POLITECNICO DI MILANO

School of Industrial and Information Engineering

Master of Science in Management Engineering

MASTER OF SCIENCE THESIS

An investigation of financial markets: from random walk to multifractals

Academic year 2018/2019

Supervisor
Prof.ssa Anna Paola Florio

Francesco Lucchini
898214

Milano, 18 December 2019

TABLE OF CONTENTS

TABLE OF FIGURES

ABSTRACT (English version)

This thesis studies the financial markets, in particular the most important models applied in finance and assumptions on which they rest, with a critical look at the characters who created them. The historical starting point are the orthodox financial models, those that have as conceptual basis Louis Bachelier and the random walk. We review the models belonging to this theoretical trend, those of Fama (1970), Markowitz (1950), Sharpe (1964), Black and Scholes (1973). The hypotheses underlying the standard theory are analyzed, in particular the Normal distribution and the hypothesis of price independence. Mandelbrot fractal geometry is introduced as a recent tool for describing markets and the most reference model is presented (Calvet, Fisher and Mandelbrot 1997a), which is based on the concepts of long-tails and long-term dependence in the financial movements. The aim is to understand the models functioning and the underlying assumptions, as well as to provide a picture of the state of economic literature in this field. The investigation methodology follows a practical approach, not lost in excessive formalisms but uses real examples with graphs and figures, referring the more technical parts to the appendices. Finally, a possible agenda for the future is presented.

Keywords: Modern Financial Theory, Random Walk, Efficient Market Hypothesis, CAPM, Portfolio Theory, Options, Fractal Geometry, Long-Tails, Long-Term Dependence, Multifractal Model of Asset Returns.

ABSTRACT (versione italiana)

Questa tesi studia i mercati finanziari, in particolare i più importanti modelli applicati nella finanza e le assunzioni su cui poggiano, con uno sguardo critico ai personaggi che hanno contribuito a crearli. Il punto di partenza storico sono i modelli finanziari ortodossi, quelli che hanno come capostipite concettuale Louis Bachelier e il random walk. Si passano in rassegna i modelli appartenenti a questo filone teorico, quelli di Fama (1970), Markowitz (1950), Sharpe (1964), Black e Scholes (1973). Si analizzano le ipotesi alla base della teoria standard, in particolare la distribuzione Normale e l'ipotesi di indipendenza dei prezzi. Viene introdotta la geometria frattale di Mandelbrot come strumento recente di descrizione dei mercati e ne viene presentato il modello di maggior riferimento (Calvet, Fisher e Mandelbrot 1997a), che si basa sui concetti di code lunghe e dipendenza a lungo termine nei movimenti finanziari. Lo scopo è capire il funzionamento dei modelli e le ipotesi alla loro base, oltre che fornire una fotografia dello stato della letteratura economica in questo ambito. La metodologia di indagine segue un approccio pratico, non si perde in eccessivi formalismi ma utilizza esempi reali con grafici e figure, rimandando le parti più tecniche alle appendici. Infine, viene presentata una possibile agenda per il futuro.

Parole chiave: Teoria Finanziaria Moderna, Random Walk, Ipotesi del Mercato Efficiente, CAPM, Teoria del Portafoglio, Opzioni, Geometria Frattale, Code Lunghe, Dipendenza a Lungo Termine, Multifractal Model of Asset Returns.

INTRODUCTION

The systematic study of the financial markets began with Bachelier’s doctoral thesis *Theorié de la Speculation* in 1900. Before that date, markets were mostly characterized by investors that based their decisions on intuition and experience. This work starts from Bachelier’s framework and traces the history of the more important financial models until the Multifractal Model of Asset Returns by Mandelbrot (Calvet, Fisher and Mandelbrot 1997a), analyzing in particular the underlying assumptions and commenting on them. Since the finance world is extremely practical, in this investigation a concrete approach is adopted, with many links to real facts. The more technical parts, also in order not to overload the reading, are included in the final appendices.

Chapter 1 is a historical excursus of modern financial theory. It describes the most important standard models, each linked to their inventors. Louis Bachelier built the theoretical framework of the models in this chapter with the introduction of the random walk, the financial equivalent of Brownian motion, which assumes that prices are independent and follow a Gaussian distribution. Later, researchers started from this base to elaborate concepts such as the Efficient Market Hypothesis (Fama) and models such as the CAPM (Sharpe), the portfolio theory (Markowitz) and the formula to evaluate options (Black and Scholes). Everyone relies on and assumes true the idea of random walk. At the end of the chapter you can find a practical case of application of the Normal distribution to the values of a stock market index and a discussion of the results.

Due to the recent global financial crisis, some criticisms have been made of the historical building of finance and in particular of the assumptions on which it rests. In Chapter 2 the most criticized hypotheses are analyzed, that is the use of the random walk as a theoretical framework and the assumptions of investors rationality and prices continuity.

The most famous opponent of orthodox finance is Benoit Mandelbrot, the inventor of fractal geometry.

Chapter 3 is an introduction to the topic of fractals, explaining how they relate to financial markets and providing the theoretical mathematical tools needed to understand this recent type of geometry. There are many images of fractal shapes and fractal cartoons are introduced, a technique of graphic construction of fractal shapes very suitable for the financial world. Particular attention is given to the concept of fractal dimension, which is the most important part of this geometry.

Chapter 4 presents the fractal reference model for the description of financial markets, the Multifractal Model of Asset Returns. This model is based on three concepts. The first, long-tails, represents a new way of seeing markets, through the power laws and the mathematics of Levy's stable distributions. The second, long-term dependence, is a new concept within the financial world, where price changes have always been considered independent. The third, the trading time, is a different way of thinking to the normal clock time and is the key to integrate the first two. Within the chapter fractal cartoons are again used to present ideas also from a graphical point of view, including the final model which is also described by a mathematical perspective.

CHAPTER 1.

A HISTORICAL EXCURSUS OF MODERN FINANCIAL THEORY

1.1 Louis Bachelier

The first attempt to create a model capable of describing the financial markets was Bachelier's Random Walk, who was the first to introduce the concept of probability into the capitals market within his doctoral thesis Théorie de la spéculation (Bachelier 1900*)*.

The thesis of Bachelier, a PhD student of the University of Paris (whose professor was the well-known Henry Poincaré), was not considered at the very first start by his contemporaries, mainly because in France uncontrolled speculation did not enjoy a good reputation and also finance in general was a subject far away from the classic intellectual interests in vogue at the *Sorbonne*. It was 50 years later that some economists took up the concepts and, based on his ideas, built the modern financial theory. Paul Cootner, an economist at MIT even called Bachelier «so exceptional for his work that it can be said that the study of speculative prices had its moment of glory when it was conceived». The socioeconomic context in which Bachelier developed his own study was that of a lively trade of government bonds, due to the fact that in the period after the French Revolution, the new republican government had issued a billion francs of perpetual bonds[1] to compensate for the nobility, and these instruments had been hugely successful in terms of trading, so much so that, by 1900, the total value of domestic and foreign bonds in circulation was 70 billion francs[2] and that therefore a very liquid futures, premium contracts and other derivatives market had developed in parallel. Bachelier wanted to find formulas to determine the price of these instruments, but to do so it needed to know how the prices of the underlying instruments (bonds) varied. He started from the following insights, taken from the first two paragraphs of Bachelier's thesis:

> The influences which determine the movements of the Stock Exchange are innumerable. Events past, present or even anticipated, often showing no apparent connection with its fluctuations, yet have repercussions on its course. Beside fluctuations from, as it were, natural causes, artificial causes are also involved. The Stock Exchange acts upon itself and its current movement is a function not only of earlier fluctuations, but

[1] Without repayment of capital, but with permanent fixed coupons.

[2] It should be noted that the French Government's budget for its expenditure was, at the same time, 4 billion francs.

> also of the present market position. The determination of these fluctuations is subject to an infinite number of factors: it is therefore impossible to expect a mathematically exact forecast. Contradictory opinions in regard to these fluctuations are so divided that at the same instant buyers believe the market is rising and sellers that it is falling. Undoubtedly, the Theory of Probability will never be applicable to the movements of quoted prices and the dynamics of the Stock Exchange will never be an exact science. However, it is possible to study mathematically the static state of the market at a given instant, that is to say, to establish the probability law for the price fluctuations that the market admits at this instant. Indeed, while the market does not foresee fluctuations, it considers which of them are more or less probable, and this probability can be evaluated mathematically.

He developed his model thanks to an analogy that he noticed between the diffusion of heat through a substance and the trend of the value of bonds, with the hypothesis that, since it was impossible to evaluate individually all the effects of the single factors that impact on the valuation of the securities (as it is impossible to estimate in precise way how the single particles of matter spread), one could more simply consider the probability structure describing the overall system to obtain probabilistic predictions of future performances. According to Bachelier, the bond market was a balanced game (zero-sum game), in which financial assets prices make a series of random movements (random walk)[3] that do not depend on past movements: price changes therefore form a sequence of independent and equally distributed random variables.

To explain this one can compare the market trend to the launch of a coin (which is in fact a balanced game with two random outcomes), whose faces correspond to the two possible price directions, "positive variation" and "negative variation". This comparison has a strong implication: suppose Mr. Finance and Mr. Economics decide to play the classic coin toss, and that at every launch the loser must give the sum of 1 euro to the winner and that obviously the two outcomes have equal probability (the coin is not tricked, so the game is fair). The expected value (mathematical expectation) of the game for both will tend to 0 after a sufficiently large number of launches, and this because at each launch will be equiprobable the output of one or other face, regardless of past outcomes: the coin indeed has no memory. If the bond market is a balanced game with the two possible outcomes corresponding to the two faces, in the vision of Bachelier, it will behave like the

[3] A mathematical treatment of the random walk model is available in Appendix 1.

coin, and it follows that also the market has no memory, therefore any change in prices is random and independent from all the previous ones.

The fact that prices make a series of random movements allows to use the concept of random walk to describe its properties. As in the random walk we have to assume that we will find the drunk on average around the same point where we left him[4], in the absence of new information that could alter the equilibrium between supply and demand, the price will fluctuate on average around the starting point although there may be larger or smaller random fluctuations. Two perspectives can be distinguished: an ex-ante, in which there is no information and in which therefore it is not possible to predict the future trend of the market, whose direction towards the top or down will be again equiprobable (such as the coin) and an ex-post, in which you become aware of the information and you can infer on any cause-effect relationships that will result in price changes (but, again, it is not possible to state with certainty which direction will be, because in this intellectual framework prices move according to the random walk) generating probabilistic statements to try to make a profit.

Bachelier went further, noting that if you represented on a chart all the changes in the price of a bond in a sufficiently long period (at least one month), these are distributed according to a normal or gaussian distribution, in which there will be at the center of the bell the many small changes in prices and in the tails the few big changes. This discovery constitutes one of his most important insights, because it allows to have simple rules to calculate the probability of any variation[5].

---

[4] The similarity between the random walk and the drunken man's walk derives from an exchange of letters in the scientific journal «Nature», volume 72, July 27, 1905 and two following volumes. Karl Pearson asked if any reader could solve the following problem: «(…) a man starts from a certain point *O* and walks *l* meters in a straight line; then he turns any corner and walk for other *l* meters, again in a straight line. The man repeats the process *n* times. What is the probability of being at a distance between *r* and *δr* from the starting point *O*, after *n* movements?». The answer came from a scientist, Lord Rayleigh, to whom Pearson replied later, ironically: «Lord Rayleigh's solution teaches us that in an open space the place where it is most likely to find a drunken man barely able to stand is near the point from which he started!».

[5] A mathematical description of the normal or gaussian probability distribution is available in Appendix 1.

1.2 Eugene Fama

For many years, the work of Bachelier was not recognized and fell into oblivion. In 1956 his studies were cited in academia by a PhD student of Paul Samuelson at MIT, Richard J. Kruizenga, in his thesis *put and Call Options: A Theoretical and Market Analysis* (Kruizenga 1956).

From that moment on, thanks to their practicality, Bachelier's ideas began to circulate more and more insistently in the economic field, until they found a proper theoretical framework in the famous *Efficient Market Hypothesis (EMH)* (Fama 1970) and in the assumption of rationality of the investors of Eugene Fama, which constitute the real basis of the modern financial theory.

"Efficient market" means that securities prices always reflect all the information available at a given time, so at any time the price is the right one, or more correctly, the price summarizes the best current overall market hypothesis with regard to the likely yield of the security considering the information available. The most immediate consequence is that the market becomes impossible to beat and therefore the only way to get a profit (at least in the short term) is to anticipate its movements in a *risky way* (in a very banal way, if you believe that the FED will lower rates, you could invest in US government bonds) or to guess and be particularly lucky, as well explains the famous metaphor of the economist Burton Malkiel in his book *"Walking on Wall Street"* (Malkiel 1973): «A blindfolded monkey throwing darts at the WSJ pages could choose a portfolio of shares that is as good as a portfolio carefully chosen by experts»[6].

According to Fama there are three forms of market efficiency hypotheses:

- weak form of efficiency: if the prices reflect all the information contained in the historical series of the prices

[6] The same point of view seems to have been adopted, a year later, by Samuelson, who in *The Journal of Portfolio Management* (Samuelson 1974) states, with no less sarcasm: «(...) a certain respect for the evidence of the facts obliges me to incline to the hypothesis that the majority of the portfolio managers would have to abandon the business - give to the repair of the hydraulic systems, teach ancient Greek or contribute to the PNL as executives of the company.».

- semi-strong form of efficiency: if the prices reflect all the information contained in the historical series of the same prices, plus any other public available information
- Strong form of efficiency: if prices reflect all the information contained in the historical price series, any other public available information, as well as any private information

As noted, strong market efficiency implies semi-strong market efficiency, which in turn implies weak market efficiency. All 3 imply (even if at different levels), as mentioned above, that it cannot be expected to allocate a portfolio with a higher expected return than the market for the same level of risk.

Fama's studies led him to believe that markets were efficient at least in a weak form.

For a practical example of market efficiency, consider the following two cases (Mandelbrot 2004):

> (…) a skillful chartist believes he has identified a certain structure in the records of the old prices– in January, we say, prices always tend to rise. Can he be enriched by this information by buying in December and reselling in January? The answer is no. If the market is large and efficient, others will also see the trend, or at least the fact that he is exploiting it. Soon, when more operators anticipate recovery in January, more people will buy in December and after in November, to beat the trend to a recovery in December. At the end, the whole phenomenon will have spread over so many months that it will cease to be relevant; the trend will have faded, killed by the same fact of being discovered.

> (…) the CEO of a multinational corporation starts collecting his rights of option, knowing that the debt is a bomb that's about to go off. How long can he take advantage of this inside information? In an efficient market, not for long. The operators will understand that the captain is abandoning the ship and imagine that something bad is going to happen. So, they'll sell, too, and the stock will go down.

The prices are then adjusted in the market every time thanks to the achievement of a new equilibrium between buyers and sellers and the following variation will have the same probability to go in one direction or in the other (balanced game); the price will always correspond to its fair value.

Rational investor means that he knows that the market price is "right" and that he then prepares forecasts of future market movements by pursuing a rational strategy.

1.3 Harry Markowitz

In 1950 Harry M. Markowitz drew up, in his doctoral thesis at the University of Chicago, the *modern theory of the portfolio*[7] (Markowitz 1952) and took the first step in the practical application of Bachelier. The basic question was whether a systematic and simple way could be found to evaluate in which shares to invest to build a good portfolio. At that time John Burr Williams′ book *Theory of Investment Value* (Williams 1938) was in vogue. Its thesis was the following: to forecast a share value one must consider the forecast of dividends adjusted by inflation rate and other factors that can alter the value. Markowitz was convinced that this method was simple, but not functional to the real world because investors do not consider only the potential profit (otherwise everyone would buy the shares of the same company, those with higher expected return), but also the risk, trying to diversify investments.

Markowitz's thesis was then that to describe a share two only measures are enough: the return and the risk. The return is the expected average prices of selling (i.e. the price at which the investor forecasts to be able to sell his security at a certain time) and it can be forecasted with classic instruments such as the fundamental analysis[8], for example by estimating the corporate earnings growth. The risk is the standard deviation (or variance) and is predicted using the normal distribution properties, according to whom the returns of the share are (assumed to be) distributed.

---

[7] The modern portfolio theory is one of the three most innovative practical elements, together with the CAPM model and the Black-Scholes formula for the evaluation of options that will be examined in the follow-up, of the second half of the 20th century and orthodox financial theory in general.

[8] Fundamental analysis is one of the main approaches to analyzing listed assets; it seeks the fair value of a security and has a long-term perspective. It consists mainly of two types of analysis: a macroeconomic type trying to find links between macroeconomic variables and market prices, using econometrics, and a company-type analysis, specific to the company in question.

For example, if the expected return of company X for year t is 7% and if in year t-1 the shares had a volatility of less than 20% in 2 out of 3 cases, we are asserting that in year t there will be a return between -13% and 27% in 67% of the cases.

To build an efficient portfolio of shares, it is not enough to include investments with higher profitability and lower risk, but it is also necessary to assess the level of correlation between the different securities: some shares could be related and tend to move in the same direction due to analogies of business or industry strategy, for example, and it would be a problem if an investor had a highly correlated portfolio that loses value at the same time. A portfolio is efficient if it produces maximum profit with the minimum risk[9]. For every given level of risk it is possible to build an efficient portfolio (therefore with the minimum possible risk) with a certain expected profitability, and if you plot all the efficient portfolios to vary the level of risk, you get a growing curve that takes the name of *efficient frontier.* The resulting function is monotonous increasing, which implies that you can increase the expected return only by increasing the risk.

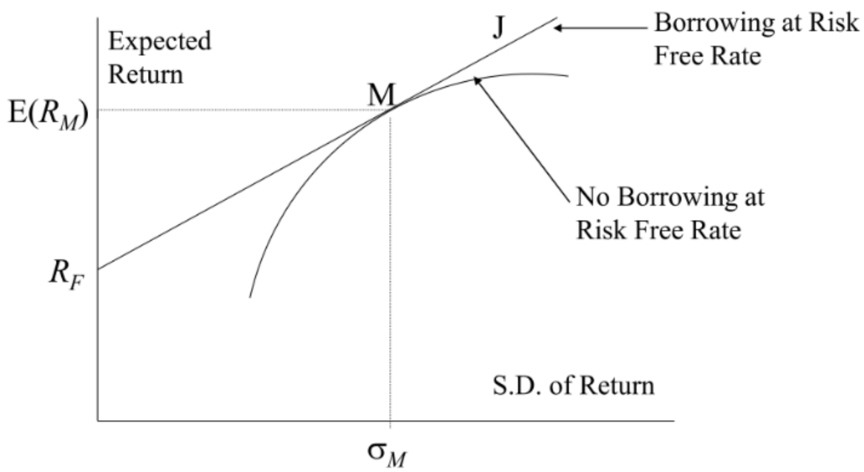

Figure 1. Efficient Frontier. From Financial Risk Management course of Politecnico di Milano.

[9] In the case of a portfolio consisting of only two securities *a* and *b*, the expected portfolio return shall be calculated as $\mu_P = \mu_a w_a + \mu_b w_b$, where μ_i are the single expected returns and w_i are the weights of the securities in the portfolio. Instead, portfolio variance is calculated as $\sigma_P^2 = w_a^2\sigma_a^2 + w_b^2\sigma_b^2 + 2w_a w_b \sigma_a \sigma_b \rho_{ab}$, where σ_i are the single standard deviations and ρ_{ab} is the correlation coefficient of securities *a* and *b*.

Portfolio theory thus provides a practical tool for the choice of investment capital allocation that a rational investor will make based on his degree of risk aversion: if Mr. Finance is very risk-averse he will choose a portfolio at the bottom left of the efficient frontier, which will ensure a low level of risk, but it is expected that he will lose some of the profits he could have made if he had been *braver* (or maybe just more *greedy*)[10].

Markowitz himself specified that his model may not always be adequate. He said indeed that it is not certain that the use of the bell-curve is justified because one cannot be certain of the normal distribution of all the shares and, secondly, building an efficient portfolio requires an accurate analysis of the expected future return of securities, otherwise the portfolio will also be inaccurate. A further problem for the years in which Markowitz developed his theory was to calculate the correlations between shares, because of the low and expensive computing power of technology at the time.

1.4 William Sharpe

This last problem was solved by William Sharpe in 1964, in an article that followed his doctoral thesis under the unofficial supervision of Markowitz himself (Sharpe 1964).

Sharpe noted that if everyone allocated their savings according to the results of portfolio theory, you would have a single efficient portfolio in which everyone would invest, which is called market portfolio, and it would then be the market itself to make all the necessary calculations at all times, providing from time to time the optimal combination[11]. Moreover, at that point the single shares should be assessed only on the basis of their comparison with the rest of the market and with the main "antagonist", that are the risk-free government securities (which are in fact the most

[10] Note that bonds with a (hypothetical) risk level of 0, the so-called risk-free securities, can also be considered in the construction of a portfolio, because if interest rates are sufficiently high, they could constitute a more efficient investment than certain shares.

