] Benoit B. Mandelbrot, Richard L. Hudson, *The (Mis)Behavior of Market. A Fractal View of Risk, Ruin and Reward*, 2004

So, after five years, we can present some fundamental works in the last part of this thesis, which could be a revolution for the whole finance world, dealing with one of the new theory of portfolio selection and the option pricing in a fractional Brownian market, using the power of *Fractal Geometry*.

Before moving to the next chapter, we are able to understand some empirical principles which will be inserted in the models we present.

Five behavioral rules of markets:

- *Markets are risky*

  Extreme asset returns are an intrinsic rule, they often happen, so one cannot exclude them using the Gaussian *bell curve*, because in our world "normal" does not mean symmetric, lack of improbable variations, everything could occur, so an exact forecast is impossible, there's always something which completely slips the human mind and genius. But knowing the violent and wild market turbulence is already useful;
- *Troubles never come alone*

  The turbulence concentrations tend to cluster. So a *bull* tendency could continue for a short or long period, but we must know that it could suddenly stop and collapse, becoming a *bear* one.

  Hence, price changes are discontinuous;
- *Markets have a personality*

  Traders must not operate in the markets only focusing attention on the exogenous factors like news and other people, because prices are formed also by endogenous factors which are typical of the intern market functioning. According to Mandelbrot, wars begin, peace comes back, firms fail, everything that influences prices comes and goes, but the fundamental process by which prices react to news does not change;
- *Markets deceive*

No one is able to forecast everything only thanks to past data. Here the metaphor of the turkey, exposed by Taleb in *The Black Swan*, sounds good. What is the faith's tendency of a turkey which is fed everyday by a human hand? We can think it is bull one, increasing daily, but a man is not a turkey, he has a brain for thinking and anything could happen, that there could be a day in which everything vanishes. Hence only past data are insufficient to forecast future, because the turkey may appreciate the human goodness, but what happens to the bull trend when the turkey is killed? Of course, faith vanishes;

- *Market time is relative*

  Mandelbrot has introduced a new concept of time. A trader must focus attention not only on the *clock time*, but also on the *trading time*, a different time from the linear one of our normal way of thinking. This time speed up the time clock during periods of high volatility and slows during periods of stability.

After all, we let the readers interpret the meaning of Taleb's phrase stated during an interview: "Hey, one single observation, OK, can destroy thousands of years of confirmation"[16].

[16] *Report*, *Top Theorists Examine Rippling Economic Turbulence,* , 21 October 2001, moderating by Paul Solman, guests N.N. Taleb and B.B. Mandelbrot. See Section "Interview" for the complete interview.

Chapter 2

FRACTAL GEOMETRY: FASCINATING CHAPTER

Fractal Geometry is the central issue of this work and it is the revolution of all sciences existing in our world. Instead of writing a summarized definition of Fractal Geometry, we prefer to present Mandelbrot's definition written in a paper 1989[17], as follows:

> Fractal geometry is a workable geometric middle ground between the excessive geometric order of Euclid and the geometric chaos of general mathematics. It is based on a form of symmetry that had previously underused, namely invariance under contraction or dilation. Fractal geometry is conveniently viewed as a language that has proven its value by its uses. Its uses in art and pure mathematics, being without 'practical' application, can be said to be poetic. Its uses in various areas of the study of materials and of other areas of engineering are examples of practical prose. Its uses in physical theory, especially in conjunction with the basic equations of mathematical physics, combine poetry and high prose. Several of the problems that fractal geometry tackles involve old mysteries, some of them already known to primitive man, others mentioned in the Bible, and others familiar to every landscape artist.

Hence, a ground named *fractal geometry*, which is a middle between the extremes of linear and *organized* geometric order and *orderly* geometric chaos.

Standard geometry has long proved to be effective in sciences. Yet there's no question that Nature fails to be locally linear. But, on the other side, complete

[17] Benoit B. Mandelbrot, *Fractal Geometry: what is it, and what does it do?*, Yale University, 1989.

chaos could not conceivably lead to a science. The monstrous beauty of this geometry is its astonishing results that often seem to involve structures of great richness, generated by algorithms of extraordinarily shortness and simplicity.
We can assert that fractal geometry is a new *geometric language*, which studies different aspects of different objects, either mathematical or natural, that are not smooth, but rough and fragmented to the same degree of all scales. The concept *scaling* is intimately connected with the main characteristic of fractals, the so-called *symmetry*, that is, “invariance under dilations and contractions”, but it will be showed in the next sections.
A goal of science is to describe nature quantitatively. To *see* is a skill that we must learn, and that we must learn what to measure. Therefore, the eye plays an important role to analyze nature first. The *Elements of Euclid* is the oldest treaty based on an almost modern mathematical thought, which is concentrated on regularity. Lines, planes and spheres best represent the reality, but we do not live in Plato’s *World of Ideas*, just look around. In 1982, Mandelbrot, in his manifest *The Geometry of Nature*, asserted that: “*Why is geometry often described as 'cold' and 'dry'? One reason lies in its inability to describe the shape of a cloud, a mountain, a coastline or a tree. Clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in a straight line”.*

The scope of this chapter is to provide a detailed explanation of everything concerning fractals, in a fascinating way, showing in part, the specific mathematical terms whose tools has proved to be attractive and fruitful, but by no means easy.

2.1 Preliminary studies on the irregularity

Fractalists' mathematical language keeps on evolving and expanding with each new use. Their straightness must be seen with their own eyes, observing reality and experiencing.

The main assumption based on Gauss and Legendre is to consider the deviation of reality from an ideal form, when they introduced the *Ordinary Least Square*. But we cannot consider irregularity as an imperfection, a mere gap from an ideal world. It is the essence of nature, and of finance too, and it will be showed in the last part of this chapter. The fundamental point of *fractal geometry* is to set regularity in what irregularity is, that is, the basic structure of what amorphous is. Mathematics has always tried to research the so-called invariances, or symmetries, the main properties that do not change from an object of study to another. These "invariances under dilations and contractions" link a set with its parts: each part is a linear geometric reduction of the whole, with the same reduction ratios in all directions.

Fractal geometry is the instrument used to individualize these repeated configurations, to analyze, quantify and manipulate them. Having an idea of a fractal, you might see a Roman broccoli, which is a fascinating natural fractal.

2.1.1 How to construct a fractal

The simplest fractals are constructed by starting from a classic geometric object like a triangle, a straight line, a spherical solid, which is called *initiator*. Moreover, it is necessary also a *generator*, a shape, which is a simple geometric form: a zig zag line, a wavy curve, or as we shall see for the diagram of prices, a bullish or bearish sequence of prices of few dollars. Hence, the first construction stage replaces each side of the initiator by an appropriately rescaled, translated and rotated version of the generator. Then, a second stage repeats the same construction with more broken line obtained at the first stage and so on, so that the process follows a *recursive rule*.

See Figure 1-C2 to have an idea on how to construct a fractal.

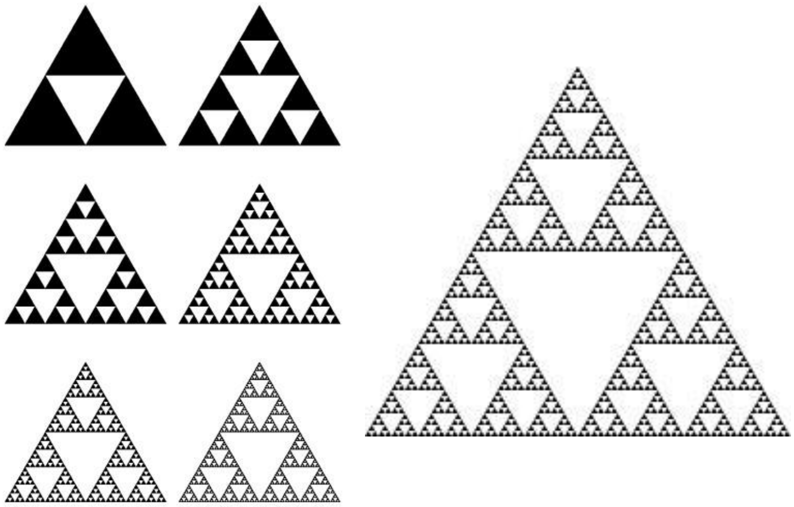


Figure 1-C2. *Sierpinski gasket***.** The initiator is a black triangle, while the first of the six triangles on the left (side) is the generator. The others show the recursive process.

One of the most popular examples of this phenomenon was invented by the Polish mathematician Waclaw Sierpinski. Any triangle can be cut into four congruent triangles, and the first step in creating Sierpinski's figure is to remove the middle triangle. The next step is to remove the middle of each of the remaining triangles. Repeating this over and over again it creates what is known as the *Sierpinski gasket*. This object has the remarkable property that doubling its size produces a figure composed of three copies of the original figure. If we double the size of one-dimension object, we obtain two copies of the original; if we double the size of two-dimension object, we obtain four copies of the original. The Sierpinski gasket has such a dimension similar to raising two to that exponent, so that we obtain three. There is no whole number with this property, so the dimension of the Sierpinski gasket lies somewhere between one and two. Specifically, it is the logarithm of three to the base of two.

Another figure is created by such an infinite process is the *Koch snowflake*. The snowflake starts with an equilateral triangle. On each of the 3 edges of that triangle, we erect a triangle one-third larger. Then we erect a triangle one-ninth larger than the original on each of 12 edges of the previous figure. So each of these edges is replaced by four smaller edges one-third longer than the original. It follows that the total perimeter is multiplied by $4/3$ at each stage. Leading this process to its limit, we have *Koch's snowflake* as follows in Figure 2-C2.

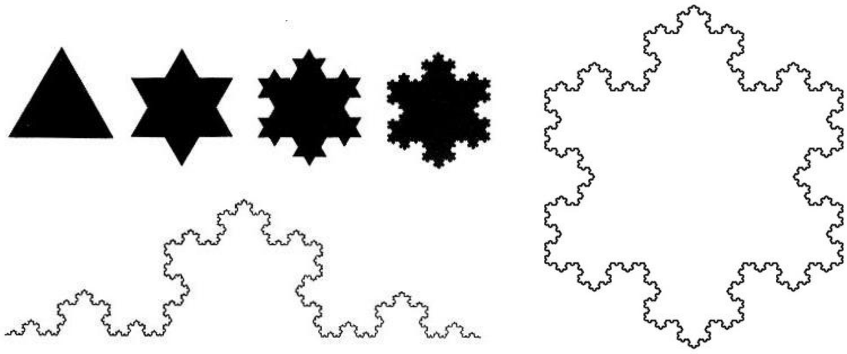


Figure 2-C2. *Koch snowflake*. The initiator is a black triangle in the upper left, while the second of the four triangles is the generator. The graph below is a side of Koch snowflake which is on the right.

Continuing to expose fractals, we close this graphical presentation of the "beauty" showing two more, *Cantor dust* and *Mandelbrot set*.

Cantor dust (or *Cantor set*), introduced by German mathematician Georg Cantor in 1883, is a set of points lying on a single line segment that has a number of remarkable and deep properties. Consider a line segment of unit length. Remove its middle third. Now remove the middle thirds from the remaining two segments; then the middle thirds from the remaining four segments and so on. If we can continue this construction infinitely, what remains is a remarkable subset of the real numbers called the Cantor set, as shows in Figure 3-C2.

Figure 3-C2. *Cantor dust*. It starts from a simple straight line and the generator is the same line. The result is the disappearance of any solid line, in fact, remains only a sprinkle of points.

The example which concludes this section is the picturesque Mandelbrot set, so named not by Mandelbrot himself but by his colleagues. This set illustrates the deep link between fractal geometry and chaos theory. Using a very simple algorithms: $z_1 = z_0 + c$, where z_0 is the starting point of the recursive process, c is a constant constrained between 3 and 4 and z_1 is the first output. To construct this, we can assign a pixel on the monitor, coloring it. The surprise is that if we observe it on a smaller and smaller scale, we will discover that the graph became more and more complicate, as it is shown in Figure 4-C2.

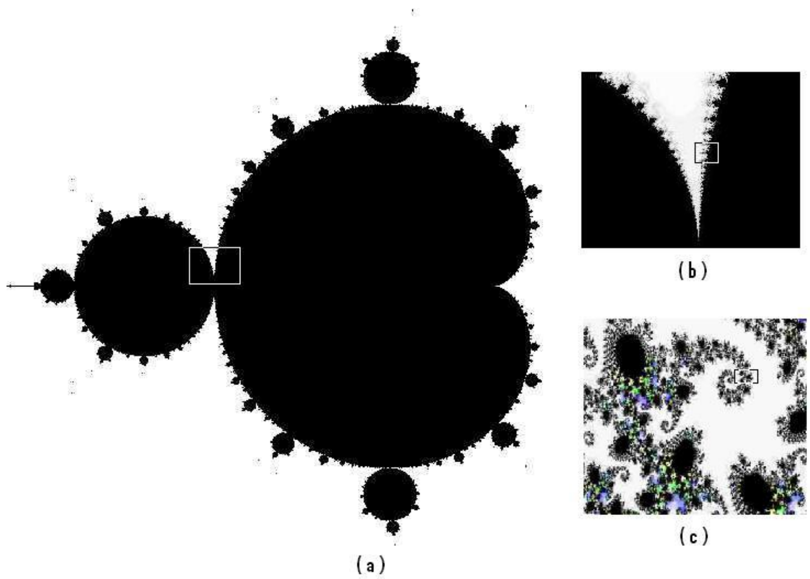


Figure 4-C2. *Mandelbrot set* (a). (b) and (c) are zoomed.

Except for the last graph of Figure 4-C2, the other fractals seem to come out from simple processes. But if we can repeat it more times with more details, it could become more complicated, for example, transforming the quoted zig zag line in a jugged curve, or shuffling at random the segment of a generator. Many fractals do exist, and all have in common the same characteristic: the *scaling*. Here each part recalls the whole under a specific ratio, and it is due to this if the fractals are defined *self-similar* and *self-affine*. Before continuing this analysis, it's important to say that we will show the mathematical approach of scaling distributions and other methods in details in the next chapter, so what we present below will be developed later.

2.1.2 Self-similarity and self-affinity

We have asserted that the main fractal characteristic is what we call *scaling*[18], meaning that each part recalls the whole under a specific ratio. When this ratio is the same in all directions, we can define a fractal *self-similar*. They are similer to the zoom of a camera which enlarges or reduces everything framing in the same size, showing at different focal lengths, something similar. Moreover, we must be careful not to confuse similarity with identity, because we refer to the fact that each part is *similar* to the whole, but not necessarily *identical* to the whole. A classic example of self-similarity by Mandelbrot is the coastline of Britain[19], which looks jagged from a satellite views, so if you zoom in to 10,000 feet, it looks pretty similarly jagged. If you zoom in to 5,000 feet, the coastline looks similarly jagged yet again, and so on. The concept of self similarity doesn't require *exactly* the same view from each altitude, but only that it is similar in its texture, irregularity, or coarseness. See Figure 5-C2 for an example of self-similar fractal.

[18] As we shall see, according to Mandelbrot, scaling is one of the two invariance principles in economics. The other is *stationarity*.
[19] B. B. Mandelbrot, *How long is the coast of Britain? Statistical self-similarity and fractional dimension*, Science, 1967.

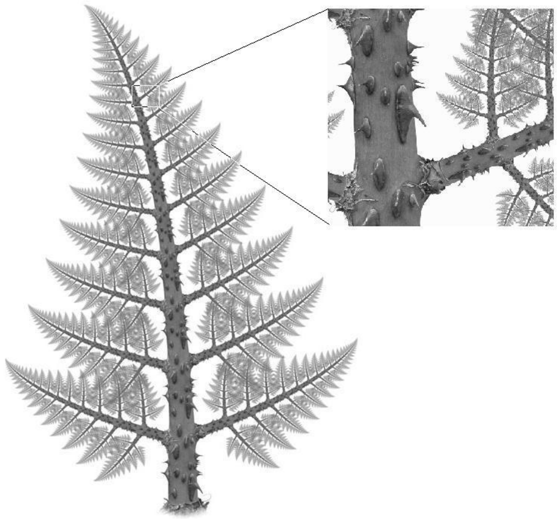


Figure 5-C2. The tree is a black-and-white digital artwork graphically constructed by iterating an arrangement of a small portion of a photograph of a rosebush (Copyright 2007 Robert Fathauer)[20].

Furthermore, when the quoted ratio changes from a direction to another, as we shall see in Bachelier's graphs or in the graphs of asset returns, the fractals are defined *self-affine*. Hence, while *self-similar* fractal structures have symmetrical properties that not only depend on the direction (they scale isotropically) but also depend on the direction in space, a *self-affine* fractal does not look similar if it is viewed from a closer distance. It *looks similar* only if one of the space coordinates is rescaled appropriately.

The roughness of *self-affine* fractals can be quantified by means of the *Hurst exponent,* $\boldsymbol{H}$, ranging from 0 to 1. A large part of Chapter 4 is dedicated to this exponent, but in Appendix we will show its estimation methods. See Figure 6-C2 for an example of *self-affine* fractal.

[20] We are grateful with Robert Fathauer to have given permission to use this image.

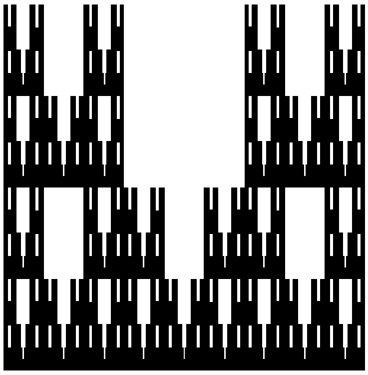


Figure 6-C2. The pieces are scaled by different amounts in the x and y directions.

We invite the reader to pay attention to the difference in meaning of the two last concepts, because, unfortunately, many probabilists persist in using *self-similar* when they really mean *self-affine*.

2.1.3 Multifractality

If fractals follow different factors of scale in several point of enlargement or reduction, they are called *multifractals* and their properties are more complicated and powerful. Many patterns that seem fractal in a first approximation prove, on a second look seem to be multifractals. What we will show afterwards is that the most general category of graphs allows the generator to include diagonal boxes with different values of the exponent H, which is constrained between $H_{\min} > 0$ and $0 < H_{\max} < \infty$. So these graphs are necessarily multibox. In other words, multifractality allows some generator intervals to be axial, hence includes graphs that combine continuous variation with jumps.

2.2 Fractal dimension

According to Mandelbrot, the main notable characteristic of fractal geometry is the particular way to conceive the dimension. For Euclid a point does not have a

dimension, a line has two and our world three, while Einstein added another one, the time.
To have a good idea of what we want to demonstrate when we talk about fractal dimension, try to imagine a cotton ball. From a great distance, it is almost visible; effectively it is a dot, without dimension according to Euclid. If we put it in our hands, it will become a three-dimensional object. Observing it in details, we realize that it is a tangle of fibers in one-dimensional space. If we look it more closely, the fibers are three-dimensional filaments. Proceeding with a microscope, we find again points without dimension. Therefore, it's natural to ask how many dimensions a cotton ball has? Zero, one or three?
It is important to consider such dimension an instrument of measure. Indeed, fractal dimension is a statistical concept that measures how a fractal fills the space as we zoom down to smaller and smaller scale. Mandelbrot realized that the coast of Britain has a fractal dimension of about 1.25. Does this make sense? The answer is *yes* because we must consider that the fractal dimension of a line is 1, while that of a square is 2, so an irregular coast is more complicated than a line, but it may never fill a bi-dimensional space like a square. Another example: if we measure the surface of a lung, we can realize that it has a big surface similar to that of tennis court, but its fractal dimension is near 3, so its nature is almost three-dimensional. Hence, a fractal dimension is the first instrument to measure how a thing is convoluted, twisted.

2.2.1 And what about the definition on fractal dimension?

Having presented literally what we mean by fractal dimension, now we present some of the most used definition of it, mentioning names such *Rényi* and *Hausdorff*. Furthermore, practically, the *box-counting dimension* and *correlation dimension* are widely used, partly due to their ease of implementation.
For some classical fractals, all these dimensions coincide, but in general they are not equivalent. Consider fractal in Figure 2-C2, the *Koch snowflake,* a careful

observer should note that the same ratio, $4/3$, is always repeated. It means that in the first recursive step and so in the others, there are four broken line on a straight line long one third of the lenght of the object. But the lenght of the curve between any two points on the Koch Snowflake is infinite.

There are two main approaches to generate a fractal structure. One is to grow from a unit object, and the other is to construct the subsequent divisions of an original structure, like the Sierpinski triangle in Figure 1-C2. Follow the second approach.

2.2.2 Similarity dimension

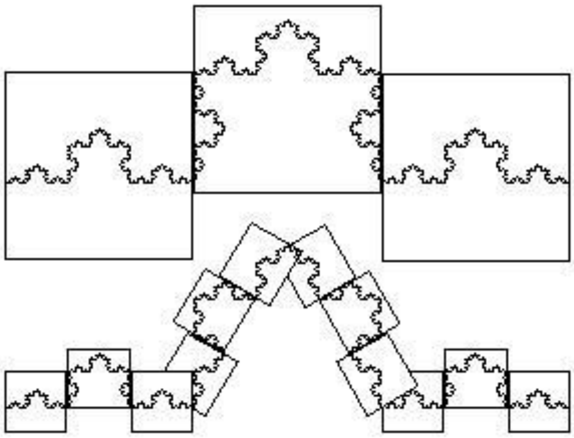
Consider an object with a linear size equal to 1 residing in Euclidean dimension D. Reducing its linear size by the factor $1/r$ in each spatial direction, it takes $N = r^D$ number of self-similar objects to cover the original one. In this case, when the ratio of enlargement or reduction is the same, fractal dimension coincides with *similarity dimension, D_s*, so that

$$D_s = \frac{\ln(N)}{\ln(1/r)} \tag{2.1}$$

where r is the ratio of enlargement or reduction and N is the number of self-similar objects to cover the original one. In the Koch case, $r = 1/3$ and $N = 4$, so that $D_s = \ln(4)/\ln(3) = 1{,}2618 \ldots$ which is more or less the Hausdorff dimension.

2.2.3 Dimension based on box counting

As the name implies, this dimension is obtained by counting the numbers of boxes of several largeness as necessary to cover a fractal form. Consider again the Koch curve, we can see that three boxes of one third-length are necessary to cover the curve, as follows in Figure 7-C2.



$r_1 = 1/3\,, N(r_1) = 3$

$r_2 = 1/9\,, N(r_2) = 12$

Figure 7-C2. Fractal dimension of Koch curve based on box counting.

Continuing the process in Figure 7-C2, we could find that the fractal dimension becomes

$$D_b = \lim_{r_n \to 0} \frac{\ln N(r_n)}{\ln(1/r_n)} \tag{2.2}$$

where $N(r_n)$ is the number of boxes of radius r_n necessary. So the dimension is again $D_b = D_s = \ln(4)/\ln(3) = 1{,}2618 \ldots$ Moreover, for the Sierpinski triangle in Figure 1-C2, it is $D_b = \ln(3)/\ln(2) = 1{,}5849 \ldots$

2.2.4 Other dimensions

Similarity and box counting dimensions are two of a multitude of variation on the notion of dimension.

The *mass dimension* is based on the idea of how the mass of an object scales with the size of the object, assuming unchanged density. Locating a dot P inside the object, near the middle and denoting by $M(r)$ the amount of mass of the object inside the circle of radius r and centered at P, if the *power law* relation

$M(r) = kr^D$ holds over some range of r values, then the mass dimension $D_m = D$.

Furthermore, when self-similar objects are described by a power law[21] like $M(r)$, Hausdorff dimension is defined by $D = \ln M(r)/\ln r$, that is, the dimension of the scaling law. However, in general, the proper dimension to use turns out to be the Minkowski-Bouligand dimension, D_{MB}. Let $F(r)$ be the area traced out by a small circle with radius r following a fractal curve, then

$$D_{MB} = \lim_{r \to 0} \frac{\ln F(r)}{-\ln r} + 2 \tag{2.3}$$

allowing that, for all strictly self-similar fractals, the Minkowski-Bouligand dimension is equal to the Hausdorff one. Otherwise, $D_{MB} > D$.

Hence, this is an instrument used to measure every kind of phenomena, also the variability of a financial diagram.

2.3 Fractals in finance

In finance, fractals have followed a tortuous route. In 1987, after the quoted crash[22], Wall Street was finding new ideas so fractals became a trend. More and more researchers dealt with them so, from then on, fractals are becoming an alternative to the obsolete financial and economic models.

Fractals in finance are used to make prediction as to the risk involved for particular stocks. To discover them, one must be a careful empiricist. Indeed, many real data seem to exhibit both global dependence and long tails, the main characteristics of financial data which have contributed to the collapse of the edifice of modern finance. It is important to dedicate some chapters only to this particular topic because of its central role played beyond the goal of this work,

[21] For a detailed close examination see Chapter 3.

[22] See Section 1.1 of Chapter 1.

mentioning revolutionary instruments such as *Fractional Brownian motion* and *Lévy distributions' family*.

Naturally, the fact that price changes follow the Gaussian distribution is not the only pillar which *modern finance* is based on. Another pillar is their independence. For many years financial analysts and researchers have inquired into the short-run dependence, asserting that in a small time interval price variations influence each other, also conjecturing the so-called *thrust effect*, that is, when a price starts to raise, an increase is more probable than a decrease. Conversely, in the middle-run, from three to eight years, what seems to happen is exactly the contrary; a share which has increased over some years has a greater probability of decrease rather than an increase.

Other economists have also conjectured the *effect of passenger mania*, that is, for years, investors could be in favour of a firm, so sales increase earnings too, but suddenly everything could change and a reverse trend could occur.

Chapter 4 will present another concept of dependence, the long-run one, which has an infinite effect.

Finally, we can assert that markets are not those conjectured by the recent theories, but are more turbulent and dangerous, so a new mathematical and statistical baggage must be introduced.

2.4 After all

Fractals can seem at random. It is often difficult to classify them on the basis of the classical geometry and analysis. They are frequently irregular and unpredictable, but the most important point is that they have a simple beginning. Every fractal is the logical expression of a simple idea, rule or mathematical relation.

The code to construct a fractal is composed of three letters, *initiator*, *generator* and *recursive rule*, similar to the genetic alphabet DNA. Remember that the beauty of Mandelbrot set, Figure 4-C2, comes out from a simple but powerful expression.

Fractal geometry has its roots in several fields. In 1733 the poet Jonathan Swift wrote:

So, Nat'ralists observe, a Flea
Hath smaller Fleas that on him prey,
And these have smaller Fleas to bit'em,
And so proceed ad infinitum.[23]

Hence, some poets had anticipated this phenomenon.

At the beginning, fractals were presented as paradoxes. Indeed, Giuseppe Peano showed how a one-dimensional line can fill the space which is bi-dimensional, the so-called *space-filling curve*[24]. In 1918 both Fatou and Julia worked on what we would think of as the more standard types of fractals, creating a shape which nowadays can be designed by the computer[25]. Furthermore, some of the modern interests in fractals among the people comes from the computer-assisted work of the IBM fractal project and 20th century mathematicians like Benoit Mandelbrot after whom the *Mandelbrot*, *set* is named. Pictures of this set are quite fascinating and show surprising features at every level of detail. Until the introduction of the computer, very little from this area of mathematics was in the form of graphics. IBM's fractal project added the pictures to what is inherently an incredibly visual field of mathematics. This added dimension of the field in turn gave rise to new discoveries and incarnations of fractals, such as mountains and clouds.

We conclude this chapter showing the last two fractals, Peano curve and Julia set, in Figure 8-C2.

[23] J. Swift, *On Poetry, a Raphsody*, 1733, vv. 337-40

[24] See Mandelbrot B.B., "Ch. 7: Harnessing the Peano Monster Curves", *The Fractal Geometry of Nature*, W. H. Freeman, 1982 or Sagan, Hans (1994), *Space-Filling Curves*, Springer-Verlag.

[25] See Evgeny Demidov, *The Mandelbrot and Julia sets Anatomy*, 2003

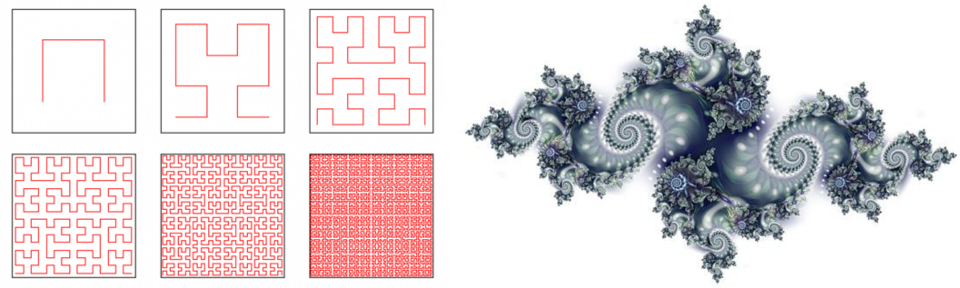


Figure 8-C2. *Peano space-filling curve* on the left and *Julia set* on the right.

What we will show in the next chapter is the beauty of the powerful Mandelbrot's discover on the new statistical methods that contribute in introducing more precision and stronger realism in financial models.

Chapter 3

MARKET TURBULENCE: DISCONTINUITY AND SCALING

For years on end, many great banks have picked up information from their operations room. But what is all this if EMH is valid? According to EMH, if information is available on the markets, it is already included in prices, so there's no opportunity of arbitraging. But why does the collect information center exist in a financial intermediary? "We do not believe in efficient markets, so we collect information", this is a declaration of a financial analyst during an interview. Indeed analysts spend part of their work-life collecting and analyzing information, studying the mutable "volatility surface" of markets. In the B&S-m, this surface should not be interesting, it should be flat. Actually, it is turbulent and wild, from which derives the major part of intermediaries' profits.