[11] From here the concept of equity indexed fund born, such as the Exchange-traded Fund (ETF).

important investment alternative to the stock market). He introduced then the formula for pricing any type of asset, the *Capital Asset Pricing Model (CAPM)*[12]:

$$E(r) = r_f + \beta \left[E(r_m) - r_f\right]$$

which states that the expected return *r* of the security can be expected to be equal to the return of a risk-free asset (government risk-free bond) plus the product between the Beta Sharpe coefficient and a market risk premium (because the market is risky and requires additional return). The coefficient Beta represents the volatility of the share with respect to the market; the market risk premium is the return the market is expected to make more with respect to a risk-free title[13].

The model therefore shows that the more you decide to risk (that is the more you accept a high volatility of the share you want to buy) the more a high yield is expected. The formula then indicates what is the minimum return you would accept to take the risk of investing in that asset.

The diffusion of this model is also due to the fact that it allows to "save" all the calculations of the correlations previewed from the model of Markowitz: if a portfolio counted 50 different shares, employing the Markowitz's model we should in fact make 1325 calculations, while employing Sharpe's only 51[14].

Some criticisms have also been made to the Sharpe model, the most important of which concerns the use, again, of the normal distribution to describe the shares prices and consequently for the calculation of the Beta parameter. If price changes were not distributed according to a gaussian, but for example according to a power law, the results would be inaccurate and potentially harmful; in particular, in such a case, an investor would build a portfolio with shares with a risk and return that would be very different from the estimated one. The problem is not so much the

[12] Only financial assets will be considered here.

[13] The coefficient Beta of Sharpe is mathematically defined like the relationship between how much it varies the stock with the market, that is the Covariance, and the Variance of the stock.

[14] The calculation of all expected returns, variances and correlations (with simple combinatorial calculation) in the first case was considered; the calculations of the market return and of the parameters Beta of Sharpe for every share in the second case was considered.

lack of accuracy in the calculations but the fact that the investor would be convinced that they are right and would act accordingly by exposing himself to great risks without his knowledge.

1.5 Fisher Black and Myron Scholes

Modern stock options are a relatively recent tool in the financial world, having been inaugurated in 1973 by the Chicago Board of Trade that established the Chicago Board Options Exchange (CBOE)[15], but became immediately one of the most appreciated financial instruments: as noted by J. Finnerty in his article *The Chicago Board Options Exchange and Market Efficiency* (Finnerty 1978), in the first month of trading at the CBOE (May 1973) 34599 contracts were traded, while already in 1976 the average monthly trades became 1.5 million at the CBOE, to which are added 800000 at the American Stock Exchange[16] for a total increase of 6600%. Much of this success can certainly be attributed to the fact that they allow to recreate classical investment positions exposing a limited amount of capital and thus reducing the risk: one of the first contracts negotiated at the CBOE concerned options that they had as underlying Xerox[17] shares and offered the opportunity to purchase 100 shares at 160 dollars within 3 months (the spot price of the shares was 149 dollars) with a value of this option of $5.5 per share (Mandelbrot 2004).

There have been many attempts by leading scholars to find a correct formula to evaluate the price of options (and premium certificates in general), from Sprenkle (1961) to Samuelson (1965), but all their results depended on arbitrary parameters, mostly related to the estimation of the price of the underlying to maturity, and therefore could not be considered universally correct.

It was Fischer Black, with the help of Myron Scholes, who solved this problem: in 1973 their famous article was published (Black and Scholes 1973) explaining how the options could be priced without considering the final value of the security,

[15] Premium contracts similar to modern options existed instead already, as seen above, were one of the topics of Theorié de la Speculation of Bachelier.

[16] Known today as NYSE American.

[17] American manufacturer of printers and photocopiers.

but only the spot price of the share, its maturity and its volatility, that is considering the probability that the option would be exercised at maturity, which happens if the value of the underlying is within a precise range (the probability, therefore, that the option was in-the-money); so they created the famous Black-Scholes formula to evaluate a European call option[18]:

$$c = S_0 N(d_1) - Xe^{-rT} N(d_2)$$

The debut of the formula in the real world took place shortly after, by his own inventors, but did not guarantee positive results, as reported by Black in *How we came up with the option formula* (Black 1989):

> Scholes, Merton and I, along with others, rushed to buy a bunch of these certificates [of the National General]. For once, we felt like we had done the right thing. Then a company called American Financial announced a bid for the National General […]. This had the effect of significantly reducing the value of certificates.

As highlighted in the original 1973 article, the model is only valid under certain “ideal conditions”:

> In deriving our formula for the value of an option in terms of the price of the stock, we will assume "ideal conditions" in the market for the stock and for the option:
>
> a) The short-term interest rate is known and is constant through time.
> b) The stock price follows a random walk in continuous time with a variance rate proportional to the square of the stock price. Thus, the distribution of possible stock prices at the end of any finite interval is lognormal. The variance rate of the return on the stock is constant.
> c) The stock pays no dividends or other distributions.
> d) The option is "European," that is, it can only be exercised at maturity.
> e) There are no transaction costs in buying or selling the stock or the option.
> f) It is possible to Borrow any fraction of the price of a security to buy it or to hold it, at the short-term interest rate.
> g) There are no penalties to short selling. A seller who does not own a security will simply accept the price of the security from a buyer, and will agree to settle with the buyer on some future date by paying him an amount equal to the price of the security on that date.

[18] For a more detailed discussion, see Appendix 1.

> Under these assumptions, the value of the option will depend only on the price of the stock and time and on variables that are taken to be known constants.

As can be seen, even the Black-Scholes formula assumes that prices move according to a random walk and that their distribution is normal[19].

Despite this, the formula spread worldwide, mainly due to its ease of use and its ability to give an economic value to the risk.

All the instruments of orthodox finance analyzed so far are based on the initial work of a bulletin board of 1900 and, as in any building, an error in the foundations would make the whole structure unstable.

1.6 Example of practical application of Normal distribution

Let us now proceed with a practical example of using Normal distribution to describe the prices of financial securities. We take a sample of historical values with weekly time frame of the index S&P 500 from Monday 05/01/1970 until Friday 28/12/2018[20] (weekly data have been taken because daily data are not provided) and, considering the weekly closing values, we carry out a natural logarithmic transformation (in order to make data more comparable and to eliminate the positive asymmetry due to the long time scale with its effects on prices, e.g. inflation) of these values.

We proceed by calculating the differences between the adjacent logarithmic values, so as to obtain the changes in value (logarithmic) from one week to the next.

Let us now assume that the price changes are, overall, a population distributed in a normal way: then we use as estimates of the population the sample average to estimate the population average, and sample variance to estimate population variance. We get the following values:

[19] The article says that the price distribution is assumed log-normal. The log-normal distribution is a probability distribution of a random variable X whose log X logarithm follows a normal distribution. In this case, the random variable X is the strike price, that is the price of the action at maturity, and the logarithm of X appears within the calculation of d_1.

[20] https://it.investing.com/indices/us-spx-500-historical-data [22/09/2019].

$\bar{x}$ = -0,0013 , $\hat{\sigma^2}$ = 0,000504 , $\hat{\sigma}$ = 0,022453

At this point we choose a particular historical weekly variation, such as that of the week that led to the black Monday of 1987 (that from 12 to 18 October 1987), which is 0.130071 (also in logarithmic values).

We calculate the number of standard deviations for which it differs from the mean value (from the center of the bell) calculating the value z-score, standardizing therefore the value of the variation:

$$z = \frac{x - \mu}{\sigma}$$

and we get z=5.85.

The weekly change considered differs from the average value of 5,85 standard deviations. Knowing that the probability that a value differs by at least 5 standard deviations is about 1 out of 3,5 million, theoretically it follows that if the prices of the securities had normal distribution, a catastrophic event such as that of 1987 should have happened less than once every 3.5 million weeks, less than a *week* in 67308 *years.* At a practical level, however, an event at the ends of the tails similar to the one just analyzed has also occur in 1929, 1998 and, more recently, 2007, which brings the counting to four times in 78 years.

Remember also that the time frame used was weekly and this could have damped some variations: in case of daily time frame in fact, even worse results could have achieved.

This example demonstrates the inability of normal distribution to accurately approximate extreme events at the ends of the tails, while it is more accurate in periods of normal (low) volatility, when the values are in the center of the bell.

In this first chapter we analyzed the general picture of orthodox financial theory according to the discoveries of its main characters, both from a theoretical point of view, with Bachelier and Fama, and from a practical point of view, with

Markowitz, Sharpe and the Black and Scholes couple. The second chapter will analyze the underlying assumptions of this theory.

CHAPTER 2.

CRITICISMS TO MODERN FINANCIAL THEORY

Over the years, there have been many criticisms of standard financial theory. We have already examined empirically how the assumption of normal distribution of prices is far from reality.

The second chapter will go into more detail, analyzing what the “new” finance with its characters considers the wrong assumptions underlying the standard theory set out in the previous part.

2.1 Prices follow a Brownian motion (random walk)

The most important assumption introduced by Bachelier is certainly that of the random walk, which implies in particular the normal distribution of prices and their independence.

Let us consider the following graphical example, referred this time to the Dow Jones index[21]:

Figure 2. Dow Jones 1916-2003 daily logarithmic values. From Mandelbrot (2004).

[21] Share index that includes the 30 leading companies listed in NYSE.

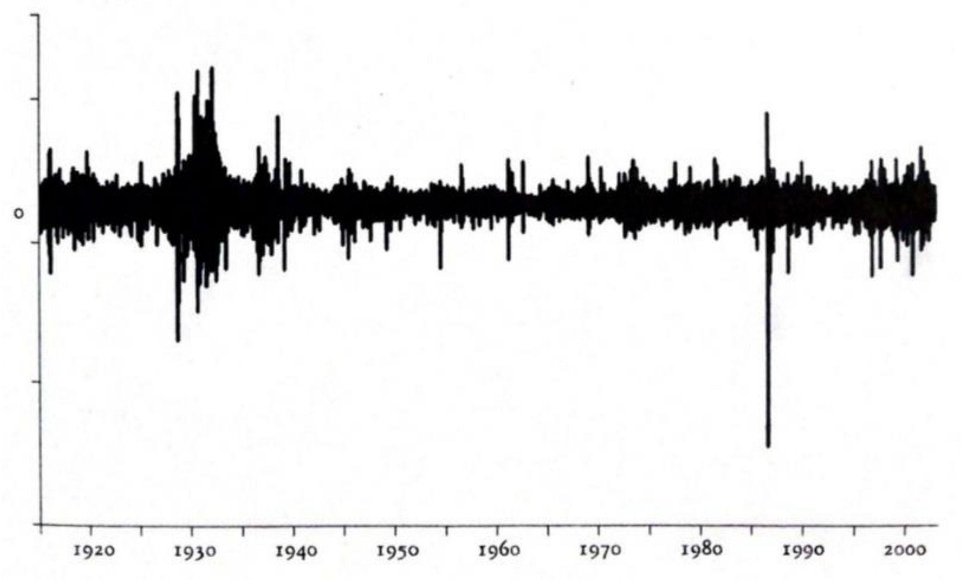

Figure 3. Dow Jones 1916-2003 daily logarithmic variations. From Mandelbrot (2004).

The first graph represents the daily values in logarithmic scale of the Dow Jones from 1916 to 2003, while the second represents the daily variations, always in logarithmic scale. Compare these two graphs with the following two, which represent for the same time period a computer-built example of what should have been the price recordings (in logarithmic scale, in order to compare very distant time periods) if they had been normally distributed:

Figure 4. Computer-built diagram of normally distributed prices. From Mandelbrot (2004).

Figure 5. Computer-built diagram of normally distributed variations. From Mandelbrot (2004).

It is immediately noticeable that, although the two daily price graphs may suggest a certain similarity, the two daily change graphs immediately refute the possibility that these prices will be distributed in a Gaussian way, as the volatility of the real index is much more pronounced than it should be to be able to use that distribution. To facilitate the vision, consider these other two graphs, representing the same previous concepts but directly expressing price changes in number of standard deviations rather than in logarithmic scale:

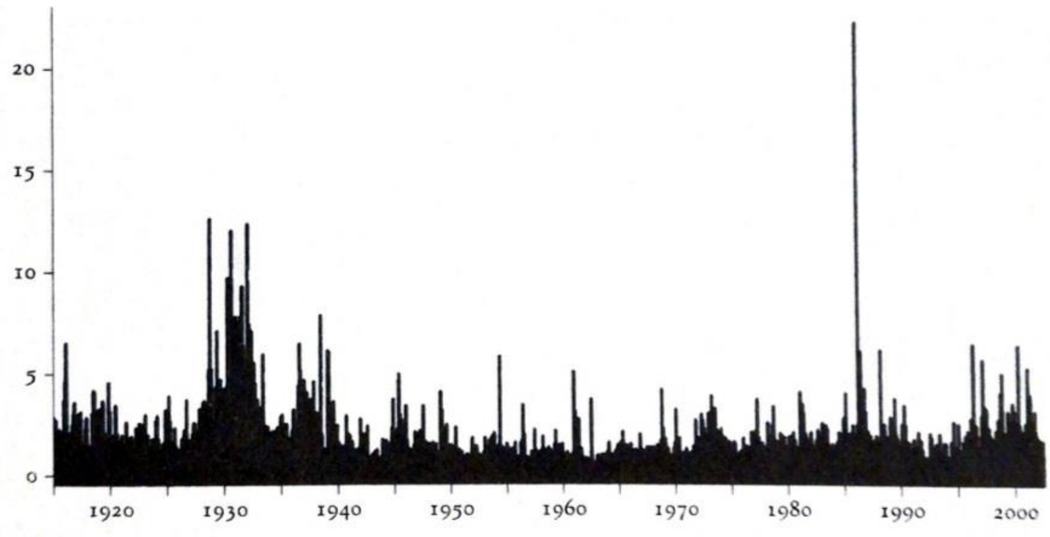

Figure 6. Price variations in number of standard deviations (not Normal). From Mandelbrot (2004).

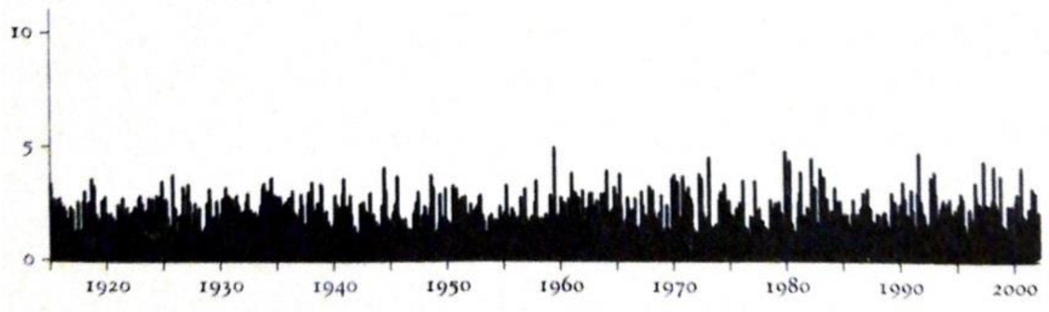

Figure 7. Price variations in number of standard deviations (Normal). From Mandelbrot (2004).

Variations in standard deviations of the real Dow Jones do not agree with the normal distribution, which considers, as seen in its treatment in Appendix 1, observation of values beyond the three standard deviations is almost impossible (as clearly shown in Figure 7 which is specially constructed on a normal distribution). The real values say that not only is it likely to exceed the 3 standard deviations, but that in times of severe market turbulence (the crisis of 1929 and 1987) it is easy to reach and exceed the 10 standard deviations, which would be statistically impossible to obtain if the assumption of normal price distribution were correct. As a more rigorous test it is also possible to consider the statistical measure of kurtosis[22] (for normal distribution equal to 3), as done by Wim Schoutens (2003) with the S&P 500 index: the mathematician noted that the daily changes in the index between 1970 and 2001 had a kurtosis of 43.36. The values of S&P 500 in the years considered therefore have a kurtosis well beyond the value of the normal, and even if the values recorded during the collapse of October 1987 are excluded, kurtosis remains elevated, equal to 7,17. The observation of these facts thus leads to the rejection of the assumption of normal price distribution.

Regarding the assumption of price independence, which is also necessary in order to study the markets in accordance with the random walk model, it is possible to cite the study by Ferson and Harvey (1991), which analyzed the stock exchanges of 16 of the world's largest economies and found that, if an index moved in a certain direction in a given month, it became slightly more likely that it would continue the same trend in the following month; the academicians therefore found evidence of a certain form of short-term dependence. Also in the medium term, a study by Fama and French (1988) identified a dependence of stock prices: according to them, approximately 10% of the yield of a given asset over a period of eight years could be attributed to their performance over the previous eight years. The analysis of the real facts through the discussed researches, once again, does not allow to assume with full confidence the independence of prices and consequently, considering also

[22] The kurtosis is one of the standard measures to describe the shape of the curve of a distribution, and in particular the degree of height of the curve. A normal curve has a kurtosis equal to 3, while greater kurtosis describe higher curves in the center and with thicker queues.

the examples regarding the not normal distribution, prices do not seem to move in accordance with the random walk[23].

2.2 The agents are rational

The second prerequisite for the importance of standard theory is undoubtedly the assumption of rationality of investors: when a person has a certain amount of information, he exploits that information to take the rational decision to bring as much wealth as possible. The definition before is true is we assume that the purpose of an investor is profit, and that is not emotionally conditioned, that is in its function of utility does not appear for example the utility of other economic entities.

To investigate the way in which individuals make decisions, a new way of approaching the study of economic agents - behavioral economics - has developed in the last 50 years. This approach has its roots in Kahneman and Tversky's 1979 work, *Prospect theory: Decision Making Under Risk.* This discipline assumes that agents do not always take their decisions in a rational way, or at least not only based on maximizing their utility function and acting in a selfish way. Empirical simulations have been developed with the objective of dividing a monetary sum between two individuals to investigate such anomalies with respect to the classical hypothesis, such as the *ultimatum game* and the *dictator game* (Camerer and Thaler, 1995):

The *ultimatum game* has two participants, the *Proposer* and the *Responder;* the first is given an amount of money and is asked to donate a part to the second, who can only decide whether to accept or refuse. If the Responder accepts the offer, they both pocket the parts decided by the Proposer, if instead he refuses both will receive nothing. If the rational individual hypothesis were true, the Proposer should choose to donate the smallest unit of money possible, and the Responder should accept any positive sum: in this way both would maximize their own utility function as

[23] It should be noted that no further evidence is needed to support this claim as, in addition to the fact that Dow Jones represents the 30 most important *blue chips* listed in NYSE and can therefore be considered a good indicator of the overall market trend, no amount of positive examples can prove with certainty the correctness of a theory, but only one opposite example shows the incorrectness.

provided by the rationality hypothesis. The empirical results showed instead that Proposer's average offer was between 30 and 40%, reaching even, in many cases, a 50-50 offer; offers lower than 20% were generally rejected. The previous results were obtained either with a small sum of 10$, or with a larger sum of 100$.

The *dictator game* is instead a variant of the previous game, developed for the first time by Kahneman and Tversky (1986) and it has the objective of verifying to what extent the generous offers obtained in the *ultimatum game* are due to the generosity of the Proposer and how much instead to the fact that the Proposer himself was afraid of being refused a too low offer, and therefore fail to pocket anything. In this game, there are still two subjects, this time called the *Allocator* and the *Recipient*; the former receives a sum of money and can decide, again, to what extent dividing it with the latter, which, on the other hand, has no decision-making power and must necessarily accept the offer that has been made to him. The simulation results showed that, in this game variant, the Allocator offers were lower than in the previous game, but still largely positive. This leads to the conclusion that the generous offers in the first game were caused by both factors mentioned above: agents may be generous to others, or they may be afraid that their offer will be rejected and not receiving anything.

In both cases, therefore, it follows that not all observed subjects can be considered rational individuals, because either act not maximizing their own utility function (when they are generous to others) or incorporate into their own utility function also the utility function of the other individual (when they are afraid of being rejected the offer).

2.3 Agents have rational expectations

Another important assumption of Orthodox theory is that of rational expectations: individuals use the information available to them efficiently, without making systematic errors in the formation of expectations regarding economic variables. An agent may make errors of assessment, but the community as a whole cannot make the wrong predictions and therefore will have correct expectations. From this it follows that the individuals behave *collectively* in the same way, that is to say with equal information they will have homogenous expectations with the same temporal horizons of reference, and it is therefore not possible the formation of financial bubbles.

Marianna Grimaldi and Paul De Grauwe developed a model (Grimaldi and De Grauwe, 2003) with the aim of understanding why bubbles occur within the currency market. They developed their model by releasing the hypothesis of rational expectations and introducing a population of agents divided into two distinct classes which, while having access to the same information, use different forecasting models and thus can generate different expectations: *fundamentalists,* who make predictions based on the fundamental analysis, and *chartists*, who instead use technical analysis trying to extract knowledge from the past trend. The fundamental exchange rate is set to 0. Different shocks are then simulated at the initial equilibrium condition of the exchange rate (that is, noise is created in the available information) which generate different solutions of (new) equilibrium[24] that can be summarized in two types. The first where, due to slight shocks to the initial conditions, the equilibrium is reached at the fundamental exchange rate (Figure 8, horizontal section) and given the name of fundamental solution. The second where, due to strong shocks to the initial conditions, the equilibrium diverges from the fundamental one (Figure 8, oblique sections) and to which the name of bubble attractor is given.

[24] The simulations have been made by the authors for one hundred thousand time periods.

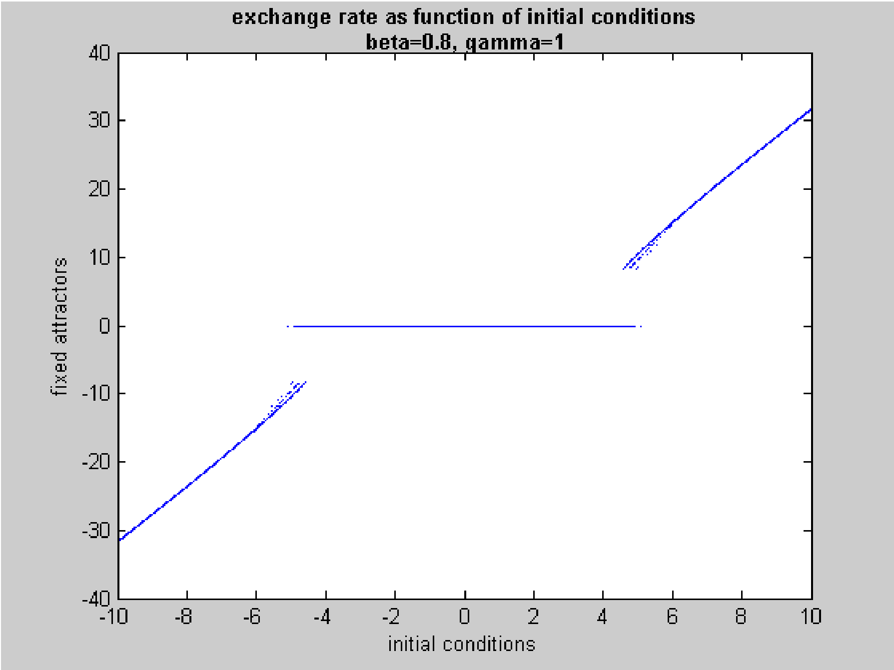

Figure 8. Exchange rate as function of initial conditions. From Grimaldi and De Grauwe (2003).

The greater the initial shock, the further the new equilibrium will be from its fundamental value. The different nature of these two types of balance can also be seen in relation to the percentage of chartists present in the market. The weight of the chartists according to the initial conditions is shown in Figure 9 below

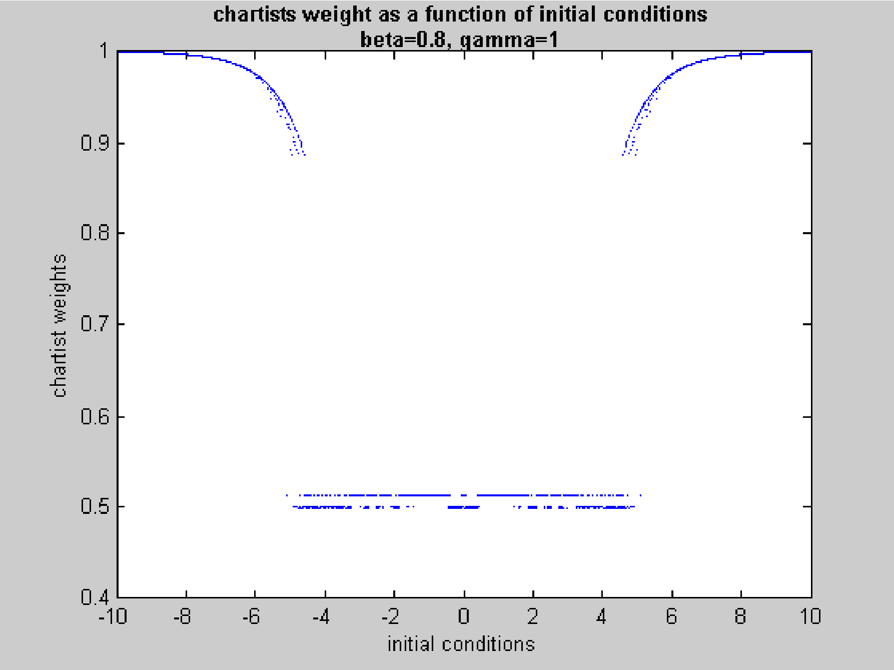

Figure 9. Chartists weight as a function of initial conditions. From Grimaldi and De Grauwe (2003).

For small initial shocks, the weight of chartists stands slightly above 50%, that is when the value of the exchange rate stands at around its fundamental value, the fundamentalists and chartists are equally present in the market. For great initial shocks instead, the weight of the chartists reaches 100% and the fundamentalists leave the market; so, when the chartists take control of a sufficiently large part of the market, the exchange rate value converges to the bubble attractor. The bubble attractor is therefore an equilibrium value that is reached when the weight of the fundamentalists in the market becomes sufficiently irrelevant, or no longer able to trigger a reversion of the rate towards its fundamental value. In this case, since the exchange rate value differs from its fundamental value, a larger or smaller magnitude of bubble can occur (the magnitude depends on the initial shock).

From the model one concludes therefore that the interactions of only two *different* classes of economic agents can generate a system not more linear, but chaotic, in which unexpected events occur (the bubbles) not justified by the fundamental values underlying the instrument. If in the real world there are at least two different classes of individuals, guided therefore by different economic dynamics of creation of expectations, the hypothesis of rational expectations must be rejected according to Grimaldi and De Grauwe.

2.4 Price change is continuous

The last hypothesis of the standard theory which will be examined in this discussion is that of the continuous price change, or the fact that stock prices or exchange rates do not rise and fall by several points at once but move smoothly from one value to the next. The usefulness of price continuity lies in the fact that it thus becomes possible to use mathematical tools such as continuous functions and differential equations to describe stocks quotations.

Mandelbrot (2004) rejected this assumption by pointing out that, for example, in the currency market, brokers often quote rounded prices, avoiding intermediate values, and this causes that in 80% of cases the quotations end 0 or 5. The same mathematician also argued that almost every day at NYSE there is an imbalance of orders for some title, that is a mismatch between supply and demand. On 8 January 2004, Reuters' press service reported, for example, eight misalignments; in this case, some important news that came to the market had caused an imbalance between supply and demand because the purchase and sales orders no longer collided, as a result of which operators had to increase or reduce the quotes up to a point of equilibrium.

In this second chapter of the thesis, the most important economic and mathematical assumptions underlying modern financial theory were examined and the main refutations in the literature were reported. The discussion will now continue by analyzing another piece of literature that helps describe financial markets, this is more recent and not yet fully covered: Mandelbrot’s fractal geometry. In particular, the third chapter will deal with a short mathematical description of this type of geometry, while in the fourth and last part these instruments will be applied to financial markets.

CHAPTER 3.

THE FRACTAL GEOMETRY

3.1 Why the fractals?

The term fractal and multifractal geometry was coined by its own inventor, Benoit Mandelbrot, in 1975, to name the new mathematical instrument capable of describing what is irregular and rough. Fractal, which derives from the Latin *fractus*, a past participle of *frangere* which means *to break*, indicates a model or a shape whose parts echo the whole: by extending a part, one finds a similarity, more or less accentuated, with the original figure.

Since its discovery, this geometry has been used in a wide range of fields, from natural ones such as fluid dynamics, hydrology and meteorology to those created by man such as the internet transmission of digital images, the measurement of fractures in metals and the analysis of brainwave recorded in EEGs. The possibility of using this mathematics to describe the trend of the financial markets, was conceived by Mandelbrot (2004) thinking of a sector that might initially seem very different, that of the blowing of the wind.

The wind is a classic example of the form of fluid current called *turbulence*, that is a sudden change in the air flow that can be caused by several factors, such as the upward currents or vortices. The following two figures show, respectively, a diagram of the variation of wind speed while turbulent currents with bursts are generated and a diagram of the course of the volatility of the stock market:

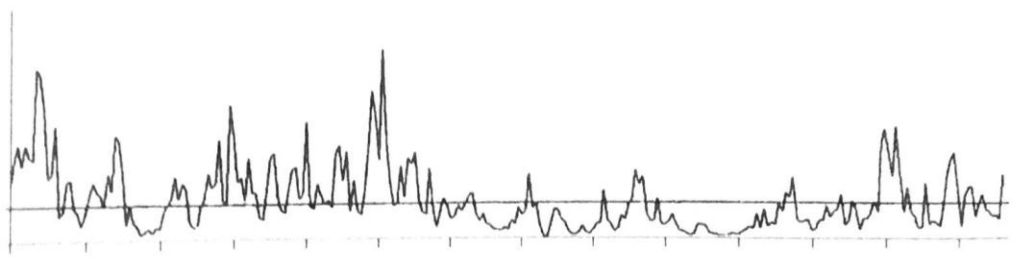

Figure 10. Variation of wind speed. From Mandelbrot (1972).

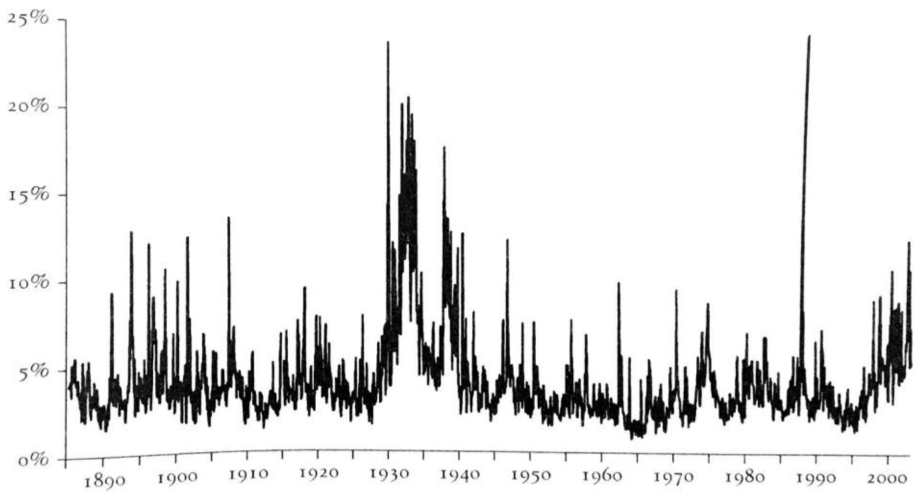

Figure 11. Variations of the stock market. From http://schwert.simon.rochester.edu/volatility.htm.

A certain similarity can be observed between the two graphs, in particular as regards the volatility of changes and the concentration of important events over time. Consider, for an example of turbulence in the financial markets, what happened on 27 and 28 October 1997, through the direct voice of the Division of Market Regulation of the SEC, the US Stock Exchange Supervisory Committee, in its report *Trading Analysis of October 27 and 28, 1997* of 1998[25]:

> On October 27 and 28, 1997, the nation's securities markets Fell by a record absolute amount on then-record trading volume. On Monday, October 27, the Dow Jones Industrial Average ("DJIA") declined 554.26 points (7.18%) to close at 7161.15. This represented the tenth largest percentage decline in the index since 1915. October 27 was also the first time that the cross-market trading halt circuit breaker procedures had been used since their adoption in 1988. At 2:36 p.m., the DJIA had declined 350 points, thereby triggering a 30-minute halt on the stock, options, and index futures markets. After trading resumed at 3:06 p.m., prices fell rapidly to reach the 550-point circuit breaker level at 3:30 p.m., thereby ending the trading session 30 minutes prior to the normal stock market close.
>
> […] While the U.S. securities and futures markets initially opened sharply lower on the morning of October 28, prices quickly turned around to close dramatically higher on then-record volume trading. For example, while the DJIA declined 187.86 points (2.62%) by 10:06 a.m., share prices thereafter began to rise sharply. By 10:20 a.m., the DJIA had recovered to within 25 points of the previous trading session's closing value, and by 10:25 a.m. the DJIA was 50 points above the previous close. By 10:34 a.m., the DJIA had risen 137.27 points (1.92%) from Monday's close, and prices continued to rise in choppy trading for the remainder of the day. As a result, the DJIA ended the day up 337.17 points (4.71%) at 7,498.32.

[25] https://www.sec.gov/news/studies/tradrep.htm.

The air turbulence and turbulence in the financial markets are clearly different as regards the underlying causes of the phenomena, but according to Mandelbrot can be described from the same mathematical point of view, as well as two different events such as elliptical orbits and the value of options according to Black and Scholes are both described using normal distribution. Fractals can thus become a tool for analyzing financial markets.

3.2 Notes on fractal geometry

Nature is not *regular:* shapes such as circles and squares or concepts such as lines, planes and spheres are the result of a human elaboration, on all the treatise Euclid's *Elements*, which gave body to the Euclidean geometry studied to this day, but these are almost never present naturally in the world.

Nature is rather *irregular:* «the clouds are not spheres, the mountains are not cones, the coasts are not circles, a bark is not regular and the lightning does not travel in a straight line» to quote Mandelbrot's manifesto book of 1982 *The Fractal Geometry of Nature*.

A fractal is a set *F* that must enjoy properties *similar* to the following:

- Self-similarity: *F* is the union of a number of parts which, enlarged by a certain factor, reproduce all *F*; in other words *F* is union of copies of itself at different scales. If the ratio according to which the parts echo the whole is the same in all directions, we are talking about *self-similar* fractals. If the ratio changes along different directions, for example the figure is reduced more in length than in width, we are talking about *self-affine* fractals. If in several points the enlargement or the reduction follow different factors of scale, we speak of *multifractals*.
- Fine structure: F reveals details at each enlargement.
- Irregularity: cannot be described as a place of points that meet simple geometric or analytical conditions (the function describing them is recursive)

- Hausdorff dimension > Topological dimension: the characteristic of fractals, from which their name derives, is that, although they can be represented in a conventional space with two or three dimensions, their size is not integer (the concept of size will be further developed in a dedicated paragraph).

Fractal geometry deals in detail with roughness[26], going to identify regularity in what is irregular, that is finding the configurations that are repeated in order to analyze, quantify and manipulate them. Configuration can take many forms: a concrete repeating module or an abstract shape, for example the probability that a certain grid box will turn black or white. The frame can be reduced or enlarged, can undergo compressions or twists, or both, with equal or different scale factors. The way these iterations take place can be deterministic or random.

3.3 Construction of simple fractals: Brown-Bachelier cartoons

The simplest fractals can be built with three things: it starts from a classical object of Euclidean geometry, which is called *initiator;* then a shape, called *generator*, is necessary, which generally can be another geometric figure; finally, the fractal construction process is needed, called *recursion rule*.

In the financial field, the simplest fractal that can be built is Brown-Bachelier cartoon: the result is a price chart in accordance with the characteristics of Brownian motion and therefore of Bachelier's random walk, but built with fractal geometry.

Below is represented step by step the method of construction: you draw a box of height and width equal to 1 (but representing it with dilated scale of width, to facilitate the visibility of the figures) and inside it a line is drawn starting from the bottom left corner and reaching the top right corner, that is from the position (0,0) to the position (1,1); the latter represents the trend line of the chart and it is the *initiator*, which in this case ensures a rising price trend. The next step is to draw a broken line, the *generator*, overlapping it on the initial straight segment,

[26] The use of the word "roughness" instead of "irregularity" derives from the speech of Benoit Mandelbrot at a TED conference in 2014.

manipulating its size so that the ends of the broken line match the ends of the trend line; in this case the *generator* is a broken 3-segment and up-down-up path type. The points where it changes direction are crucial factors for the result, as well as the number of segments that make it up. The *recursion rule* is to replace, step by step, the generator (that is the broken one) to every segment that is in the diagram, having the only foresight to always match the ends (in descending intervals to make it possible you need to turn the broken line). The process is repeated, and after a sufficient number of iterations an irregular and rough curve begins to be obtained, which begins to look like a price diagram, as the following figures show:

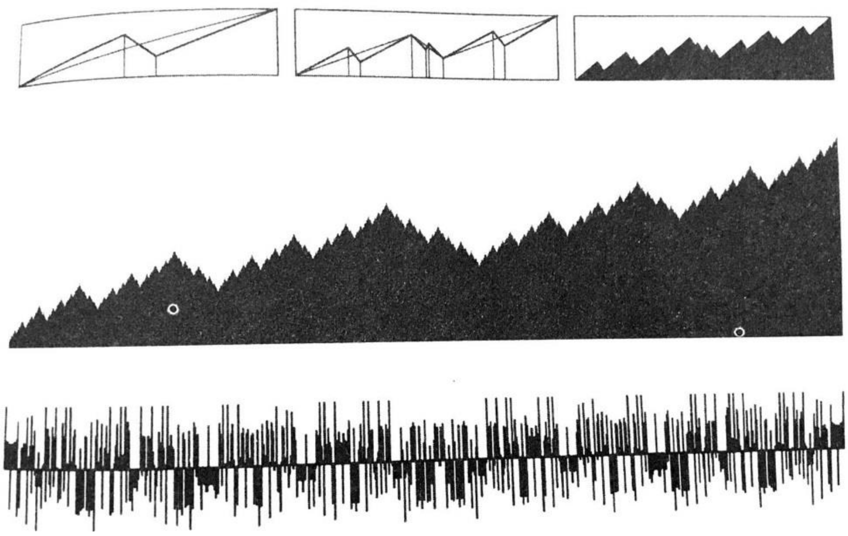

Figure 12. Brown-Bachelier cartoon. From Mandelbrot (2004).

The first line of figures represents the process of fractal construction, the second line is the enlarged final result and the third line shows how the final chart actually has price changes in accordance with Bachelier's model.

As realistic as the previous figures may seem, they were built by deciding a priori the type of variations that would have the prices: the generator was repeated in fact according to the same up-down-up path at each iteration. To make the result even more realistic, it is possible to add to the process an element of randomness, as provided for by the price independence in Bachelier's model: the randomized structure is added by modifying the generator, letting it be the fate to decide, at each iteration, what order have the three segments of the broken line, generating different combinations such as up-up-down, down-up-up or down-down-up. The following figures show such random construction process, but assuming only three variants of the generator (down-up-up, up-down-up and up-up-down):

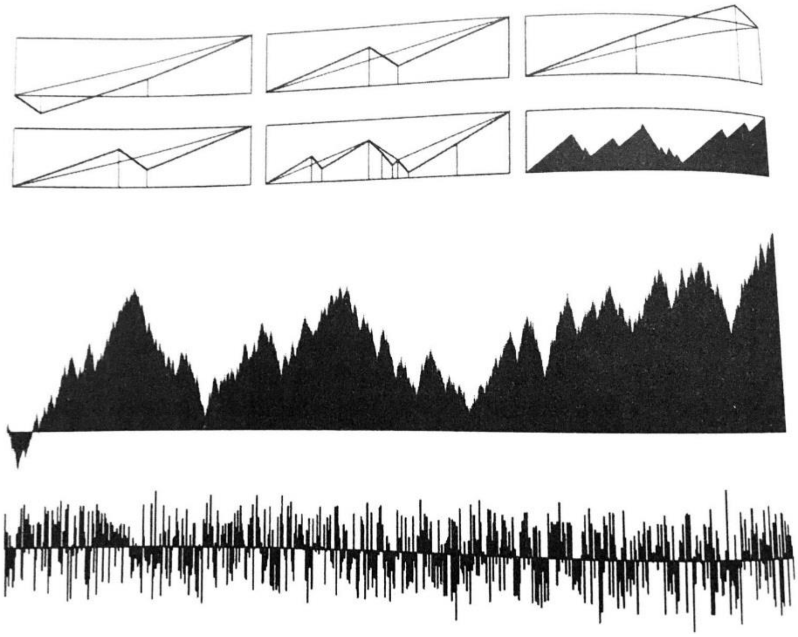

Figure 13. Brown-Bachelier cartoon with randomness. From Mandelbrot (2004).

As before, the figures initially describe graphically the construction process (the first two lines), after which the enlarged final result and finally the graph of the price changes. It is interesting to note that the price diagram this time also includes negative changes, below the initial value, and that the general structure of the changes graph has remained almost unchanged from the first to the second case, result due to the fact that in agreement with the normal distribution assumed by Bachelier's model, the prices never move beyond a certain number of standard deviations.