Hence, markets are turbulent, not as peaceful as the majority of the models shown, also for the non-rationality of investors who often operate without a logical scheme, without a theoretical utility function. Nowadays there's an important research field, which is trying to confute this secular theory, called *behavioral finance*[26], whose treatment is beyond the aim of this work. Besides, all investors are not equal, homogeneous, as the classical theory foresees, having different time horizons, different approaches to shares, bonds, derivatives and others financial instruments. Moreover, another wrong assumption is the continuity of asset returns. All financial graphs show signs of discontinuity. Now we are in a époque where the motto *Natura non facit saltum* of the text *Principles of Economics* of Alfred Marshall, 1890, is no longer valid. Price jumps are sometimes significant and sometimes unimportant. Discontinuity is the most important ingredient of financial markets, so it must be introduced in the models. A result of this leads us not to support the Wiener Brownian motion showed in the Section 1.3.1 of Chapter 1.

Therefore, the fractal approach to finance and economics is based on two features: the first is the importance of invariances and the possibility of identifying *stationary* and *scaling* as invariance principles in economics; the second is the recognition that the probability theory is really versatile, and capable to define three states of randomness.

In a way, this chapter will show an important concept, recognized but always neglected by economists in their models, that is, *discontinuity* which is intimately connected with concentration and cyclicity. Moreover, we present a detailed mathematical presentation on *stable non-Gaussian random processes*, introducing *Lévy's processes*.

[26] This particular analysis applies scientific research on human and social, cognitive and emotional factors to better understand economic decisions by consumers, borrowers, investors, and how they affect market prices, returns and the allocation of resources. See D. Kahneman, A. Tversky (1979), *Prospect Theory: An Analysis of Decision under Risk, Econometrica* 47 (2): 263–291.

3.1 Market turbulence and randomness

Asserting that the track of settlements and diagram of prices, can be described as random processes is not the same to say that is irrational rather than unpredictable. Financial prices are unpredictable and uncontrollable. The best thing we can do is to evaluate the probability that a certain event happens. Probability is the only instrument we have at our disposal in financial markets which are faster and faster, more and more variable. *A posteriori* it's simple, easy and manic to explain the link between causes and effects, as everyone can do this, but it is too late, in the meanwhile one could gain or lose money. The central aim of Section 3.2 is to present the three *states of randomness*, which, according to Mandelbrot, correspond to three well-known distributions: the *Gaussian*, the *log-normal* and the *scaling with infinite variance*. They differ deeply from one another, from the viewpoint of the addition of independent addends in small or large numbers. The three states are: *mild*, *slow* and *wild*.

3.1.1 States of Randomness: the probabilistic approach

Why must we consider three states of randomness? When we face with a new phenomenon of fresh dataset, the first step is to identify its state of randomness. Mandelbrot's states of randomness come from some physical concepts, that is, he recalls that gases, liquids and solids are distinguished through two criteria: flowing versus non-flowing, and having a fixed or variable volume. Therefore, the three states of randomness are also defined by two mathematical criteria. Given a sum of N independent and identically distributed random variables, those criteria depend on two notions: *portioning* and *concentration*.

Take the random variables U defined by the tail probabilities, $Pr(U > u) = P(U)$, where u is a cutoff, *portioning* concerns the relative contribution of the addends U_n to the sum $\sum_1^N U_n$. Moreover, we must consider also the *concentration ratio* of the largest addend to the sum, that is, *concentration* is the situation that prevails when this ratio is high. Conversely, the situation that prevails when no addend predominates is called *evenness*. Hence, the contrast between concentration and evenness leads to three principal categories:

- *Mild randomness* corresponds to short- and long-run dependence;
- *Slow randomness* corresponds to short-run concentration and long-evenness;
- *Wild randomness* corresponds to short- and long-run concentration.

Consider these in details.

Long-run, $N \to \infty$, is enough to distinguish between the "wild" state of randomness and the remaining states, called *pre-Gaussian*. An example of long-run portioning is the largest daily price which increases over a significant long period of time. Pre-Gaussian yields approximate equality in the limit, expressing that the largest addend is relatively negligible. On the contrary, wild randomness yields an increasing concentration, so the largest addend remains non-negligible. Hence, we can summarize the main idea in this way: the *portioning in the long-run* can take two different forms: *even*, with concentration converging to 0 as $N \to \infty$ and *concentration* with the largest addend remaining of the order of magnitude of the sum.

On the opposite, *short-run*, $N = 2$ or "a few", concerns with the case where, given two independent and identically random variables U' and U'', we must understand whether these parts of U $(= U' + U'')$ are more or less equal or dissimilar. Consider the example on page 124 of *Fractal and Scaling in Finance*:

> Suppose you find out that the annual incomes of two strangers on the street add up to \$2,000,000. It is natural and legitimate to infer that the portioning is concentrated, that is, there is a high probability that the bulk belongs to one or the other stranger. The \$2,000,000 total restricts the other person's income to be less than \$2,000,000, which says close to nothing. The possibility of each unrelated strangers having an income of about \$1,000,000 strikes everyone as extraordinarily unlikely, though perhaps less unlikely that if the total were not known to be \$2,000,000.

And about the last case, the *middle-run*? According to Mandelbrot, an economic long-run only matters when it approximates the middle-run reasonably, or at least provides a convenient basis for corrective terms leading to a good middle-run description.

Furthermore, the state of randomness differs according to how fast the generalized inverse function[27] P^{-1} decreases as its argument tends to 0, that is, according to how fast the moments U^q increase as $q \to \infty$[28]. Hence if the tails are very heavy, the value of a variable is very large in absolute value, and if the longer tails are heavy, less and less moments do exist.

Trying to define pre-Gaussian randomness, its fluctuation is averaging, so the law of a large number (LLN) shows that sample averages converge asymptotically to population expectations; its fluctuation is Gaussian, that is, the central limit theorem (CLT) shows that the fluctuations are asymptotically Gaussian; its fluctuation is Fickian, that is, the central limit theorem also shows that they are proportional to $\sqrt{N}$. The last result will be useful when we talk about the pure Brownian motion because of its equal distance in time which is independent[29].

All these properties fail when we deal with distributions that have an infinite population variance and where the dependence is not short-range or local but long-range or global. This is the case of the scaling distribution with $\alpha < 2$, called *α-stable*, which deserves a particular treatment because of their complexity. Besides, they are not pre-Gaussian but define the *wild* state of randomness.

And what about the non-Gaussian limit? Each value $\alpha < 2$ defines the domain of attraction, that is, if U_n is in the domain of universality of $\alpha < 2$, the limit is a random variable called *L-stable*. Indeed, in contrast to the width of the domain of the Gaussian, each of those domains is extremely narrow and reduces the variable for which $Pr\{U > u\} \sim u^{-\alpha} L(u)$, where $L(u)$ is logarithmic or slowly

[27] The inverse probability function is the quantile function, that is, the function that, for each value constrained between 0 and 1, gives the x-axes where the distribution has that value.

[28] Mandelbrot, B. Benoit, *Fractals and Scaling in Finance*, 1997.

[29] We consider both LLN and CLT because long-run deals with the behavior of a distribution as N increases.

varying in the sense that for all h, $\lim_{u\to\infty} L(hu)/L(u) = 1$. Let us make an analytical clarification: we take a sequential sum $\sum_1^N U_n$ for a sequence of independent and identically distributed random variables U_n, with $1 \leq n < \infty$. It is possible to choose the sequence A_N and B_N so that $A_N\{\sum_1^N U_n - B_N\}$ converges to two distinct limits, random and non-random, that are:

- If we chose $A_N = 1/N$ and $B_N = 0$, for LLN the limit is $E[U]$, that is non-random;
- If we chose $A_N \sim 1/\sqrt{N}$ and $B_N = NE[U]$, for CLT the limit is Gaussian, and so random.

Those two ways define the pre-Gaussian behavior.
But in the absence of a varying term $L(u)$, the choice of A_N is $A_N = N^{1/\alpha}$ for all α, therefore $\sum_1^N U_n$ is asymptotically of the order of $N^{1/\alpha}$. Then, the choice of B_N is $NE[U]$ when $1 < \alpha < 2$, and 0 when $0 < \alpha < 1$.

&&& Pre-Gaussian category &&&
To understand better what *concentration* and *portioning* mean, we must refer to detailed concept of probability theory. The quoted residual pre-Gaussian category[30] can be subdivided into two categories: the first is based on *concentration in mode* and the second is based on asymptotic *concentration in probability*, which will lead to the "tail-preservation" relation $P_N(u) \sim NP(u)$.
The last relation comes from a simple concept: let us consider the random variables U_j, with $1 \leq j \leq N$, independent and identically distributed, with the tail probability $P(u)$, and $\tilde{P}_N(u)$ is the tail probability of $\widetilde{U}_N = \max_j(U_j)$, so $1 - \tilde{P}_N(u) = \{1 - P(u)\}^N$ and, in the tail, $\tilde{P}_N \sim NP$. The concept of concentration is related to that of location of the most probable value called "modes". Those

[30] We define this as *residual* because of the characteristic of the *wild* state, that is, concentration long-run portioning.

locations lead to a criterion based on the convexity of $\log P(u)$, which will serve to define "concentration versus evenness in mode"[31]. Indeed, let us consider the probability density of the variables U' and U'', $P(u)$ and their convolution[32] denoted by $P_2(u) = \int P(u)P(x-u)\,du$, when u is known, the conditional probability density of u' is given by the so-called *portioning ratio*:

$$\frac{P(u')P(u-u')}{P_2(u)} \tag{3.1}$$

where the denominator is a constant and it remains to study the numerator, whose integral is dominated by values that the conditional density takes in intervals near the modes. Leaving out all analytical demonstrations, we can write that

- when the graph of $\log P(u)$ is cup-convex, ∩, the portioning ratio is maximum for $u' = u/2$ and portioning is even in terms of the mode;
- when the graph of $\log P(u)$ is cap-convex, ∪, the portioning ratio is minimum for $u' = u/2$ and portioning is concentrated in terms of the mode;
- the last case, when the graph of $\log P(u)$ is straight, the addend are exponential, and the portioning ratio is a constant.

But, according to Mandelbrot, these implications have many flaws. Moreover, if there's concentration in mode, that is, $P(u')P(u-u')$ is maximum for u' near 0 and u' near u, apart from u' near $u/2$, and considering the following convolution's partition (given a value of $\tilde{u}$ that satisfy $\tilde{u} < u/2$) :

[31] Mandelbrot, B.B., *Fractals and Scaling in Finance*, 1997, page 131
[32] In statistics, probability distribution of the sum of two independent and identically distributed variables is the convolution of each distributions.

$$P_2(u) = \int_0^x P(u)P(x-u)\,du = \left\{\int_0^{\tilde{x}} + \int_{\tilde{x}}^{x-\tilde{x}} + \int_{x-\tilde{x}}^{x}\right\} P(u)P(x-u)$$
$$= I_L + I_O + I_R \qquad (3.2)$$

we can define *short-run concentration in probability* if the middle interval $(\tilde{u}, u-\tilde{u})$ has the following properties as $u \to \infty$:

- the relative probability in the middle interval $I_O/P_2(u)$ tends to 0;
- the relative length of the middle interval $(u-2\tilde{u})u$ does not tend to 0.

Moreover, *short-run concentration* in probability prevails when $\log P(u)$ is slow to vary, to decrease and cap-convex, its derivative $P'(u)/P(u)$ tends to 0 as $u \to \infty$. Therefore, the tail-preservation concept is related to the fact that, being $P(u)$ and $P_N(u)$ respectively the tail probabilities of U and of a sum of N variables with the same distribution, we obtain:

$$P_N(u) = NP(u) \qquad (3.3)$$

&&& Extreme randomness &&&

Wild randomness means that the largest of many addends is of the same order of magnitude as their sum. Hence, we must consider the moments of the quoted probability density:

$$E[U^q] = \int_0^\infty u^q P(u)du \qquad (3.4)$$

which have a global maximum for some value $\tilde{u}_q$ defined by the equation

$$0 = \frac{d}{du}[q\log u + \log p(u)] = \frac{q}{u} - \left|\frac{d\log p(u)}{du}\right|$$

where the dependence of $\tilde{u}_q$ on q is ruled by the convexity of $\log p(u)$.
Consider the following three cases:

- when $\log p(u)$ is rectilinear, the $\tilde{u}_q$ are uniformely spaced;
- when $\log p(u)$ is cap-convex, $\tilde{u}_q/q$ is decreasing;
- when $\log p(u)$ is cup-convex, $\tilde{u}_q/q$ is increasing.

But, knowing the $\tilde{u}_q$ is not enough, we must also know that $u^q P(u)$ is in the neighborhood of $\tilde{u}_q$, and this distribution is approximated to Gaussian's as follows:

$$\log[u^q P(u)] = \log p(u) + qu = constant - \left(u - \tilde{u}_q\right)^2 \tilde{\sigma}_q{}^{-2/2}$$

so the moments $E[U^q]$ are in the interval $\left[\tilde{u}_q - \tilde{\sigma}_q, \tilde{u}_q + \tilde{\sigma}_q\right]$, and Gaussian's moment are characterized by *delocalization*, which means that its *q*-intervals overlap for all the values of σ. Moreover, the log-normal's moments are *uniformly localized* because of its non-overlapping of the *q*-intervals, when it is skew. In addition, these moments are *asymptotically localized* when the *q*-intervals cease to overlap for a high *q*.

3.1.2 States of Randomness: the seven based on the probabilistic approach

To understand better how to identify a state of randomness, we present from Mandelbrot's list seven states of randomness, which help us to classify them thanks to the concepts exposed in the last section.
We have mentioned that we can discover them through a careful attention to the moments' probability distribution and its inverse function.
Therefore, the rate of increase as function of q of the moments $E[U^q]$ or the scale factor $\{E[U^q]\}^{1/q}$ are the most important instruments.
Note the following classification from mild to wild state of randomness:

- *Proper mild randomness*. Evenness of the short-run portioning for $N = 2$, P^{-1} increases near $x = 0$ no faster than $|\log x|$, or $\{E[U^q]\}^{1/q}$ increases near $q \to \infty$ no faster than q. [Example: the Gaussian];
- *Borderline mild randomness*. Concentrated short-run portioning for $N = 2$, but even beyond a finite cutoff. [Example: the exponential $P(u) = e^{-u}$, which is the limit case of a particular function called Gamma[33], for $\gamma < 1$];
- *Slow randomness with finite and delocalized moments*. P^{-1} increases faster than $|\log x|$, but no faster than $|\log x|^{\Gamma/w}$, with $w < 1$, or $\{E[U^q]\}^{1/q}$ increases faster than q but no faster than a power $q^{1/w}$. [Example: $P(u) = exp\{-u^w\}$, with $w < 1$].
- *Slow randomness with finite and localized moments.* P^{-1} increases faster than any power $|\log x|^{1/2}$, but less rapidly than any function of the form $exp(|\log x|^\gamma)$, with $\gamma < 1$, or $\{E[U^q]\}^{1/q}$ increases faster than any power of q but remains finite. [Example: the log-normal and $P(u) = exp\{-(\log u)^\lambda\}$, with $\lambda \leq 1$]
- *Pre-wild randomness.* P^{-1} increases more rapidly than any function of the form $exp(|\log x|^\gamma)$, with $\gamma < 1$, but less rapidly than $x^{-1/2}$, or $\{E[U^q]\}^{1/q}$ is infinite when $q \geq \alpha > 2$. [Example: the scaling $P(u) = u^{-\alpha}$, with $\alpha > 2$. The power U^q becomes a *wild* random variable if $q > \alpha/2$.]
- *Wild randomness.* It is characterized by infinite variance $E[U^2] = \infty$, but $E[U^q] < \infty$ for small $q > 0$. [Example: the scaling $P(u) = u^{-\alpha}$, with $\alpha < 2$].
- *Extreme randomness.* It is characterized by $E[U^q] = \infty$ for all $q > 0$. [Example: $P(u) = 1/\log u$, whose concentration converges to 1 as $N \to \infty$]

[33] For a complex number γ, with positive real part, the Gamma function is defined by $\Gamma(\gamma) = \int u^{\gamma-1} e^{-u} du$

The last state is never encountered in practice, therefore, the Section 3.2 is dedicated to the scaling distributions called, α-stable, which are the best approximation of reality when they are applied in finance and represent better the last but one state, the *wild* one.

3.1.3 States of Randomness: from an extreme to another

This section represents a summary of the three states of randomness in a more literal way.

&&& Mild Gaussian world &&&

If we take a coin and play heads or tails dividing this game in sets, the amount of money which we can win or lose, varies a lot over the sets, but it is often equal to zero. If we construct a diagram to indicate the number of times we win, the diagram will have the form of a curve, where the smaller winnings are around the mean, zero, while the rare ones are at the extremes. This distribution is described by only two parameters, mean μ and variance σ^2, or standard deviation σ and has the characteristic function, $f(x)$, shown in the equation (3.1)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \tag{3.5}$$

where x is the variable in point, like a height or a Intellective Quotient. So x determinates in what curve's point we are. If we are around the mean, the probability is very high, conversely, it is very low, and the standard deviation measures the tails' thickness. Mean and standard deviation are the only parameters necessary to know everything about the population. So we can understand that a single throwing is insignificant for the *bell curve*, contradicting the already cited Taleb's phrase, that is, one single observation can destroy thousands of years of confirmation.

&&& Slow Log-normal world &&&

This world is characterized by a slow randomness. In finance, the logarithmic transformation of prices is always used because a logarithm changes the scale of a number so that, instead of concentrating on the size of the number, we can easily compare it with the other closed numbers. For example, a price jump of one dollar, from 10 to 11, or from 1000 to 1001, is the same on the dollar scale, but the logarithmic scale shows that the first is more important than the second.

Accordingly, in probability theory, a log-normal distribution is a probability of a random variables whose logarithm is normally distributed. So if y is the random normally distributed variable, then $x = exp\{y\}$ has a log-normal distribution; likewise, if y is log-normally distributed, then $\log y$ is normally distributed. Many economists and analysts deal with log-normal because of its simple properties. In fact, a variable might be modelled as log-normal if it can be thought of as the multiplicative product of many independent random variables each of which is positive. In finance, a long-term discount factor can be derived from the product of short-term discount factors.

Its probability density is shown by the equation (3.6).

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left\{-\frac{(\ln x - \mu)^2}{2\sigma^2}\right\} \tag{3.6}$$

where $x > 0$

&&& Wild Cauchy world &&&

Augustin-Louis Cauchy, a French mathematician of the 19th century, conceived a new way to observe the world totally opposite to Gaussian's. According to

Mandelbrot, to imagine this world. We should think of a patched archer who is going to shoot an arrow at the target placed on the wall in front of him. His patched eyes allow him to shoot in a random way. For many times he doesn't hit the target, of course. Besides, in fifty per cent of the cases, he does not hit the wall. Supposing that he continues to shoot for many sets and that we register his shot, for any set we can calculate the error average and the standard deviation, but he does not live in a "bell curve" world, so his errors are not mild. Moreover, trying to calculate the mean after each shot, we can realize that, while in the Gaussian world each shot gives an insignificant contribute to the mean, in the wild one a single shot can overwhelm one-thousand ended up near the target. So they never position around a mean and never have a constant deviation from this. Consequently, the errors are not mean-reversing, they have an infinite mean and standard deviation and so on the other moments.

The reduced probability density of Cauchy is represented by the equation (3.7), as follows:

$$f(x) = \frac{1}{\pi(1 + x^2)} \tag{3.7}$$

The graph is similar to Gauss distribution, but with longer and thicker tails.

3.2 How make the rare events more probable

Section 3.1 has shown that dealing with Gaussian distributions we cannot consider the best state of randomness, when we have worries about financial problems. For the reason explained in Section 1.2.2 of Chapter 1 and illustrated by Figure 2-C1, we can present a family of particular probability distribution, called *α-stable*, which is also named *Lévy distributions*.

Paul Pierre Lévy (1886 - 1971) was a French mathematician who was active especially in probability theory, introducing martingales, Lévy flights, Lévy processes, Lévy measures, Lévy constant and Lévy distribution. All these works begin from what we define α*-stable* distributions, which are non-Gaussian. Indeed, Gaussian distribution and processes have long been well understood and their utility both as stochastic modelling constructs and analytical tools is well-accepted. However, they do not allow for large fluctuations and are often inadequate for modelling high variability. Moreover, non-Gaussian stable models do not share such limitations. In general, the upper and lower tails of their marginal distributions decrease like a power function whose rate of decays depends on the parameter α, which takes a value constrained between 0 and 1. The smaller α is the slower the decay is and the heavier the tails are. In addition, these distributions have an infinite variance and when $\alpha \leq 1$, they have an infinite mean as well.

Before dealing with them, we have to present what the power laws are.

3.2.1 The power of power laws: *scaling*

If our eyes observe more carefully a financial price series, we will find out that it does not follow the Gaussian bell curve, and its tails follow a *power law*. *Power laws* are mathematical relationships between two quantities. For example, let us consider the surface of a plot of land, it increases in proportion to the side of the second power; if the side doubles, surface quadruples; if it triplicates, surface increases nine times. In economics, an Italian economist Vilfredo Pareto discovered these laws about a century ago, which describe the distribution of income, according to which, the majority of wealth is concentrated in the hands of very few individuals. So adapting a power law to price series, we could make rare events more probable. Indeed, the behavior of these large events is related to the study of *theory of large deviations*, also called *extreme value theory*, which considers the frequency of extremely rare events, like stock market crashes and large natural disasters.

A *power law* is any polynomial relationship that exhibits the property of *scale invariance* and is represented by the equation (3.8). Consider a random variable U and its tail probability specified by the tail distribution $P(u) = Pr\{U > u\}$, then the last relation expresses the following power law:

$$P(u) = Pr\{U > u\} = Cu^{-\alpha} \tag{3.8}$$

To understand the meaning of the term scaling, we must consider scaling under conditioning. Suppose that $P(u)$ becomes known, this U is at least equal to w. So U becomes a conditioned random variable W and the tail distribution is

$$P_W(u) = Pr\{W > u\} = Pr\{U > u|U > w\} = \frac{P(u)}{P(w)} \tag{3.9}$$

Taking the tail distribution $P(u) = Cu^{-\alpha} = (u/\tilde{u})^{-\alpha}$, when $w > \tilde{u}$, conditioning yields $P_W(u) = (u/w)^{-\alpha}$, which is functionally identical to $P(u)$[34]. This is the main property, that is, *scale invariance*.

Now consider the logarithmic transformation $V = \ln U$, we obtain

$$Pr\{V > v\} = exp\{-\alpha(v - \tilde{v})\} \tag{3.10}$$

Conditioned by $V > w > \tilde{v}$, the tail distribution becomes

$$P_w(v) = Pr\{V > v|V > w\} = exp\{-\alpha(v - w)\} \tag{3.11}$$

which is identical to $Pr\{V > v\}$, except for a change of *location* rather than *scale*. Hence, a system ruled by scaling rather than Gaussian processes can represent better market's activity, which is formed by more and more individuals who

[34] This law is well-known as Pareto law. According to his work on income, $P(u)$ is the percentage of individuals with an income above u, while $\tilde{u}$ is the minimum income.

operate through several strategies and thoughts. Indeed, the most important feature of the scaling distribution is the length of its tail, not its extreme skewness.

But now we want the reader to pay attention to the slight difference between *Pareto law* and *scaling distribution*, each of which can be either "uniform" or "asymptotic". Uniform scaling distribution is concerned with two state variables $\tilde{u}$ and α, so that

$$P(u) = \begin{cases} (u/\tilde{u})^{-\alpha} & \text{when} \quad u > \tilde{u} \\ 1 & \text{when} \quad u < \tilde{u} \end{cases} \tag{3.12}$$

where, according to Pareto law, $\tilde{u}$ is a minimum income, while the exponent $\alpha > 0$ quantifies the notion of inequality of distribution. The corresponding density $P(u) = -\,dP(u)/du$ is

$$P(u) = \begin{cases} \alpha(\tilde{u})^{\alpha} u^{-(\alpha+1)} & \text{when} \quad u > \tilde{u} \\ 0 & \text{when} \quad u < \tilde{u} \end{cases} \tag{3.13}$$

But what about its asymptotic behavior? The term asymptotic means that

$$P(u) \sim (u/\tilde{u})^{-\alpha}, \text{ as } u \to \infty \tag{3.14}$$

where the sign $\sim$ means "behaves like". And so

$$\frac{P(u)}{(u/\tilde{u})^{-\alpha}} \to 1, \text{as } u \to \infty \tag{3.15}$$

Another kind of power law distribution with an exponential cutoff is (3.16)

$$P(u) = -dP/du = ku^{-\alpha-1}e^{-\lambda u} \tag{3.16}$$

This distribution does not scale and consequently is not asymptotically a power law[35]. However, it does approximately scale over a finite region before the cutoff, so this can be considered such a common alternative to the asymptotic power-law distribution because of its capturing of finite-size effects. But it is demonstrated that the parameter λ is very small, and in addition, some crucial properties of the asymptotic Pareto distribution correspond to $\lambda = 0$.

However, the *asymptotic scaling distribution* must be at least approximately correct for large u.

One of the most important laws is that of Zipf, a linguist of the 1930s who discovered this law analyzing the frequencies of words in a book. In many papers and books, many words are repeated and what Zipf did was to choose a book and to count up how many times a word is repeated. The next step was to order these according to their frequency, giving a rank to each of them, that is, 1 to the most repeated word and so on. Finally, drawing the graph, we realize that the frequencies' curve does not decrease in a regular way, but at the beginning it falls in a dizzy manner and then decreases more slowly. The Zipf law is shown by the equation (3.17).

$$Q(r) = Fr^{-1/\alpha} \tag{3.17}$$

where $Q(r)$ is the probability distribution with respect to the rank r, F is a constant and $1/\alpha$ is the critical factor of the power law, that is the larger α, is the richer the vocabulary is and so, the curve decreases more slowly.

Let us try to apply this experiment to the price series of The Coca-Cola Co. The result is shown in Figure 1-C3, which is a classical form of *scaling*.

[35] See Section 3.3.5 for the truncation of probability distributions.

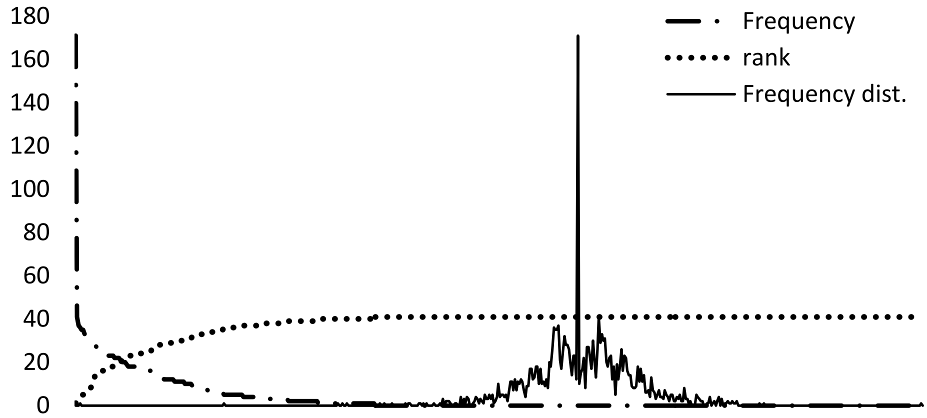


Figure 1-C3. Frequency distribution (solid thin line), ordered frequencies (dashed line) and rank (dotted line). Weekly log-return from 1962 to 2009, 2089 data.

A variant of Figure 1-C3 can be presented by *Pareto chart* exposed in Figure 2-C3, also referred to log-returns of The Coca-Cola Co. over the same period.

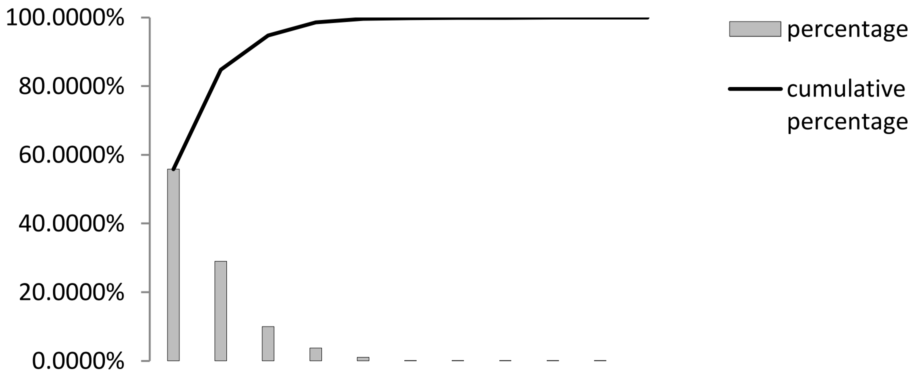


Figure 2-C3. Pareto chart. In this case, the classes are only thirteen. Weekly log-return from 1962 to 2009, 2089 data.

So Zipf's result is similar to that of Pareto's, even if the playing field is different. But how to estimate the exponent α? We will deal with this problem in Chapter 5.

3.3 Stable non-Gaussian random processes and Lévy processes

The *theory of errors* presupposes that every kind of measurement error has a stable "bell curve" distribution, that is, we can sum up the measurement errors generated by several independent fonts and this sum has again the Gaussian distribution. Also if mean and variance vary, the distribution remains the same. Also Cauchy expressed the same concept, but this time, on the other extreme side. The sum of the results of the distribution of two shooters is always the same, that is, Cauchy distribution, is also stable. But there is a complete family of these *stable distributions*, named *L-stable* by Mandelbrot and referring to the mathematician Paul Pierre Lévy.

On one hand, the Gaussian bell curve seem to be egalitarian, while, on the other hand, the Cauchy one is too dictatorial, because few data can dominate the whole. Lévy has played an important role in *probability theory*, indeed, his distributions allow to connect these two extremities.

Let us try to write some fundamental definitions to better understand.

3.3.1 Univariate stable distribution

Univariate stable distribution is characterized by four parameters. These are the index of stability α, the scale parameter γ, the skewness parameter β and the shift parameter δ. There are four equivalent definition of them.

DEFINITION 1. A random variable X is said to have a stable distribution if for the positive numbers A and B, there is a positive number C and a real number D so that

$$AX_1 + BX_2 \stackrel{\text{def}}{=} CX + D \tag{3.20}$$

where X_1 and X_2 are independent copies of X and "$\stackrel{\text{def}}{=}$" denotes equality in distribution. If $D = 0$, X is called *strictly* stable, and also *symmetric* stable if the distribution of X is also symmetric.

DEFINITION 2. A random variable X is said to have a stable distribution if for any $n \geq 2$, there is a positive number C_n and a real number D_n so that

$$X_1 + X_2 + \cdots + X_n \stackrel{\text{def}}{=} C_n X + D_n \qquad (3.21)$$

where $X_1, X_2, \cdots, X_n$ are independent copies of X.
Feller[36] claims that in (3.20) we have necessarily

$$C_n = n^{1/\alpha}$$

for some $0 < \alpha \leq 2$.
While the first two definitions concern the stability property, meaning that the family of stable distributions is preserved under convolution, the third will concern the role of a stable distribution in the context of the central limit theorem[37], and the last one specifies the characteristic function of a stable random variable.

DEFINITION 3. A random variable X is said to have a stable distribution if it has a domain of attraction, i.e., if there is a sequence of i.i.d. random variables $Y_1, Y_2, \cdots, Y_n$ and a sequence of positive numbers $\{d_n\}$ and real numbers $\{a_n\}$, so that

$$\frac{Y_1, Y_2, \cdots, Y_n}{d_n} + a_n \stackrel{d}{\Rightarrow} X \qquad (3.22)$$

where the sign "$\stackrel{d}{\Rightarrow}$" denotes the convergence in distribution

---

[36] Feller, W., *An Introduction to Probability Theory and Its Applications*, Vol. 2, Wiley, New York, 1971
[37] The CLT is explained in Section 3.1.1

DEFINITION 4. A random variable X is said to have a stable distribution if there are parameters $0 < \alpha \leq 2$, $\sigma \geq 0$, $-1 \leq \beta \leq 1$, and δ real, so that its characteristic function has the following form:

$$\varphi(\theta) = \begin{cases} \exp\left\{-\gamma^{\alpha}|\theta|^{\alpha}\left(1 - i\beta(\text{sign}\theta)\tan\dfrac{\pi\alpha}{2}\right) + i\delta\theta\right\} & \text{if } \alpha \neq 1 \\ \exp\left\{-\gamma|\theta|\left(1 - i\beta\dfrac{2}{\pi}(\text{sign}\theta)\ln|\theta|\right) + i\delta\theta\right\} & \text{if } \alpha = 1 \end{cases} \tag{3.23}$$

where α is the index of stability, the same of the power law, and

$$\text{sign}\theta = \begin{cases} 1 & \text{if } \theta > 0 \\ 0 & \text{if } \theta = 0 \\ -1 & \text{if } \theta < 0 \end{cases}$$

and γ, β and δ are unique (β is irrelevant when $\alpha = 2$).

Leaving off several steps, now we present the so-called Lévy representation of the characteristic function (3.24).

$$\varphi(\theta) = \exp\left\{i\delta\theta - \gamma^{\alpha}|\theta|^{\alpha}\left\{1 - \frac{i\beta\theta}{|\theta|}\tan\frac{\pi\alpha}{2}\right\}\left|\cos\frac{\pi\alpha}{2}\right|\right\} \tag{3.24}$$

where all the parameters have the same previously range of values but when $1 < \alpha < 2$ the *L-stable* variable U has $\mathrm{E}[U] < \infty$ but $\mathrm{E}[U^2] = \infty$ and so on the q-moments, $\mathrm{E}[U^q] = \infty$. Moreover, when $0 < \alpha < 1$ also the first moment is infinite.

Therefore, we can observe Table 1-C3 for a short explanation of the four parameters, Table 2-C3 to understand how this distribution varies when α and β vary, and Figure 3-C3 and Figure 4-C3 for a graphic representation.

Variable	Description
α	Peakedness index
β	Skewness index
γ	Scale parameter
δ	Location or shift parameter

Table 1-C3. Lévy distributions' four parameters and a short explanation.

Variable	Distribution
$0 < \alpha < 1$	$\mathrm{E}[U^q] = \infty \;\; q = [1;\, \infty)$
$1 < \alpha < 2$	$\begin{cases} \mathrm{E}[U] < \infty \\ \mathrm{E}[U^q] = \infty \;\; q > 1 \end{cases}$
$\beta = 0$	Skewness distribution
$\alpha = 1\, \beta = 0$	Cauchy distribution; very thick tails
$\alpha = 2\, \beta = 0$	Standard Gaussian bell curve

Table 2-C3. Several parameters' values and the corresponding distribution.

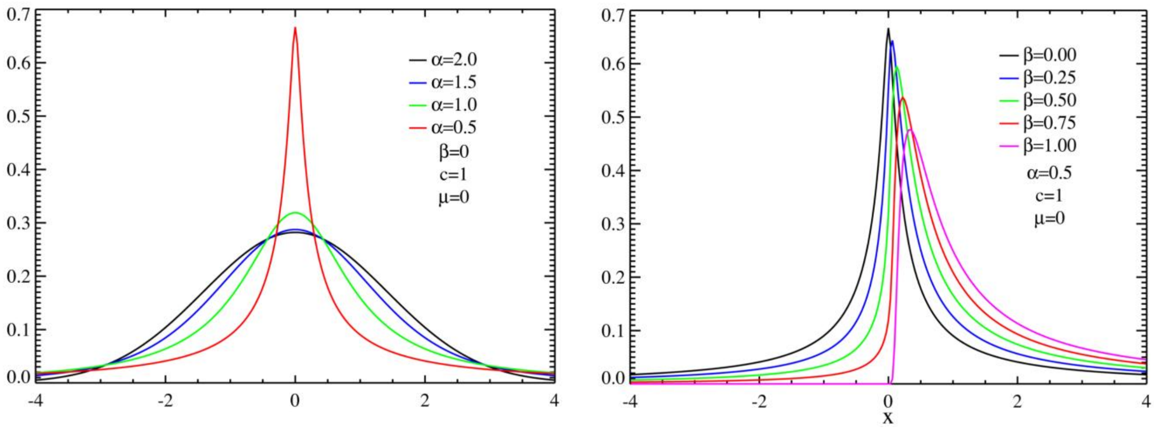


Figure 3-C3. Simulation of Lévy distribution for several α parameters (on the left) and for several skewness parameters (on the right).

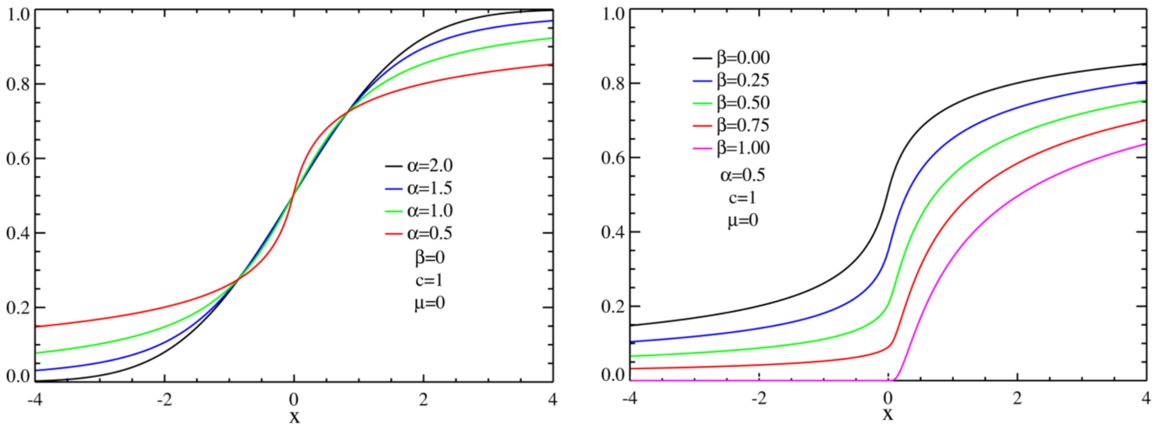


Figure 4-C3. Simulation of cumulative Lévy distribution for several α parameters (on the left) and for several skewness parameters (on the right).

3.3.2 Multivariate stable distribution

Multivariate stable distributions are the distributions of a stable random vector. It is well-known that any linear combination of Gaussian components is again a Gaussian random vector. The same is true for a scale distribution but it turned out that the converse is not always true, because the converse holds when the linear combination are either strictly or when $\alpha \geq 1$.[38] The definition of stability is the same exposed in Section 3.3.1, but, in this case, we have X as a vector of stable random variables in $\mathbb{R}^d$.

By the definition (3.23), the equation (3.25) has the characteristic function

if $\alpha \neq 1$

$$\Phi_\alpha(\theta) = \exp\left\{-\int_{S_d} |(\theta,s)|^\alpha \left(1 - i\,\mathrm{sign}\big((\theta,s)\big)\tan\frac{\pi\alpha}{2}\right)\Gamma(ds) + i(\theta,\delta^0)\right\}$$

if $\alpha = 1$

$$\Phi_\alpha(\theta) = \exp\left\{-\int_{S_d} |(\theta,s)|^\alpha \left(1 - i\,\frac{2}{\pi}\mathrm{sign}\big((\theta,s)\big)\ln(\theta,s)\right)\Gamma(ds) + i(\theta,\delta^0)\right\} \tag{3.25}$$

[38] See Samorodnitsky G., Taqqu S. Murad, *Stable Non-Gaussian Random Processes – Stochastic Models with Infinite Variance*, Chapman & Hall, New York, 1994

where Γ is a finite measure on the unit sphere of $\mathbb{R}^d$ called *spectral measure* and replaces both the scale and skewness parameter. Besides, δ^0 is the shift vector which play a role similar to the shift parameter in the univariate case.

The reason why we have written and exposed this complex equation is that the concept of Gaussian covariance is now replaced by two other equations as a measure of bivariate dependence, *covariation* and *codifference*.

Leaving out the complex formulas, we present literally those two concepts.

The first is designed to replace the covariance (the case $\alpha = 2$) when $1 < \alpha < 2$. It shares some properties of the covariance. Unfortunately, it is neither symmetric nor additive in its second component. However, it becomes additive only when the random variables are independent.

Like the covariation, the codifference is reduced to the covariance when $\alpha = 2$ and while the covariation may not be defined for $\alpha \leq 1$, and the codifference is defined for all $0 < \alpha \leq 2$.[39]

3.3.3 Generalization of the concept of L-stability

We have seen that if we consider two independent Gaussian random variables, G' and G'', of zero mean and of a standard deviation equal to σ'^2 and σ''^2, respectively, their sum, $G' + G''$, is also a Gaussian variable of mean square equal to $\sigma'^2 + \sigma''^2$. In particular, the "reduced" Gaussian variable is a solution to

$$s'U + s''U = sU \tag{3.26}$$

where s is a function of s' and s'' given by the auxiliary relation

$$s^2 = s'^2 + s''^2$$

[39] For an analytic view, see Samorodnitsky G., Taqqu S. Murad, *Stable Non-Gaussian Random Processes – Stochastic Models with Infinite Variance*, Chapman & Hall, New York, 1994

So s, s' and s'' are scale factor which correspond to the root-mean-square in the Gaussian world. But they also express a kind of L-stability or invariance under addition, that is, the equations (3.20) and (3.21). So while the Gaussian is the only solution of (3.26) for which the second moment is finite, when the variance is infinite, like for L-stable distribution, (3.26) possesses many other solutions. This was shown by Cauchy, who considered the random variable U for which

$$\Pr\{U > u\} = \Pr\{U < -u\} = 1/2 - (1/\pi)\tan^{-1} u$$

so that its density has this form

$$d\Pr\{U < u\} = \frac{1}{\pi(1+u^2)}$$

For this law, all the moments are infinite, where the auxiliary relation takes the form

$$s = s' + s''$$

where the scale factor is not defined by any moment.

But, in 1925, Lévy discovered the solution of (3.26), that is, the equation (3.24) but in a logarithmic transformation

$$\log \int_{-\infty}^{\infty} \exp(iu\theta) d\Pr\{U < u\} = i\delta\theta - \gamma|\theta|^{\alpha}\left\{1 - \frac{i\beta\theta}{|\theta|}\tan\frac{\pi\alpha}{2}\right\} \quad (3.27)$$

for which everything has been exposed in the previous sections is valid. It is called “reduced” when $\gamma = 1$ and $\delta = 0$. Hence, the auxiliary relation is now

$$s^{\alpha} = s'^{\alpha} + s''^{\alpha}$$

More generally, suppose that U' and U'' are stable, with the same values of α, β and $\delta = 0$, but different γ, ($\gamma'$and γ''), the sum $U' + U''$ is also stable with α, β, $\gamma = \gamma' + \gamma''$ and $\delta = 0$ ($\delta = \mathrm{E}[U]$ and is finite if and only if $1 < \alpha \leq 2$).

Moreover, considering U_n in n independent stable variables with those parameters, (3.27) becomes

$$n \log \varphi(\theta) = iN\delta\theta - N\gamma|\theta|^{\alpha} \left\{1 - \frac{i\beta\theta}{|\theta|} \tan \frac{\pi\alpha}{2}\right\} \tag{3.28}$$

which differs from (3.27) only for the location δ and scale γ parameters which are multiplied by N, which are the equivalent of the mean and of the standard deviation of the Gaussian distribution, respectively, but with different meanings. The result does not change if δ and γ have several values,

$$\sum_{j=1}^{n} \log \varphi_j(\theta) = i\theta \sum_{j=1}^{n} \delta_j - |\theta|^{\alpha} \sum_{j=1}^{n} \gamma_j \left\{1 - \frac{i\beta\theta}{|\theta|} \tan \frac{\pi\alpha}{2}\right\} \tag{3.29}$$

The property of invariance under addition is at the base of the interest towards those distributions because they can be applied to financial data. Indeed, whatever time interval is, price changes between two intervals can be seen as the sum of the same price changes between two sub-intervals. Therefore, if asset returns are random L-stable i.i.d variables, then daily, weekly and monthly price changes have the same distribution except for the scale and the location parameters.

3.3.4 Fractal dimension of symmetric L-stable distribution

Considering Y as a variable obtained by dividing the variable X for a scale factor $n^{1/\alpha}$, the sum of n independent copies of Y is

$$Y_1 + Y_2 + \cdots + Y_n = Yn^{1/\alpha} = n^{1/\alpha}\left(X/n^{1/\alpha}\right) = X$$

where the fractal dimension is obtained by

$$D = \frac{\ln(n)}{\ln\left(n^{1/\alpha}\right)} = \alpha$$

Hence, we have the Gaussian distribution $D = 2$ like that of a plane, and the Cauchy one $D = 1$ like that of a straight line, while for the other value of $\alpha \in (1, 2]$, we have an infinite number of distributions.

3.3.5 Truncation and multiplicative exponential decay

This section is dedicated to a clarification of many concepts exposed previously.

A scaling variable is said to be "truncated", if it does not exceed a finite maximum $u_{\max}$. On certain markets the distribution tails are shorter than the ones implied by the scaling distribution. So we define U as a L-stable distribution that was truncated to $u_{\max} < \infty$ and we consider the sum of N such variables. For small values of N, the distribution of the normalized sum is unaffected by $u_{\max}$. In addition, for a large N, a different normalization converges to the Gaussian. Therefore, the tails become increasingly short as N increases. Hence, truncating a distribution causes the fall of some values of N between those two zones, which is called "transient" and hard to control.

The density (3.15) with the exponential decay $\lambda > 0$ shares the main virtue of the truncation with $u_{\max} < \infty$. Assuming $\lambda > 0$ the q-th moment is infinite when $q > \alpha$.

3.4 An empirical analysis of asset returns

To verify the hypothesis of invariance, we consider the daily, weekly and monthly price variations of the index S&P 500 from January 3, 1950 to December 4, 2009. Now we try to make an experiment. Observe carefully Figure 5-C3. Is there something strange?

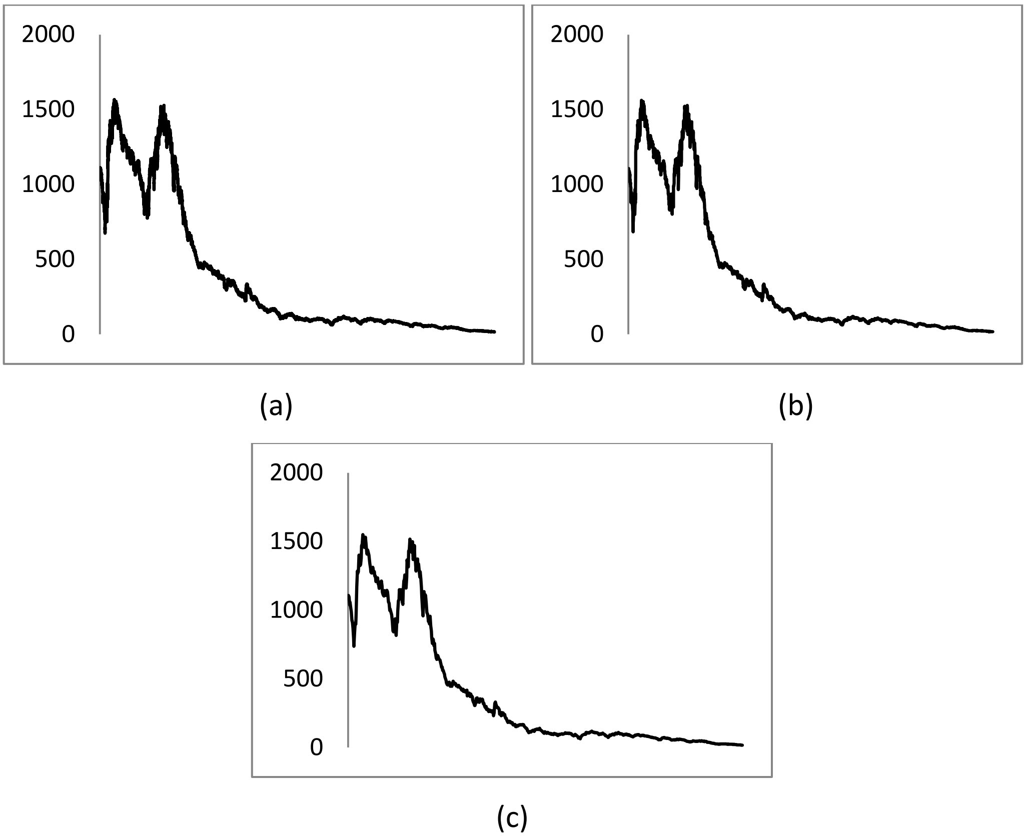


Figure 5-C3. Scaling in finance. Daily (a), weekly (b) and monthly (c) prices of index S&P 500 from 1950 to 2009.

The three graphs are similar but they differ on scale. This is a way to represent the scaling, that is, the *scale invariance*.

In addition, we observe Table 3-C3 and Table 4-C3. These tables are constructed by calculating the actual mean (in bold on the diagonal) and then multiplying them for coefficients, in parenthesis, obtained by relating the several time horizons. For example, the coefficient $(1/4)$ in the cell (5 days-1 day) is obtained by relating 1 to 5. The same procedure for Table 4-C3 but now with the variance of price changes.

Price changes' means			
	1 day	**5 days**	**20 days**
1 day	**2.78** (1)	13.91 (5)	55.65 (20)
5 days	2.67 (1/5)	**13.36** (1)	53.46 (4)
20 days	2.90 (1/20)	14.51 (1/4)	**58.03** (1)

Table 3-C3. *Scale invariance*: about means

Price changes' variances			
	1 day	**5 days**	**20 days**
1 day	**0.94** (1)	4.70 (5)	18.80 (20)
5 days	0.86 (1/5)	**4.31** (1)	17.26 (4)
20 days	0.89 (1/20)	4.44 (1/4)	**17.75** (1)

Table 4-C3. *Scale invariance*: about variances

We can realize that the actual mean and variance and the fitted ones coincide. Indirectly we can verify the constancy of the exponent α. Indeed, considering α based on the form of the distribution, if this remain constant, we realize that α has not changed. To understand better this concept, see Table 5-C3 which is obtained by dividing the amount of observation in the distribution's tails for the their total number. For example, the probability that two observations are in the tail is constrained from 0.25 to 0.27 percent, so they are almost equal. These conclusions imply that short-run investors run the same risk of an investor who operates on a longer time horizon considerate.

Probability to find n observation	probability 20 days	from 1 day to 20 days	from 5 days to 20 days
1	0.1391%	0.1326%	0.1280%
2	0.2782%	0.2653%	0.2560%
3	0.4172%	0.3979%	0.3840%
4	0.5563%	0.5306%	0.5120%
5	0.6954%	0.6632%	0.6400%
6	0.8345%	0.7959%	0.7680%

Table 5-C3. Probability that an event is in the distribution's tails over different time intervals. Note the similarity among these probabilities.

This concept introduces the central topic of the next chapter which will deal with discontinuity according to the Fractional Brownian motion.

Chapter 4

FRACTIONAL BROWNIAN MOTION, LONG-RUN DEPENDENCE

In 1965, Mandelbrot realized that asset return are dominated by a global dependence. Here, the term global is referred to long-run. What Mandelbrot did is to introduce *infinite memory* into statistical modelling. In particular he introduced the so-called fractional Brownian motion (fBm), a process having one significant parameter: the Hurst or Hölder exponent H satisfying $0 < H < 1$.
The Wiener Brownian motion, shown in Section 1.2.1, is the atypical special case corresponding to the value $H = 1/2$. The link between this exponent and Pareto's distribution deals with the relationship between *long-run dependence* and *tails' peakedness*. In fact, while for the WBm model of Bachelier we have that $\Delta x \sim \sqrt{\Delta t} = \Delta t^H$ whose exponent is time invariant and $H = 1/2$, for the FBM model of Mandelbrot we have $\Delta x \sim \Delta x^{1/\alpha}$ or $\Delta x \sim \Delta t^H$ whose exponent is again time invariant but $H \neq 1/2$. Indeed, according to a metaphor suggested by the statistical physics of magnets, an infinite range dependence controlled by

power-law expressions is the rule of a system similar to actual interactions, which only occur between immediate neighbors[40].

Moreover, we can pass from *unifractality* to *multifractality*, the models that show the exponent H not depending on time called *uniscaling*, while, *multiscaling* allows the exponent to depend on t. In other words, large or small values of $H(t)$ express, respectively, that $x(t)$ varies slowly or rapidly near the instant t. Therefore, in our opinion, Mandelbrot has introduced a revolutionary concept of time differing from the clock time, and related to trade, that is why it is called *trading time*. This special time ruled by a devil staircase[41] is called *fractal time*, because it is a function of physical time which is reduced to a series of mutually independent jumps of widely varying size. For this reason, the graphs generated by fBm are *self-affine*.

The central aim of this chapter is to show the main properties of the fBm, introducing a new way to conceive financial reality, in order that to present new methods in pricing derivatives and in the selection portfolio theory.

4.1 Fractal Market hypothesis

The *Fractal Market Hypothesis* (FMH) clashes with the *Efficiency Market Hypothesis* (EMH) shown in Section 1.2.4. FMH emphasizes the impact of the liquidity and the investment horizons on the investors. Its aim is to provide a theoretical model able to describe price changes closer to reality.

A short-run investor, who suffers a loss due to his time horizon, will be assisted by a long-run investor for whom a negative event is partly negative. Consequently the market is *self-stabilizing*. This is the conclusion exposed in Section 3.4 and shown by the Table 5-C3, according to which market operators share the same risk level, an adjustment is made for the scale of the investment

[40] Mandelbrot, B.B., *Fractals and Scaling in Finance*, 1997, page 36

[41] The exponent α of the Lévy motion is fed in by choosing a Lévy staircase of dimension $\alpha/2$, so the two main parameters which combine long tails and long-run dependence are the α exponent which is twice the α exponent if the Lévy staircase, and the exponent H.

horizon. Hence, the shared risk explains why the frequency distribution of returns look the same at different investment horizon. This hypothesis is called *Fractal* because of its property of *self-similarity*[42].

Conversely, the instability of markets comes from the breaking down of this fractal structure. A break-down occurs when investors with long investment horizons either stop participating in the markets or become short-run investors.

The fractal statistical structure exists because it is a stable structure. We have seen how a price diagram decreases according to a power law. Hence, each price formation depends on the previous one. If one generation is negative, then its negativity can continue for a certain time interval. However, the more heterogeneous investment horizons are, more the panic investment at one horizon can be absorbed by the other ones as a buying or selling opportunity.

Indeed, when the investment horizon becomes uniform, the market falls, where discontinuities appear in price sequence. For Gaussian, a big change is made up of many small changes. But, during panic and stampede phases, the market often skips over prices. The discontinuities, which are the result of lack of liquidity, cause big changes, and fat tails appear in the frequency distribution of returns.

Another explanation is that, if the information received by the market is important to both the short- and long-runs, also the liquidity can be affected.

Many examples can explain these facts, so see Figure 1-C4 for a graphic view.

[42] See Peters, E. Edgar, *Fractal market analysis: applying chaos theory to investment and economics*, 1994

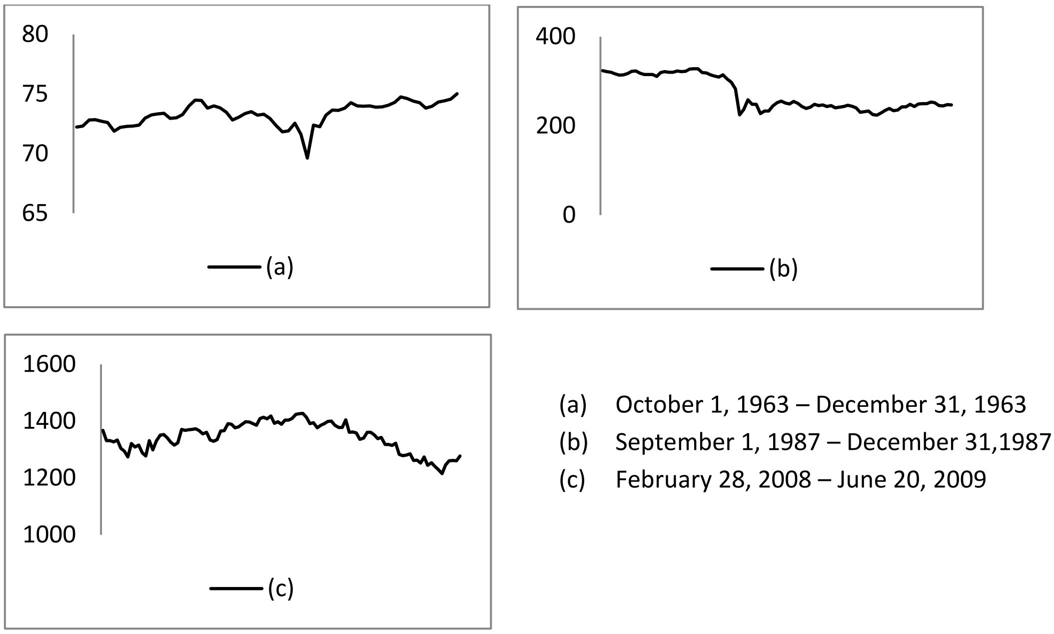


Figure 1-C4. S&P 500 index over the three described periods.

So the main concepts of FMH are:

- A market is stable when it is composed by a large number of investors who operate on a different investment horizon;
- A technical analysis influences the short-run decision, while in the long-run one the fundamental analysis predominates;
- Instability is caused by a momentary exchange of investment horizons, which leads to a lack of traders who absorb the negotiations assuring the liquidity;
- Price reflects a combination of evaluations of technical and fundamental analysis. Short-run price variations could be more volatile than the long-run ones.

4.2 Brownian motion

This section makes a short review of the Wiener Brownian motion, but now according to a probability approach.

Let us consider $B(t)$ as the Brownian motion. It is defined for equally-spaced time steps Δt, so that all the increments $\Delta B(t)$ are independent, isotropic and random.

Independent means that the value of the current increment does not affect the next one. *Isotropic* means that the increments are equally probable to occur in all directions. *Random* means that future increments are unpredictable. An example of WBm is in Figure 2-C4.

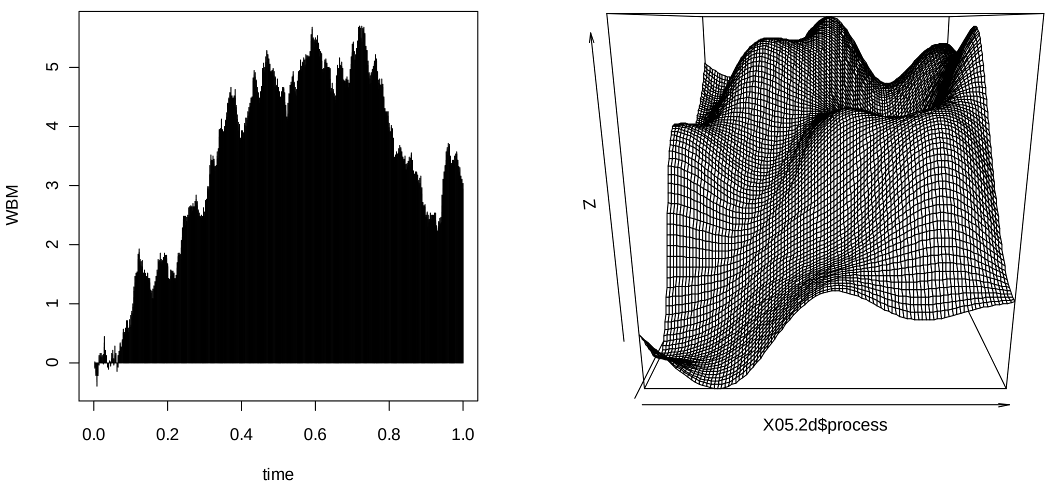


Figure 2-C4. A simulation of a one-dimensional (on the left) and two-dimensional (on the right) Wiener Brownian motion.