The addition of a randomness factor within the fractal model allowed a result more similar to the real price diagrams.

3.4 Image gallery

Examples of more complex, natural and artificial fractals will be proposed, as this will be a preliminary to the understanding of the fractal dimension paragraph, the central topic of all the fractal geometry, which will follow.

The apparent complexity of fractal forms fascinated Leonardo da Vinci long ago, who in his studies of anatomy tried, without much success, to analyze the structure of human lungs, whose bronchi are one of the most fascinating examples of fractals in nature. Another classic example of natural fractals is the network created by the neurons in the brain, in which the self-similarity at different scales is clearly visible[27].

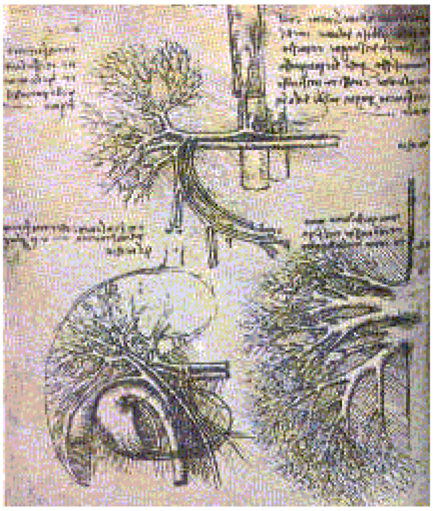

Figure 14. Human lungs. From the "anatomical studies" of Leonardo da Vinci.

[27] Current scientific literature states that the human body uses fractal geometry to create complex shape thanks to the simplicity of its recursion rules, which allows to optimize the available resources.

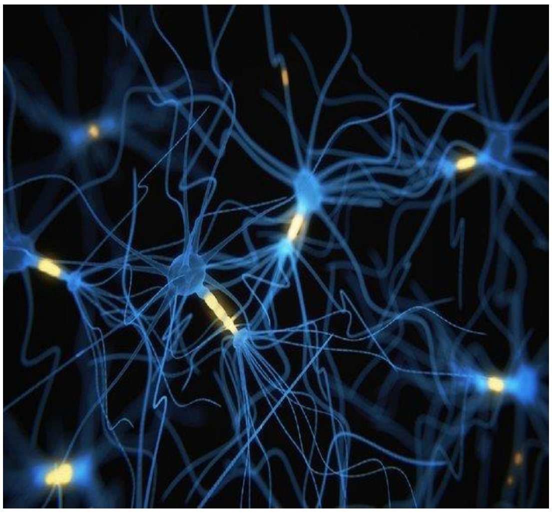

Figure 15. Neurons. From https://gds.it/speciali/scientist-tecnica/2017/06/28/-taggati-i-neuroni-in-attivita-per-immortalare-i-i-pensieri-03f9e924-6bda-40ea-991d-4eeff3f3c10d/.

Other natural objects structured according to fractal rules are the mountain ranges and the fern leaf, in which once again it is easy to identify the self-similarity: if the image as a whole is zoomed in somewhere, you get an enlargement that looks like the previous image.

Figure 16. Mountain range. From https://footage.framepool.com/it/shot/121095720-mackenzie-country-alpi-neozelandesi-nubinacciose-innevato.

Figure 17. Fern leaf. From https://pixers.it/adesivi/foglia-di-felce-16186465.

With fractal mathematics it is possible to construct complex artificial structures of great mathematical interest, some of which are the solution of mathematical paradoxes that have long remained without an answer. Below are some of them, briefly analyzing also the related construction process.

3.4.1 Sierpinski gasket

The gasket takes its name from the Polish mathematician Waclaw Sierpinski who studied a set of bizarre shapes that compress curves of infinite length within figures of finite size.

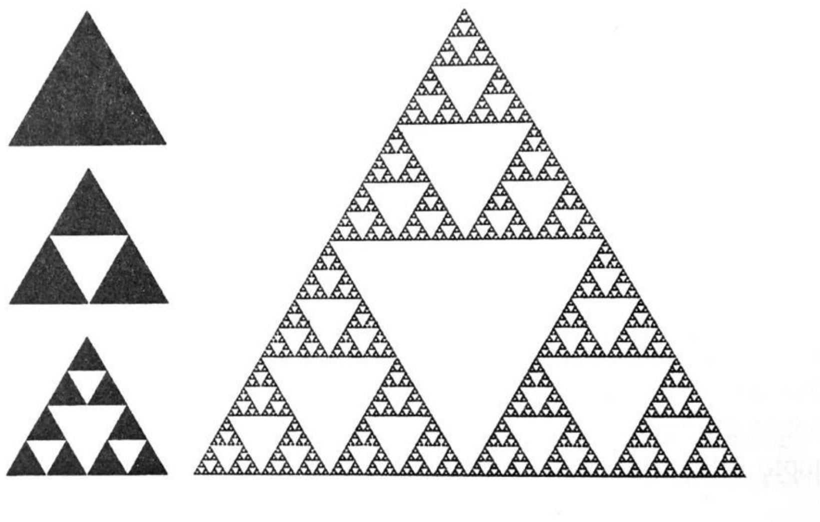

Figure 18. Sierpinski gasket. From Mandelbrot (2004).

There are different construction processes of the gasket, but the simplest is the one highlighted in the figure above: it starts from the black triangle in the upper left (initiator) and a new triangle is constructed in which the triangle formed by the union of the midpoints of the sides is removed (generator). The process is repeated each time a new triangle is encountered (recursion rule). The result is a pitted triangle that takes the name of gasket, which is formed by copies of itself to scales, where the scale factor remains equal to each dimension; a fractal of this type, as seen above, is defined as self-similar.

3.4.2 Sierpinski tetrahedron

The Sierpinski tetrahedron represents the three-dimensional case of the previous plane figure. For its construction, tetrahedra are used instead of triangles. The result is once again a fractal which repeats itself exactly on each scale without deformation, so it is also self-similar.

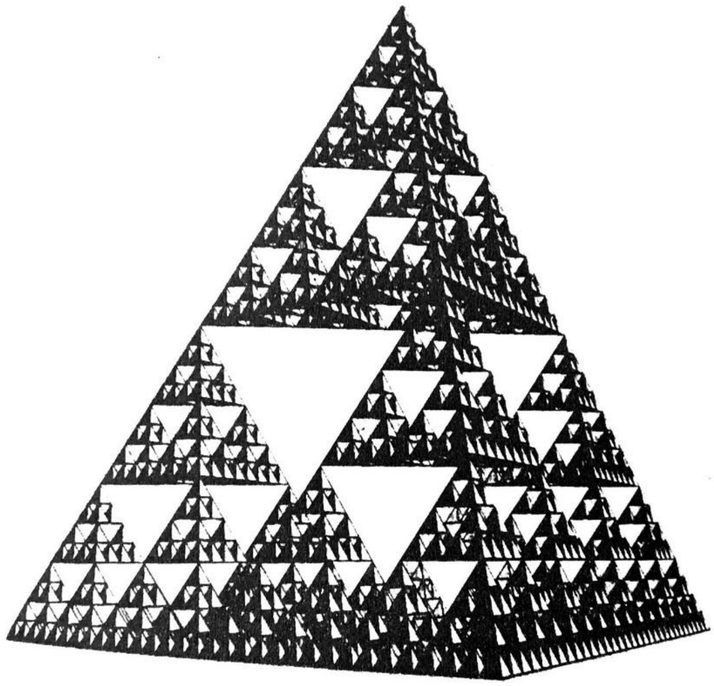

Figure 19. Sierpinski tetrahedron. From Mandelbrot (2004).

3.4.3 Cantor set

It is a subset of the range [0,1] of real numbers introduced by German mathematician George Cantor. The Cantor set is recursively defined starting from the interval segment [0,1] and removing a central open segment from each interval at each step. At the first step the sub-interval (1/3,2/3) is removed from [0,1] and the interval remains at two intervals [0,1/3] ∪ [2/3,1]. At the second step a central open segment is removed in both of these ranges (one-third the length of the segment, as in the first step) and four even smaller intervals are obtained. The Cantor set consists of all the points of the starting interval [0,1] that are never removed from this recursive process: in other words, the set that remains after iterating this process infinite times. Its peculiarity lies in the fact that it starts from a set of topological size equal to 1 and comes, after infinite iterations, to a set having a null topological dimension.

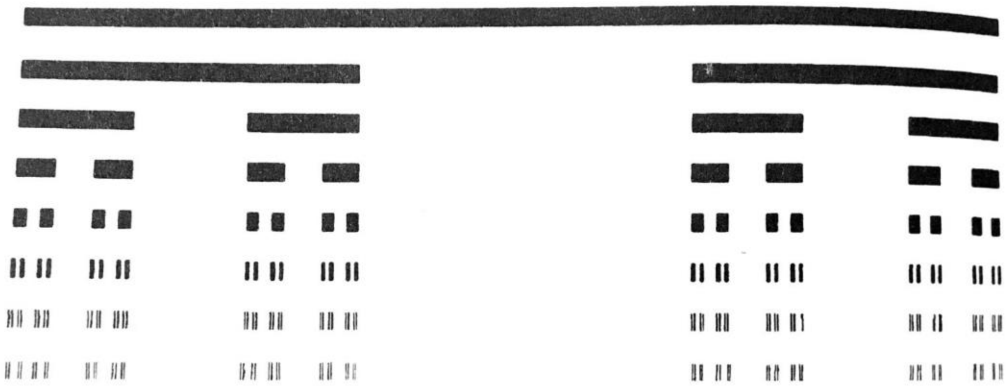

Figure 20. Cantor set. From Mandelbrot (2004).

The Cantor set is also found in similar versions in nature, for example in the emission spectrum of certain organic compounds.

3.4.4 Koch curve

The Koch curve is another self-similar fractal, in which the ratios of its parts do not undergo deformations with respect to the original figure. It is obtained starting from a segment of finite length, dividing that segment into 3 equal parts and replacing the central segment with two equal upward segments that form the two sides of an equilateral triangle. For each of the new segments that are formed, the process is repeated. A figure is obtained whose length increases by a ratio of 4 to 3 with the previous one, having therefore an infinite length.

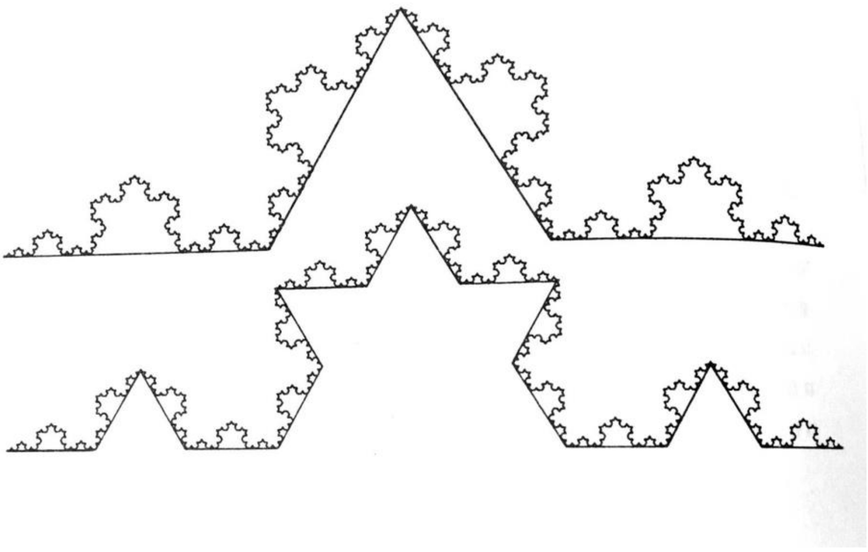

Figure 21. Koch curve. From Mandelbrot (2004).

The peculiarity of this figure is due to the fact that, even if it is of infinite length, it is continuous; it is also not derivative at any point, as it is not possible to draw a tangent at a point.

If instead of starting from the initial segment, you start from an equilateral triangle and apply the same procedure as before on each side, you get something similar to a snowflake, which takes the name of *Koch snowflake.* The singularity of this figure is that an infinite perimeter contains a finite area.

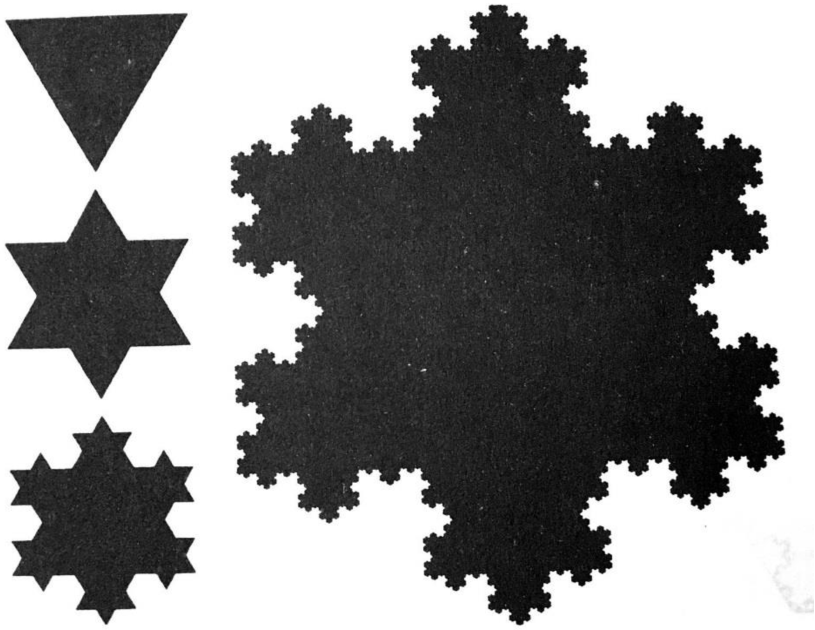

Figure 22. Koch snowflake. From Mandelbrot (2004).

3.4.5 Random fractal curves

Previous fractal figures are created in a non-stochastic way. If you add the case in recursive processes, you get more complex figures that can very realistically reproduce objects already present in nature. The following two figures are an example.

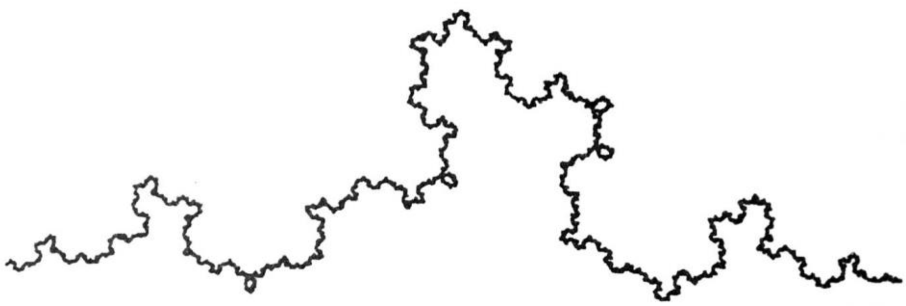

Figure 23. Random fractal curve (1). From Mandelbrot (2004).

Figure 24. Random fractal curve (2). From Mandelbrot (2004).

The first figure is a particular Koch curve created randomly: the initiator remains the same segment as before, but here, at each iteration, the middle third of the segment is replaced by a generator with the sides of the equilateral triangle that can be randomly turned upward or downward. By going through the process, you get a figure very similar to a coastline.

The second figure has similarities with the Cantor set: you start with a square, divide it into 125 squares and choose a random number to darken. Then the process is repeated in each of the darkened squares. The result is an irregular cluster of black dots, which looks like a galaxy cluster diagram.

3.4.6 Mandelbrot set and Julia set

As a last graphical example of fractal two closely connected sets will be presented, the Mandelbrot set and the Julia set. The mathematical function that describes both of them, a feedback loop, is as follows:

$$z_{n+1} = z_n^2 + c, \quad \text{con} \quad z, c \in \mathbb{C}$$

The Mandelbrot set is obtained by fixing $z_0=0$ and varying *c* in the complex plane. If the succession does not diverge to infinity, it is said that *c* belongs to the set; it is therefore the set of complex numbers *c* for which the succession is limited. If the boundary of the figure is represented on a screen, it assumes a fractal conformation (see the figure later) whose complexity increases with the enlargement applied. In summary, if you indicate with M the Mandelbrot set, it is defined by

$$M = \begin{cases} z_0 = 0 \\ z_{n+1} = z_n^2 + c \end{cases}$$

The Julia set J_c *is* obtained by fixing *c* and varying z_0 in the complex plane. If the succession does not diverge to infinity, it is said that *z* belongs to the set. It is made up of all those points whose behavior is chaotic, that is to say as a result of small arbitrary disturbances, can change dramatically[28]. It is defined by

$$f_c(z) = z^2 + c \quad ,$$

so for each complex number *c* you get a different Julia set.

To represent such sets graphically, you can color with a dark shade on a computer screen the points that converge and assign a different color to the points that diverge to the infinite, with a shade that varies depending on how fast they do it. The following figures show the Mandelbrot set and two Julia sets, obtained respectively with c=0,285+0,013i and c=-0.835-0,2321i.

[28] Its complement in the complex plane is the Fatou set, which is then the set of those points whose behavior is more stable.

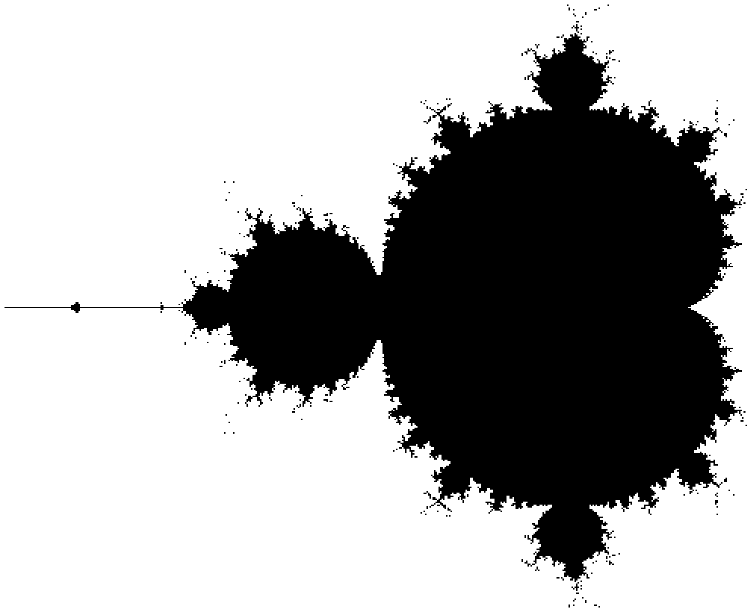

Figure 25. Mandelbrot set. From Mandelbrot (2004).

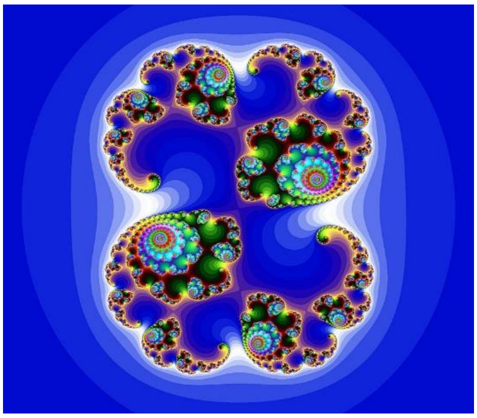

Figure 26. Julia set (1). From Wikipedia.

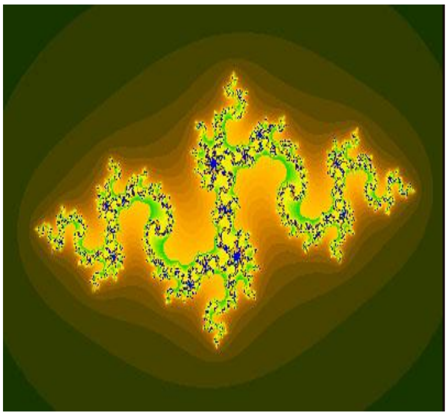

Figure 27. Julia set (2). From Wikipedia.

The relationship between Mandelbrot and Julia set is as follows:

- if c is chosen inside the Mandelbrot set (in the black part, where it converges) then the corresponding Julia set J_c will be connected[29]
- if c is chosen out of the Mandelbrot set then the corresponding Julia set is not connected
- If c is chosen on the boundary of the Mandelbrot set then the corresponding Julia set is reduced to a skeleton that has no internal part, that is it has no area, but is still connected.