4.2.1 Fractal properties and the probabilistic approach

According to the concepts exposed in Chapter 2, observing WBm at different scales, we can realize that it is also a fractal. At each level of magnification, the trace of the function has approximately the same jaggedness, so the function is *self-similar*.

To show this property, we refer to the approach of Feder[43]. Consider a simple random walk in which all steps are of the same size and each step may be either up or down. At each step, the probability of moving up or down is the same. This random walk is illustrated in Figure 3-C4.

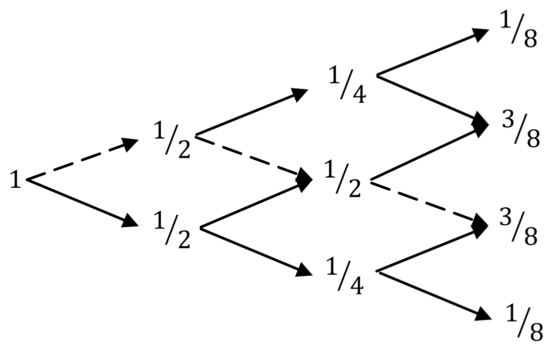


Figure 3-C4. A possible moving for a simple random walk of three steps. The dotted line shows one possible path for the random walk.

Considering U as the random variable which represents the number of steps in our random walk, we can model U with a binomial distribution which describes the likelihood of each possible outcome of an experiment. Here the experiment consists of n independent trials and each trial has the same probability p of success. It is the same concept according to which a coin tossing is based on. For example, toss a coin 20 times, with the probability of tossing heads $1/2$ for each trial. The probability of each possible number u of successes, represented in Figure 4-C4, is

$$P(u) = \binom{n}{u} p^u q^{n-u} \tag{4.1}$$

where the expected value and the variance are, respectively,

[43] Feder, J., *Fractals*, Plenum Press, New York, 1988

$$\mathrm{E}(U) = np$$
$$\mathrm{Var}(U) = npq$$

where $q = (1 - p)$.

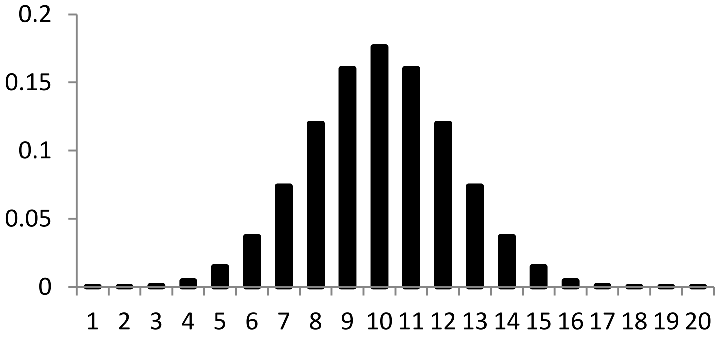


Figure 4-C4. A binomial distribution with $n = 20$ and $p = 0.5$

So, for our random walk, we have

$$P(u) = \binom{n}{u}\left(\frac{1}{2}\right)^{n} \tag{4.2}$$
$$\mathrm{E}(U) = \frac{1}{2}n$$
$$\mathrm{Var}(U) = \frac{1}{4}n$$

Consequently, the next expectation is equal to 0 and the variance of U is proportional to n, so the standard deviation is proportional to $n^{1/2}$. Since the variance of the simple random walk scales with the number of steps taken, it is self-similar.

Observing carefully the Figure 4-C4, we realize that the shape of the binomial's probability distribution seems to be similar to the Gaussian one, if the number of trials, n, is large, the binomial distribution is approximately equal to the normal one, as follows in Figure 5-C4.

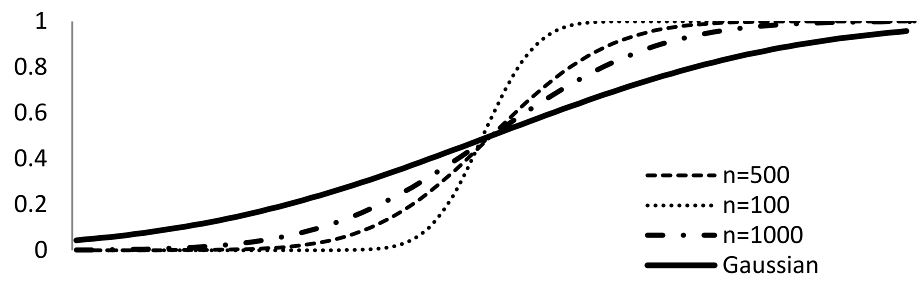


Figure 5-C4. Binomial distribution's convergence to Gaussian one.

Therefore, now we show the property of self-similarity for a more complex model of Brownian motion. Feder[44] describes a random walk in which the step length has a normal distribution.

Retrieve the equation (3.5). We consider ε as the length of each positive or negative step and τ and $\mathcal{D}$ represent respectively the length of each time step and the *diffusion coefficient*[45], then (3.5) can be rewritten as

$$\mathrm{P}(\varepsilon,\tau) = \frac{1}{\sqrt{4\pi\mathcal{D}\tau}}\exp\left\{-\frac{\varepsilon^2}{4\mathcal{D}\tau}\right\} \tag{4.3}$$

where the expected value and the variance are:

$$\mathrm{E}[\varepsilon,\tau] = 0$$
$$\mathrm{Var}[\varepsilon,\tau] = 2\mathcal{D}\tau$$

To show the property of a self-similarity, now we consider the value of ε at every bth time step, so the probability distribution becomes:

[44] Feder, J., *Fractals*, Plenum Press, New York, 1988

[45] This is a comlex concept so we say only that it is a factor of proportionality representing the amount of substance diffusing across a unit area through a unit concentration gradient in unit time. For more details see Heitjans, P., Karger, J., *Diffusion in condensed matter: Methods, Materials, Models*, Birkhauser, 2005

$$\mathrm{P}(\varepsilon, b\tau) = \frac{1}{\sqrt{4\pi\mathcal{D}b\tau}} \exp\left\{-\frac{\varepsilon^2}{4\mathcal{D}b\tau}\right\} \qquad (4.4)$$

where the new variance is

$$\mathrm{Var}[\varepsilon, \tau] = 2\mathcal{D}b\tau$$

Hence something changes in the WBm. In fact, for every Δt time steps, we have a vertical range ΔB, whose average will be $b^{1/2}\Delta B$. But now it is necessary to make a clarification. The vertical range of this random walk does not scale directly with the horizontal range, but rather with the square root of the horizontal one, so it is more accurately described by the term *self-similarity* as shown in Figure 6-C4.

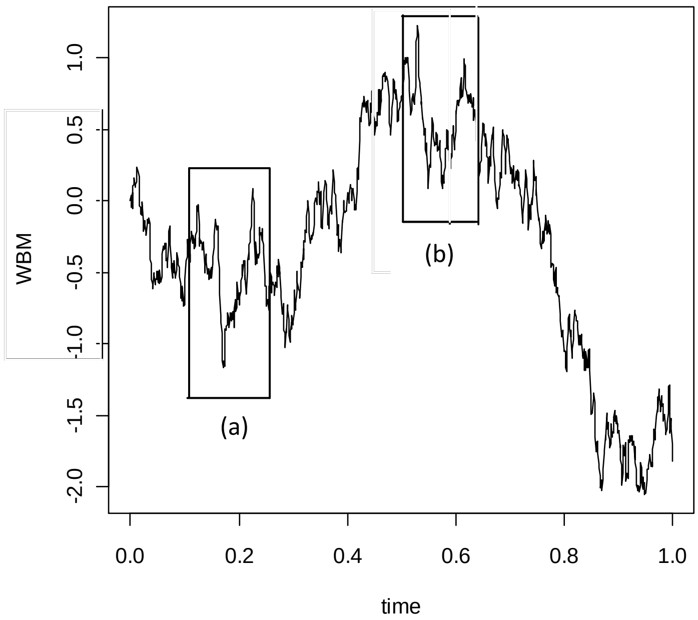


Figure 6-C4. Self-similar Brownian motion. Note that (a) and (b) are similar from a direction to another, so self-similarity.

Instead of considering the WBm in terms of increments ε, Wiener considers the random function $B(t)$ which describes the position itself.
Before obtaining the probability density function and then dealing with its increments, now we have to present the property of WBm.
Indeed, we can consider its covariance property, that is,

$$\mathrm{E}[B(t)] = 0 \quad \forall t \in \mathbb{R}$$
$$\mathrm{E}[B(t)B(s)] = \frac{1}{2}[|t| + |s| - |t-s|] = s = \min(s,t) \quad \forall (t > s > 0) \in \mathbb{R}$$

On the contrary, if we deal with its increments to obtain the probability density function, and consequently its expected value and variance, we can replace ε and τ respectively with $B(t) - B(s)$ and $|t-s|$ in the equation (4.4):

$$\mathrm{P}\big(B(t) - B(s)\big) = \frac{1}{\sqrt{4\pi\mathcal{D}|t-s|}} \exp\left\{-\frac{\varepsilon^2}{4\mathcal{D}|t-s|}\right\} \tag{4.5}$$
$$\mathrm{E}[B(t) - B(s)] = 0$$
$$\mathrm{Var}[B(t) - B(s)] = 2\mathcal{D}|t-s|.$$

Wiener showed that the function $B(t) - B(s)$ defined by

$$B(t) - B(t_0) \sim \varepsilon |t-s|^{1/2} \tag{4.6}$$

where ε is a normally distributed random variable with zero mean and unit variance, has a probability density function, an expected value and a variance that coincide with the above expressions, which we can rewrite in another way:

$$\mathrm{E}[B(t+s) - B(t)] = 0$$
$$\mathrm{Var}[B(t+s) - B(t)] = |t-s|.$$

So we realize that WBm's increments are stationary.

This stochastic process has the following properties:

- *Continuity*: a continuous function is the one where, intuitively, small changes in the input result in small changes in the output;
- *non-differentiability*: if a function is differentiable, it may not have any edges;
- *infinite variation*: a plot of the WBm function can have an infinite vertical range within a finite length of time[46];
- *independence of increments*, because of the Gaussian distribution of ε;
- *uniformity of increments*, that is, the graph is like the White Noise one.

Figure 7-C4 makes a graphic summary of the above concepts.

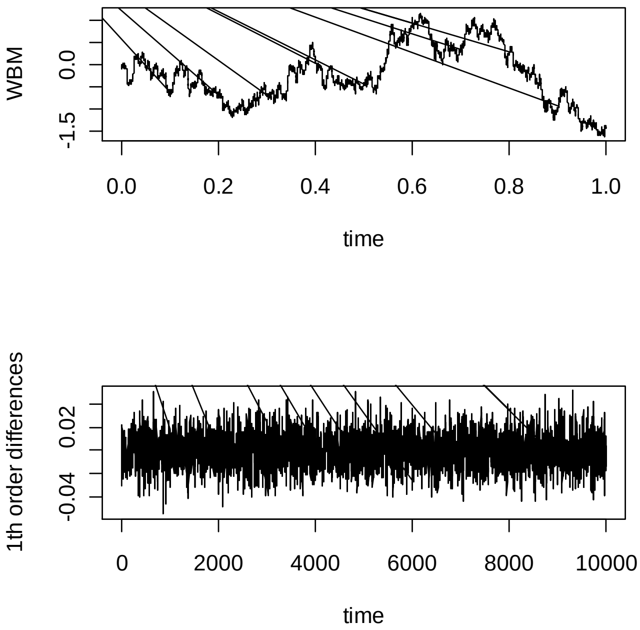


Figure 7-C4. The top graph illustrates a graph of WBm. The second shows the corresponding increments over successive small intervals of time. Note that the 1^{st} order differences' graph is similar to White Noise's.

[46] See Breiman, Leo, *Probability*, Addison-Wesley Pub. Co., 1968

4.3 Fractional Brownian motion

The *fractional Brownian motion* (FBM), $B_H(t)$, is a generalization of the WBm, so we can write the following relationship:

$$\mathrm{E}[B_H(t)] = 0 \quad \forall t \in \mathbb{R}$$
$$\mathrm{E}[B_H(t)B_H(s)] = \frac{1}{2}[|t|^{2H} + |s|^{2H} - |t-s|^{2H}] = t^{2H} \quad \forall t, s \in \mathbb{R}$$

From the covariance property, it turns out that fBm itself is not a stationary process. But if we consider its increments, it represents a random process with Gaussian increments that satisfy the following diffusion rule, for all t and s and for $0 < H < 1$:

$$B_H(t) - B_H(s) \sim \varepsilon |t-s|^H \tag{4.7}$$
$$\mathrm{E}[B_H(t) - B_H(s)] = 0$$
$$\mathrm{Var}[B_H(t) - B_H(s)] = 2\mathcal{D}\tau(|t-s|/\tau)^{2H}$$

or, equally,

$$\mathrm{E}[B_H(t+s) - B_H(t)] = 0$$
$$\begin{aligned}\mathrm{Var}[B_H(t+s) - B_H(t)] &= \mathrm{E}\big[\big(B_H(t+s) - B_H(t)\big)\big(B_H(t+s) - B_H(t)\big)\big] \\ &= \mathrm{E}\left[\big(B_H(t+s)\big)^2\right] + \mathrm{E}\left[\big(B_H(t)\big)^2\right] - 2\mathrm{E}[B_H(t+s)B_H(t)]^{2H} \\ &= t^{2H} + s^{2H} - 2 \cdot \frac{1}{2}[|t|^{2H} + |s|^{2H} - |t-s|^{2H}] \\ &= |t-s|^{2H} \qquad \forall t, s \in \mathbb{R}\end{aligned}$$

The first and second moments do no depend on time but only on the length of the increment, so fBm has stationary increments. But how we will see in the next section, increments of fBm are not independent like the classical WBm, that is, the case $H = 1/2$.

4.3.1 Exponent H and long-run statistical dependence

The exponent H has two thoroughly disparate historic roots. Firstly, it refers to the initial letter of the hydrologist H.E. Hurst (1880-1978), who dealt with a difficult problem of civil engineering. Secondly, it has deep roots in pure mathematics, namely, in the work of L.O. Holder (1859-1937). So seeing that the initial letter is the same, we prefer to call it with the letter "H".

The revolution which occurred in finance is that this exponent measures the persistent and the anti-persistent level in a series. In fact, although many economists have always dealt with a short-run dependence, visible in a simple auto-correlogram showed in Figure 8-C4, the exponent H gives a measure of the *long-run dependence*, according to which there are correlations which decrease very slowly and never seem to vanish altogether. This concept bring us to the well-know "butterfly effect".

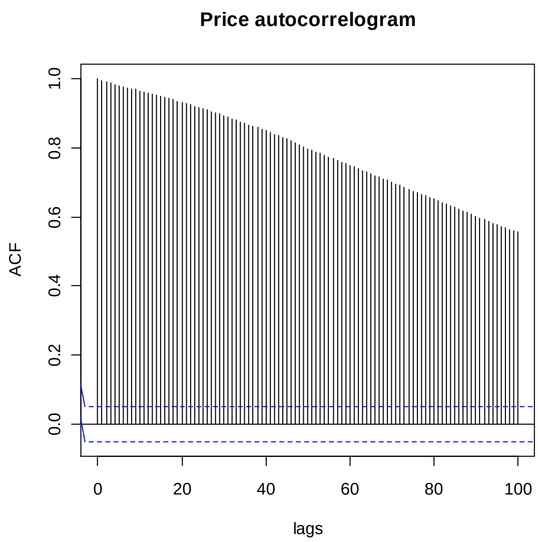


Figure 8-C4. Auto-correlogram of the first 100 Goldman Sachs prices from January 2, 2004 to December 12, 2009. The stronger correlations are the short-run ones between closer periods.

The long-run dependence comes from an important property of fBm which concerns the quantities $[B_H(0) - B_H(-T)]/T$, called *past* average, and $[B_H(T) - B_H(0)]/T$ called *future* average. Both are Gaussian random variables, and their correlation is

$$C = \frac{1}{2}\frac{(2T)^{2H} - T^{2H} - T^{2H}}{(T^H)^2} = 2^{2H-1} - 1 \tag{4.8}$$

which is independent of T because of its self-affine characteristic. Also in the case $H = 1/2$, the mild randomness, C is equal to zero. Instead, $C > 0$ in the "persistent" case $1/2 < H < 1$ and $C < 0$ in the "anti-persistent" case $0 < H < 1/2$. In both cases, FBM is neither a martingale nor a Markov process. See Figure 9-C4 and Figure 10-C4 for a graphic view.

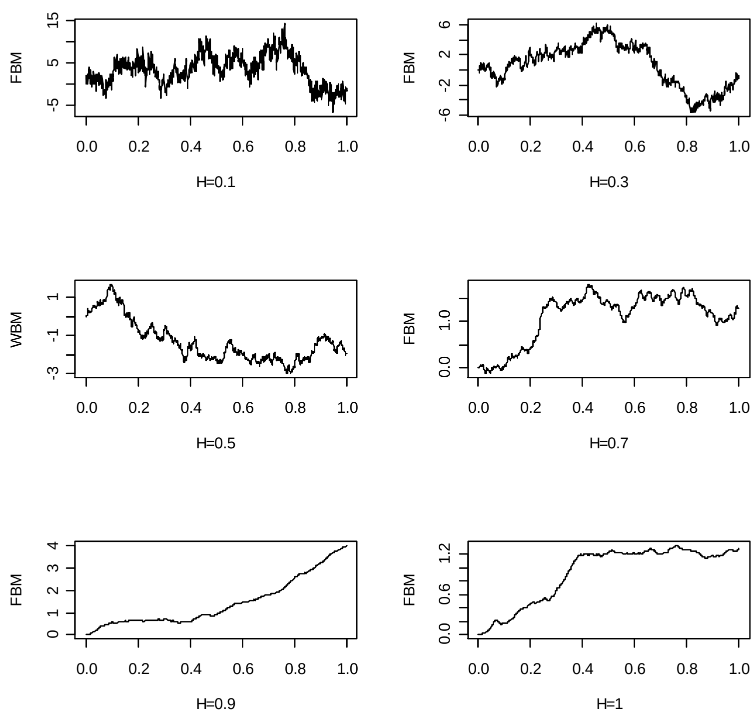


Figure 9-C4. Simulations of fBm from the anti-persistent case, $H = 0.1$, to the persistent one, $H = 1$.

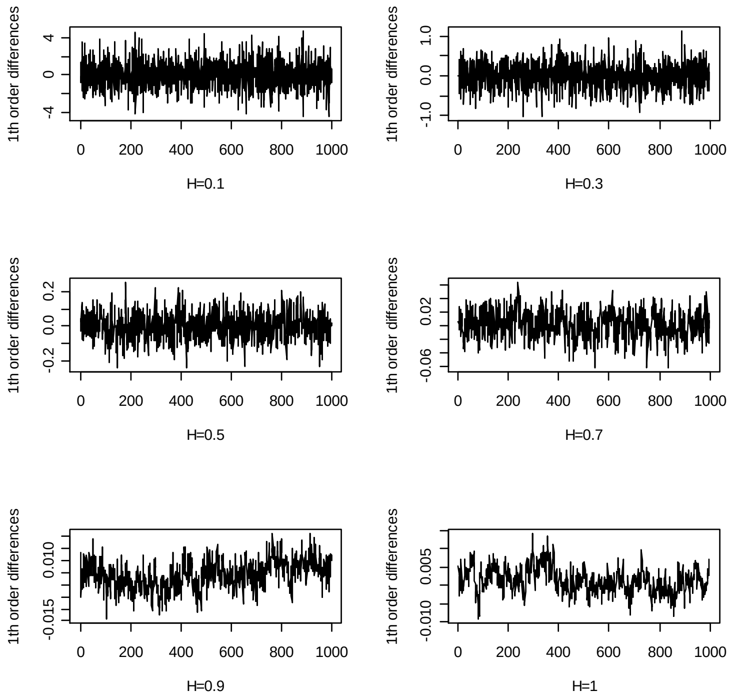


Figure 10-C4. First order differences of the fBm exposed in Figure 9-C4. Note the graphic difference between the pure WBm, $H = 0.5$, and the anti- and persistent cases.

According to Mandelbrot, $B_H(t)$ is continuous and non-differentiable, and the spectral density of the fractional Gaussian noise, $B'_H(t)$, is proportional to f^{-B}, with the exponent $B = 2H - 1$ ranging between 1 and -1. These phenomena are denoted by physicists as "$1/f$ noise" or "pink noise" (Figure 11-C4).

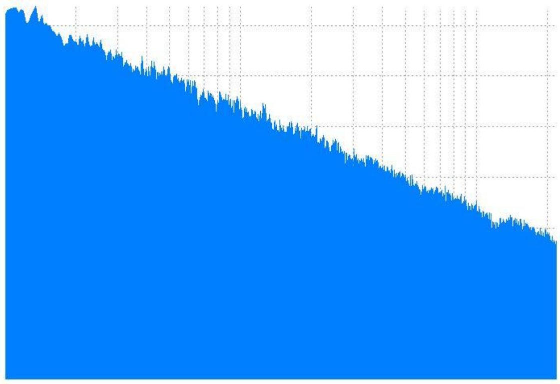


Figure 11-C4. Spectrum of a pink noise approximation. Note the similarity with the auto-correlogram in Figure 8-C4

We refer to the reader to see Appendix A for the methods used to estimate the Hurst exponent. Table 1-C4 shows an empirical analysis of financial series on the exponent H.

	H		***H***
GENERALI	0.74	**FINARTE CASA D'ASTA**	0.57
PIRELLI & R	0.61	**FIAT**	0.57
MEDIOLANUM	0.83	**OLIDATA**	0.80
ACOTEL GROUP	0.68	**BANCA PROFILO**	0.64
LOTTOMATICA	0.47	**ENGINEERING**	0.75
LUXOTTICA	0.62	**BANCO SANTANDER**	0.63
LA DORIA	0.90	**BENETTON GROUP**	0.57
TENARIS	0.56	**ENEL**	0.56
TOS'S	0.51	**ERG**	0.50
FASTWEB	0.40	**BREMBO**	0.63

Table 1-C4. Estimates of the exponents H of 20 assets calculated on a sample of 1099 data, from December 28, 2004 to March 25, 2009

Therefore, the different range of possible Hurst parameters divides the family of fBm into three groups that can be distinguished by these typical criteria:

- $H < 1/2$, the anti-persistent case;
- $H = 1/2$, the pure WBm;
- $H > 1/2$, the persistent case.

4.3.2 Exponent H and auto-covariance property

As shown in Figure 8-C4, the increments of the different processes can also be characterized by using the auto-covariance properties. A stochastic process has a short memory provided by its auto-covariance function and it declines at least exponentially when the lags are increased. An intermediate memory exist, if its auto-covariance function only declines hyperbolically but the infinite sum of all the absolute values of auto-covariances still exists. If the latter condition is no longer satisfied, a long memory exists[47].

Now we examine the auto-covariance function of the stationary process of fBm's increments. Considering $\gamma_H(\tau)$ as the auto-covariance function and τ as a generic time, we have

$$\begin{aligned}\gamma_H(\tau) &= \mathrm{E}\big[\big(B_H(t+\tau) - B_H(t+\tau-1)\big)\big(B_H(t) - B_H(t-1)\big)\big] \\ &= \mathrm{E}[B_H(t+\tau)B_H(t)] - \mathrm{E}[B_H(t+\tau)B_H(t-1)] \\ &- \mathrm{E}[B_H(t+\tau-1)B_H(t)] + \mathrm{E}[B_H(t+\tau-1)B_H(t-1)] \\ &= \frac{1}{2}[(t+\tau)^{2H} + t^{2H} - \tau^{2H} - (t-1)^{2H} + (\tau+1)^{2H} \\ &- (t+\tau-1)^{2H} - t^{2H} - (t+\tau-1)^{2H} - t^{2H} + (\tau+1)^{2H} \\ &+ (t+\tau-1)^{2H} + (t-1)^{2H} - \tau^{2H}] \\ &= \tfrac{1}{2}[(\tau+1)^{2H} - 2\tau^{2H} + (\tau-1)^{2H}].\end{aligned}$$

According to Rostek[48], the last term is an approximation of the second derivative of the function $f(\tau) = \tau^{2H}$, called central finite difference. So for a large τ, the auto-covariance function behaves like the second derivative

$$f''(\tau) = 2H(2H-1)\tau^{2H-2}$$

[47] This detailed explanation is given by Stefan Rostek in his PhD thesis: *Option Pricing in Fractional Brownian markets*, April 2009

[48] See note 8.

According to the theory of infinite sum, the infinite sum of values of the second derivative only exists for $H < 1/2$, but is unlimited for $H > 1/2$. So we can realize that the anti-peristent fBm has an intermediate memory, while the persistent one has a long-memory, as shown in Figure 12-C4 and Figure 13-C4.

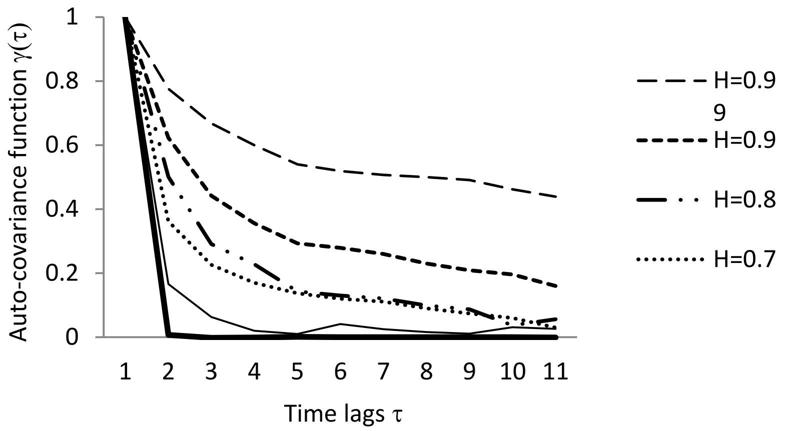


Figure 12-C4. Auto-covariance function of fBm for the case of persistence.

Figure 12-C4 shows that all the curves are bounded by an upper and a lower limiting curve. The upper boundary is the line of total persistence, while the lower one is the case of a serial independence.

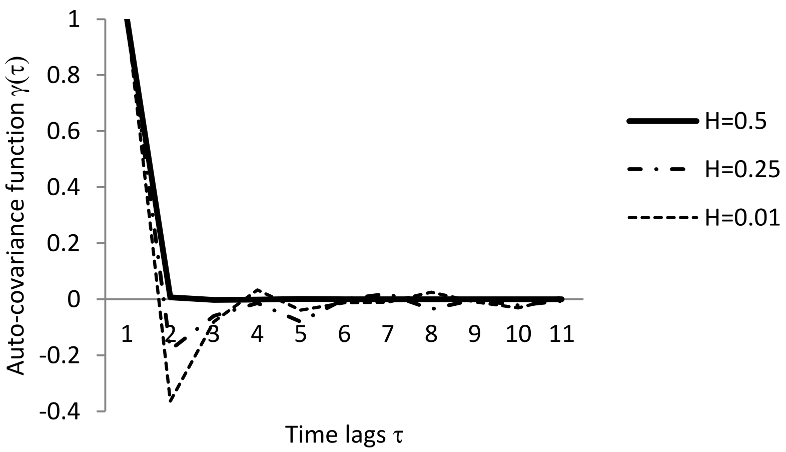


Figure 13-C4. Auto-covariance function of fBm for the case of anti-persistence.

Figure 13-C4 shows the anti-persistent case where there is again the case of a serial independence. The other cases present a particular auto-covariance function that differs from zero in only two cases, when there is either a total over a lap between the two increments or in the case of two neighboring increments of equal length. For non-overlapping increments $(\tau > 1)$ all the curves tend towards zero as the distance grows.

4.3.3 *Trail* and *Graph* dimensions

Upon the Hurst exponent, we can define two types of fractal dimensions. Indeed, while self-similar fractals have a unique fractal dimension, Mandelbrot showed that self-affine fractals demand at least two, *graph dimension* D_G and *trail dimension* D_T where $D_G = 2 - H$ and $D_T = 1/H > D_G$ which refer to different geometric objects.

The value D_G is the box dimension of the graph of $X(t)$, while the value D_T corresponds to the box dimension of a trail. If a one-dimensional WBm $X(t)$ is combined with another WBm $Y(t)$, the process becomes a two dimensional Brownian motion, like that in Figure 2-C4. So the value D_T is the fractal dimension of the three dimensional graph of coordinates t, $X(t)$ and $Y(t)$ and the projected trail of coordinates $X(t)$ and $Y(t)$[49]. In addition, D_G is applied to the coordinates t and $X(t)$, or to t and Y(t).

A FBM with $H \neq 1/2$ can be only embedded in a space of dimension $\mathrm{E} \geq \max \ (2, 1/H)$.

4.3.4 The revolutionary *trading time* and *Noah* and *Joseph effects*

According to Mandelbrot, if we work on a *trading floor* we can experiment with a different concept of time. In fact, in some days a trader is so busy not to realize that the time is flying, but in others where time seems to stop. This is *trading time* which differs from the natural clock time, so modelling a series on this time

[49] For a WBM with $H = 0.5$ we get $D_T = 2$ and $D_G = 1.5$.

is more realistic. To do this we must consider the compound processes. A process is "compound", "decomposable" or "separable" if their variation can be separated into the combination of two distinct contributions. The first is a trading time θ, a random non-decreasing function of clock time t. Feller[50] defines $\theta(t)$ as a directing function. The second, called compounding function, makes prices as function of a *trading time*. This will be presented in a graph for a better interpretation.

Given a price function $P(t)$ and an arbitrary choice of $\theta(t)$, we can define a function $X(\theta)$ that

$$X[\theta(t)] = P(t) \tag{4.9}$$

which is a self-affine process. It is shown that if this process is continuous the trail dimension does not change, conversely, if it is discontinuous both trail and graph dimension change.

But what does trading time do? What is his main contribution to Brownian motion?

Given two non-overlapping time increments $d't$ and $d''t$, the corresponding increments $d'B(t)$ and $d''B(t)$ are independent. But when B follows a trading time θ, that is, a non-linear one, a non-decreasing function of t, $B(t)$ is replaced by $B^*(\theta) = B[t(\theta)]$ whose increments exhibit a very strong dependence also if they remain white, that is, uncorrelated[51].

Moreover, this time is called fractal because it is ruled by a Lévy devil staircase[52].

More general, we must refer to a multifractal structure which involves a non-decreasing multifractal random function with an infinite number of parameters

[50] Feller, W., An Introduction to Probability Theory and its Application, vol. II New York, 1950, pp. 347

[51] It is now necessary to clarify that uncorrelation does not imply independence in our real world, only in the Gaussian one, in which it define the presence or not only of a "linear" dependence.

[52] The term "staircase" refers to the presence off lat steps in the Lévy motion. A fleeting instant of clock time allows trading time to change by a positive amount, generating the price jumps.

whose increments are called multifractal measures. A price model of *multifractal time* presents the following characteristics:

- price increments do not follow a Gaussian distribution;
- concentration of the volatility: the periods of big price variation are clustered and followed by the more placid ones
- the moments show a structure of scale invariance, characterized by parameter α;

These concepts can be summarized in two Mandelbrot's effects: *Noah* and *Joseph effects*. The first refers to the biblical story of the Flood. As the Genesis reports, when Noah was six-hundreds years, God ordered that the Flood would purify our world dominated by wickedness. This tragic background is also in the financial markets, better represented by discontinuity. The second effect, instead, refers to the story of Joseph, a Jewish slave, who interpreted the Pharaoh's dream of even fat years followed by the seven lean ones. So the last defines the tendency of the persistence of a time series. Summarizing, while *Noah effect*, measured by the exponent α, deals with the dimension of an event, *Joseph* is measured by the exponent H, depends on the precise order of events. In this way, Mandelbrot constructed a model in which he combined long-tails and long-dependence.
In some cases these two effects are so correlated that we can write

$$H = \frac{1}{\alpha} \tag{4.10}$$

According to this relationship (4.10), we can describe the excess of reaction leading the investors to sell a big amount of assets after a period of excessive optimism which helps the formation of the "speculative bubbles". The greatness of this excess of reaction can be estimated either through the evaluation of the demand and supply, or through the exponent α intrinsic to market data. But it is

not a consolation because no one is able to forecast the time of the next day, so we cannot forecast the instant in which the bubble breaks.
So in a market where we combine discontinuity and long-tails, it is impossible to foresee the future.

4.3.5 The property of uni- and multi-scaling

The property of uniscaling is referred to the equation (4.7), according to which the scale factors based on moments to satisfy

$$\{\mathrm{E}[B_H(t+T)-B_H(t)]^q\}^{1/\mathrm{q}}=cT^H \qquad (4.11)$$

for all $q>-1$ to avoid the infinite case, and where c is a constant. We realize that every moment is independent of q, so we can write that $\Delta B_H \sim \Delta t^H$.
From this last result we obtain that

$$H \sim \frac{\log|\Delta B_H|}{\Delta t}$$

The case where the moments depend on q is called *multiscaling*. Multifractal objects are more complicated and are characterized by many exponents, therefore, the moments of a multifractal measure ΔM take the form

$$\mathrm{E}[(\Delta M)^q]=\Delta t^{\tau_D(q)+1}$$

and

$$\{\mathrm{E}(\Delta M)^q\}^{1/\mathrm{q}}=\Delta t^{\sigma_D(q)}$$

where $\sigma_D(q)=\frac{1+\tau_D(q)}{q}$, $\tau_D(q)$ is the moment exponent function which is a list of principal exponents.

Summarizing:

- The *uniscaling cases*. Multifractal time reduced to clock time, is $\tau_D(q) + 1 = q$, implying uniscaling, since $\sigma_D(q)$ is independent of q;
- The *multiscaling cases*. $\tau_D(q)$ must satisfy two conditions:
  a) $\mathrm{E}[(\Delta M)^0] = 1$, that is, $\tau_D(0) = -1$
  b) $\mathrm{E}[\Delta M] = \Delta t$. that is, $\tau_D(1) = 0$

  However, the graph of $\tau_D(q) + 1$ is not a straight line. So $\sigma_D(q)$ decreases as $q \to \infty$.

Trying to construct a multifractal FBM. We consider $\bar{H}$ as a vector of the exponents of the form

$$\bar{H} = 0.25 + 0.25 \sin 6\pi\bar{x} \tag{4.12}$$

where $\bar{x}$ is a sequence of numbers constrained between 0 and 1.

So this multifractional Brownian motion is represented by the Figure 14-C4.

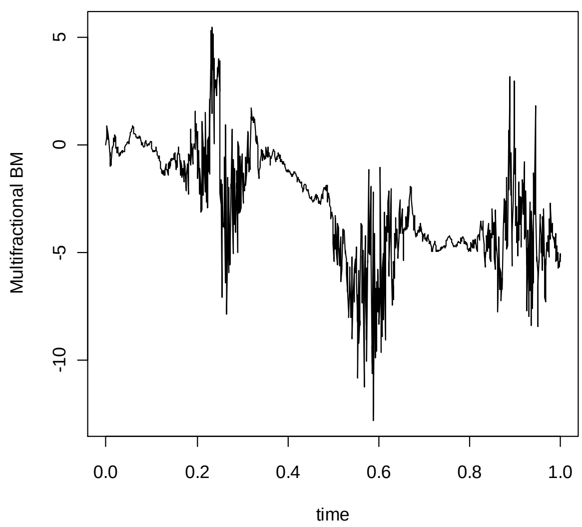


Figure 14-C4. Multifractional Brownian motion according to the vector $\bar{H}$, equation (4.12). Note the characteristic of *self-affine*.

Chapter 5

PORTFOLIO SELECTION FOR STRONG FLUCTUATING ASSETS

All the concepts exposed in the previous chapters are now presented in an empirical application. This chapter and the next last one deal with an application of both α-stable distributions and fractional Brownian motion to finance. Indeed, we have seen that financial markets are turbulent, so that the extreme events, taking place over several periods, produce variations we can define "outliers". But this term is correct only if we refer to the Gaussian world where they are more and more improbable. According to Mandelbrot, power laws have an extraordinary power, so they manage to make the quoted outliers more probable. For example, we consider the equation (3.11), $P(u) = (u/\tilde{u})^{-\alpha}$. For a minimum value $\tilde{u}$ and for $\alpha = 3$, we can calculate the probability to have an event similar to that on October 19, 1987, -22%. According to a Gaussian distribution, the probability of this negative event is equal to 10^{-107}, while taking arbitrarily $\tilde{u} = 4.68\%$, we have $P(22\%) = (22\%/4.68\%)^{-3} = 0.96\%$[53].

[53] According to Taleb, the exponent α is equal to 3 for financial markets. Moreover, $\tilde{u}$ is taken to be four times the standard deviation which is 1.17%. In addition, time interval of Dow Jones Industrial Average prices is from October, 2 1928 to November, 11 2009.

Therefore, this reasoning allows us to consider the mean as insignificant, and every variation deviates from it in a wilder and wilder manner. Hence, we must consider the α-stable distributions with infinite variance, changing the methods used to select a portfolio because of the *wild* state of randomness.

After presenting a long-run strategy able to choose the risky assets to insert them in the portfolio, this chapter discusses the optimal allocation problem with respect to α-stable distributed returns. In particular, we take into consideration some portfolio selection models based on a different risk measure, so we construct the *dispersion matrix* using the covariations, a statistical instrument which correspond to the covariance for $\alpha = 2$. Then, we compare the optimal allocation obtained with the Gaussian and the stable distributional assumption for the risky returns.

The practical and theoretical appeal of the stable non-Gaussian approach is given by its attractive properties which are the same as the normal one. In fact, stable distributions have a domain of attraction, so according to the Central Limit Theorem (CLT), any distribution in the domain of attraction of a specified stable distribution has properties close to those of stable distribution. In addition, another aspect is the already quoted *invariance under addition*.

Finally, we derive the generalized equilibrium relationship between risk and return again under the assumption that the changes of price follow a symmetric α-stable distribution, with $1 < \alpha < 2$, introducing the *stable* CAPM with a generalized coefficient β.

5.1 Efficient frontier with fluctuating assets

The graph which shows the non-stationary of variance, Figure 2-C1, demonstrates that a distribution with finite variance cannot be used to foresee or apply to several methods. The assumption of Gaussian distribution is not verified empirically. Mandelbrot and Fama[54] proved that an empirical distribution of price changes such as stocks, foreign currencies, etc…, are closer to stable distributions rather than to the "bell curve". Assuming that $1 < \alpha < 2$ implies that mean is assumed to be finite but it is not necessary in this way for the variance. Therefore, the classical mean-variance approach, exposed in Section 1.4.1.3, is no longer valid. If a model does not consider the possibility of strong prices' fluctuations, it of course gives a wrong estimate of the quantity of risk and of the risk premium.

However, some techniques have to change. The *covariance* operator gives way to a more generalized one, the *covariation*, explained in Section 3.3.2, which corresponds to the covariance if and only if $\alpha = 2$, while in the other cases, we have to calculate the α-moments of the distributions.

5.1.1 Efficient frontier with risky assets

In our context, the measure of risk is the scale parameter of an appropriate multivariate symmetric stable distribution.

Let's consider $\bar{R}$ as the vector of considered asset returns and $\mathrm{E}[\bar{R}] = \mu^0$. We assume that $\bar{R} - \mu^0$ follow a $S\alpha S$[55] distribution with $\alpha > 1$. The symmetry assumption allows positive and negative returns to be weighted in the same way. Therefore, hypothetical investor has preferences which are well-represented by a utility function defined over the mean and the scale of a portfolio return $r_P = \sum_{i=1}^{n} \omega_i r_i$, where r_i is the return of the asset i and ω_i is the amount invested in the asset i.

So we have:

$$\mathrm{E}[\bar{r}_P] = (\omega, \mu^0) \qquad (5.1)$$

[54] Mandelbrot, B.B., *New Methods in Statistical Economics*, Journal of political economy, 56, 1963; Fama, E.F., *The Behavior of Stock Market Prices*, Journal of business, 38, January 1965.

[55] $X \sim S\alpha S$ means that X follows a stable distribution $S_{\alpha,\beta}(\gamma, \delta)$ with $\beta = \delta = 0$ to indicate symmetry.

$$\|r_P\|_\alpha = \left(\int_{S_n} |(\omega, s)|^\alpha \Gamma(ds) \right)^{\frac{1}{\alpha}} \tag{5.2}$$

where ω is the n-vector of the portfolio weights and μ^0 is the n-vector of asset return means.

Then, respecting the relationship between risk and return, according to which we cannot increase an expected return without increasing risk, we have to solve the following optimization problem:

$$\min_{\omega \in R^n} \int_{S_n} |(\omega, s)|^\alpha \Gamma(ds) \tag{5.3}$$

$$\textit{subject to}:$$

$$(\omega, \mu^0) = \bar{r}_P$$

$$(\omega, \bar{e}) = 1$$

where $\bar{r}_P$ is a given value of portfolio return and $\bar{e}$ denotes a n-vector of those.

Press and Arad[56] showed that the efficient set is convex, meaning that the efficient frontier is the locus of all the convex combinations of any efficient portfolio.

5.1.2 The whole model and its discretization

Several authors have applied those theories to finance. In this research field an important role is played by the work of Ortobelli, Rachev and Schwartz[57] who

[56] Press, S.J., *Multivariate Stable Distributions*, Journal of multivariate analysis, 2:444-462, 1972; Arad, W.R., *The Implications of a Long-Tailed Distribution Structure To Selection and Capital Asset Pricing*, PhD thesis, Princeton University, January 1975

[57] Ortobelli, S., Rachev, S., Schwartz, E., *The problem of optimal asset allocation with stable distributed returns*, UC Los Angeles, Finance, Anderson Graduate School of Management, January 2000

show the significant differences in the portfolio allocation when the data fit the stable non-Gaussian or the normal distribution.

To understand how the optimization problem (5.3) comes out, we have to consider the following form of the characteristic function for $(1 < \alpha < 2)$

$$\Phi_{\bar{R}}(t) = \mathrm{E}[\exp(it'\bar{R}\,)] = \exp\left[-(t'Qt)^{\alpha/2} + it'\mu^0\right] \tag{5.4}$$
$$= \exp\left[-\int_{S_n} |t's|\,\Gamma(ds) + it'\mu^0\right]$$

where $Q = \left[\frac{R_{ij}}{2}\right]$ is a positive definite $(n \,\mathrm{x}\, n)$-matrix, μ^0 is the mean vector, and $\Gamma(ds)$ is the spectral measure with support concentrated on $S_n = \{s \in \mathbb{R}^n : \|s\| = 1\}$.

The term R_{ij} is defined by

$$\frac{R_{ij}}{2} = \left[\tilde{r}_i, \tilde{r}_j\right]_\alpha \left\|\tilde{r}_j\right\|_\alpha^{2-\alpha} \tag{5.5}$$

where the covariation $\left[\tilde{r}_i, \tilde{r}_j\right]_\alpha$ between two joint symmetrically α-stable random variables $\tilde{r}_i = r_i - \mu_i$ and $\tilde{r}_j$ is given by

$$\left[\tilde{r}_i, \tilde{r}_j\right]_\alpha = \int_{S_2} s_i \left|s_j\right|^{\alpha-1} \mathrm{sgn}\left(s_j\right)\Gamma(ds) \tag{5.6}$$

which is equal to the equation (5.2). In particular,

$$\|\tilde{r}_i\|_\alpha = \left(\int_{S_2} |s_i|^\alpha\,\Gamma(ds)\right)^{\frac{1}{\alpha}} = \left(\left[\tilde{r}_i, \tilde{r}_j\right]_\alpha\right)^{\frac{1}{\alpha}}$$

5.1.3 Estimating the parameters

The characteristic function (5.4) is used to estimate the parameter μ and Q, where the estimator of μ is given by the vector $\hat{\mu}$ of a sample average. According to Taqqu and Samorodnitsky[58], for every $1 < p < \alpha$, we can write

$$\frac{\left[\tilde{r}_i, \tilde{r}_j\right]_\alpha}{\left\|\tilde{r}_j\right\|_\alpha^\alpha} = \frac{\mathrm{E}\left[\tilde{r}_i \tilde{r}_j^{\langle p-1 \rangle}\right]}{\mathrm{E}\left[\left|\tilde{r}_j\right|^p\right]} \tag{5.7}$$

where the scale parameter σ_j can be written $\sigma_j = \|\tilde{r}_i\|_\alpha$.

Using the *moment method* suggested by Taqqu and Samorodnitsky[59] in the symmetric case, $\beta = 0$,

$$\sigma_j^p = \|\tilde{r}_i\|_\alpha^p = \frac{\mathrm{E}\left[\left|r_j - \mu_j\right|^p\right] p \int_0^\infty u^{-p-1} \sin^2 \mathrm{u} \mathrm{du}}{2^{p-1} \Gamma(1 - p/\alpha)}$$

So It follows from (5.7)

$$\frac{R_{ij}}{2} = \sigma_j^2 \frac{\mathrm{E}\left[\tilde{r}_i \tilde{r}_j^{\langle p-1 \rangle}\right]}{\mathrm{E}\left[\left|\tilde{r}_j\right|^p\right]}$$

Finally, the estimator $\hat{Q} = \left[\frac{\hat{R}_{ij}}{2}\right]$ is correct for the covariation matrix Q

$$\frac{\hat{R}_{ij}}{2} = \sigma_j^2 \frac{\sum_{k=1}^N \tilde{r}_i^{(k)} \left|\tilde{r}_j^{(k)}\right|^{\langle p-1 \rangle}}{\sum_{k=1}^N \left|\tilde{r}_j^{(k)}\right|^p} \tag{5.8}$$

[58] Lemma 2.7.16 in Samorodnitsky, G., Taqqu S. M., *Stable non-Gaussian random processes – Stochastic models with infinite variance*, Chapman & Hall, New York, 1994

[59] Property 1.2.17 in Samorodnitsky, G., Taqqu S. M., *Stable non-Gaussian random processes – Stochastic models with infinite variance*, Chapman & Hall, New York, 1994

where p represents the rate of convergence of the empirical matrix $\hat{Q}$ to the unknown, to be estimated, matrix Q, which will be faster if p is as large as possible[60].

5.1.4 The optimization problem

The previous section allows us to present our asset allocation.

This section discusses the *stable dispersion measure* which is similar to the equation (5.8). Indeed, this measure can be seen as a generalization of the classic standard deviation.

Let us consider the following stable risk measure

$$\sigma_{\omega,r} = \sqrt{\omega' Q \omega} \tag{5.9}$$

where $\omega' Q \omega$ is the weighted dispersion matrix which can be estimated by calculating the elements, q_{ij}, as follows:

$$\bar{q}_{ij} = (\bar{q}_{jj})^{\frac{1}{2}} \mathrm{A}(1) \frac{1}{N} \sum_{k=1}^{N} \tilde{r}_{i,k} \operatorname{sgn}(\tilde{r}_{j,k}) \tag{5.10}$$

$$\bar{q}_{ij} = \left(\mathrm{A}(p) \frac{1}{N} \sum_{k=1}^{N} \left| \tilde{r}_{j,k} \right|^{p} \right)^{2/p} \tag{5.11}$$

where

$$\tilde{r}_{j,k} = (r_j - r_f) - \mathrm{E}(r_j - r_f)$$

$$\mathrm{A}(p) = \frac{\Gamma(1 - p/2)\sqrt{\pi}}{2^p \Gamma\left(\frac{1+p}{2}\right) \Gamma(1 - p/\alpha)}$$

[60] For a proof see Rachev, S., *Probability metrics and the stability of stochastic models*, New York: Wiley, 1991.

$$p \in [0, \alpha) \approx \frac{\alpha}{3}$$ [61]

and α is the mean of the index of the stability of a return vector and α_i is the index of the stability of the i-th asset estimated, using a maximum likelihood estimator.

So, the problem (5.3) assumes the following form

$$\min_{\omega \in R^n} \sqrt{\omega' \hat{Q} \omega} \tag{5.12}$$

$$\begin{gathered} subject\ to: \\ (\omega, \mu^0) = \bar{r}_P \\ (\omega, \bar{e}) = 1 \end{gathered}$$

or

$$\max_{\omega \in R^n} (\omega, \mu^0) \tag{5.13}$$

$$\begin{gathered} subject\ to: \\ \sqrt{\omega' \hat{Q} \omega} = \hat{Q} \\ (\omega, \bar{e}) = 1 \end{gathered}$$

where the short sales are allowed.

The optimal portfolio weights ω take the following form:

$$\omega = \hat{Q}^{-1} (\mu - r_f e) \frac{m - r_f}{A - 2Br_f + Cr_f^2} \tag{5.14}$$

where $\mu = \mathrm{E}[r]$; $m = \omega'\mu + (1 - \omega')r_f$; $e = [1, \dots, 1]'$; $A = \mu' \hat{Q}^{-1} \mu$; $B = e' \hat{Q}^{-1} \mu$; $C = e' \hat{Q}^{-1} e$.

[61] The parameter p is computed in order to minimize the rate of convergence of asset return series. According to Ortobelli, a good approximation is $\alpha/3$.

So we can note that the *dispersion frontier* is obtained in the same way as the *mean-variance one*.

5.2 The stable *Capital Asset Pricing Model*

By observing carefully the optimal weights (5.14), we can realize that every optimal portfolio can be seen as the linear combination between the market portfolios

$$\widehat{\omega}' r = \frac{(\mu - r_f e)\hat{Q}^{-1} r}{B - C r_f}$$

According to Sharpe equilibrium model, the expected return of the asset i is given by

$$\mathrm{E}[r_i] = r_f + \beta_{i,m}\left[\mathrm{E}(\widehat{\omega}' r) - r_f\right]$$

where $\beta_{i,m} = \frac{\omega' \hat{Q} e^i}{\omega' \hat{Q} \omega}$, with e^i, the vector with 1 in the i-th component and zero in all the other components.

Instead, Gamrowski and Rachev[62] propose a generalization of Fama's *stable CAPM* assuming $r_i = \mu_i + b_i Y + \varepsilon_i$, for $i = 1, \dots, N$, where ε_i and Y are α-stable distributed and $\mathrm{E}(\varepsilon | Y) = 0$.

So, we obtain the following formula for the stable CAPM:

$$\mathrm{E}[r_i] = r_f + \tilde{\beta}_{i,m}\left[\mathrm{E}(\widehat{\omega}' r) - r_f\right]$$

where

---

[62] Gamrowski, B., Rachev, S., A testable version of the Pareto-stable CAPM, Mathematical and computer modeling, 29, 61-81, 1999

$$\tilde{\beta}_{i,m} = \frac{1}{\alpha[\widehat{\omega}'\tilde{r},\widehat{\omega}'\tilde{r}]_\alpha}\frac{\partial[\widehat{\omega}'\tilde{r},\widehat{\omega}'\tilde{r}]_\alpha}{\partial\widehat{\omega}_i} = \frac{[\tilde{r}_i,\widehat{\omega}'\tilde{r}]_\alpha}{[\widehat{\omega}'\tilde{r}\ ,\ \widehat{\omega}'\tilde{r}]_\alpha}$$

In addition, the coefficient can be estimated as shown in (5.7) having

$$\beta_{i,m} = \frac{\widehat{\omega}'\hat{Q}e^i}{\widehat{\omega}'\hat{Q}\widehat{\omega}} = \frac{1}{\sigma_{\widehat{\omega}'r}}\frac{\partial\sigma_{\widehat{\omega}'r}}{\partial\widehat{\omega}_i} = \tilde{\beta}_{i,m}$$

5.3 The empirical results

The empirical analysis concerns the construction of a portfolio composed of 20 risky assets, selected with respect to a particular long-run strategy. The analysis's results allow us to show that the stable frontier is more risk preserving than the Gaussian efficient one. In fact, for a given expected return, or conversely for a given dispersion value, the stable frontier exploits more investment opportunity because, as we can see, the locus of all the convex combinations of any efficient portfolio is larger than the Gaussian one.

5.3.1 Selecting risky asset

When an investor decides to invest in stock market, the first step is to choose and apply a strategy based on his own time preferences. What we present now is a long-run strategy because of the indexes used to evaluate a company.

Let us consider only two economic indexes, *earning/price ratio* and *return on investment*.

The first index is really important, because it allows us to buy risky assets at an undervalued price. Let us suppose that the spot price of a firm is €12 and that this firm is able to obtain an earning of € 1.20 per share, so the gain amounts to 10%. Comparing this with the risk-free return of a ten-yearly Treasury bond which we approximate to 6%, we can realize a risk premium of 4%. But to foresee the earning of a firm is not easy, so to compare several firms we use the following ratio:

$$\frac{EBIT}{actual\ value}$$

where *EBIT* is the acronyms of *Earning Before Interest and Tax* and *actual value* of a company is the sum of *equity market value* (or revalued asset) and *net financial debt*.

The choice of EBIT rather than the classic earning and the choice of the actual firm value rather than a market capitalization has a fundamental meaning. In fact, the actual firm value adds to the market capitalization the net financial debt used to generate operating profit. In addition, EBIT allows us to compare firms with different tax rates and debt levels.

The second index, return on investment has the following form:

$$\frac{EBIT}{working\ capital + fixed\ assets}$$

The choice of EBIT has been previously explained. The denominator's construction is based on the following reasoning: a company has to fund not only credits and warehouse goods, using the working capital, but also a tangible asset thanks to which it can develop its business.

Table 1-C5 shows how to construct a cash flow.

How to estimate a cash flow

(+) revenue	
(-) expenses	
= EBITDA	(Earnings Before Interest Tax Depreciation and Amortization)
(-) amortization	
= EBIT	(Earning Before Interest and Tax)
(-) tax	
= NOPAT	(Net Operating Profit After Tax)
(+) amortization	
= net tax EBITDA	
(+/-) working capital variation	
(+/-) fixed investment/disinvestment	
= UCF	Operating Cash Flow or Unlevered Cash Flow
(-) financial cost	
(-) Repayment of loans	
(+) New investment	
= LCF	Available Cash Flow or Levered Cash Flow

Table 1-C5. The procedure used to estimate the unlevered (available to stakeholders) and levered (available to shareholders) cash flow.

But how are both the indexes related?

A company which is able to obtain great earnings and whose *earning power*[63] is very high, the re-investment policy can realize excellent results, increasing its evaluation.

As shown by a work of Joel Greenblatt[64], this strategy is long-run because its results are obtained after ten or more years, so the market is able to identify an undervalued price only after some years.

With respect to those two indexes, we can select companies through several steps:

[63] The ability of a company to make a profit on its operations which can be good estimate with our *return on investment*.

[64] Greenblatt, J. teaches Finance at Columbia Business School and is the author of "*The little book that beats the market*".

1. Taking a list of companies with homogeneous market capitalizations (we have taken into account companies with a market capitalization of minimum € 200 millions, € 1 billion for extra-European companies);
2. Classifying them according to our first index and giving rank one to the best of them, rank 2 to the second and so on;
3. Classifying them again according to our second index and giving rank one to the best of them and so on;
4. Inserting in portfolio the best twenty or thirty companies corresponding to the sum of the two ranks. For example, for the company A the rank based on our first index is 5, while the other based on our second index is 34, so the company A is in the 39th position;
5. Buying only 5-7 assets, using a percentage of our money;
6. Repeating step 5 every two or three months, obtaining a portfolio composed of 20-30 assets after nine or ten months;
7. Managing the portfolio substituting all the assets every year;
8. Being patient.

Table 2-C5 shows the companies selected thanks to the previous procedure.

Company	Market	Share per € 1000	Capitalization (in €)	(1) EBIT/actual value	(2) Return on Investment	Field	Rank (1)	Rank (2)	Sum rank (1) and rank (2)
Telegate AG	DEU	104.17	204 M	22%	872%	Service Company	10	6	16
Poyry Oyj	FIN	84.03	696 M	12%	2139%	Service Company	20	1	21
Gestevision Telecinco SA	ESP	95.6	2547 M	15%	852%	Media	17	7	24
Global Payments Inc.	USA	30.98	2633 M	9%	1672%	Service Company	23	2	25
Endo Pharmaceuticals	USA	69.8	1681 M	17%	599%	Biotechnology	15	10	25
Sohu.com Inc.	USA	24.32	1583 M	9%	1659%	Data Storage Service	23	3	26
AmerisourceBergen Corp.	USA	53.49	5390 M	11%	1072%	Biotechnology	21	5	26
FTI Consulting, Inc.	USA	33.28	1558 M	9%	1562%	Service Company	23	4	27
Hewitt Associates, Inc.	USA	34.67	2699 M	11%	777%	Service Company	21	8	29
Apollo Group, Inc.	USA	23.85	6499 M	12%	644%	Schools	20	9	29
GameStop Corp.	USA	70.24	2348 M	19%	297%	Consumer Technology	13	17	30
GEA Group AG	DEU	63.15	2911 M	16%	309%	Service Company	16	15	31
Questar Corporation	USA	32.46	5376 M	12%	465%	Petroleum/natural gas	20	12	32
Terra Industries Inc.	USA	42.9	2329 M	29%	169%	Chemicals	4	28	32
EVS Broadcast Equipment SA	BEL	20.88	650 M	11%	367%	Audio e video equipment	21	14	35
Herbalife Ltd.	USA	32.32	1886 M	11%	309%	Personal care products	21	16	37
EMCOR Group, Inc.	USA	55.81	1182 M	24%	161%	Building	8	30	38
King Pharmaceuticals, Inc.	USA	113.56	2188 M	12%	281%	Biotechnology	20	19	39
IMS Health, Inc.	USA	67.2	2717 M	10%	255%	Data Storage Service	22	21	43
Lockheed Martin Corporation	USA	18.76	20314 M	15%	188%	Defence and Aerospatiale	17	26	43

Table 2-C5. Portfolio.

5.3.2 Stable and Gaussian frontier: a comparison

According to the optimization problems (1.29) and (5.12), respectively the Gaussian and stable ones, this empirical analysis indicates that the stable non-Gaussian allocation is more risk preserving than the normal one.

Let us consider $\bar{x}^t S \bar{x}$ and $\omega' \hat{Q} \omega$ respectively the variance-covariance matrix and the dispersion one. In particular, these are (20×20)-matrixes. We consider μ^0 as the mean vector.

Table 3-C5 presents the estimation of all used parameters.