3.5 The fractal dimension

Several definitions of dimension have been introduced in the history of mathematics, over all the vector space and topological dimension. The *dimension of a vector space* is the number of linearly independent vectors that make up the base, that is the number of generators. It corresponds to the expectations of the common sense, for example a point has dimension 0, a straight line has dimension 1, a plane has dimension 2 and a space has dimension 3. This first definition however is not suitable for example to describe curves that fold on themselves. To obviate this (and others) problems was formalized the concept of *topological dimension*[30] that allows to describe figures and shapes that do not change topologically when a without tears deformation is carried out and is based on the concept that the dimension of a set equals the number of independent parameters necessary to describe a point on it; for example, the dimension of a plane is two because to describe a point on it two parameters are sufficient, that is the Cartesian coordinates. There are therefore different definitions of dimensions, due to different purposes and points of view.

Imagine a ball of string, with a diameter of 10 cm and a thickness of string equal to a few millimeters and consider it from the point of view of a Euclidean vector space, that is the mathematical space to which we are commonly used to. If you look at it from far away it will look like a point, that is it will have a null

---

[29]A set is called connected if it cannot be represented as the union of two or more open not empty and disjointed sets.

[30] A mathematical description of the topological dimension and how it can be calculated is available in Appendix 2.

dimension. If you get close enough to touch it, it takes on three-dimensional shape and dimension. If you look even closer, the filaments that make up the string seem to have a dimension of 1. By looking at these filaments under the microscope, you find that they are actually three-dimensional. If you had a sufficiently powerful microscope, you could look at atoms as one-dimensional points. On the other hand, if you measure the length of a coastline, you can obtain different measures that grow with the increase in the accuracy of the measure (think of the difference that can be obtained by measuring the coast of Italy on a map using a ruler whose minimum dimension is 1 mm rather than 1 cm). The dimension of an object therefore depends on the point of view from which you look at it, or on the instrument with which you measure it. The ways of conceiving the dimension previously analyzed are not suitable for describing these measures, and the concept of fractal dimension has therefore been introduced.

The fractal dimension is a number that provides a statistical index of complexity, comparing how much detail in a fractal figure changes with the scale at which it is measured, and is a fractional number. This is the reason why the fractals have taken their name because they are forms or figures with a fractional dimension. The definition of fractal dimension is not unique, indeed there are several specific definitions. The most important are the correlation dimension, the Hausdorff dimension and the Minkowski-Bouligand dimension[31].

3.5.1 Correlation dimension

It is linked to the concept of how many copies (c) of a d-dimensional object are needed to enlarge or reduce the same object of a times. For example, to double ($a=2$) a segment ($d=1$) you need two segments ($c=2$), to double ($a=2$) a sheet of paper ($d=2$) you need four sheets ($c=4$) to double up it twice in all directions. The generic formula that binds this relationship is therefore:

$$c=a^d.$$

If you solve the previous equation for dimension d, you get:

[31] For self-similar fractals, all dimension types coincide.

$$d = \frac{log(c)}{log(a)}$$

where *d* is the correlation dimension. This dimension is one of the simplest and can therefore only be used with self-similar fractals.

For the purposes of fractal geometry, it is more convenient to consider *a* as the inverse of the reduction or dilation ratio of the unit of measurement, and *c* as the number of units of measurement necessary to measure the figure. As an example, consider the calculation of the correlation dimension of the Koch curve in the following figure:

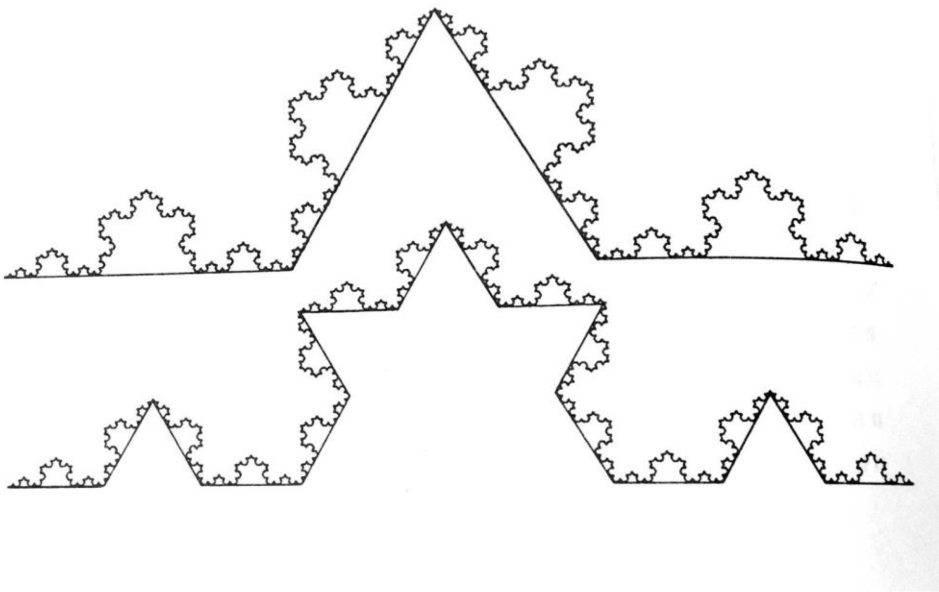

Figure 28. Koch curve (2). From Mandelbrot (2004).

If you measure with a ruler as long as 1/3 of the width of the object, that is as a section of the broken line of the first Koch curve in the figure, you get a measure of 4.

If you redo the same procedure with a ruler this time equal to 1/9, you get a measure of 16, as in the second case; you notice then that the ratio is always 4 to 3.

The correlation fractal dimension of the Koch curve is therefore $d = \frac{log(4)}{log(3)}$ =1,2618. This result means that the curve is larger than a one-dimensional line, but smaller than a two-dimensional plane; intuitively it makes sense because, being the curve folded, it fills the space more than a segment, but less than a plane.

3.5.2 Hausdorff dimension

The Hausdorff dimension is especially important because it falls within the very definition of fractal, as seen above. For a figure to be considered fractal, its Hausdorff dimension must be strictly larger than its topological dimension.

To define in a practical way the Hausdorff dimension of an *X* form, one must consider the number $N(r)$ of maximum *r* radius balls necessary to completely cover *X*. Decreasing r, $N(r)$ increases. If $N(r)$ grows in the same way as $1/r^d$ when *r* tends to 0, then it is said that *X* has dimension ***d*** [32].

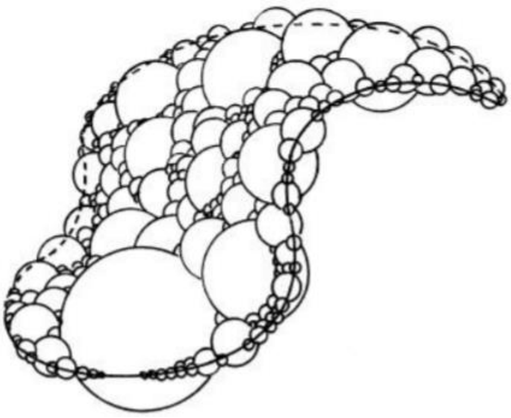

Figure 29. Balls of different radius in a curve. From Magnanini, R., Dispense per il corso Istituzioni di Analisi Superiore 1, Università degli Studi di Firenze.

3.5.3 Minkowski-Bouligand dimension

To calculate the Minkowski-Bouligand dimension of a fractal, the *box-counting* method is used, i.e., you insert the fractal into a container (a portion of the plane if it is two-dimensional, a portion of space if it is three-dimensional) and the container is divided into cells of width ε; the number of cells necessary to fully cover the fractal is N_ε. The dimension is obtained by progressively reducing the width of ε and comparing it with the increase of N_ε:

[32] The figure shows how balls of different radius can fill the surface of a curve in space.

$$dim_{M.B.} = \lim_{\varepsilon \to 0} \frac{log(N_\varepsilon)}{log(1/\varepsilon)}.$$

If the limit does not converge, then we are talking about the upper and lower dimension of the cells that correspond respectively to the upper and lower limit above calculated. The Minkowski-Bouligand dimension is therefore well defined only if the upper and lower cell dimension exists and is equal. The higher dimension is also called *Kolmogorov dimension.*

There is a relationship between Hausdorff dimension, upper dimension and lower dimension, and it is as follows:

$$dim_{Haus} \leq dim_{lower} \leq dim_{upper}.$$

The box-counting method will now be applied to measure the dimension, once again, of the Koch curve (Mandelbrot 2004):

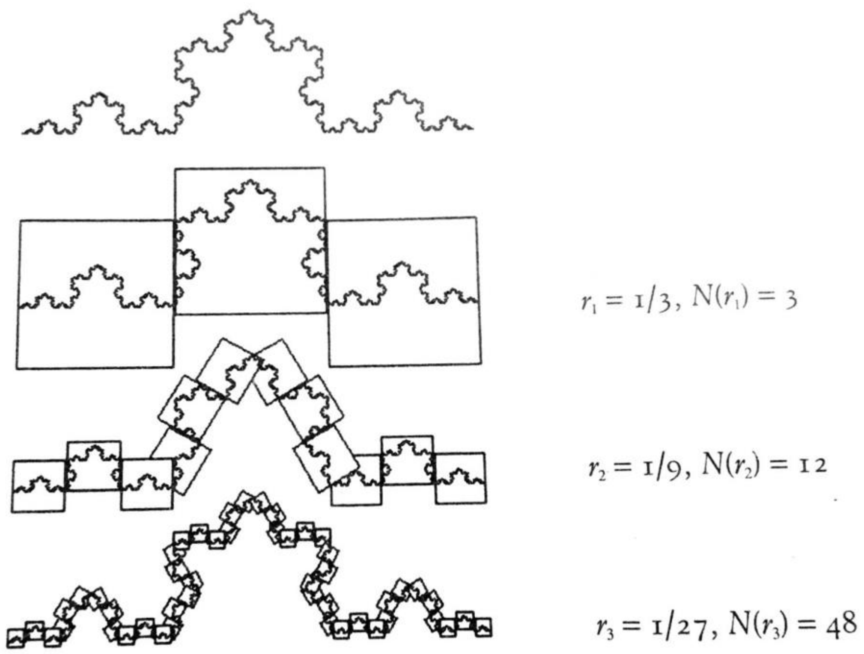

Figure 30. Box-counting on Koch curve. From Mandelbrot (2004).

From the figure you can see that, if you start by covering the figure with cells 1/3 of the curve wide, 3 are necessary. If you reduce the cells 1/3 more (that is 1/9 compared to the original curve), you need 12. Continuing in the same way, with cells 1/27 of the initial curve wide, the necessary cells become 48.

The structure is regular, in fact at each step if the cell dimension decreases by 1/3, the number of cells needed to cover the object increases by 4 times. Moving to the limits with the previous definition, you get a dimension of 1,2618 as in the case of the correlation dimension because, as mentioned before, the Koch curve is a self-similar fractal and therefore all dimension definitions coincide.

CHAPTER 4.

FRACTAL ANALYSIS IN THE FINANCIAL MARKETS

This part of the discussion will cover the events that led Benoit Mandelbrot to discover the possibility that the financial markets have fractal characteristics, going on to analyze the various components that, in the final section of this part, they will converge within a fractal market description model[33].

In particular, the model will be based on three concepts: long-tails, long-term dependence and trading time.

4.1 The long-tails

Long-tails are a property of statistical distribution[34]. A long-tails distribution is characterized by a density function that allows values even very far from its average value. Graphically, it is characterized by a bell higher in the center and thicker in the tails than the classic Gaussian distribution and has therefore a greater kurtosis. In practice, this characteristic implies that extreme values are much more probable and have great importance within the sample, much more than in the case of a Normal. Long tails are generated by a power law relation[35].

As an initial example of power laws and long-tails, we consider three famous cases where distributions with these features have been found.

4.1.1 Cauchy blindfolded archer

Augustin-Louis Cauchy, a 19th century French mathematician, disagreed with the prevailing opinion of the era that consisted in the use of Normal distribution to describe almost all phenomena. Imagine an archer standing in front of a target hanging from an endless wall; the archer is blindfolded, so the result of the shot is random. Consider several long-sufficiently sets of shots.

If the shots followed the distribution of the Gaussian curve, for the most part they would be very close to the target and in minimal part very far, and for each set it would be possible to calculate the average error and the standard deviation. The

[33] This part relies heavily on Mandelbrot's original work, both in its mathematical form (Calvet, Fisher and Mandelbrot 1997a) and in its more literary form (Mandelbrot 2004).

[34] Tails represent the part of the distribution values far from the mean value, i.e. from the center of the curve.

[35] For an in-depth discussion of the power laws, see Appendix 3.

archer of Cauchy, on the other hand, lives in a world in which the errors are not slight: often the arrow ends hundreds of meters far from the target, or even travels almost parallel to the wall ending to kilometers of distance.

In a Gaussian context after a number of shots the average result will have stabilized on a certain value and even the most wrong shots will give a negligible contribution to the sample; it will be practically impossible for the next shot to change the mean significantly. The case of Cauchy is different, because a very wrong shot could make a greater contribution than a slightly wrong previous hundred shots; the results in such a world never stabilize on a predictable value nor does the variation in results stabilize around an average value. Such a distribution is called *Cauchy's*.

Cauchy distribution is characterized by undefined expected value and variance, is a stable distribution and is a power law for sufficiently large values of its probability density function[36]. Its tails are therefore fat, and this implies that variability is wild: «errors are not sand grains, but are a combination of sand, pebbles, boulders and mountains»[37]. One extreme value can affect thousands.

4.1.2 Zipf words

George Kingsley Zipf was a Harvard teacher who wrote a book, *Human Behavior and the Principle of Least Effort,* in which he described the power laws as a relationship present not only in the social sciences (scientists already used them, for example, to relate the number of earthquakes with their intensity), but also in all kinds of human-related activities. In particular, he found an application in estimating the richness of vocabulary in a text. You proceed as follows: you choose a text and count how many times a word appears. After having sorted the words according to their frequency, the most common word is assigned the grade 1, the next the grade 2, and so on. You then draw the graph of the word frequencies according to their degrees. The resulting curve drops at the beginning vertiginously and then decreases more slowly, creating a hyperbola branch. By comparing curves

[36] Stable distributions will be analyzed later.

[37] Mandelbrot (2004).

relative to different texts, you can compare the richness of vocabulary of one to the other.

Zipf also empirically wrote the equation to describe a similar relationship, which is now known as Zipf distribution (or even Zeta distribution)[38]:

$$P(x) \approx Fx^{-1/\alpha}$$

where *P* is the probability distribution function, *x* the degree, *F* a constant that Zipf estimated equal to 1/10 and $1/\alpha$ the critical factor of the power law. This law is generic; if you consider it with reference to the count of the words in a text, the greater the value of α the richer is a vocabulary, that is, the curve that expresses the frequency of each word with respect to its degree drops more slightly, so that rare words appear more often than they would otherwise. In his example Zipf estimated that the value of α was equal to 1. It is also a stable distribution and being a power law it has also fat tails and presents a wilder variability than the Gaussian distribution.

4.1.3 The distribution of Pareto income

The Italian economist Vilfredo Pareto was fascinated by problems such as how to achieve wealth or how it is distributed within society and decided to measure the gap between rich and poor. He collected many data on wealth and incomes of different eras and countries and by analyzing these data discovered that, by reporting income level data on one axis and the number of people with that income on the other, similar figures were obtained for almost all places and centuries. Consider the following graph, taken from the most famous Pareto's book[39]:

[38] Zipf distribution is the discrete case of Pareto distribution.

[39] Pareto (1905).

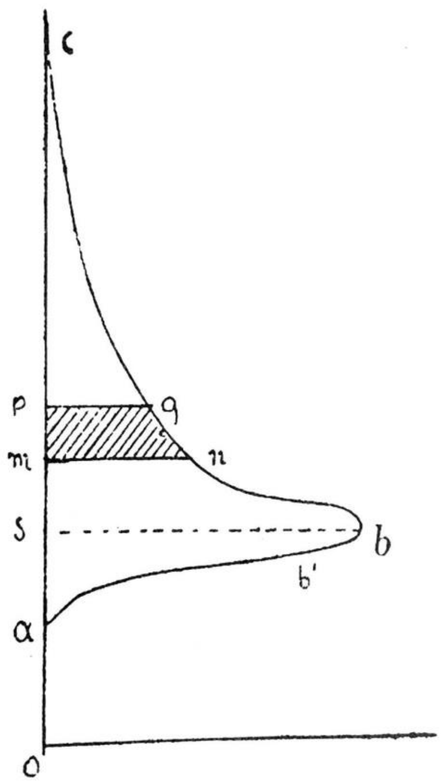

Figure 31. Income curve. From Pareto (1905).

the income level is represented on the vertical axis, while the number of people on the horizontal axis. The hatched area represents the number of people with an income between *m* and *p*. It may be noted that society cannot be considered as a social pyramid, where the level of income increases or decreases in a linear way between the different classes; it is an asymmetrical curve, very wide in the right side where most people are located and very narrow in the left side where there are few people disproportionately rich. According to Pareto (1905) the external shape of the curve remains unchanged over time; it is a social law that occurs at the cost of social revolutions:

> It is not only the accumulation of inferior elements in a social stratum that harms society, but also the accumulation in lower strata of elect elements that are prevented from rising. When at one time the higher strata are filled with decayed elements and the lower states are filled with elective elements, social equilibrium becomes extremely unstable

> and a violent revolution is imminent. One can, in some respects, compare the social body to the human body, which readily perishes when the elimination of toxins is prevented.

The corresponding probability distribution is valid not only for income, but also for numerous other social laws and is called *Paretian* or *Pareto distribution* and in the simplest form, with respect to income, is:

$$P(x) = (\frac{x}{m})^{-\alpha}.$$

P(x) indicates the percentage of persons with income above *x*, *m* is the minimum income in the analyzed society. α is the exponent of the power law, which Pareto estimated for incomes equal to 3/2. For example, if you want to know the percentage of people with more than 20 times the minimum income, just calculate $P(20 \cdot MinYield) = (20)^{-3/2}$=1.11%.

If you report the income data on the logarithmic chart, you get a downward straight line: we are therefore still in the presence of a power law, which implies that even the distribution of Pareto has long-tails (on the left side, as shown in Figure 31) and this practically implies that the elite of the richest people possess most of the wealth of the planet (as we know to be true even today).

The Pareto distribution is stable for $0<\alpha<2$.

4.1.4 Stable distributions to analyze the real world

Previous cases allowed us to begin to see how long-tails, and therefore the power laws, describe more realistically the realities associated with human activities than the "classic" Gaussian distribution. Until the second post-war period, however, a common thread had not yet been found explaining why different distributions such as those of Cauchy, Zipf and Pareto had in common the fact of being power laws.

Mandelbrot identified the link in the mathematics of Levy's stable distributions, a topic already known in academia, but which had not yet found a practical application. Let's see what it is.

A probability distribution is called *stable* (or even α*-stable of Levy*) if the linear combination of two independent random variables with the same distribution also has the same distribution. A random variable is stable if its distribution is stable. Examples of this type are Gaussian's, Cauchy's, Pareto's and Zipf's.

It is therefore a family of distributions that can be described with 4 parameters:

- α: is the stability parameter of the distribution and is a measure of the concentration (thickness) of the tails.
- β: is a measure of the degree of asymmetry of the curve.
- *c:* scale parameter, determines the size of the overall probabilities.
- μ: positioning parameter.

α is the most important parameter, varies in the range (0;2] and is the same α as that previously encountered in the three analyzed cases; it is in fact the exponent of the power laws: this is why stable distributions are useful to mathematically describe the reality, because they are able to easily describe the power laws and therefore the long-tails, which as we have seen often appear in the real world. As this parameter decreases, the thickness of the tails increases. In addition, all stable distributions, except Normal, are power laws for sufficiently large values.

The mean and the variance are not defined respectively for $\alpha \leq 1$ and for $\alpha < 2$: in such a case the contribution of the tails can be so high that the two moments never stabilize around a value and they can therefore intuitively be considered infinite.

The most important values of α are:

- α=1, which identifies the Cauchy distribution and Zipf distribution (in the case of words).
- α=1/2, indicating Levy distribution.
- α=3/2, which refers to the Pareto distribution (in the case of income).
- α=2, owned by the Normal distribution.

In addition, the subset $1 < \alpha < 2$ identifies a number of distributions called *Pareto-Levy's*. They are the most useful distributions to describe the financial markets.