	α	Scale parameter	Standard deviation	
telegate AG	1.442	0.008%	2.140%	**p**
Poyry Oyj	1.314	0.013%	4.718%	0.492
Gestevision Telecinco SA	1.539	0.011%	2.216%	
Global Payments Inc.	1.546	0.009%	2.174%	**A(p)**
Endo Pharmaceuticals	1.589	0.009%	2.130%	0.922
Sohu.com Inc.	1.472	0.026%	3.466%	
AmerisourceBergen Corp.	1.508	0.006%	1.682%	**A(1)**
FTI Consulting, Inc.	1.599	0.011%	2.463%	0.566
Hewitt Associates, Inc.	1.451	0.007%	2.062%	
Apollo Group, Inc.	1.441	0.014%	2.933%	**N**
GameStop Corp.	1.576	0.022%	3.071%	1268
GEA Group AG	1.478	0.017%	2.784%	
Questar Corporation	1.504	0.014%	2.792%	
Terra Industries Inc.	1.492	0.032%	3.912%	
EVS Broadcast Equipment SA	1.406	0.013%	6.989%	
Herbalife Ltd.	1.358	0.014%	3.034%	
EMCOR Group, Inc.	1.452	0.019%	3.071%	
King Pharmaceuticals, Inc.	1.444	0.011%	2.489%	
IMS Health, Inc.	1.359	0.007%	2.207%	
Lockheed Martin Corporation	1.546	0.006%	1.707%	
α **mean**	**1.472**			

Table 3-C5. Stable and Gaussian parameters' estimation.

Having estimated all the parameters, now we can construct both the efficient frontiers which are represented by Figure 1-C5.

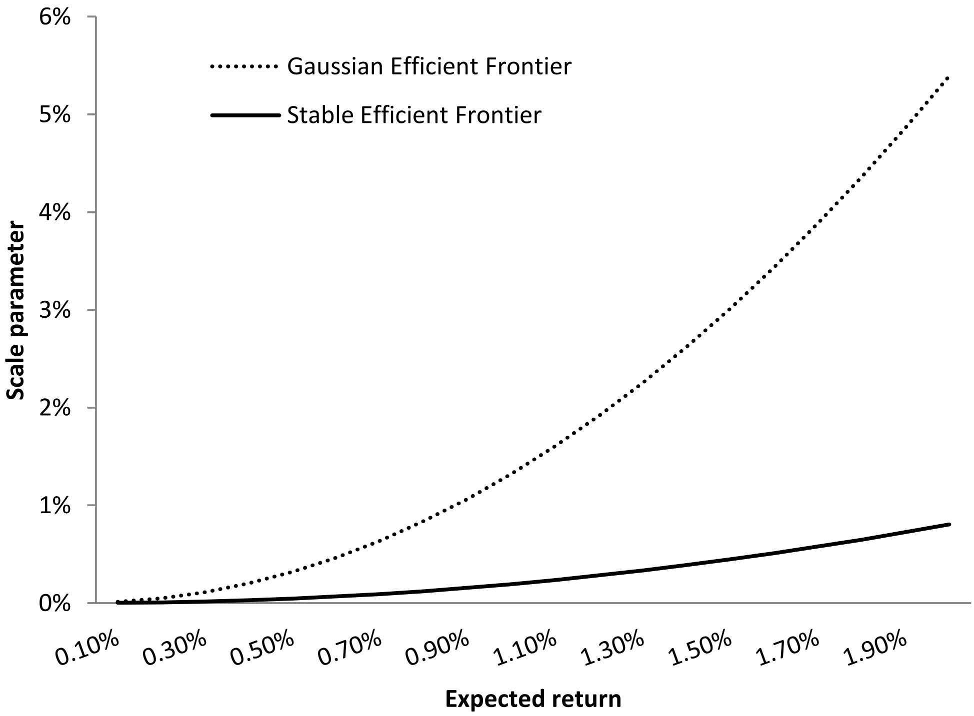


Figure 1-C5. Stable frontier versus Gaussian frontier with short sales and without the risk-free asset.

Observing carefully Figure 1-C5, we realize that there is an area between the two frontiers that becomes greater and greater for higher expected returns.
In fact, to have a better idea of the stable frontier's risk preserving we can look at the Table 4-C5.

GAUSS-STABLE Comparison				
Expected value	**0.1%**	**1%**	**5%**	**10%**
Gaussian scale parameter	**0.014%**	**1.324%**	**34.117%**	**137.032%**
Stable scale parameter	**0.002%**	**0.195%**	**5.142%**	**20.723%**
Amount of risk in excess	**0.012%**	**1.129%**	**28.974%**	**116.310%**

Table 4-C5. Comparison of both the frontiers for some given expected values.

In this case, according to the equation (5.5), the Gaussian scale parameter is equal to the standard deviation divided by $\sqrt{2}$, so we can compare the two risk measures.

5.4 Fascinating and powerful conclusion

The previous analysis is very important. The plot of the Gaussian and stable efficient frontiers in the same mean-scale space, displayed by Figure 1-C5, shows that for a given expected portfolio return, the associated risk in the stable model is lower than its counterpart in the Gaussian model. We can conclude that the portfolio selection, obtained in this way, allows us to take into consideration the so-called *outliers*, that is, the extreme asset return thanks to the stable non-Gaussian distribution family. In Chapter 3 we have explained the characteristics of these distribution.

Summarizing these characteristics which are really important to demonstrate our thesis, according to which the stable asset allocation is better than the Gaussian one, we can say that

- they are characterized by the tendency to follow trends and cycles, with sudden changes of direction;
- they are more heavy-tailed than the normal ones and with the maximum value, around which the majority of the observations is concentrated;

- they have finite mean and infinite variance for $1 < \alpha < 2$;
- they are characterized by *invariance under addition* which replicate the *scaling* properties of the financial market;

These properties can be found in the financial market, where it is shown that changes of price do not follow a normal distribution; that "one single observation can destroy thousands of years of confirmation"[65]; that daily, weekly and monthly prices have similar graphic form; and finally, that the assumption of stationary in variance is not a real assumption.

In conclusion, our study could be a useful and convenient tool for fund managers and investors who try to maximize their trade-off between risk and return.

[65] See for example the *Black Thursday*, October, 19 1987 when the Dow Jones fell by 22%.

Chapter 6

OPTION PRICING WITH RESPECT TO THE FRACTIONAL BROWNIAN MOTION

In 1973, Fisher Black and Myron Scholes first articulated a formula able to price derivatives in their paper "*The pricing of options and corporate liabilities*". This work presents a mathematical model of the market where the stock's price follows a stochastic process. In the same year, Robert C. Merton published a paper expanding the mathematical understanding of the option pricing model and coined the term thanks to which we can refer to, the *Black and Scholes model*. The model, presented in Section 1.3, is criticized by many other economists, especially for the reasons exposed in Chapter 1. Because of its incompatibilities with real market data, the main critics are interested in the stochastic process of Brownian motion. Firstly, price changes cannot follow a Gaussian distribution as shown by the real data. Secondly, the process of observable market values seems to exhibit a serial correlation, allowing us to distinguish the persistent case from the anti-persistent one, as shown in Chapter 4. Therefore, a model which can better represent reality is the so-called fractional Brownian motion, the real candidate. This motion was introduced by Mandelbrot and Van Ness in 1968 and is again a Gaussian stochastic process. It is able to capture *long-range dependencies* or *persistence* and the characteristic of *self-similarity* of financial series thanks to the *Hurst exponent*, lying between one and zero.

According to the theories explained in Chapter 4, now we want to apply the respective theory of option pricing to the fractional Brownian motion, which corresponds to the pure Brownian motion for $H = 1/2$.

This chapter, together with the fifth one, can be seen as the core of the thesis and presents the option pricing model according to the continuous time fractional Brownian market introduced by the work of Stefan Rostek and Rainer Schöbel[66]. We will price options using conditional expectation of fBm and a conditional version of the fractional Itô theorem. The reason is to obtain a closed-form solution for the price of a European option written on a stock, following a fBm with arbitrary Hurst parameter, and we will show the reasons for excluding arbitrage.

6.1 Binomial approximation of an arithmetic fBm

As shown in Chapter 4, a pure Wiener Brownian motion (WBm) can be approximated to a binomial distribution by increasing the number of variables, independently and identically distributed at random. But for the fBm the quoted result is achieved through a more refined procedure. According to Mandelbrot and Van Ness[67], the fractional Brownian motion, B_t^H, with Hurst parameter H can be regarded as a moving average of a two-sided classical Brownian motion B_s:

$$B_t^H = c_H \left[\int_{\mathbb{R}} \left((t-s)_+^{H-\frac{1}{2}} - (-s)_+^{H-\frac{1}{2}} \right) dB_s \right] \tag{6.1}$$

[66] Rostek, S. and Schöbel, R., *Risk Preference Based Option Pricing In a Fractional Brownian Market*, Eberhard Karls Universität Tübingen, May 2005.

[67] Mandelbrot, B.B and van Ness, J.W., Fractional Brownian Motion, Fractional Noises and Applications, SIAM Rev 10(4), p 422-437, 1968

where c_H is a normalizing constant.

As shown in Equation (6.1), a fBm is related to an infinite past, so there are some problems if we want to model a process going infinitely back to the past, when any future step depends on the whole history. Therefore, to render this procedure easier, we look at a process starting at a fixed point in time $t = 0$, in order to consider a fBm as a finite Brownian integral, given by Norros and Valkeila[68]. Their derivation of this finite interval representation of fBm concerns the following formula:

$$B_t^H = \int_0^t z(t,s)dB_s \tag{6.2}$$

where

$$z(t,s) = c_H \left[\frac{t^{H-\frac{1}{2}}}{s} (t-s)^{H-\frac{1}{2}} - \left(H - \frac{1}{2}\right) s^{\frac{1}{2}-H} \int_s^t u^{H-\frac{3}{2}} (u-s)^{H-\frac{1}{2}} du \right]$$

So the last formula allows us to consider a fractional Brownian motion as a weighted sum of Brownian increments going finitely back to the past, which are variables independently and identically distributed at random $\xi_i^{(n)}$, with zero mean and unit variance,

$$B_t^{(n)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \xi_i(n) \tag{6.3}$$

[68] Norros, I., Valkeila, E., Virtamo, J., An Elementary approach to a Girsanov Formula and Other Analytical Result on Fractional Brownian Motion, Bernoulli 5, p. 571-587, 1999

where in the n-th approximation step each unit time interval is divided into n discrete steps and time t is rounded down onto the next n-th part by replacing it by $[nt]/n$.

To understand how the last sum approximates to the standard Brownian motion, we can observe the following first two moments:

$$\mathrm{E}\left[B_t^{(n)}\right] = \mathrm{E}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} \xi_i(n)\right] = \frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} \mathrm{E}[\xi_i(n)] = 0 \tag{6.4}$$

$$\mathrm{Var}\left[B_t^{(n)}\right] = \mathrm{E}\left[\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} \xi_i(n)\right)^2\right] = \frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} \mathrm{E}\left[\left(\xi_i(n)\right)^2\right] = \frac{[nt]}{n} \to t \tag{6.5}$$

Sottinen[69] gives us a more complex version of equation (6.2), where he introduces a continuous time weighted kernel

$$z^{(n)}(t,s) = n\int_{s-\frac{1}{n}}^{s} z\left(\frac{[nt]}{n}, u\right) du$$

achieving

$$B_t^{H(n)} = \int_o^t z^{(n)}(t,s) d\, W_s^{(n)} = \sum_{i=1}^{[nt]} \int_{\frac{i-1}{n}}^{\frac{i}{n}} z\left(\frac{[nt]}{n}, u\right) du \frac{1}{\sqrt{n}} \xi_i^{(n)} \tag{6.6}$$

which equals $B_t^{(n)}$, WBm, for $H = 1/2$.

Therefore, we obtain a binomial random walk which is a sum of uncorrelated random variables and which approximates to the pure Brownian motion.

[69] Sottinen, T., *Fractional Brownian Motion, Random Walks and Binary Market Models*, Financ Stochast 5(3), p.343-355, 2001

6.2 Fractional Brownian motion and its conditional moments

The next step that allows us to obtain a closed-formula of option pricing is to use the information of all the steps up to a certain time t. So considering the approximation of fBm, equation (6.6), a better representation of a finite-interval of classical Brownian motion is given by its conditional moments, the conditional expectation and variance. Considering $\mathfrak{F}_t$ as the information set to time t, we have

$$\begin{aligned}\mathrm{E}[B_t^H \mid \mathfrak{F}_t] &= \mathrm{E}\left[\int_0^t z(T,s)dB_s|\mathfrak{F}_t\right] \\ &= \int_0^t z(T,s)dB_s + \mathrm{E}\left[\int_t^T z(T,s)dB_s|\mathfrak{F}_t\right] \\ &= \int_0^t z(T,s)dB_s\end{aligned}$$

obtaining, according to Sottinen, we obtain

$$\hat{B}_{T,t}^H = \mathrm{E}[B_T^H \mid \mathfrak{F}_t] = \sum_{i=1}^{[nt]} z^{(n)}\left(T, \frac{i}{n}\right)\frac{1}{\sqrt{n}}\xi_i^{(n)} \qquad (6.7)$$

where we use the same coefficients as B_T^H only for summing them to time t. In this way, we can construct a conditional binomial tree which is again a representation of the evolution of fBm to time T, where each node indicates the mean of all the terminal nodes descending from it. In addition, it is interesting to investigate the second moment, the conditional variance defined by

$$\hat{\sigma}_{T,t}^2 = \mathrm{E}\left[B_t^H - \hat{B}_{T,t}^H \middle| \mathfrak{F}_t\right]^2$$

According to Rostek, comparing the values of the conditional variances based on a different number of approximating steps with the continuous time limit case, we can realize that the speed of convergence differ eminently. This speed depends both on the Hurst parameter and on the relation between the lengths of the observation and the prediction interval.

So we can conclude with some important characteristics of a fBm. This process is no longer a *martingale* because of the lack of equality between the future prediction and the present value. Besides, the prediction not only depends on the past, but also on all historic random realizations, so the process is neither Markovian. So in the next sections we will give a detailed explanation to solve this problem when we deal with financial models.

6.3 Binomial approximation of a geometric fBm

We know that the stochastic process of a spot price S_t is defined by the following differential equation

$$dS(t) = \mu S(t)dt + \sigma\mu S(t)dB_t^H$$

where dB_t^H is the increment of fBm.

Now let us consider the following recursion rule allowing us to generate a recursive multiplicative tree, starting with a value S_0:

$$S_n = S_{n-1} \circ (1 + \mu_n + X_n)$$

where $\circ$ is either the ordinary product or the discrete Wick product[70], and X_n is the approximation of fBm increment as the equation (6.7) shows.

[70] This type of product vanishes if the two factors have at least one gene rating random variable, ξ, in common. If none of the generators coincide, the Wick product equals the ordinary one. In formula $\prod_{i\in A}\xi_i \diamond_d \prod_{j\in B}\xi_j = \begin{cases}\prod_{i\in A\cup B}\xi_i & \text{if } A\cap B=\emptyset\\ 0 & \text{otherwise}\end{cases}$

But these types of products give us two different price processes. Indeed, consider the equation (6.7) and its expression in the following way

$$B_T^H = \sum_{i=1}^{[nt]} k^{(n)}(T,i)\xi_i^{(n)}$$

where $k^{(n)}(T,i) = z^{(n)}\left(T,\frac{i}{n}\right)\frac{1}{\sqrt{n}}$.

Let us consider only two steps of recursion, using $t_1 = 1$ and $t_2 = 2$ without a drift μ. So we have

$$B_0^H = 0 \ , \qquad B_1^H = k(1,1)\xi_1 \ , \qquad B_2^H = k(2,1)\xi_1 + k(2,2)\xi_2$$

and

$$dB_1^H = k(1,1)\xi_1 \ , \qquad dB_2^H = \big(k(2,1) - k(1,1)\big)\xi_1 + k(2,2)\xi_2$$

With $S_0 = 1$, the ordinary product, OP, give us the following price process of the geometric fBm:

$$\begin{aligned} S_0^{(OP)} &= 0 \\ S_1^{(OP)} &= S_0^{(OP)}(1 + dB_1^H) = 1 + k(1,1)\xi_1 \\ S_2^{(OP)} &= S_1^{(OP)}(1 + dB_2^H) \\ &= 1 + k(2,1)\xi_1 + k(2,2)\xi_2 + k(1,1)k(2,2)\xi_1\xi_2 \\ &+k(1,1)\big(k(2,1) - k(1,1)\big)\xi_1^2 \end{aligned}$$

where ξ_1^2 is equal to 1, so the associated term contributes to the drift of the process, and the expected value is not zero.

While using the Wick product, WP, we realize that it eliminates the squared term and so the drift remains unchanged. This is the reason why we introduce this product.

$$
\begin{aligned}
S_0^{(WP)} &= 0 \\
S_1^{(WP)} &= S_0^{(WP)} \diamond_d (1 + dB_1^H) = 1 + k(1,1)\xi_1 \\
S_2^{(WP)} &= S_1^{(WP)} \diamond_d (1 + dB_2^H) \\
&= (1 + k(1,1)\xi_1) \diamond_d \big(1 + \big(k(2,1) - k(1,1)\big)\xi_1 + k(2,2)\xi_2\big) \\
&= 1 + k(2,1)\xi_1 + k(2,2)\xi_2 + k(1,1)k(2,2)\xi_1\xi_2
\end{aligned}
$$

The binomial tree modelling the classical Brownian motion is important to price derivatives only if we suppose the absence of arbitrage. But we will see that for $H \neq 1/2$, the absence of arbitrage is no longer valid in the fractional background, unless we impose some restrictions.

6.3.1 Arbitrage in fractional framework and its exclusion

According to the binary fBm explained in the previous section, the assets only change their value at discrete point in time $0 = t_0 < t_1 < \cdots < t_n = T$.

Let us consider a riskless asset $A_j^{(n)}$, where (n) is the n-th approximation, and its dynamics

$$A_j^{(n)} = \big(1 + r^{(n)}\big)A_{j-1}^{(n)}$$

and a risky stock

$$S_j^{(n)} = \left(1 + \mu^{(n)} + X_j^{(n)}\right) \circ S_{j-1}^{(n)}$$

where, according to Rostek, $X_j^{(n)}$ has the following representation:

$$X_j^{(n)} = \sigma\left(B_{t_j}^{H(n)} - B_{t_{j-1}}^{H(n)}\right) \tag{6.8}$$
$$= \sigma\sqrt{n}\left(\sum_{i=1}^{j}\int_{\frac{i-1}{n}}^{\frac{i}{n}} z\left(\frac{j}{n},s\right)ds\xi_i^{(n)} - \sum_{m=1}^{j-1}\int_{\frac{m-1}{n}}^{\frac{m}{n}} z\left(\frac{j-1}{n},s\right)ds\xi_m^{(n)}\right)$$

According to Sottinen, we can rewrite this replacing

$$k(j,i) = k^{(n)}(j,i) = \sqrt{n}\int_{\frac{i-1}{n}}^{\frac{i}{n}} z\left(\frac{j}{n},s\right)ds$$

and so

$$f_{j-1}(\xi_1,\cdots,\xi_{j-1}) = \sum_{i=1}^{j-1}\left(k(j,i) - k(j-1,i)\right)\xi_i$$

hence, the equation (6.8) becomes

$$X_j^{(n)} = \sigma\left(k(j,j)\xi_j + f_{j-1}(\xi_1,\cdots,\xi_{j-1})\right)$$

So only two possible values exist of the binary random variable $X_j^{(n)}$ for each step j, which are defined by $d_j^{(n)}$, for $\xi_j^{(n)} = -1$, and $u_j^{(n)}$, for $\xi_j^{(n)} = +1$.

Finally, we can derive the no-arbitrage relation necessary to ensure any time step so we have

$$S_j^{(n)}(\xi_j = -1) - S_{j-1}^{(n)} < rS_{j-1}^{(n)} < S_j^{(n)}(\xi_j = +1) - S_{j-1}^{(n)} \tag{6.9}$$

which indicates that the return of the risky asset exceeds the riskless interest rate in the case $X_j^{(n)} = u_j^{(n)}$, while it drops below the riskless rate in the case $X_j^{(n)} = d_j^{(n)}$.

According to the equation (6.8), its increments are

$$S_j^{(n)} - S_{j-1}^{(n)} = \mu S_{j-1}^{(n)} + X_j^{(n)} S_{j-1}^{(n)}$$

and so, the “no-arbitrage relation” is

$$d_j^{(n)} < r - \mu < u_j^{(n)} \tag{6.10}$$

Rostek provides an explicit arbitrage possibility. Let us suppose that the difference $r - \mu$ is negative and that the binomial tree is strictly upward to step $j > N_H$, where N_H represents a critical step number depending on Hurst parameter and tending to infinity as $H \to 1/2$, that is

$$\left(\xi_1, \xi_2, \cdots, \xi_{j-1}\right) = (+1, +1, \cdots, +1)$$

obtaining

$$d_j^{(n)} = f_{j-1}(+1, +1, \cdots, +1) - k(j, j)$$
$$u_j^{(n)} = f_{j-1}(+1, +1, \cdots, +1) + k(j, j)$$

and

$$r - \mu < 0 \le d_j^{(n)} < u_j^{(n)}$$

so the relation (6.10) is violated. An investor could buy one stock at step $j-1$ at price S_{j-1} and borrow the same amount paying the riskless interest rate r. In the worst case, ξ_j is <1 and the stock moves downward taking the value $S_j = S_{j-1}(1+\mu+d_j)$. Being $\mu + d_j > r$, we have a gain. On the other hand, if $r-\mu$ is positive, the opposite trading strategy is allowed, that is, the *short-selling* one.

In the next section we will introduce some restriction to eliminate arbitrage possibility.

6.4 Restriction in eliminating the arbitrage possibility

The previous section shows us that an arbitrage is possible in a fractional framework. But a careful reader can realize that the quoted arbitrage is allowed only if an investor is as fast as the market so, the one-step buy-and-hold strategy is possible only if an investor can react quickly.

Therefore, if we introduce a minimal delay between two consecutive transactions of the same investor, we have a modified framework where investors cannot react as fast as the market free of arbitrage.

Look at the Figure 1-C6 to understand better.

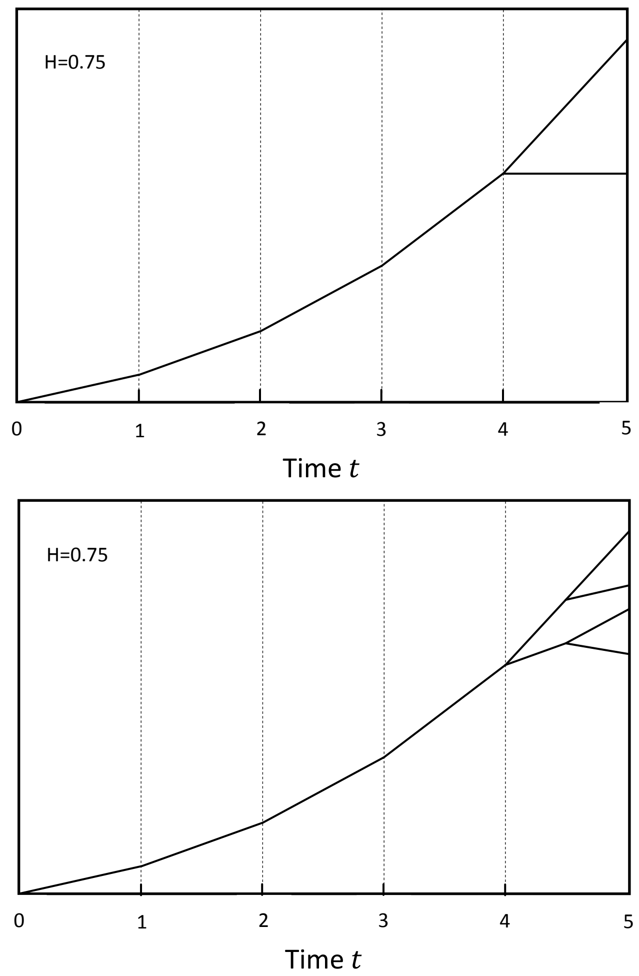


Figure 1-C6. Value of the fractional price process (y-axes). The lower chart shows the exclusion of arbitrage by restricting trading strategies. Investors can only make transactions at nodes on a dashed line.

The above chart in the Figure 1-C6 shows that a buy-and-hold strategy is possible by buying the stock in time $t = 4$ and selling in $t = 5$. While in the lower chart, an investor cannot exploit the transaction because in $t = 5$ he could encounter a loss.

So, even if an investor sells immediately after having bought an asset, he would have missed a number of transaction nodes caused by the multitude of investors. If we introduce a great number of intra-interval nodes, the arbitrage opportunity vanishes and this number depends on both the amount of historic available information and Hurst parameter. An high level of persistence and more information about the past increase the number of necessary steps.

6.5 Different arbitrage strategies

6.5.1 Arbitrage according to Shiryayev

We have seen that if we adopt the ordinary product, the fBm is not a semimartingale because of the presence of a drift.

Shiryayev[71] analyzes a financial model where the drift of the risky asset is equal to the interest rate of the riskless asset and the volatility is equal to one.

We have a bond A_t and a stock S_t following

$$dA_t = rA_t dt \tag{6.11}$$

$$dS_t = rS_t dt + S_t \delta B_t^H \tag{6.12}$$

From the chain rule formulated as a differential equation given by

$$d\mathrm{F}(s, B_s^H) = \frac{\partial F}{\partial s} ds + \frac{\partial F}{\partial B_s^H} \delta B_s^H$$

it follows that the explicit equations of the basic market assets have the form

$$A_t = A_0 e^{rt}$$

$$S_t = S_0 e^{rt + B_t^H}$$

where A_0 and S_0 are assumed to be one. So the value of the portfolio X_t^π based on the weighted-strategy $\pi = (\beta_t, \gamma_t)$ is

[71] Shiryayev, A.N., *On arbitrage and replication for fractal models*, research report 20, MaPhySto, Department of Mathematical Sciences, University of Aarhus, Denmark, 1998

$$X_t^\pi = \beta_t A_t + \gamma_t S_t \tag{6.13}$$

where the strategy is called *self-financing*, if

$$dX_t^\pi = \beta_t dA_t + \gamma_t dS_t$$

So we have

$$\beta_t = 1 - e^{2B_t^H}$$
$$\gamma_t = 2\left(e^{B_t^H} - 1\right)$$

Substituting the last result in (6.13), we obtain

$$X_t^\pi = e^{rt}\left(e^{B_t^H} - 1\right)^2$$

Applying the chain rule, we obtain

$$dX_t^\pi = re^{rt}\left(e^{B_t^H} - 1\right)^2 dt + 2e^{rt+B_t^H}\left(e^{B_t^H} - 1\right)\delta B_t^H \tag{6.14}$$

where the second term can be seen as follows:

$$2e^{rt+B_t^H}\left(e^{B_t^H} - 1\right)\delta B_t^H = \gamma_t S_t \delta B_t^H \tag{6.15}$$

and the first one

$$re^{rt}\left(e^{B_t^H} - 1\right)^2 dt = \gamma_t S_t r dt + \beta_t r B_t^H dt \tag{6.16}$$

Then, combining the equations (6.14), (6.15) and (6.16), we obtain

$$\begin{aligned} dX_t^\pi &= \gamma_t S_t (rdt + \delta B_t^H) + \beta_t r B_t^H dt \\ &= \beta_t dA_t + \gamma_t dS_t \end{aligned}$$

Hence, the strategy is *self-financing* because its initial capital needed is zero and the successive portfolio value is non-negative. So this configures an arbitrage strategy.

6.5.2 Arbitrage according to Bender

In contrast to Shiryayev, Bender[72] gives a different approach using the Wick integration. The starting point is similar to Shiryayev's, that is,

$$dA_t = rA_t dt \tag{6.17}$$

$$dS_t = \mu S_t dt + S_t \sigma B_t^H \tag{6.18}$$

where now the novelty is that Bender allows for an arbitrary constant drift μ and volatility σ.

So the results of the differential equation are

$$\begin{aligned} A_t &= A_0 e^{rt} \\ S_t &= S_0 \exp\left(\mu t - \frac{1}{2}\sigma^2 t^{2H} + \sigma B_t^H\right) \end{aligned} \tag{6.19}$$

and the strategy is given by

$$\begin{aligned} \beta_t &= 1 - \exp(-2rt + 2\mu t - \sigma^2 t^{2H} + 2\sigma B_t^H) \\ \gamma_t &= 2S_0^{-1}\left(\exp\left(-rt + \mu t - \frac{1}{2}\sigma^2 t^{2H} + \sigma B_t^H\right) - 1\right) \end{aligned}$$

[72] Bender, C., *Integration with respect to fractional Brownian motion and related market models*, University of Konstanz, Department of Mathematics and Statistics: PhD thesis, 2003

which is again an arbitrage opportunity.

6.6 Option pricing according to Hu and Øksendal

Based on the equation (6.19), Hu and Øksendal[73] deal with the fractional Brownian Black-Scholes market defined by the equations (6.17) and (6.18).