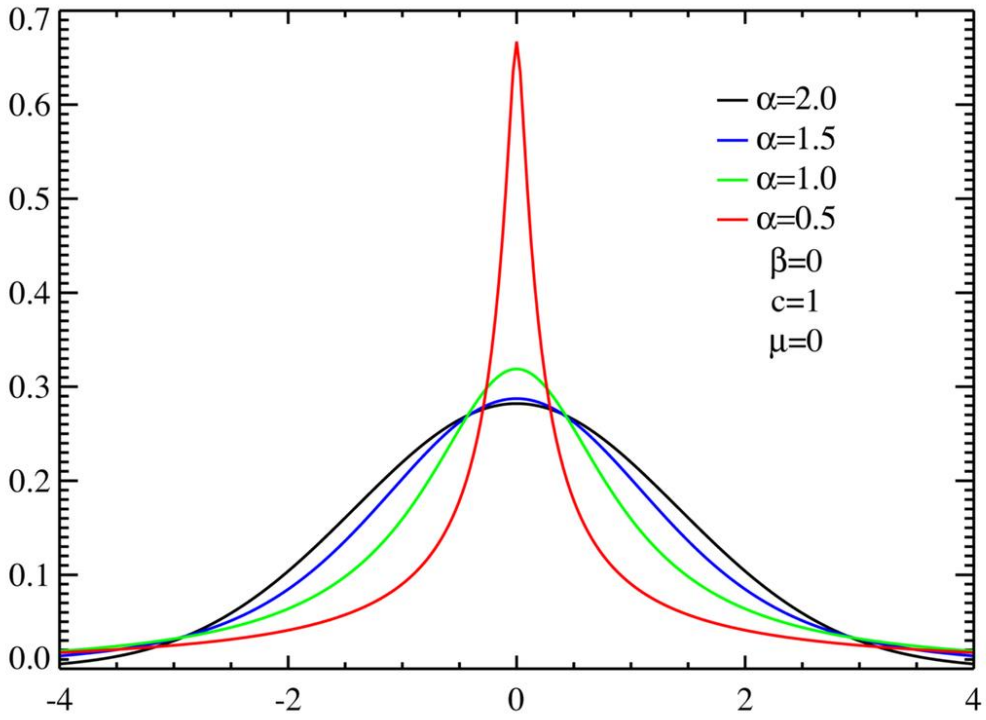

Figure 32. Shape of some stable distributions. From Wikipedia.

Another important peculiarity is that stable distributions are attractors for the sum of independent and identically distributed variables and this allows to integrate the *central limit theorem*[40] for those variables that have $\sigma \to \infty$: the sum of a sufficiently large number of variables of this type can be approximated with a stable distribution different from Gaussian as usually done.

Stable distributions allow therefore to join mathematically, in a unifying concept, ideas apparently disconnected to each other, through the manipulation of four simple parameters and in particular of the parameter α that regulates the power laws. If the power laws are a "way" of seeing the world, this type of distributions constitute the instrument able to mathematically model them.

[40] The central limit theorem states that the sum or average of a large number of independent and equally distributed random variables is approximately normal, regardless of the underlying distribution.

4.1.5 The mystery of the cotton

Summing up what we has been seen so far in this chapter, long-tails are a consequence of power laws, which have always been observed in many natural and social phenomena. Stable distributions allow to describe these concepts with a single mathematical tool of 4 parameters. But how do the financial markets relate to all this?

Mandelbrot, thanks to a series of fortuitous circumstances, found that the income chart discussed above was very similar to the chart that correlated cotton prices in the US with their frequency over a very long period (more than one hundred years). In the cotton market there were series of small price movements together with a few huge movements (too large to fit into the *Normal* case) as well as the Zipf text contained a great amount of rare words and a few common words, and just as in the world there are entire areas of poor people and a few billionaires. The connection was the power-law distributions, so the long-tails.

Moreover, the variance behaved in a strange way, that is, it never stabilized on a single value, but it varied in an irregular way, it was indefinite. This suggested the use of stable distributions.

By plotting changes in cotton prices and their frequency in logarithmic scale, Mandelbrot (2004) obtained the following diagram:

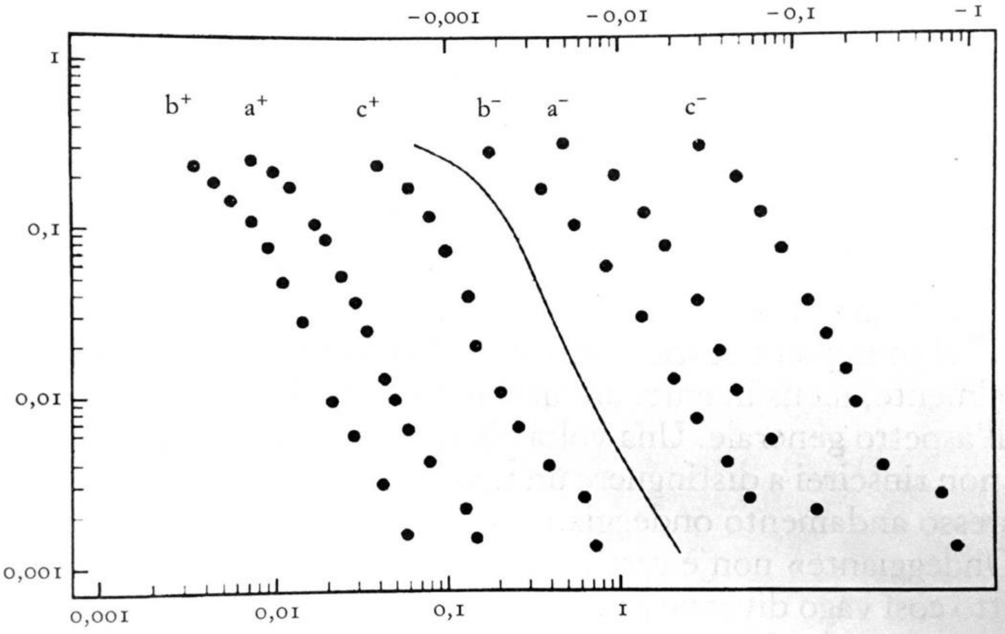

Figure 33. Frequency changes in cotton price in log scale. From Mandelbrot (2004).

The horizontal scale shows the positive (sign +) and negative (sign -) logarithmic variations at different time scales: *a* indicates the daily variations from 1900 to 1945, *b* indicates the daily changes from 1944 to 1958, *c* indicates the monthly changes from 1888 to 1940. The vertical scale shows their frequency.

The resulting forms seem to be straight lines, and this is the confirmation of the presence of a power law[41]. By measuring the slope of these lines, the author found a value of -1,7 and therefore both the exponent of the power law and the stability parameter of the stable distribution associated with cotton were equal to α=1,7 (because as seen in the paragraph of stable distributions, it is the same α). The variations in cotton process were therefore halfway between the wild variations in Pareto income (α=3/2) and the slight variations in Gaussian (α=2).

A first important result is that the variance of cotton price movements for such a parameter value is indefinite because α<2, and therefore it is not possible to use orthodox finance instruments such as Markowitz's portfolio theory to describe them.

A second result is the presence of long tails which implies a wide variability and therefore discontinuity in the value of prices, as it was expected from the value

[41] As shown in Appendix 3.

of α. Variations may be extreme, or at least more extreme than expected by standard models. In other words, the result of a single very bad day can erase the results of many other gone well. This is something to consider when deciding what to invest in.

It can also be seen that the forms are completely similar even to different temporal scales, denoting self-similarity. Changes in prices are invariant at a time scale. This leads us back to fractal geometry, as it is able to describe shapes similar to themselves at different scales. In this case, however, the factor of scale is time, not space. It is possible therefore to apply a fractal scaling to the variations of the prices, as it will be shown in the final model.

In conclusion, the first ingredient of the model are the long-tails of financial markets, caused by power laws that can be parameterized by stable distributions. The fractal scaling feature over time justifies the use of a multifractal model[42].

4.1.6 Brownian cartoons for the case of long-tails

As last part of this chapter, we take again the Brown-Bachelier cartoons seen in the third part, but this time inserting the discontinuity caused by the long tails and obviously maintaining the fractal construction process (scaling). The following figure illustrates the construction process:

[42] Remember that a form is multifractal if the scale factor in which the parts echo the whole is different along different directions.

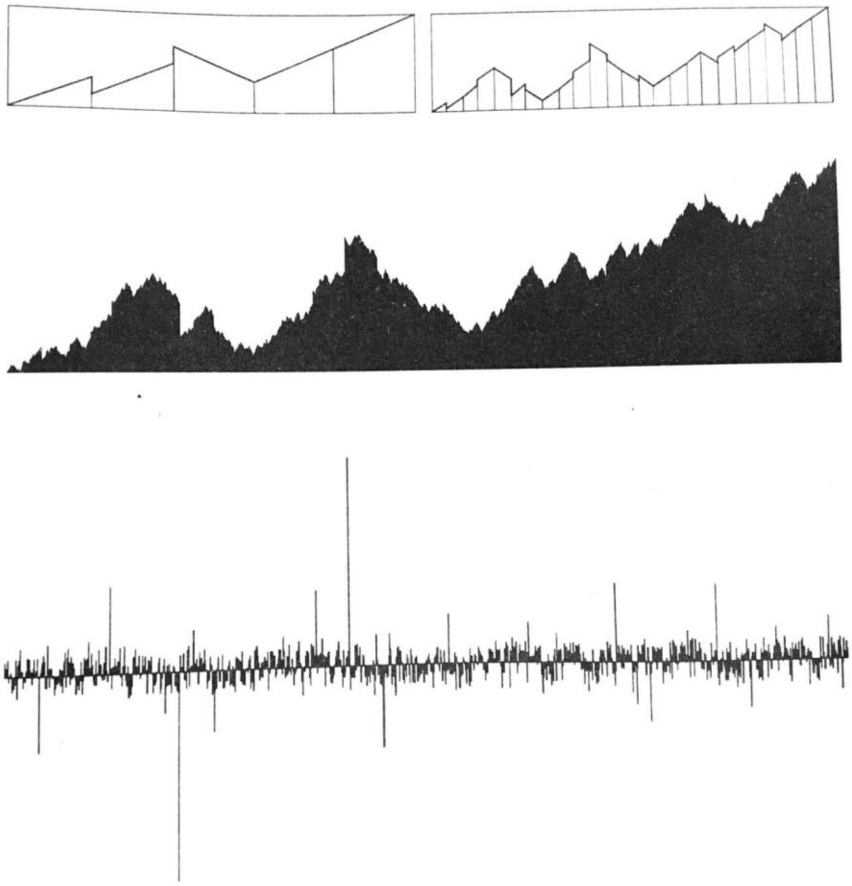

Figure 34. Cartoon with discontinuity. From Mandelbrot (2004).

The initiator is, as in the original case, an ascending line starting from the point (0,0) and reaching the point (1,1)[43]. The generator consists of five inclined intervals, to which are added two vertical discontinuities that represent respectively a sudden collapse and a market boom, result of the thick queues. By going through the process as we know, you get the fractal cartoon in the middle. Finally, the graph of variations is shown.

The result is a diagram showing a kind of invariance of scale and wild variation similar to that of cotton prices. The cartoon is beginning to be much more realistic, especially with regard to the graph of variations, where more or less long series of slight variations and stable market mix with sudden wild variations that result in collapses or boom. Indeed, that is exactly what happens in real financial markets.

[43] Remember that the scale of the width of the container is dilated to facilitate visibility, but it is a square.

4.2 The long-term dependence

4.2.1 Dependence in hydrology

The idea that financial markets may be dependent on long term came to Mandelbrot from the work that Harold Edwin Hurst did in Cairo to design reservoirs for storing Nile water. During the 20th century the British Empire wanted to map and control the Nile in order to exploit its economic potential. In previous centuries, the river alternated periods of flood with dry periods and the consequences were periods of prosperity or famine. The objective pursued by Hurst was to control this variability, within the framework of a project called *century storage* that aimed to accumulate the water needed to cope with the maximum possible drought.

Until then, the engineers designing the dams assumed that the annual changes in the flow of water were statistically independent and with normal distribution, that is, as in the case of the launch of a coin. Hurst instead discovered that the statistic range[44] of the Nile was much larger than it should have been if the assumption of independence had been true. The problem was not the individual flow values (once the original sequence was destroyed, the data was well matched to a bell curve), but the sequence with which they were presented: «Although many natural phenomena have an almost normal distribution, this is so only when the order of succession is ignored. When records of a natural phenomenon extend over long periods, both the average and the standard deviation vary considerably from one period to another. The tendency to present in groups makes the average and the standard deviation calculated over a small interval of years more variable than in random distributions»[45].

In particular, the range did not increase as a function of the square root of the observations as in the head or cross game, but as a power of ¾ of the standard deviation. There was a long-term dependence, so the sequence of data was important: if it rains for many years, the water level in a reservoir rises. If these rainy years are followed by a time of moderate weather, the tank remains full due

[44] The range is the simplest index of statistical variability and is given by the difference between the maximum value of a distribution and the minimum value.

[45] Hurst (1951).

to the effect of the previous years. If later a period of drought comes, the water level in the tank drops, but in each period will still contain more water than it would contain without the rainy years. Past events therefore continue to have a degree of dependence on future events.

The formula that Hurst found is:

$$R = \sigma(\frac{N}{2})^H$$

where *R* indicates how large a reserve tank should be to avoid flooding or drought. *R* depends on the standard deviation of rainfall from one year to the next (σ), on the number of years (*N*) and on the Hurst coefficient (*H*), in this case approximately equal to ¾, which will be examined in detail below.

4.2.2 Dependence in finance

Returning to the world of finance, Mandelbrot was the first to hypothesize the presence of long-term memory in the financial markets. A simple example of this (Mandelbrot 2004) is the reliance of IBM in 1982 on the process of building microprocessors and building software for the emerging personal computer, respectively to Intel and Microsoft, until then semi-unknown companies. The story is well known: Intel and Microsoft have grown enormously, while IBM had to divest some business units, but the stock prices of the three companies are still intertwined. The price of each one influences that of the others because the profits or losses of one influence the business of the others. An event 40 years ago, the birth of Intel and Microsoft under the initial auspices of IBM, still resonates today in the price of its shares.

The standard model of Brownian motion, in which events are assumed independent of each other, states that a Brownian particle in the water moves away from the starting point according to the rule of square root of time (as in the game of coin, where in place of time the square root of the number of shoots is considered): a particle that moves for 25 seconds will move five times more away than one that travels for a single second. Applying this concept to finance guarantees the possibility of finding a range in which an economic variable can most likely move. If the variable has a long-term dependency, it no longer oscillates within this interval (the hypothesis of independence is missing), but maybe if one day it increases, it will tend to increase for a while longer. In head or cross game, that would only happen if the coin was rigged.

It is possible to describe this tendency to present a long-term memory through the *H* Hurst exponent already seen above. The concept is always linked to Brownian motion, but in this case the distance travelled is proportional to some power of the time that does not necessarily have to be ½. This power corresponds to *H* and can be any fractional value between 0 and 1. Because of this, Mandelbrot named this series of interdependent increments as *fractional Brownian motion*.

The Hurst exponent quantifies the tendency of a time series to regress towards the mean or to tend towards a certain direction, then identifies its autocorrelation:

- $0 < H < ½$ indicates that the time series is anti-persistent, so the values will not deviate much from the initial value, i.e. after an increase is more likely to occur a decrease, rather than a further increase.
- $H = ½$ is the standard Brownian exponent, in which the values of the time series are independent.
- $½ < H < 1$ indicates that the time series is persistent, in which values tend to continue in the same direction for a while, creating nonperiodic finite cycles.

The following figures show, with extreme examples, how the random time series evolve as *H* changes*:*

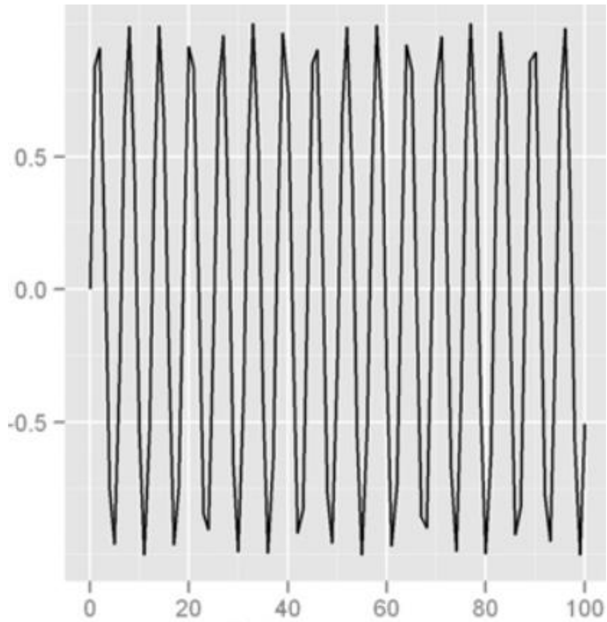

Figure 35. Anti-persistent series, H=0,043. From http://analytics-magazine.org/the-hurst-exponentnent-predictability-of-time-series/.

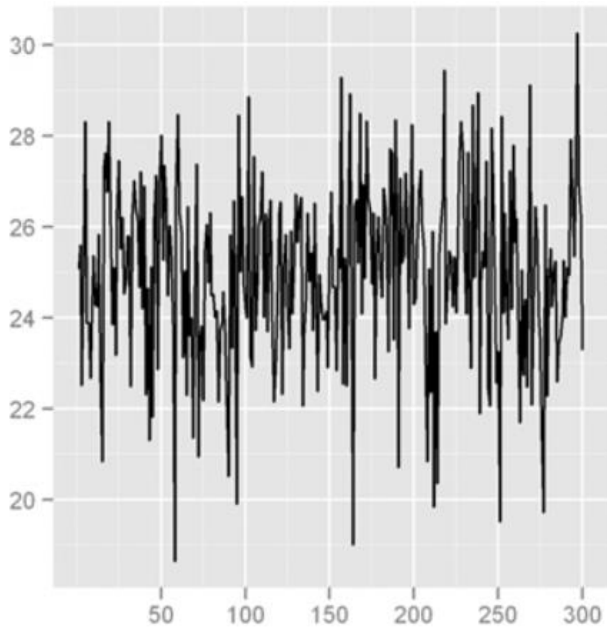

Figure 36. Brownian series, H=0,53. From http://analytics-magazine.org/the-hurst-exponentnent-predictability-of-time-series/.

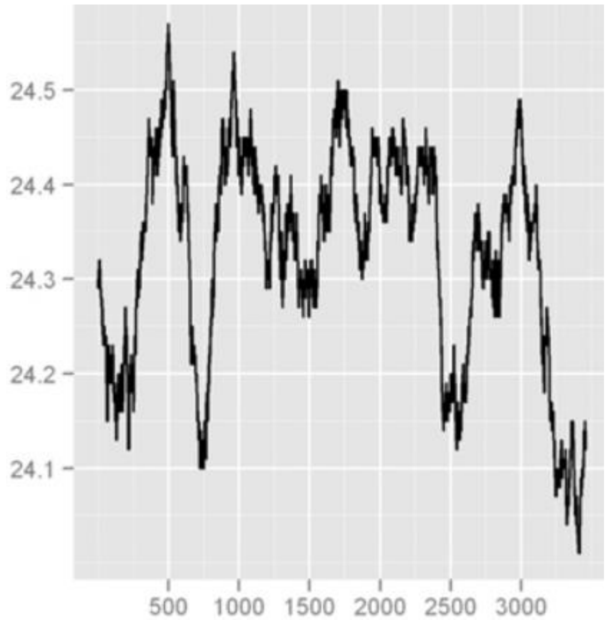

Figure 37. Persistent series, H=0,95. From http://analytics-magazine.org/the-hurst-exponentnent-predictability-of-time-series/.

Current literature suggests that a high value of H is a symptom of a market in which the emotional behavior of the crowd, which tries to follow the trend, has strong consequences, such as the formation of speculative bubbles. *H* values close to ½ indicate a very random market, in which the standard market model works quite well.

Peters (1991) found high values of *H* for technology companies such as Apple (0.75), Xerox (0.73) and IBM (0.72). More traditional companies like Texas State Utilities have more Brownian values (0.54).

Another study (Garcin 2016) investigated *H* values for major currency pairs, finding values very close to 1/2 as GBP/EUR (0,505), USD/EUR (0,5) and CHF/EUR (0,505). Slightly less Brownian values were found instead in those pairs containing a weaker currency, such as SEK/EUR (0,511) and SDG/EUR (0,51).

Fractal geometry can also be applied to long-term dependency, as the Hurst exponent is closely linked to the notion of fractal size (*D*). For self-similar time series, where $1 < D < 2$, you have that $D = 2 - H$. Fractals therefore allow to identify the presence of memory in the time series. For time series that identify instead multidimensional processes *H* and *D* are not so closely connected. In particular, the Hurst exponent describes the structure of the series for asymptotically longer periods, while the fractal dimension for asymptotically shorter periods.

Note that a peculiar property of the major part of long-term dependency processes is that there are always small apparent configurations, which may appear or disappear in a short time, but which do not represent a trend and therefore cannot be predicted. It is a fractal characteristic, because if you take as an example the figure above with the case of persistent series and you enlarge one of its parts, you get a figure that echos the whole one, but that represents a shorter period of time.

4.2.3 The Rescaled Range Analysis (R/S)

Summarizing what we have seen so far, the first two components of the multifractal model of financial markets are the discontinuity caused by the long-tails of distribution and long-term dependency.

The first is described by the coefficient α, which represents both the exponent of the power laws and the stability parameter of the stable distributions; as seen, a low value of α indicates a more risky market, in which price fluctuations can be wilder, while a high value indicates a market more similar to that of the classic model based on the launch of a coin.

The second is represented by *H*, the Hurst coefficient, which is a measure of the long-term memory of the time series, and it is the exponent of the fractional Brownian motion; a value of *H* equal to ½ represents the standard Brownian case, while higher values indicate persistence in the increments of motion and lower values indicate the opposite, that is anti-persistence.

To distinguish the effect of these two parameters in a series and to calculate the Hurst exponent, Mandelbrot and Wallis (1969) have developed a non-parametric statistical test[46] called *Rescaled Range Analysis* (R/S)[47].

The underlying idea is that the effect of long tails depends on the relative size of each event, while the dependency effect depends on the precise time sequence. If you destroy the ordered sequence by mixing the data, you eliminate any

[46] A non-parametric statistical test is a type of test that does not require assumptions about the distribution of the analyzed data.

[47] Resulting from Hurst's previous work in hydrology.

dependency and only discontinuity remains visible. To complete the test, simply compare the two sequences. If there is a difference it means that the precise sequence had value and there is therefore a certain degree of dependence. If there is no difference, the degree of dependence is negligible. In both cases a measure of dependence is obtained, that is the value of *H*. In some cases, there is a dual relationship between the two parameters and you have that $H = \frac{1}{\alpha}$; in the coin toss, for example, we know that H=1/2 and α=2 because it is an independent standard Brownian motion with Gaussian distribution.

In order to estimate the Hurst coefficient, the dependence of the *rescaled range*[48] on the time period *n* of observation must first be estimated. A time series of total length *N* is divided into a number of smaller intervals of length *n=N*, *N/2*, *N/4*, …. The mean rescaled range (R/S value) is then calculated for each value of *n*.

For a finite time series of length *n* with observed values $X=X1, X_2, \dots, X_n$, the R/S value of the rescaled range is calculated as follows:

1. Calculate the average of the values observed over the whole period;

$$m = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

2. Create a series expressing the deviation from the mean of the observed value in each sub-time interval *t*;

$$Y_t = X_t - m \quad for \quad t = 1, 2, ..., n.$$

3. Calculate the cumulative series of previous deviations, *Z*;

[48] The rescaled range is a measure of statistical variability. Its purpose is to provide an assessment of how the apparent variability of a series changes with the length of the time period considered.

$$Z_t = \sum_{i=1}^{t} Y_i \quad for \quad t = 1, 2, ..., n.$$

4. Calculate the range *R*;

$$R(n) = max(Z_1, Z_2, ..., Z_n) - min(Z_1, Z_2, ..., Z_n).$$

5. Calculate the standard deviation *S*;

$$S(n) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - m)^2}.$$

6. Calculate the rescaled range *R(n)/S(n)*.

To find the value of *H*, we use the following power law relation:

$$\frac{R(n)}{S(n)} = an^H$$

where *H* is the Hurst exponent, *a* is a constant and *n* the length of the considered time period.

Using a logarithmic transformation *H* can be estimated using the following regression as a function of *log(n)*, where *H* represents the slope of the regression line:

$$log(\frac{R(n)}{S(n)}) = log(a) + H\, log(n).$$

Mandelbrot (1972) states that R/S analysis is more appropriate than autocorrelation and variance analysis because it allows to consider distributions with infinite variance (stable distributions except Normal). Moreover, he claims that time series have infinite memory and that R/S analysis can detect this characteristic[49].

4.2.4 Brownian cartoons with long-term dependence

As previously done with the Brownian motion and with the long tails, also in this case you can use a fractal construction process to obtain a cartoon that considers the long-term dependence (Mandelbrot 2004). In the graphs in Figure 38 the initiator is still a growing trend line and the generator is a broken line of type up-down-up. In the case of the standard Brownian cartoon (i.e. the Brown-Bachelier cartoon of the third part) the width of each interval of the broken line was equal to the square root of the height, that is at the height elevated to ½. The exponent ½ is the Hurst coefficient. In fact, the standard Brownian motion has independent increases that are denoted by such values of *H*.

[49] Some criticisms have been made of the infinite memory assumption and the use of R/S analysis in general. For a report of these criticisms and some subsequent confutations see Nawrocki (2000).

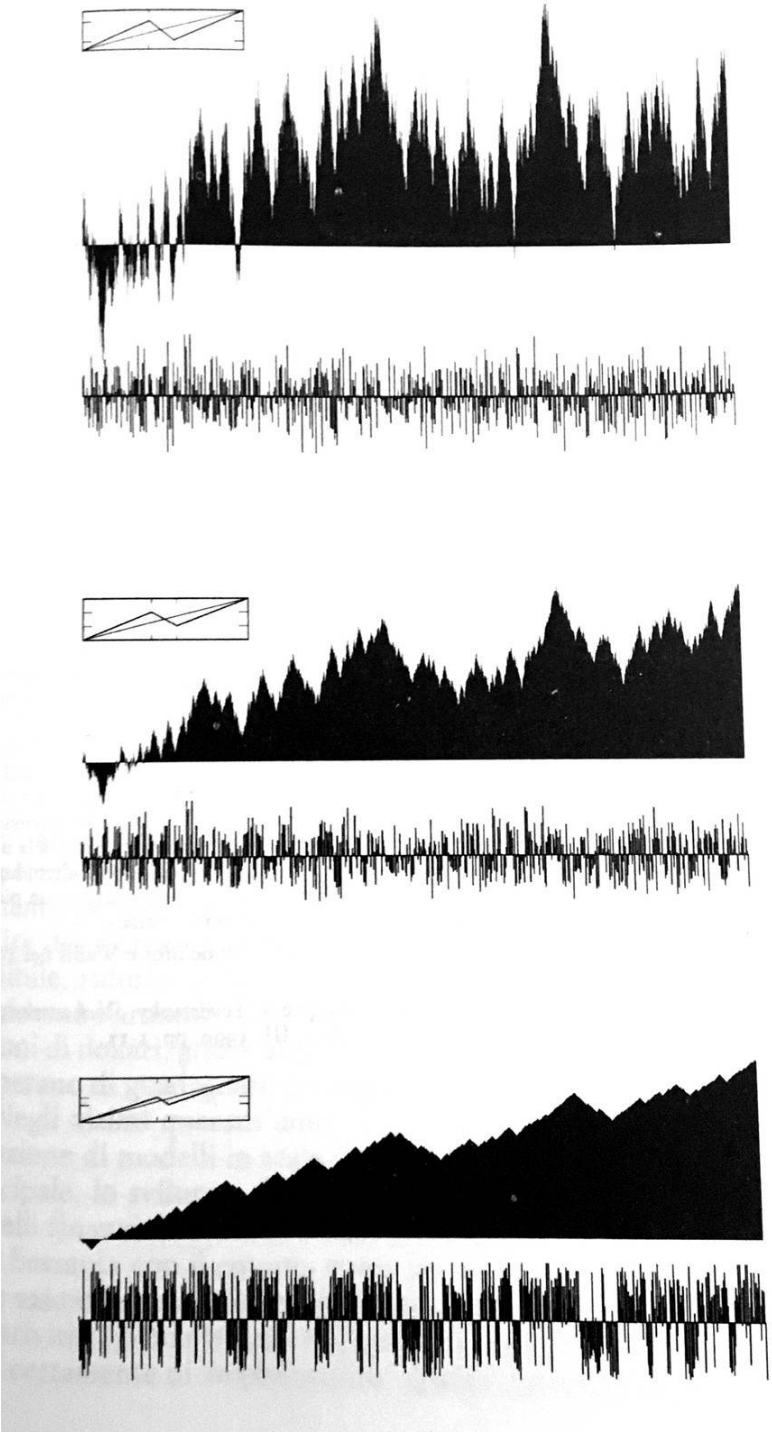

Figure 38. Cartoons with long-term dependence as H varies. From Mandelbrot (2004).

Of the three cases in the figure, in the first group the height of the range is equal to the width elevated to an exponent less than ½ and is therefore an anti-persistent time series, as you can see from the graph of the trend where the opposite increments always tend to compensate. The second group is the standard Brownian case, in which H=1/2 and the increments are independent; it's the Brown-Bachelier cartoon. In the third group the range height is equal to the width elevated at an exponent H>1/2 and the time series is therefore persistent, as shown by the graph of variations where you see that increments in the same direction tend to follow each other for a while.

4.3 The trading time

The last component of the model to be presented in the next chapter is the trading time. This concept is linked to the fact that in the financial markets the activity is concentrated in some periods, generating moments in which high volatility accumulates followed by times of low volatility. The trading time differs from the normal clock time because in some points undergoes a temporal acceleration and in others a slowdown. In particular, it is obtained by applying to the clock time a multifractal process, described below, which is able to imitate the volatility characteristic to gather in some periods.

Examples of concentration of elements in some points are recurring in nature. The strong gusts of wind, for example, tend to concentrate for a limited time, then are substituted by light breezes. Apart from the concept of time, it is possible to find analogous behaviors in the distribution of the oil or gold deposits, with some zones having them in abundance and others having none at all.

4.3.1 The multiplicative cascade

The multifractal process that allows the transformation of clock time into trading time is called *multiplicative cascade* and is based on a long series of multiplications[50].

As an example of such a process (Mandelbrot 2004), it is useful to consider time as a form of matter. As such, we can replace the time with a gold field, for the purposes of explanation. Gold, as mentioned before, is not distributed uniformly

[50] The process is multifractal because, as seen in the third part, a multifractal has the characteristic of reproducing itself in different parts following different scale factors. In addition, Mandelbrot (1997) found that multifractals reproduce well both long tails and long-term dependency.

around the world, but is concentrated in some areas, as well as in financial markets the activity is concentrated in some periods. Suppose we take a map of South Africa, in particular a cross-section of the ground cut along an axis oriented from west to east. Let's start with a low-resolution map, such as to divide the country into two areas, the western and the eastern. Approximately 60% of the gold fields are in the west, while 40% are in the east. If we look more closely and divide into two each part, we note that 60% of the western gold is in the west quarter, which therefore contains 36% of the total. In the second quarter there is therefore 40% of western gold, equal to 24% of the total. Continuing to look more closely, to divide the areas into smaller and smaller parts and to multiply in cascade the obtained values to get a percentage of the total, you get the following graph:

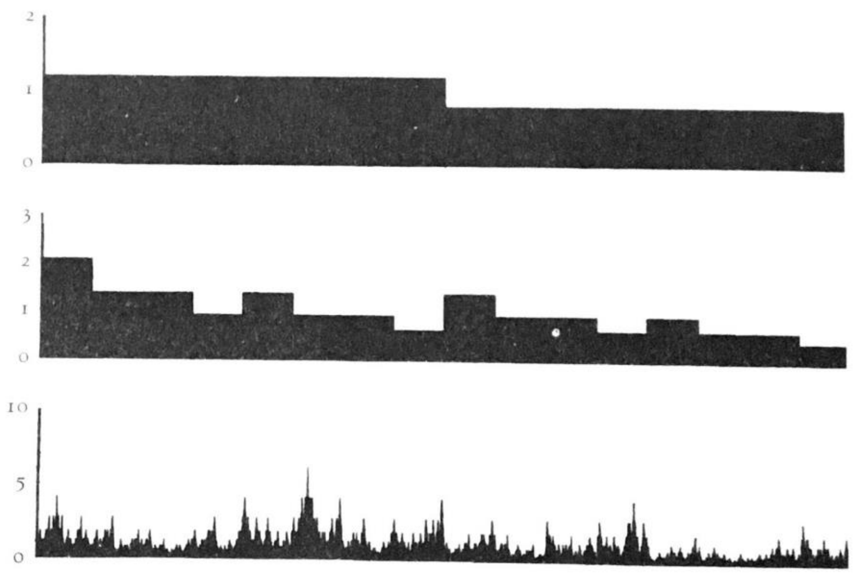

Figure 39. Multiplicative cascade in case of gold distribution. From Mandelbrot (2004).

It is a very irregular distribution, with some parts rich in gold and some very poor.

If the form of subdivided matter is not gold, but time, the trading time is obtained, characterized by an acceleration in the peaks and a slowing down in the depressions of the graph.

4.4 Multifractal Model of Asset Returns (MMAR)

In this last paragraph we will present the multifractal model for the description of the financial markets, as originally conceived by Mandelbrot (Calvet, Fisher and Mandelbrot 1997a).

As anticipated, the MMAR is a set of three main components: it incorporates the effects of long tails and long-term dependency and uses the trading time to combine them.

It will be described according to two perspectives: the simple one of the previously used cartoons and the one a little more complex of a rigorous mathematical model. The first is a simplified version of the second one, but it shares the same basic ideas.

4.4.1 Multifractal Model of Asset Returns: the multifractal cartoon

The multifractal cartoon is more complicated than the previous ones; it is a compound cartoon. It includes three generators: a parent generator and a mother generator that produce a new child generator that takes some characteristics from the former and some from the latter. Consider the following figure illustrating the construction process:

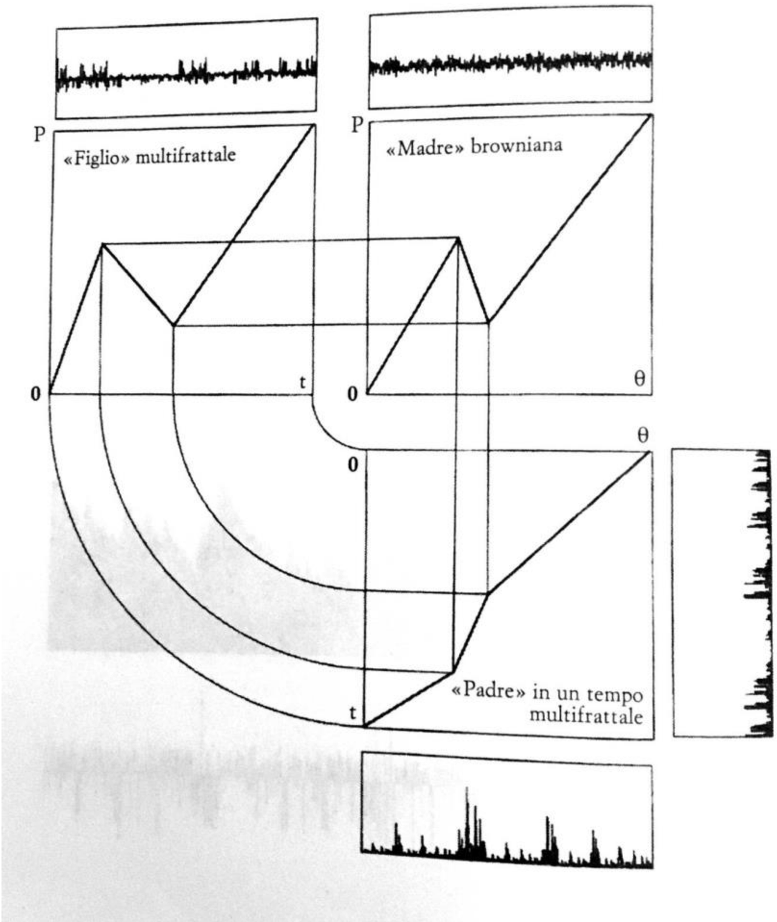

Figure 40. Multifractal cartoon. From Mandelbrot (2004).

The axes are marked with the letters *t*, θ and *P*; *t* indicates the clock time, *P* the price of a financial asset and θ the trading time.

The father takes the clock time and turns it into trading time, as we have seen earlier, by applying a multiplicative cascade. The mother instead takes the clock time and turns it into a price. The joint action of the two generates a child who takes the trading time from the father and converts it into a price according to the rules provided by the mother. The last step of the construction process is to use the child generator to get a fractal price diagram. The result, as you can see in the upper left of the figure, is a graph of variations that incorporates the effects of long tails (variability is wild) and long-term dependency (variability accumulates in some periods), obtained thanks to the use of the trading time.

4.4.2 Multifractal Model of Asset Returns: the mathematical model

The MMAR is a continuous-time multifractal process that incorporates long tails and long-term dependency shown by many financial time series. It is built composing a fractional Brownian motion with a multifractal process of time deformation, which transforms the clock time into trading time (the multiplicative cascade). The deformation of time produces long memory in volatility and implies that the statistical moments of price returns vary like a power law, that is that they have long tails (Calvet and Fisher 2001).

The model is characterized by a form of temporal invariance called *multiscaling* that applies to the moments of the prices distribution. This multiscaling property generates the multifractality of the model, as it is applied with different scale factors at different points.

There are two assumptions:

- Unlike long tails that characterize stable distributions as described in Mandelbrot (1963), the MMAR does not necessarily imply that variance is infinite.
- Unlike the fractional Brownian motion (Mandelbrot and Van Ness 1968), the MMAR shows long-term dependence in the absolute values of price increments (clustering effect), while price increments themselves may be unrelated.

The construction of the model (Calvet, Fisher and Mandelbrot 1997a) is as follows.

Let *P(t)* be the price of a financial asset at time *t*. We introduce the following notation:

$$X(t) = ln\,P(t) - ln\,P(0)$$

and we assume the following:

- **Assumption 1.** X(t) is a compound process:

$$X(t) \equiv B_H [\theta(t)]$$

where $B_H(t)$ is a fractional Brownian motion with Hurst exponent *H*, $\theta(t)$ is the stochastic trading time and *t* denotes the clock time.

- **Assumption 2.** The trading time $\theta(t)$ is the cumulative distribution function (CDF) of a multifractal measure defined in [0,T].

  In particular, $\theta(t)$ is a multifractal process with continuous and non-decreasing path and steady increments.

- **Assumption 3.** $B_H(t)$ and are $\theta(t)$ independent.

  Trading time $\theta(t)$ plays a crucial role in the MMAR. We note first that $\theta(0) = 0$ almost certainly since $X(0) = 0$ by definition. The Assumption 2 requires that $\theta(t)$ is the CDF of a random self-similar measure, such as the multiplicative cascade. The trading time $\theta(t)$ presumably makes the price X(t) multifractal, and we expect that $\tau_\theta(q)$ and $\tau_X(q)$, i.e. the time-scaling and prices-scaling functions are highly correlated. This intuition leads to the following theorem.

  Theorem. Under the assumptions [1] – [3], the X(t) process is multifractal, with scaling functions $\tau_X(q) \equiv \tau_\theta(H_q)$ and steady increments.

The above construction allows to generate a large class of multifractal processes, therefore to generate multiple financial graphs with the hypotheses analyzed during this section. By applying the Montecarlo method and creating a sufficiently large number of random fractal price diagrams, statistical tests can be carried out in order to obtain useful information to try to understand the market[51], although there is still a shortage of statistical methods applicable to models that are invariant both to scale and time[52].

---

[51] The MMAR is therefore an alternative to more traditional models such as ARCH.

[52] The first application of this model was to the Deutsche Mark/US Dollar pair and gave good results. See Calvet, Fisher and Mandelbrot (1997b) for details.

CONCLUSION

As stressed during this discussion, there is no doubt that the standard models describing how the markets work present many shortcomings, even conceptual ones. These can - at least - be reduced with the use of multifractal models, incorporating in particular the ideas of wild price changes (long tails) and long-term dependence. However, the application of fractal geometry to financial markets is a young and still incomplete science. For the future, the hope is that this trend will be taken forward and that we will be able to gain a deeper understanding of the markets, in such a way that money is kept more safe and that the economy can prosper more by anticipating and defusing in time the financial crises' traps. To make this possible, research will have to focus on some aspects that, however, are still without an answer such as:

- α and H.

  The two parameters expressing, respectively, the length of the tails and the degree of long-term memory do not yet have a univocal way to be calculated. The research carried out so far has found, for the same share, different values depending on the method employed.
- Risk.

  The most commonly used risk measurement tool is the standard deviation of Gaussian distribution. This parameter has been used in a wide range of fields, from portfolio management (Markowitz) to risk management of banking activities (Value at Risk, VaR). If the Gaussian distribution is not suitable for describing markets, the use of such a measure of risk is harmful and should be replaced with something else that incorporates the evidence of wild variations and long memory.
- Options.

The Black and Scholes formula incorporates the assumption of Normal distribution and therefore uses the standard deviation as a measure of risk. Here, too, the same reason of the previous point applies.

At the current state of art the fractal geometry is more a help for not losing dramatically huge fortunes in times of wild volatility, rather than a tool to earn money in the market (Mandelbrot 2004). As a consequence, to date, the most rational behavior is to integrate the two visions of the financial markets, that of Bachelier (1900) and that of Mandelbrot (1963).