Supposing that the value of the portfolio is given by the following stochastic process

$$Z_t^\pi = \beta_t A_t + \gamma_t \diamond S_t \tag{6.20}$$

now the authors replace the property of *self-financing* with another concept, that is, a portfolio is said to be *Wick self-financing* if its process satisfies:

$$\begin{aligned} dZ_t^\pi &= \beta_t dA_t + \gamma_t \diamond dS_t \\ &= \beta_t r A_t dt + \mu\gamma_t \diamond S_t dt + \sigma\gamma_t \diamond S_t dB_t^H \end{aligned} \tag{6.21}$$

where $\diamond$ indicates that the differential equation is represented in the Wick sense. From equation (6.20) we have

$$\beta_t = \frac{Z_t^\pi - \gamma_t \diamond S_t}{A_t}$$

which we substitute in (6.21) obtaining

$$dZ_t^\pi = rZ_t^\pi dt + \sigma\gamma_t \diamond S_t \left(\frac{\mu - r}{\sigma} dt + dB_t^H\right) \tag{6.22}$$

Consider now a new probability measure denoted by $\tilde{P}^H$, determined by the Girsanov change of measure and defined

[73] Hu, Y., Øksendal, B., *Fractional White Noise Calculus and Applications to Finance*, Infin Dimens Anal Qu 6(1), p.1-32, 2003

$$\tilde{B}_t^H = \frac{\mu - r}{\sigma} t + dB_t^H$$

as a fractional Brownian motion under the new probability measure. According to the fractional Girsanov theorem[74], we can rewrite (6.22) as follows

$$dZ_t^\pi = rZ_t^\pi dt + \sigma\gamma_t \diamond S_t d\tilde{B}_t^H \tag{6.23}$$

Hu and Øksendal proceed by multiplying both sides of (6.23) by e^{-rt}, and from the fractional version of Itô's Lemma, we have

$$d(e^{-rt}Z_t^\pi) = e^{-rt}dZ_t^\pi - re^{-rt}Z_t^\pi dt$$

Therefore, integrating both sides from 0 to T, the authors derive a stochastic integral having zero expectation:

$$e^{-rt}Z_T^\pi = Z_0^\pi + \int_0^T e^{-rt}\sigma\gamma_t \diamond S_t d\tilde{B}_t^H$$

and

$$\mathrm{E}_{\tilde{P}^H}[e^{-rt}Z_T^\pi] = Z_0^\pi$$

The importance of this result is that π cannot be an arbitrage strategy because a positive value of Z_0^π contradicts the condition of zero or the negative initial investment.

[74] The proof of this theorem goes beyond the aim of this work. For a detailed analysis see Section 2.5 of Rostek Stefan, *Option Pricing in Fractional Brownian Markets*, Springer, April 2009

Hence, in this way the authors have shown that this procedure represents a martingale measure and that it is the only one. Respecting this unique measure, they derive the pricing formula for the European options at time 0:

$$C_0^H = S_0 N(d_1^*) - Ke^{-rt} N(d_2^*) \tag{6.24}$$
$$P_0^H = Ke^{-rt} N(-d_2^*) - S_0 N(-d_1^*) \tag{6.25}$$

where

$$d_1^* = \frac{\ln\left(\frac{S_0}{K}\right) + rT + \frac{1}{2}\sigma^2 T^{2H}}{\sigma T^H}$$
$$d_2^* = \frac{\ln\left(\frac{S_0}{K}\right) + rT - \frac{1}{2}\sigma^2 T^{2H}}{\sigma T^H}$$

Obviously, for $H = 1/2$, we obtain the well-known Black-Scholes-Merton formula presented in Chapter 1. Generally, the formula is also valid for a different time interval, but it is necessary to replace the terms T^{2H} and T^H with respectively $(T^{2H} - t^{2H})$ and $(T^H - t^H)$.

But the derived pricing options aroused many critics concerning the economic meaning of the Wick product. Indeed, it is a product of random variables that needs to know the prospective holdings for all the possible states of nature at a precise point in time, in order to calculate the realization of the portfolio value at that point in time.

Several authors have tried to solve this problem supposing that the dynamics of this process S_t can be interpreted as the fundamental firm value, distinguished carefully from the observable market price. In fact, the latter is assumed to be the outcome of a statistic test function applied to the distribution of the stochastic process. Moreover, other concepts have been introduced, like a *mixed fractional Brownian motion* and *market imperfections*.

The next section presents the risk preference-based option pricing.

6.7 Risk preference-based option pricing

This section presents a work of Rostek and Schöbel who deal with a market where randomness follows a fractional Brownian motion. The price process evolves continually, but the novelty is represented by the introduction of a minimal amount of time between two transactions by the same investor. This restriction allows us to ensure the absence of arbitrage. Besides, in a fractional framework the well-known no-arbitrage pricing approach based on dynamical hedge is inappropriate. So, Rostek solved this problem by introducing risk preferences.

Let us fix current time t and consider a simulation of fBm. Then group each path $w_1, w_2, \ldots$ having identical trajectories to the time t into one class $[w_1]_t$. So we can define the conditional distribution of B_t^H within its observed equivalence class, which is normal with the following moments:

$$\mathrm{E}[B_T^H|\mathfrak{F}_t^H](w_1) = B_t^H(w_1) + \int_{-\infty}^{t} g(T,t,s)(w_1)ds = B_t^H(w_1) + \hat{\mu}_{T,t} \quad (6.26)$$

$$\mathrm{Var}[B_T^H|\mathfrak{F}_t^H](w_1) = \mathrm{E}\left[\left(B_T^H - \hat{B}_T^H\right)^2 \middle| \mathfrak{F}_t^H\right](w_1) = \rho_H(T-t)^{2H} = \hat{\sigma}_{T,t} \quad (6.27)$$

where

$$g(T,t,s) = \frac{\sin\big(\pi(H-1/2)\big)\,(B_s^H - B_t^H)}{\pi(t-s)^{H+\frac{1}{2}}(T-s)}$$

$$\rho_H = \frac{\sin\big(\pi(H-1/2)\big)}{\pi(H-1/2)}\frac{\Gamma(3/2-H)^2}{\Gamma(2-2H)}$$

Now we have the instruments through which we can derive the option pricing.

Let us assume risk-neutral investors possessing and using information about the past and the discounted conditional expected value of a call option based on the observation of $[w_1]_t$:

$$C_{T,H}(t) = e^{-r(T-t)}\mathrm{E}[\max(S_T - k)|\mathfrak{F}_t^H] \tag{6.28}$$

According to Rostek, the pricing problem is solved by defining a suitable measure under which expectations are taken. Hence, he proposes to use the measure Q^{SV} satisfying

$$\mathrm{E}_{Q^{SV}}\left[e^{-r(T-t)}S_T\middle|\mathfrak{F}_t^H\right] = S_t \tag{6.29}$$

which is the unbiased average risk neutral measure.
This measure is based on the following equilibrium respecting the risk-neutrality: the investor should remain indifferent in buying the stock and holding the amount S_t of the riskless asset. Then we can rewrite (6.27) in the following way:

$$\mathrm{E}_{Q^{SV}}[S_T|\mathfrak{F}_t^H] = S_t e^{-r(T-t)} \tag{6.30}$$

According to Rostek, to exploit the equation (6.28) we must consider the conditional distribution of S_T whose moments are given by the equations (6.27) and (6.27). Then applying the conditional version of the fractional Itô theorem[75], we have

$$\ln(\hat{S}_T) = \ln \hat{S}_t + \mu(T-t) - \frac{1}{2}\rho_H\sigma^2(T-t)^{2H} + \sigma(\hat{B}_T^H - \hat{B}_t^H)$$

[75] For a proof of the Conditional fractional Itô theorem, see Section 5.3 of Rostek Stefan, *Option Pricing in Fractional Brownian Markets*, Springer, April 2009.

So we realize that the logarithm of the conditional process $\hat{S}_T$ is normally distributed with the following mean m, and variance v:

$$m = \widehat{\mathrm{E}}\left[\ln\left(\hat{S}_T\right)\right] = \mathrm{E}\left[\ln\left(\hat{S}_T\right)\middle|\mathfrak{F}_t^H\right](w_1) \tag{6.31}$$
$$= \ln S_t + \mu(T-t) - \frac{1}{2}\rho_H\sigma^2(T-t)^{2H} + \sigma\hat{\mu}_{T,t}$$
$$v = \widehat{\mathrm{E}}\left[\ln\left(\hat{S}_T\right) - m\right]^2 = \mathrm{E}\left[\left(\ln\left(\hat{S}_T\right) - m\right)^2\middle|\mathfrak{F}_t^H\right](w_1) \tag{6.32}$$
$$= \rho_H\sigma^2(T-t)^{2H}$$

So S_T is normally distributed with mean M and variance V:

$$M = \exp\left(m + \frac{1}{2}v\right) = S_t e^{\mu(T-t)+\sigma\hat{\mu}_{T,t}} \tag{6.33}$$
$$V = \exp(2m+2v) - \exp(2m+v) = S_t^2 e^{2\mu(T-t)}\left(e^{\rho_H\sigma^2(T-t)^{2H}} - 1\right) \tag{6.34}$$

Now, using (6.33) and inserting it into (6.30) we obtain:

$$S_t e^{\mu(T-t)+\sigma\hat{\mu}_{T,t}} = S_t e^{-r(T-t)}$$

and so

$$\bar{\mu}(T-t) = r(T-t) - \sigma\hat{\mu}_{T,t} \tag{6.35}$$

where $\bar{\mu}$ represents the adjusted drift divided into the return received from the riskless asset and in addition, a historically shift of the distribution. It is important to remember that in the Markovian case of classical Brownian motion, the drift of the process is equal to the free-risk return.

But what does the adjustment mean? What is the financial meaning?

Ex ante, an investor has an idea on how to determine a price evolution. So if he realizes a positive evolution, he adjusts positively the stock's future distribution. A posteriori, he could observe a mispricing between his prediction and the current value so he has to correct his overestimated prediction.
Hence, inserting (6.35) into (6.31) and (6.32) we have

$$m = \ln S_t + r(T-t) - \frac{1}{2}\rho_H \sigma^2 (T-t)^{2H} \tag{6.36}$$

$$v = \rho_H \sigma^2 (T-t)^{2H} \tag{6.37}$$

Finally, according to this density of the conditional process $\hat{S}_T$ which is the conditional density of S_T based on the observation $[w_1]_t$, that is,

$$f(x)|_{[w_1]_t} = \frac{1}{x\sqrt{2\pi v}} \exp\left\{-\frac{1}{2}\frac{(\ln x - m)^2}{v}\right\} \mathrm{I}_{[x>0]}$$

now we are able to write the Rostek and Schöbel formula to price the European option:

$$\begin{aligned} C_{T,H}(t) &= e^{-r(T-t)} \mathrm{E}_{Q^{(T,t)}}[\max(S_T - K)|\mathfrak{F}_t^H] \\ &= S_t e^{m - \frac{1}{2}v - r(T-t)} N(d_1) - K e^{-r(T-t)} N(d_2) \end{aligned}$$

where

$$d_1 = \frac{m + v - \ln K}{\sqrt{v}}$$

$$d_2 = \frac{m - \ln K}{\sqrt{v}} = d_1 - \sqrt{v}$$

Then, substituting (6.36) and (6.37) we obtain the pricing formula for the fractional European call and put:

$$C_{T,H}(t) = S_T N(d_1^H) - Ke^{-r(T-t)} N(d_2^H)$$
$$P_{T,H}(t) = Ke^{-r(T-t)} N(-d_2^H) - S_T N(-d_1^H)$$

where

$$d_1^H = \frac{\ln\left(\frac{S_t}{K}\right) + r(T-t) + \frac{1}{2}\rho_H \sigma^2 (T-t)^{2H}}{\sqrt{\rho_H}\sigma(T-t)^H}$$
$$d_2^H = \frac{\ln\left(\frac{S_t}{K}\right) + r(T-t) - \frac{1}{2}\rho_H \sigma^2 (T-t)^{2H}}{\sqrt{\rho_H}\sigma(T-t)^H} = d_1^H - \sqrt{\rho_H}\sigma(T-t)^H$$

So also the *fractional Put-Call Parity* is valid, that is,

$$C_{T,H}(t) - P_{T,H}(t) = S_t - Ke^{-r(T-t)}$$

For $\rho_H = 1$ and $H = 1/2$ we obtain the classical Black-Scholes-Merton option pricing.

Table 1-C6 shows the increasing parameter according to which we can write the fractional Greeks, exposed in Table 2-C6.

Parameter	Definition
T	Time to maturity
r	Free-risk interest rate
σ	Volatility
S_0	Spot price
K	Strike price
D	Dividend
H	Hurst parameter
ρ_H	variance of conditional fBm

Table 1-C6. Parameters which the fractional European options depend on.

Fractional Greeks	Derivation
$\Delta_H = \frac{\partial C_H}{\partial S}$	$N(d_1^H)$
$\Gamma_H = \frac{\partial C_H}{\partial S^2}$	$\frac{\varphi\left(d_1^H\right)}{S_t\sqrt{\rho_H}\sigma(T-t)^H}$
$\Theta_H = \frac{\partial C_H}{\partial T}$	$H\frac{S_t\varphi\left(d_1^H\right)\sqrt{\rho_H}\sigma}{(T-t)^{1-H}} + rKe^{-r(T-t)}N\left(d_2^H\right)$
$\varrho_H = \frac{\partial C_H}{\partial r}$	$K(T-t)e^{-r(T-t)}N\left(d_2^H\right) - (T-t)S_tN\left(d_1^H\right)$
$\Lambda_H = \frac{\partial C_H}{\partial \sigma}$	$S_t\varphi\left(d_1^H\right)\sqrt{\rho_H}\sigma(T-t)^H$

Table 2-C6. The fractional Greeks and their analytical representation.

6.8 Option pricing comparison

This section gives a summary of the previous derived option pricing and some examples. See Table 3-C6 to have a schematic representation of the three considered option pricing.

<table>
<tr><th colspan="2">Option pricing according to</th></tr>
<tr>
<td>

Black-Scholes-Merton

$$c = S_0 N(d_1) - K e^{-rT} N(d_2)$$
$$p = K e^{-rT} N(-d_2) - S_0 N(-d_1)$$

where

$$d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}$$

$$d_2 = \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}}$$
$$= d_1 - \sigma\sqrt{T}$$

</td>
<td>

Hu and Øksendal

$$C_0^H = S_0 N(d_1^*) - K e^{-rt} N(d_2^*)$$
$$P_0^H = K e^{-rt} N(-d_2^*) - S_0 N(-d_1^*)$$

where

$$d_1^* = \frac{\ln\left(\frac{S_0}{K}\right) + rT + \frac{1}{2}\sigma^2 T^{2H}}{\sigma T^H}$$
$$d_2^* = \frac{\ln\left(\frac{S_0}{K}\right) + rT - \frac{1}{2}\sigma^2 T^{2H}}{\sigma T^H}$$

</td>
</tr>
<tr>
<td colspan="2">

Rostek and Schöbel

$$C_{T,H}(t) = S_T N(d_1^H) - K e^{-r(T-t)} N(d_2^H)$$
$$P_{T,H}(t) = K e^{-r(T-t)} N(-d_2^H) - S_T N(-d_1^H)$$

where

$$d_1^H = \frac{\ln\left(\frac{S_t}{K}\right) + r(T-t) + \frac{1}{2}\rho_H \sigma^2 (T-t)^{2H}}{\sqrt{\rho_H}\sigma(T-t)^H}$$
$$d_2^H = \frac{\ln\left(\frac{S_t}{K}\right) + r(T-t) - \frac{1}{2}\rho_H \sigma^2 (T-t)^{2H}}{\sqrt{\rho_H}\sigma(T-t)^H}$$
$$= d_1^H - \sqrt{\rho_H}\sigma(T-t)^H$$

</td>
</tr>
</table>

Table 3-C6. Comparison of the three derived option pricing.

Let us give an example to show a *numerical* difference among those three option pricing.

Consider the value in Table 4-C6.

Parameter	Definition
T	1 year
r	2%
σ	20%
S_0	100
K	100
D	0
H	0.1;0.5;0.9
ρ_H	0.639887 ($H = 0.1$); 1 ($H = 0.5$); 0.365709 ($H = 0.9$)

Table 4-C6. Value of option pricing parameter.

Table 5-C6 shows the results of the three option pricing exposed in Table 3-C6.

<table>
<tr><th colspan="4">Option pricing according to</th></tr>
<tr><td colspan="2" rowspan="2">Black-Scholes-Merton
$C_t = 2.543321$
$P_t = 2.044569$</td><td colspan="2">Hu and Øksendal</td></tr>
<tr><td>Anti-persistent case $H = 0.1$
$C_t^H = 7.172147$
$P_t^H = 6.673395$</td><td>Persistent case $H = 0.9$
$C_t^H = 2.543321$
$P_t^H = 2.044569$</td></tr>
<tr><td colspan="4">Rostek and Schöbel</td></tr>
<tr><td colspan="2">Anti-persistent case $H = 0.1$
$C_{T,H}(t) = 5.790906$
$P_{T,H}(t) = 5.292154$</td><td colspan="2">Persistent case $H = 0.9$
$C_{T,H}(t) = 1.645797$
$P_{T,H}(t) = 1.147045$</td></tr>
</table>

Table 5-C6. Application of the three derived option pricing.

Observing carefully the Table 5-C6, we can realize that, according to Rostek and Schöbel's option pricing, when the Hurst parameter increases, the option price decreases. But this relation depends on the maturity time as shown in the Figure 2-C6, Figure 3-C6 and Figure 4-C6.

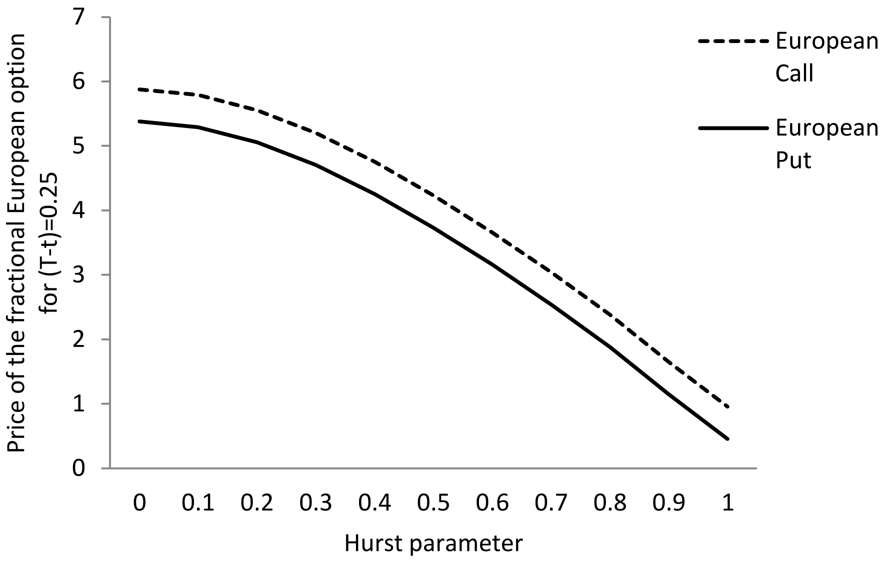


Figure 2-C6. Maturity effect on the relation between the Hurst parameter and the price of the fractional European put and call options for maturity $T = 0.25$.

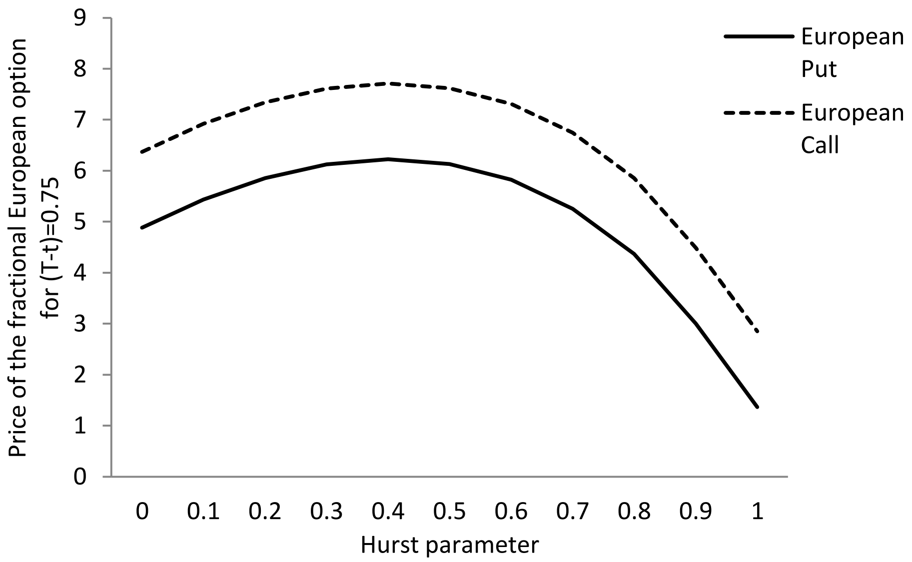


Figure 3-C6. Maturity effect on the relation between the Hurst parameter and the price of the fractional European put and call options for maturity $T = 0.75$.

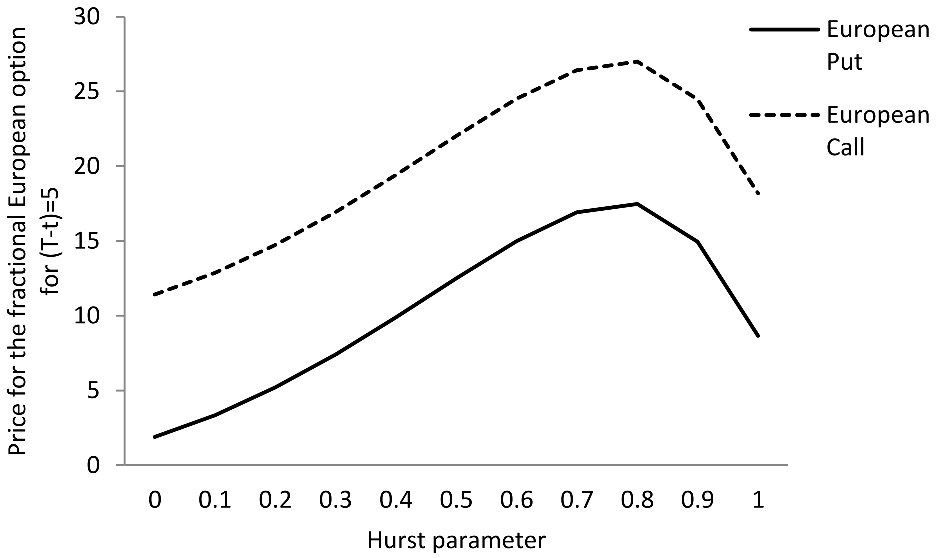


Figure 3-C6. Maturity effect on the relation between the Hurst parameter and the price of the fractional European put and call options for maturity $T = 5$.

In conclusion, we can present schematically the influence of the Hurst parameter on the price of a fractional European option for different maturity intervals τ:

- For $\frac{1}{4} < \tau < 1$, the maximum of the call value and Hurst parameter lie in the anti-persistent area;
- For $\tau = 1$, H is equal to $1/2$, so the case of a serial independence yields the highest call price;
- For $\tau > 1$, the maximum of the call value and Hurst parameter lie in the persistent area.

Conclusion

This thesis is concerned with the two most used but also most criticized statistical instruments input in financial models: the pure *Wiener Brownian motion* and the *Gaussian distribution*.
Indeed, the majority of economic and financial models have in common the Gaussian distribution, where a statistical distribution puts us in a symmetric world and where 68 per cent of the observations are around the mean, while the rare events may occur every 100000 years. Hence, a financial crisis is not included in the above models because they are valid only for a "peaceful" world.
Not until we consider non-financial data, like the height of men and women, we realize that a person three meters tall is really a rare event, and in addition that the average height is concentrated around the mean.
But is it the same about financial data? what is a crisis? or rather, what effects does it produce?
History teaches us that everything can change in a day, that suddenly an index price can fall down by 20 per cent in a few hours. Besides, such a jump corresponds to a really slow recovery or even a non-existent one.

In the beginning of Chapter 1 we have presented the most important and secular financial models taught in all the universities in the world.
After presenting their incompatibilities with the reality, Chapter 2 introduces the *fractal geometry* and its power. A *fractal* is an object that can be *self-similar* or *self-affine* upon different scale. According to Mandelbrot, this *scaling property* is well represented by a particular distribution's family, the *α-stable non-Gaussian*

one which has a closer approach to reality. Chapter 3 deals with these distributions whose power is to follow trends and cycles, with sudden changes of direction; to be more heavy-tailed than the normal ones, with the maximum value around the majority of the observations is concentrated on; to have finite mean and infinite variance for $1 < \alpha < 2$; and to be invariant under addition (the scaling property). The true power is to make the rare events more probable than the Gaussian "bell curve" (See Table 1-Con).

The Black Monday 19/10/1987		$\boldsymbol{r = -25.63\%}$	
$\boldsymbol{\tilde{r}\ min}$	$\boldsymbol{P(r > \tilde{r}) = (r/\tilde{r})^{-3}}$	$\boldsymbol{N(z)}$	$\boldsymbol{10^{-107}}$
$\boldsymbol{-0.01\%}$	0.0000000059%		
$\boldsymbol{-0.05\%}$	0.0000007423%	$mean\ \mu$	0.0184%
$\boldsymbol{-0.10\%}$	0.0000059385%	$std\ dev\ \sigma$	1.1655%
$\boldsymbol{-0.50\%}$	0.0007423135%	$z = (r-\mu)/\sigma$	−22.01%
$\boldsymbol{-1\%}$	0.0059385082%		
$\boldsymbol{-3\%}$	0.1603397208%		
$\boldsymbol{-5\%}$	0.7423135223%		
$\boldsymbol{-8\%}$	3.0405161872%		

Table 1-Con. Dow Jones Index on October 19, 1987. Gauss versus power laws. Note that for different minimum value, the power law makes the negative event more probable than the Gaussian distribution.

But non-Gaussian distribution and *discontinuity* are not the only problems. Do you really think that the next price is totally uncorrelated or independent of the previous ones? Look at the price correlogram of an asset (Figure 8-C4).

This question is partly solved by considering the *fractional Brownian motion* playing an important role in the option pricing. Black, Scholes and Merton have derived a complex model able to price derivatives, based on some abnormal assumptions, like that of independence of price changes and the so-called *Efficient Market Hypothesis*. These assumptions are replaced respectively by the fractional Brownian motion, which is always a Gaussian motion but now with a *long-run dependence* defined by the *Hurst exponent* ($0 < H < 1$; for $H = 1/2$ its increments are independent), and the *Fractal Market Hypothesis* which emphasizes the impact of the liquidity and of the investment horizons on the investors.

This thesis reaches its maximum in the two final chapters on the empirical application of the quoted concepts.

After presenting a *long-run strategy* able to choose the risky assets to insert in the portfolio, Chapter 5 deals with the *optimal allocation problem with respect to* α*-stable distributed returns*. In particular, we take into consideration portfolio selection models based on a different risk measure, constructing the *dispersion matrix* using the covariations. Comparing the optimal allocation under Gaussian and stable distributional assumption for the risky returns, we realize that the second is more risk preserving, and it includes the whole concepts which these non-Gaussian distributions are characterized by.

The last Chapter is concerned with the option pricing with respect to the fractional Brownian motion. Under the assumption of risk-neutral investors, we introduce the *option pricing depending on the Hurst exponent*. But this motion has particular consequences, because its application in financial models allows the arbitrage opportunity. To solve this problem, the work of Professor Rostek Stefan plays an important role. His models take into consideration that even if arbitrage opportunities exist, an investor could not be as fast as the market. Modifying the binomial tree and considering the *conditional fractional Brownian motion*, we realize that the arbitrage vanishes. So this option pricing is an interesting object for our purpose, because we introduce a serial correlation into financial models obtaining solutions that are in a closed-form, easy to handle and in line with economic intuition.

Appendix

METHODS TO ESTIMATE THE HURST PARAMETER

In 1951, after having analyzed more than 800 years of records, Hurst defined a method of studying natural phenomena such as the flow of the Nile River. Indeed, he observed that the flow of the Nile River was not at random, but patterned. Some years later, precisely in 1968, Benoit B. Mandelbrot defined this pattern as fractal. To measure the bias of the fractional Brownian motion, they both introduced a parameter, called H by Mandelbrot himself, which is the core of the previous chapters. As shown in the Chapter 2, also the fractal dimension is also based on this parameter. Hence, it gives a measure of the smoothness of a fractal object.

We have explained that, regarding the range of this parameter, that is, $0 < H < 1$, we realize that a low H indicates a high level of roughness and so a fractal dimension near to 2, has derived from the relation $D = 2 - H$. On the contrary, the high value of it shows the high level of smoothness, and so a fractal dimension near to 1 is the Euler dimension of a straight line.

These two parameters, H and D, can be used to analyze stock market data thanks to their capability of distinguishing persistent historical data from the anti-persistent one. Besides low H values represents higher noise, more random-like, or volatile data, so representing a higher risk. Also the opposite is true.

In this appendix we reveal two of the most important and used estimation methods of the Hurst parameter, that is, the Rescaled Range analysis method, or R/S analysis, and the Dispersional analysis, also known as the Aggregated Variance method.

A.1 Rescaled Range analysis

This analysis is very simple. In fact, unlike the other usual statistical test, this does not demand a data organization. The R/S analysis implies measures over intervals of different length, the difference between the maximum values and the minimum ones only if they are higher or lower than the predictable difference when each datum is independent of the previous one. If the two amounts are different, then order of the data becomes important. Indeed, an uninterrupted sequence of gains or losses pushes the extreme values farther than those which can be made casually.

The analysis is composed of seven steps:

1. Calculate the mean, μ, over the whole of the available data;
2. Sum the difference from the mean to get the cumulative total of each point in time, $\mathrm{V}(N,k)$, from the beginning of the period up to any time;
3. For every maximum and minimum value, $\mathrm{V}_{\max}(N,k)$ and $\mathrm{V}_{\min}(N,k)$, for $0<k<N$, calculate the range $\mathrm{R}(\tau)$;
4. Calculate the sample standard deviation, $S(\tau)$, of the values over the period τ, according to the local mean $\mu(\tau)$;
5. Then calculate the ratio $R/S=\mathrm{R}(\tau)/S(\tau)$;
6. Repeat the procedure, step $1 \dashv 5$, determining R/S for each non-overlapping segment of the dataset. Then, we take the averaged R/S value;

7. In conclusion, plot *log-log plot* that is fit linear regression Y on X where $Y = \log R/S$ and $Y = \log N$ and where the Hurst exponent is the *slope* of the regression line.