APPENDIX 1

I. Random walk

A random walk is the formalization of the idea of taking successive steps in random directions and, mathematically, represents the simplest stochastic process (i.e. it is a Markov process[53]) of which the Wiener process is the most well-known representation. It is often approached to the Brownian motion, as the random walk is an example of a theoretical model within its intellectual framework.

Here is proposed a representation as developed by Bachelier, in his *Theorié de la Speculation* of 1900 (Bachelier 1900).

Be *t* a time variable and *x* the random variable that expresses the price of a financial security.

Divide the time into very small intervals Δt, then it is possible to consider that the price varies of the fixed quantity, and very small, Δx. Suppose that, at any time *t*, the prices $x_{n-2}, x_{n-1}, x_n, x_{n+1}, \cdots$, which differ from each other of the quantity Δx, have respective probabilities $p_{n-2}, \cdots, p_{n+2}$.

Knowing the probability distribution at time *t*, you can easily know the one at $t + \Delta t$. Suppose, for example, that the price x_n occurred at time *t*. At time $t + \Delta t$ the price could be x_{n+1} or x_{n-1} . The probability p_n, that the price x_n occurred at time *t*, can be decomposed into two probabilities, at time $t + \Delta t$. The price x_{n-1} with probability $p_n/2$ and the price x_{n+1} with probability $p_n/2$.

If x_{n-1} had occurred in $t + \Delta t$, it would be because, in *t*, the price x_{n-2} or the price x_n would have occurred. The probability of the price x_{n-1} , at the time $t + \Delta t$ would then be $(p_{n-2} + p_n)/2$. That of the price x_n would be, at the same time, $(p_{n-1} + p_{n+1})/2$. That of the price x_{n+1} would then be $(p_n + p_{n+2})/2$, etc...

During the time interval Δt the price x_n has, in some way, transmitted to the price x_{n+1}, the probability $p_n/2$; x_{n+1} transmitted towards the price x_n, the probability $p_{n+1}/2$.

[53] A Markovian process is a random process in which the probability of transition, that is the probability of transition to a certain system state, depends only on the status of the immediately preceding system, and not by how it came to this state. It is characterized by absence of memory.

If $p_n > p_{n+1}$, the change in probability is $(p_n - p_{n+1})/2$ from x_n towards x_{n+1}.

«It can therefore be said that each price radiates, during a time element, a quantity of probability to its price neighbors proportional to the difference in their probabilities. The previous theory, by analogy with certain physical theories, can be called *Probability Radiation Law*» concludes Bachelier.

The previous treatment is developed according to a change in the prices of a discrete quantity in the various *t*. For a treatment in the continuous case, you can use Fourier's heat equation (see pp. 19-21 of Bachelier, 1900).

II. Normal distribution

The normal distribution, or gaussian, is a distribution of continuous random variable and it's well known because it adapts well to the distribution of numerous natural phenomena and its mathematical use is simple.

Its probability density function is: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

It depends on only two parameters, the mean and the variance, and is indicated with $N(\mu, \sigma^2)$.

The mean value μ represents the value of the data around which the function is symmetrically divided, while the variance σ^2 indicates the degree of dispersion of the sample data from the mean value. σ takes the name of standard deviation. Growing values of σ^2 correspond to flatter bell curves.

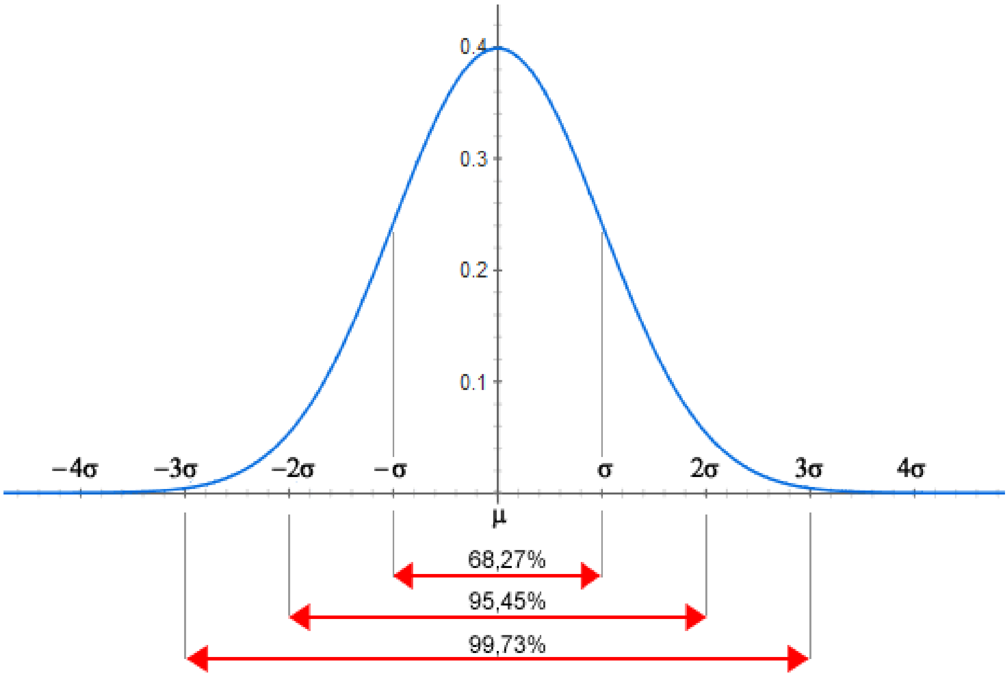

Figure 41. Cumulative distribution function of a Normal. From Wikipedia.

The following properties are valid:

- It is "stable", according to the expression coined by Paul Levy, that is the linear combination of a normally distributed variable, is also a normal variable and the sum of more normal random variables has also normal distribution.
- For the central limit theorem, with *n* equally distributed and independent random variables, for $n \to \infty$ the sample average tends to be distributed as a normal variable.

For ease of probability calculation, it is used to standardize the normally distributed variable X in another normal variable Z of mean 0 and variance 1, which is called standard normal variable. The random variable Z is defined as:

$$Z = \frac{X - \mu}{\sigma}$$

In this way you can use the probability tables, which display values for a standard normal variable.

An interesting feature of the normal is that almost all data observations are within 3 standard deviations, i.e. in the range $\mu \pm 3\sigma$. As you notice from the image above the probability that the data will deviate from the mean for x standard deviations (with x=1,2,3) is:

- 68.27% that data are observed within 1 standard deviation
- 95.45% that data are observed within 2 standard deviations
- 99.73% that data are observed within 3 standard deviations

It follows that when using the normal to describe the distribution of financial securities prices, one expects abnormal events to occur in the distribution queues (such as financial crises) although they may exist, are so infinitely likely to occur that it is negligible.

III. Efficient Market hypothesis (EMH)

Below a mathematical model describing the complete reflection of information in the prices of any listed assets is represented, as set out by Eugene Fama in his doctoral thesis (Fama 1970):

$$E(\tilde{p}_{j,t+1} \mid \phi_t) = [1 + E(\tilde{r}_{j,t+1} \mid \phi_t)] \cdot p_{j,t}$$

where

$p_{j,t}$= share *j* price in *t*,

$p_{j,t+1}$= share *j* price in *t+1 (*with reinvestment of each intermediate revenue flow from the security),

$r_{j,t+1}$= one-period percentage return defined as $\frac{p_{j,t+1} - p_{j,t}}{p_{j,t}}$,

ϕ_t= information set taken fully reflected into the price of the security in *t*.

The tilde indicates that p_{t+1} and r_{t+1} are random variables in *t*, not previously known.

The equation implies that, as written by Fama, «whatever expected return model is assumed to apply, the information in ϕ_t is fully utilized in determining equilibrium expectations returns. And this in the sense in which ϕ_t is "fully reflected" in the formation of the price $p_{j,t}$». The model therefore indicates that the expected price in *t+1*, given the information known in *t*, is equal to the price in *t* capitalized by a rate *r* (tilde) not yet known in *t*, but that depends only on the current price and the one in *t+1*, which varies with the variation of the information received in *t*.

IV. Black-Scholes' formula

As already mentioned, the formula, which allows to calculate the value of a European type call option, is:

$$c = S_0 N(d_1) - X e^{-rT} N(d_2)$$

where the first addend is the product between the today value S_0 of the underlying asset (called spot price) and the probability $N(d_1)$ of a variable d_1 normally distributed; the second addend represents the product between the exercise price *X* (strike price) of the option (discounted to the present value with a continuously compound rate *r*) and the probability $N(d_2)$ of a variable d_2 also distributed normally[54]. Formulas for calculating *d1* and d_2 are:

$$d_1 = \frac{ln(\frac{S_o}{X}) + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}.$$

[54] Note that d_1 and d_2 are normally distributed as price changes are assumed to be normal variables.

To find the corresponding value of a put option, you can use the put-call parity theorem, which relates the values of the two types of options knowing the spot price and the strike price referred to the underlying:

$$p + S_0 = PV(X) + c.$$

APPENDIX 2

I. Topological dimension

Let X be a topological space. An open cover $\{U_i\}$ of X is a collection of open sets U_i whose union is all X. Its refinement is another open cover $\{V_j\}$ such that each V_j is contained in at least one U_i.

The topological dimension of X is the smallest integer n such that each open cover X has a refinement where each point is contained in at most $n+1$ sets.

The following are two different ideas for calculating the topological dimension:

- Poincarè idea: it is a recursive method applicable to any connected topological space. Placed, by convention, equal to zero the point dimension, a space has dimension n if the connection property can be destroyed "by subtracting" to space a subspace of dimension $n-1$. The line *(*with the usual topology) will thus have dimension 1, because you can "disconnect" it by eliminating a point; the plane will have dimension 2, because you can "disconnect" it by eliminating a line, and so on.
- Lebesgue idea: we start from the observation that if you want to cover a straight line with open intervals of finite length, you have to make overlaps. However, given any such covering and eliminating any unnecessary intervals, it is always possible to ensure that each point of the line is covered by no more than two intervals. Repeating the same reasoning in the plan, using some disks instead of the previous intervals, we realize that it is not possible to limit to two the number of disks covering a generic point. In order to make the construction possible, it must be agreed that, somewhere, the number of overlapping disks rises to at least three. The general result affirms that in n-dimensional Euclidean space, covers consisting of open spheres of finite diameter and such that each point of space is covered by no more than k spheres, exist if $k=n+1$, so the dimension is $n=k-1$.

APPENDIX 3

I. Power laws

A power law is a mathematical relationship like:

$$f(x) = ax^{-\alpha}$$

α is called scale exponent.

a. In the probabilities

Within the probability distribution field, a distribution that obeys the power law is called power law distribution.

The peculiarity of these distributions is the invariance of scale, that is they do not change any scale you consider. Suppose income distribution follows a law of power. To say that distribution is scale invariant means that if every four individuals with an annual income of ten thousand euros exists one with an income of twenty thousand, then there will be one person who earns 2 billion for every four people who earn one.

The power law relationship generates long-tails in the distribution.

The scale invariance is a necessary and sufficient condition to be a power law, so vice versa if we know that a distribution is power law, we know that it is invariant to scale. In mathematical terms a distribution *p(x)* is invariant to scale if it is true that *p(bx)=f(b)p(x)*, with *f(b)* multiplicative constant.

Their generic form is $p(x)\ proportional\ to\ L(x)x^{-\alpha}$, with $L(x) = c$ if *L(x)* is constant.

For any value of α the distribution diverges when $x \to 0$, therefore it is possible to impose a minimum value x_{min}. Given the α exponent, when *x* is a continuous variable then the normalization constant *C* is given by*:*

$$C = 1/\int_{x_{min}}^{\infty} x^{-\alpha} dx$$

from which

$$p(x) = \frac{\alpha - 1}{x_{min}} (\frac{x}{x_{min}})^{-\alpha}$$

which is the CDF of a power law distribution.

An estimator for α can be found using the method of the maximum likelihood.

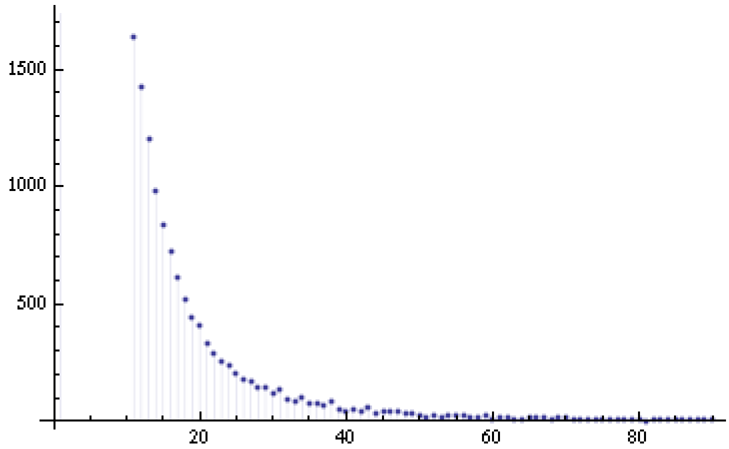

Figure 42. Power law distribution with α=3. From Wikipedia.

b. How to recognize a power law distribution

A peculiarity of power law distributions is that if they are represented on log-log scale they are straight lines, because with a logarithmic transformation the power law relation becomes linear, in fact:

$$log\, p(x) = log\,(C) - \alpha\, log\,(x)$$

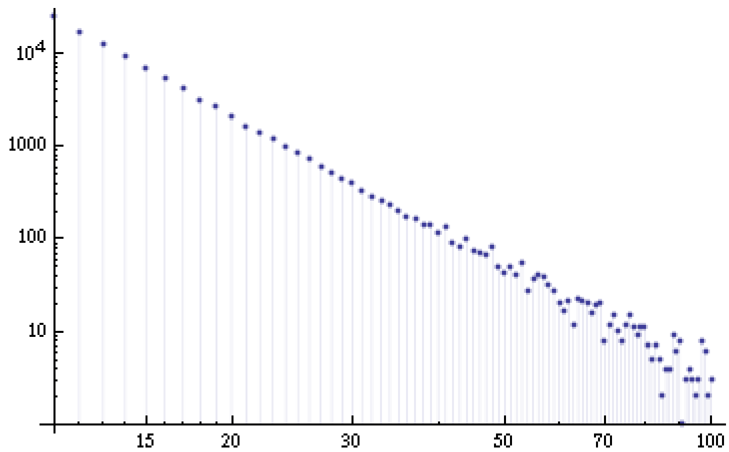

Figure 43. Power law distribution with α=3 represented in log-log scale. From Wikipedia.

It is possible therefore to approximately identify a power law also in visual way, passing to the logarithms and verifying if the resulting form is a straight line.

BIBLIOGRAPHY

Bachelier, L. (1900), *Théorie de la spéculation*, in *Annales scietifiques de l'Ecole Normale Superieure*, 17 (3), pp. 21-86.

Baciotti, A., *Quando diciamo che lo spazio ha dimensione tre, cosa intendiamo?*, Dipartimento di Matematica del Politecnico di Torino.

Black, F. (1989), *How we came up with the option formula*, in *Journal of Portfolio Management,* 15 (2), pp. 4-8.

Black, F., e Scholes, M. (1973), *The pricing of options and corporate liabilities*, in *Journal of Political Economy,* 81 (3), pp. 637-654.

Calvet, L. and Fisher, A. (2001), *Multifractality in Asset Returns: Theory and Evidence*.

Calvet, L., Fisher, A. and Mandelbrot, B. (1997a), *A Multifractal Model of Asset Returns,* in *Cowles Foundation Discussion Paper*, 1164.

Calvet, L., Fisher, A. and Mandelbrot, B. (1997b), *Multifractality of Deutschemark / US Dollar Exchange Rates*, in *Cowles Foundation Discussion Paper*, 1165.

Camerer, C.F., and Thaler R.H. (1995), *Anomalies: Ultimatums, dictators and manners*, in *Journal of Economic perspectives,* 2, pp. 209-219.

Fama, E. (1970), *Efficient Capital Markets: A Review of Theory and Empirical Work*, in *The Journal of Finance*, 25 (2), pp. 383-417.

Fama, E., and French, K. (1988), *Dividend yields and expected stock returns*, in *Journal of Financial Economics*, 22 (1), pp. 3-25.

Ferson, W., and Harvey, C. (1991), *The Variation of Economic Risk Premiums*, in *Journal of Political Economy*, 99 (2), pp. 385-415.

Finnerty, J. (1978), *The Chicago Board Options Exchange and Market Efficiency*, in *Journal of Financial and Quantitative Analysis*, 13 (1), pp. 29-38.

Garcin, M. (2016), *Estimation of time-dependent Hurst exponents with variational smoothing and application to forecasting exchange rates*, in *Archives-ouvertes.fr*.

Grimaldi, M., and De Grauwe, P. (2003), *Bubbling and Crashing Exchange Rates*, in *CESifo Working Paper*, 1045.

Hurst, E. H. (1951), *Long-term storage capacity of reservoirs*, in *Transactions of the American Society of Civil Engineers*, 116, pp. 770-799.

Kahneman, D., and Tversky, A. (1986), *Rational Choice and the Framing of Decisions*, in *The Journal of Business*, 59 (4), pp. 251-278.

Kruizenga, J. (1956), *Put and call options: a theoretical and market analysis*, MIT Press, Cambridge (Mass.).

Malkiel, B. (1973), *A Random Walk Down Wall Street*, W. W. Norton & Company Inc., New York.

Mandelbrot, B. (1963), *The Variation of Certain Speculative Prices*, in *The Journal of Business*, 36 (4), pp. 394-419.

Mandelbrot, B. (1972), *Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence*, in Rosenblatt and Van Atta (edited by), *Statistical Models and Turbulence*, Springer, pp. 333-351.

Mandelbrot, B. (1997), *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*, Springer, New York.

Mandelbot, B., and Hudson R. (2004), *The Misbehavior of Markets: A fractal view of risk, ruin and reward*, Basic books.

Mandelbrot, B. and Van Ness, J. (1968), *Fractional Brownian Motions, Fractional Noises and Applications*, in *SIAM review*, 10 (4).

Mandelbrot, B., and Wallis, J. R. (1969), *Computer experiments with fractional Gaussian noises*, in *Water Resources Research*, 5, pp. 228-267.

Markowitz, H. (1952), *Portfolio Selection*, in *The Journal of Finance*, 7 (1), pp. 77-91.

Nawrocki, D. (2000), *R/S Analysis and Long Term Dependence in Stock Market Indices*, in *Managerial Finance*, 21 (7).

Pareto, V. (1905), *Manuale di Economia Politica*, Società Editrice Libraia, Milano.

Peters, E. E. (1991), *Chaos and order in the capital markets*, John Wiley and Sons.

Samuelson, P. (1965), *Rational Theory of Warrant Pricing*, in *Industrial Management Review*, 6, pp. 13-31.

Samuelson, P. (1974), *Challenge to judgment*, in *The Journal of Portfolio Management Fall*, 1 (1), pp. 17-19.

Schoutens, W. (2003), *Lévy Processes in Finance: Pricing Financial Derivatives*, John Wiley & Sons Inc.

Sharpe, W. (1964), *Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*, in *Journal of Finance*, 19, 425-442.

Sprenkle, C. (1961), *Warrant prices and Indications of Expectations*, in *Yale Econ. Essays*, 1, pp. 179-232.

Williams, J. (1938), *The Theory of Investment Value*, Harvard University Press, Cambridge (Mass.).

SITOGRAPHY

http://www.mat.unimi.it/users/alzati/Geometria_Computazionale_98-99/apps/frattalip/teoria.html [11/10/2019].

http://www.miorelli.net/frattali/matematica.html [11/10/2019].

http://www.performancetrading.it/Documents/McAnalisi/McA_Analisi_Fondamentale.htm [21/09/2019].

https://cla.unisalento.it/c/document_library/get_file?folderId=22481013&name=D [08/10/2019].

https://en.wikipedia.org/wiki/Efficient-market_hypothesis [17/09/2019].

https://en.wikipedia.org/wiki/Fractal_dimension [14/10/2019].

https://en.wikipedia.org/wiki/Hausdorff_dimension [14/10/2019].

https://en.wikipedia.org/wiki/Hurst_exponent [10/11/2019].

https://en.wikipedia.org/wiki/Julia_set [12/10/2019].

https://en.wikipedia.org/wiki/Mandelbrot_set [12/10/2019].

https://en.wikipedia.org/wiki/Minkowski%E2%80%93Bouligand_dimension [15/10/2019].

https://en.wikipedia.org/wiki/Power_law [13/11/2019].

https://en.wikipedia.org/wiki/Stable_distribution [09/11/2019].

https://it.wikipedia.org/wiki/Aspettative_razionali [09/10/2019].

https://it.wikipedia.org/wiki/Dimensione_topologica [14/10/2019].

https://it.wikipedia.org/wiki/Insieme_di_Cantor [12/10/2019].

https://it.wikipedia.org/wiki/Legge_di_potenza [13/11/2019].