Finally, in formula:

$$R/S = \frac{\max\limits_{0<k\leq N} \sum_{j=1}^{k}(r_j - \mu) - \min\limits_{0<k\leq N} \sum_{j=1}^{k}(r_j - \mu)}{\sqrt{\frac{1}{N}\sum_{j=1}^{k}(r_j - \mu)^2}}$$

where r represents the log-return of a price series.

Consider a dataset composed of three asset of three firms which operate in different fields, Fiat, Fastweb and Ubi Bank. Each sample is of 154 weekly price, from January, 1st 2007 to January 5th, 2010.

Table 1-A shows the estimation of the Hurst parameter obtained through the previous procedure.

	Hurst Parameter	Fractal Dimension	Description
FIAT	0.70	1.30	*Persistent*
FASTWEB	0.55	1.45	*Low persistent*
UBI BANK	0.61	1.39	*Low persistent*

Table 1-A. R/S analysis. Estimates of Hurst parameter and fractal dimension.

Moreover, Figure 1-A, Figure 2-A and Figure 3-A show the respective graphs.

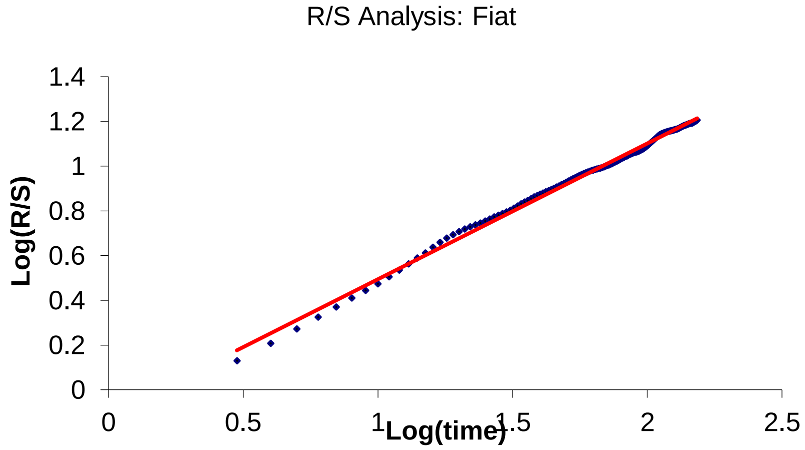


Figure 1-A. R/S analysis. Log-log plot of Fiat with the tendency line (red).

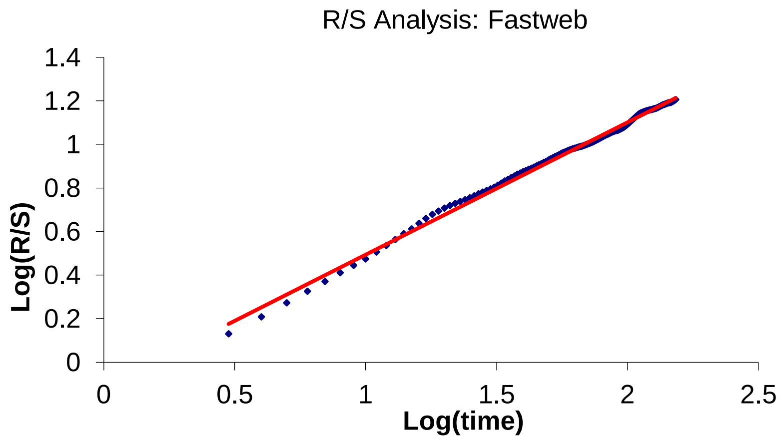


Figure 2-A. R/S analysis. Log-log plot of Fastweb with the tendency line (red).

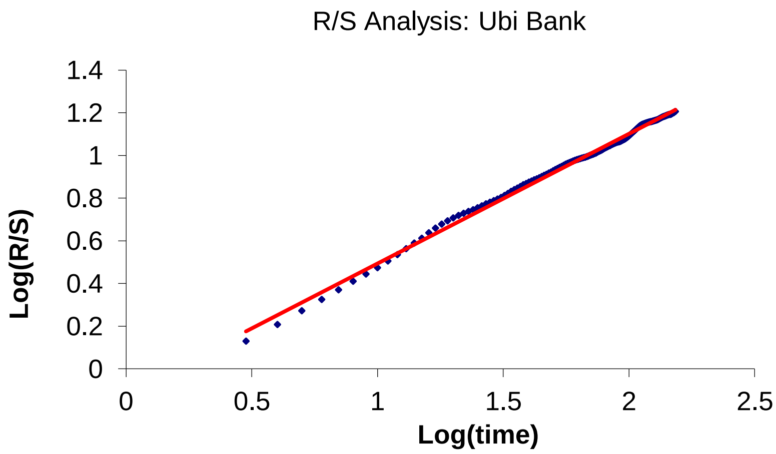


Figure 3-A. R/S analysis. Log-log plot of Ubi Bank with the tendency line (red).

A.2 Dispersional analysis

This method averages the different fractional Brownian motion, ξ_H over bids of width τ and calculates the variance of the averaged dataset.

Four steps are needed to apply this method:

1. Set the bin size to $\tau = 1$;
2. Calculate the standard deviation of N data points and record the point $(\tau, \tau\sigma_\tau)$;
3. Average neighbouring data points and store in the original dataset

$$\xi_H(i) \leftarrow \frac{1}{2}[\xi_H(2i-1) + \xi_H(2i)]$$

   and rescale N and τ appropriately

$$N \leftarrow \frac{N}{2}$$
$$\tau \leftarrow 2\tau;$$

4. Perform linear regression on the log-log plot

$$\log(\tau\sigma_\tau) = H\log\tau + C;$$

   the slope is the estimate of H.

According to Blok[76], this method performs significantly better than *Rescaled Range analysis*.

Consider the same information set used to apply R/S analysis.

Table 2-A shows the estimation of the Hurst parameter obtained through the previous procedure.

[76] Blok, J. Hendrick, *On the nature of the stock market: simulation and experiments*, PhD thesis, The University of British Columbia, November 2000.

	Hurst Parameter	Fractal Dimension	Description
FIAT	0.52	1.478888418	*Low Persistent*
FASTWEB	0.41	1.592144008	*Low Anti-persistent*
UBI BANK	0.28	1.721130723	*High Anti-persistent*

Table 2-A. *Dispersional analysis*. Estimates of Hurst parameter and fractal dimension.

The reader can observe carefully that according to this method, the results are different, making riskier the last two assets.

Moreover, Figure 1-A, Figure 2-A and Figure 3-A show the respective graphs.

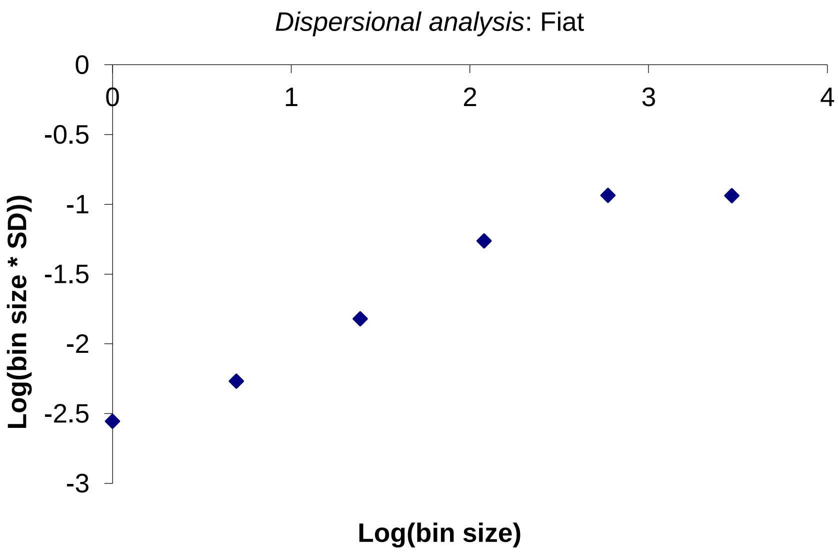


Figure 4-A. *Dispersional analysis*. Log-log plot of Fiat.

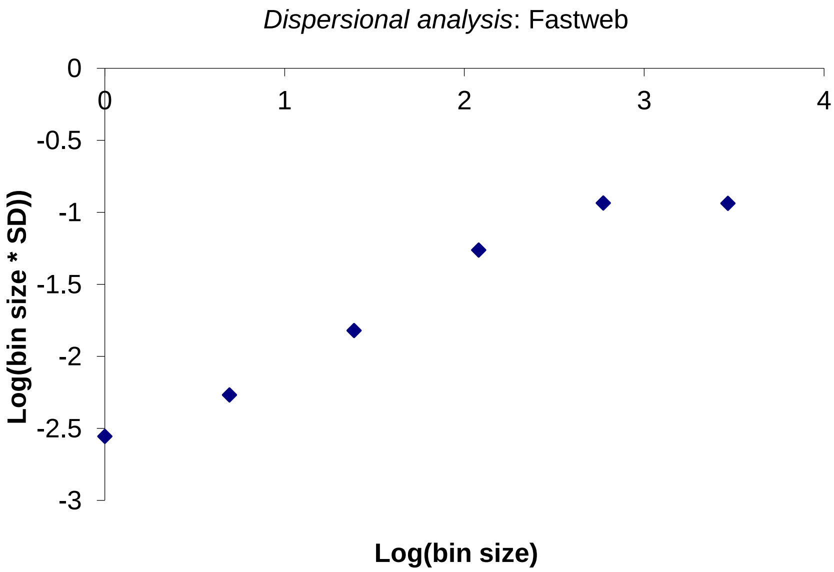


Figure 5-A. *Dispersional analysis*. Log-log plot of Fastweb.

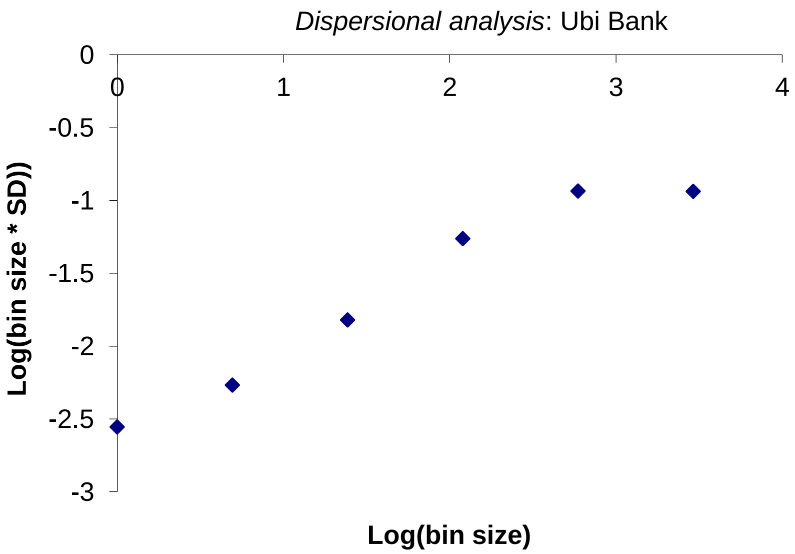


Figure 6-A. *Dispersional analysis*. Log-log plot of Ubi Bank.

References

ABRAMOWITZ MILTON and STEGUN IRENE, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards, 1964

ALAIN LE MÉHEUTÈ, *Fractal Geometries: Theory and Apllications*, CRC Press, Inc., Boca Raton, 1990

ALTINKAYA, A. MUSTAFA, DELIC, H., SANKUR, B., ANARIM, E., *Frequency Estimation of Sinusoidal Signals in Alpha-Stable Noise Using Subspace Techniques, Bogazici University*, Istanbul, Turkey, 2000

AMENC, N., LE SOURD, V., *Rating the Ratings, Edhec Risk and Asset Management Research Centre*, April 2005

APPLEBAUM, DAVID, *Lévy Processes- From Probability to Finance and Quantum Groups*, December 2004

ARAD, W.R., *The Implications of a Long-Tailed Distribution Structure To Selection and Capital Asset Pricing*, PhD thesis, Princeton University, January 1975

BALANKIN, A. S., OSWALDO, M.M., *Devil Staircase Like Behavior of the Range of Random Time Series Associated with the Tangled Nature of Evolution*, México D.F., 2005

BAXTER, MARTIN, *Lévy Simple Structural Models*, Nomura Fixed Income Quant Group, 22 December 2006

BARTH W., *Fraktale, Long Memory und Aktienkurse – eine statistische analyse fur den deutschen aktienmarkt*. Eul-Verlag, Bergisch-Gladbach, Koln, 1996

BELKACEM, L., LÉVY VÉHEL, WALTER, C., CAPM, *Risk and Portfolio Selection in "Stable" Markets*, INRIA, 1996

BELKACEM, LOTFI, *How to Select Optimal Portfolio in α-stable Markets*, INRIA, 1997

BENDER, C., *Integration with respect to Fractional Brownian Motion and Related Market Models*, University of Konstanz, Department of Mathematics and Statistics: PhD thesis, 2003

BERTOIN JEAN, *Lévy Processes*, Cambridge University Press, 1996

BLACK, F and SCHOLES, M., *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy 81 (3): 637–654, 1973

BLOK J. HENDRIK, *On the Nature of the Stock Market: Simulations and Experiments*, The University of British Columbia, November 2000

BLOK, J. HENDRICK, *On the Nature of the Stock Market: Simulation and Experiments*, PhD thesis, The University of British Columbia, November 2000

BOUCHAUD, J.P., SORNETTE D., POTTERS, M., *Option Pricing in the Presence of Extreme Fluctuations*, 1996

BOUCHAUD, J.P., SORNETTE, D., WALTER, C., AGUILAR, J.P., *Taming Large Events: Optimal Portfolio Theory For Strongly Fluctuating Assets*, World Scientific Publishing Company, Vol. 1, No. 1, 25-41, 1998

BREIMAN, LEO, *Probability*, Addison-Wesley Pub. Co., 1968

BRENNAN, WILL, Fractals In Finance and Risk, (power point presentation)

CAGLAR, M., *Simulation of Fracrtion Brownian Motion with Micropulses*, Adavances in Performance Analysis, 3:43-49, 2000

CALVET, L. and ADLAI, F., *Forecasting Multifractal Volatility*, Journal of Econometrics, 105, 27-58, 2001

CAUCHY, AUGUSTIN- LUOIS, *Cours d'Analyse de l'Ecole Royale Polytechnique*, Imprimerie Royale, 1821, reissued by Cambridge University Press, 2009

COEURJOLLY, Jean-Francois, *Simulation and Identification of the Fractional Brownian Motion: A Bibliographical and Comparative Study*, June 7, 2000

COOPERATIVE PHENOMENA GROUP, *Projects and Methods: Self-Affine Fractals*, Department of Physics, University of Oslo, 1998

DAI, W., HEYDE, C.C., *Itô's Formula With Respect to Fractional Brownian Motion and Its Application*, Journal of Applied Mathematics and Stochastic Analysis 9, Number 4, 439-448, 1996

DAVIS MARK H.A., A. ETHRIDGE, *Louis Bachelier's Theory of Speculation*, Imperial College, 2006

DEMIDOV EVGENY, *The Mandelbrot and Julia Sets Anatomy*, 2003

Economics Focus: *The Great Depression*, The Economist, 17 September 1998

EVERTSZ, CARL J.G., *Fractal Geometry of Financial Time Series*, Center for Complex System and Visualization, University of Bremen, Germany, pp. 609-616, 1995

FABBRO, G., *Analisi Frattale dei Mercati Finanziari*, Tesi di laurea, 2001

FAMA EUGENE F. and KENNETH R. FRENCH, *Common Risk Factors in the Returns on Stocks and Bonds*, Journal of Financial Economics 17, 1993, pp. 3–56

FAMA EUGENE F. and KENNETH R. FRENCH, *The Cross-Section of Expected Stock Returns,* Journal of Finance 47, 1992, pp. 427–66

FAMA, E.F., *The Behavior of Stock Market Prices*, Journal of business, 38, January 1965

FATHAUER ROBERT, *Art inspired by Mathematics*, 2007

FEDER, J., *Fractals*, Plenum Press, New York, 1988

FELLER, W., *An Introduction to Probability Theory and Its Applications*, Vol. 2, Wiley, New York, 1971

FELLER, W., *An Introduction to Probability Theory and its Application*, vol. II New York, 1950

FELLER, W., *An Introduction to Probability Theory and its Application*, vol. I New York, 1950

FRAME MICHAEL, MANDELBROT BENOIT B., NEGER NIAL, *Fractal Geometry*, 2003

GALBRAITH JAMES K. and TRAVIS HALE, *Income Distribution and the Information Technology Bubble*, University of Texas, Inequality project working paper

GAMROWSKI, B., RACHEV, S., *A Testable Version Of the Pareto-Stable CAPM*, Mathematical and computer modeling, 29, 61-81, 1999

GARY L. WISE and ERIC B. HALL, *Counterexamples in Probability and Real Analysis,* Oxford University Press, New York 1993

GLEICL JAMES, *Chaos: Making a New Science*, New York: Penguin Book, 1987

GREENBLATT, JOEL, *Il Piccolo Libro che Batte il Mercato Azionario*, CHW Edizioni, 2007

GRIMMET GEOFFREY, STIRZAKER DAVID, *Probability and Random Processes*, Oxford University Press, Third Edition, 2001

GRUBER, JASON, *Stock Market Data Analysis Using Rescaled Range (R/S) Analysis Method*, Computer Science Department Indiana, University of Pennsylvania, (power point presentation)

HALL, P., *On Some Simple Estimates of an Exponent of Regular Variation*, Journal of the Royal Statistical Society, 1982, Series B (Methodological) pp 37–42

HEITJANS P., KARGER J., *Diffusion in condensed matter: Methods, Materials, Models*, Birkhauser, 2005

HU, Y., ØKSENDAL, B., *Fractional White Noise Calculus and Applications to Finance*, Infin Dimens Anal Qu 6(1), p.1-32, 2003

HULL JOHN C., *Options, Futures and Other Derivatives*, Prentice Hall International, May 2006

KABASINSKAS, A., RACHEV, T.S., SAKALAUSKAS, L., SUN, W., BELOVAS, I., *Alpha-Stable Paradigm in Financial Markets*, June 5, 2008

KAHNEMAN D., A. TVERSKY, *Prospect Theory: An Analysis of Decision under Risk,* Econometrica 47 (2), 1979, pp. 263–291

KEYSERLING ARNOLD and R.C.L., *Change and Choice: A Compendium of Ancient and Modern Knowledge*, School of Wisdom, 1994

LEIPNIK, ROY B., *On Lognormal Random Variables: I - The Characteristic Function*, Journal of the Australian Mathematical Society Series B, vol. 32, 1991, pp 327–347

MAGNANI, OLIVIERI, ROSSI E TORRIGIANI, *Matematica Finanziaria* - ed. Monduzzi, II ed. 1998

MALKIEL BURTON G., *Efficient Market Hypothesis*, The New Palgrave: A dictionary of Economics, v. 2, pp. 120-23

MANDELBROT B.BENOIT, “Ch. 7: Harnessing the Peano Monster Curves", *The Fractal Geometry of Nature*, W. H. Freeman, 1982

MANDELBROT B.BENOIT, *Fractal Geometry: what is it, and what does it do?*, Yale University, 1989

MANDELBROT B.BENOIT, *Fractals and Chaos: The Mandelbrot Set and Beyond,* Springer, New York, 2004

MANDELBROT B.BENOIT, *Fractals and Scaling in Finance - Discontinuity, Concentration, Risk*, 1997

MANDELBROT B.BENOIT, *How long is the coast of Britain?*, Science, 1967

MANDELBROT B.BENOIT, HUDSON RICHARD L., *The (Mis)Behavior of Market. A Fractal View of Risk, Ruin and Reward*, 2004

MANDELBROT B.BENOIT, The *Fractal Geometry of Nature*, W.H. Freeman, 1983

MANDELBROT, B.B and VAN NESS, J.W., *Fractional Brownian Motion, Fractional Noises and Applications*, SIAM Rev 10(4), p 422-437, 1968

MANDELBROT, B.B., FISHER, A., CALVET, L., A Multifractal Model of Asset Returns, Yale University, October 1996

MANDELBROT, B.B., *Gli Oggetti Frattali. Forma, Caso e Dimensione*, Torino, Einaudi, 1987

MANDELBROT, B.B., *Multifractals and* $1/f$ *Noise: Wild Self-Affinity in Physics*, Springer, New York, 1999

MANDELBROT, B.B., *New Methods in Statistical Economics*, Journal of political economy, 56, 1963

MANDELBROT, B.B., *Survey of Multifractality in Finance*, Cowles Foundation for Research in Economics at Yale University, October 1999

MANDELBROT, B.B., *The Variation of Certain Speculative Prices*, The Journal of Business, Vol. 36, No. 4, pp. 394-419, October 1963
MANZO G., *Analisi Statistica del Rischio della Mondadori S.p.A.*, Tesi di laurea, 2007
MARINUCCI, D., ROBINSON, P.M., *Alternative forms of fractional Brownian motion,* Journal of Statistical Planning and Inference 80 111-122, 1999
MARKOWITZ HARRY, *Portfolio Selection: Efficient Diversification of Investment*, New York: John Wiley & Sons, 1959
MERTON, ROBERT C., *Theory of Rational Option Pricing*, Bell Journal of Economics and Management Science 4 (1): 141–183, 1973
MURA, ANTONIO, *Moto Browniano Frazionario (Dipendenza long-range, modelli FARIMA)*, Dipartimento di Fisica, INFN Bologna, 2005
NAKAO, HIROYA, Multi-scaling Properties of Truncated Lévy Flights, Physics Letters A 266, pp. 282-289, 28 February 2000
NEWMAN, M. E. J. and CLAUSET A. and SHALIZI, C. R., *Power-law distributions in empirical data*, 2007
NEWMAN, M. E. J., *Power laws, Pareto distributions and Zipf's law*, Contemporary Physics, 2005, pp 323–351.
NORROS, I., VALKEILA, E., VIRTAMO, J., *An Elementary approach to a Girsanov Formula and Other Analytical Result on Fractional Brownian Motion*, Bernoulli 5, p. 571-587, 1999
O'CONNOR, JOHN J. and ROBERTSON, EDMUND F., *Paul Pierre Lévy*, MacTutor History of Mathematics archive, May 2000
ORSINGHER ENZO, BEGHIN LUISA, *Probabilità e Modelli Aleatori*, Aracne Editrice, 2006
ORTOBELLI, S., RACHEV, S., BIGLOVA, A., STOYANOV, S., *Comparison Among Different Approaches for Risk Estimation in Portfolio Theory*, 2000
ORTOBELLI, S., RACHEV, S., SCHWARTZ, E., *The Problem of Optimal Asset Allocation With Stable Distributed Returns*, UC Los Angeles, Finance, Anderson Graduate School of Management, January 2000
PEARSON KARL, *The Problem of Random Walk*, Nature, 27 July 1905, p. 1865
PETERS, E. EDGAR, *Fractal market analysis: applying chaos theory to investment and economics*, 1994
PITMAN, W. JAMES, *Basic Properties of Brownian motion*, Lecture notes, 2003
PRESS, JAMES S., *Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference*, Robert E. Krieger Publishing Company, Malabar, Florida, 1982

PRESS, JAMES S., *Estimation in Univariate and Multivariate Stable Distribution*, Journal of the American Statistical Association, Vol. 67, No. 340, pp. 842-846, December 1972

PRESS, S.J., *Multivariate Stable Distributions*, Journal of multivariate analysis, 2:444-462, 1972

RACHEV, S., MITTNIK, S., DOGANOGLU, T., *Portfolio Selection in the Presence of Heacy-Tailed Asset Returns*, Chapter 5, Deutsche Forschungsgemeinschaft, 2002

RACHEV, S., *Probability Metrics And the Stability of Stochastic Models*, New York: Wiley, 1991

RACHEV, S., XIN, H. *Test For Association of Random Variables in the Domain of Attraction of Multivariate Stable Law*, Probability and Mathematical Statistics, Vol. 14, Fasc. 1, pp. 125-141, 1993

ROSS, M. SHELDON, *Calcolo delle Probabilità*, Seconda edizione, Apogeo Education, 2007

ROSS, WESTERFIELD, JORDAN, *Is Beta Dead?, Appendix 10A*, Corporate Finance Fundamentals, 8e, 2009

ROSTEK STEFAN, *Option Pricing in Fractional Brownian Markets*, Springer, April 2009

ROSTEK, S. and SCHÖBEL, R., *Risk Preference Based Option Pricing In a Fractional Brownian Market*, Eberhard Karls Universität Tübingen, May 2005

ROTHEMUND PAUL W.K., PAPADAKIS NICK and WINFREE ERICK, *Algorithmic Self-Assembly of DNA Sierpinski Triangles*, PLoS Biology, v. 2, issue 12, 2004

SAGAN, HANS, *Space-Filling Curves*, Springer-Verlag, 1994

SAMORODNITSKY GENNADY, TAQQU S. MURAD, *Stable Non-Gaussian Random Processes – Stochastic Models with Infinite Variance*, Chapman & Hall, New York, 1994

SAMUELSON, P.A., *Proof that properly anticipated prices fluctuate randomly*, Industrial Management Review, 1965

SAMUELSON, P.A., The Journal of Portfolio Management, 1974

SHARPE, WILLIAM F., *A Simplified Model for Portfolio Analysis*, Management Science, January 1963, pp. 277-93

SHARPE, WILLIAM F., *Capital Asset Prices – A Theory of Market Equilibrium Under Conditions of Risk*, Journal of Finance XIX, 1964, pp.425-42

SHIRYAYEV, A.N., *On Arbitrage and Replication For Fractal Models*, research report 20, MaPhySto, Department of Mathematical Sciences, University of Aarhus, Denmark, 1998

SOBEL ROBERT, *Panic on Wall Street: A Classic History of American's Financial Disaster – with a new exploration of the crisis of 1987*, E.P. Dutton, May 1988

SOLMAN, TALEB, MANDELBROT*, Report*, *Top Theorists Examine Rippling Economic Turbulence,* 21 October 2001, moderating by Paul Solman, guests N.N. Taleb and B.B. Mandelbrot

SOTTINEN, T., *Fractional Brownian Motion, Random Walks and Binary Market Models*, Financ Stochast 5(3), p.343-355, 2001

STIGLITZ JOSEPH, *The Ruin of Russia*, Comment, Education Guardian.co.uk

STRECKER, JAYMIE, *Fractional Brownian Motion Simulation: Observing Fractal Statistics in the Wild and Raising Them in Captivity*, Independent Study Thesis, The College of Wooster, 2004

SUEISHI, N. and NISHIYAMA, Y., *Estimation of Lévy Processes in Mathematical Finance: A Comparative Study*, 1980

SWAMEE, P.K., *Near Lognormal Distribution*, Journal of Hydrologic Engineering, ,2002, pp. 441-44

SWIFT JONATHAN, *On Poetry, a Rhapsody*, 1733, vv. 337-40

SWISHCHUK, ANATOLIY, *Lévy Processes-From Probability to Finance*, Mathematical and Computational Finance Laboratory, Department of Mathematics and Statistics, (power point presentation) February 3, 2005

TALEB, N.N., *The Black Swan*, 2007

TALEB, NASSIM N., *The Fourth Quadrant: A Map of the Limits of Statistics*, 2008

TAQQU, M.S., *Fractional Brownian Motion and Long-Range Dependence*, to appear in: P. Doukan, G. Oppenheim and M.S. Taqqu (Editors), Long-range Dependence: Theory and Applications (Birkäuser, Basel and Boston, 2002)

VITALE COSIMO, Introduzione alla statistica per applicazioni economiche, Vol I: *Statistica* descrittiva, Edizioni Scientifiche Italiane, 2002

VITALE COSIMO, Introduzione alla statistica per applicazioni economiche, Vol II: Probailità e *Statistica*, Edizioni Scientifiche Italiani, 2002

WANG, K.C., *Long-time correlation effects and biased anomalous diffusio*n, Phys. Rev.A 45 No 2, 833-837, 1992

WELLNER SHORACK, *Empirical Processes with Application to Statistics*, 1986

WEST, B.J. and GRIGOLINI, P., *Fractional differences, derivatives and fractal time series*, in: R. Hilfer (Editor), *Applications of Fractional Calculus in Physics (World Scientific, Singapore)*, pp. 171-201, 2000

WILLIAMS DAVID, *Probability with Martingales*, Cambridge University Press, 1991

ZELEN MARVIN and SEVERO C. NORMAN, *Probability Functions*, Chapter 26, 1964

ZIEMBA, W.T., VICKSON, R.G., *Stochastic Optimization Model in Finance*, Economic Theory and Mathematical Economics, Academic Press, New York, 1975

ZIPF, GEORGE K., *Human Behavior and the Principle of Least-Effort*, Addison-Wesley, 1949

Sitography

www. academic2.american.edu

www.berkeley.edu/users/pitman

www.brighthub.com

www.classes.yale.edu/fractals

www.composite.about.com

www.edge.org

www.financial-dictionary.com

www.finanze.net

www.formulavincente.com

www.ilsole24ore.it

www.long-memory.com

www.math.toronto.edu

www.mathfinance.cn

www.mathworld.wolfram.com

www.miorelli.net/frattali

www.performancetrading.it

www.saperinvestire.it

www.scholar.google.com

www.wikipedia.org