Statistical Physics in the Modeling of Financial Markets
M1 Project

Esteban Guevara Hidalgo
Erasmus Mundus Master in Complex Systems
esteban_guevarah@yahoo.es

August 10, 2011

Contents

Chapter 1

Introduction

During the past years, physicists have achieved important results in the field of phase transitions, statistical mechanics, nonlinear dynamics, and disordered systems. Concepts like power law, scaling, random processes and unpredictable time series are present in these fields, and they are used as interpretation of the underlying physics. By other hand, financial markets exhibit several properties that characterize complex systems. They are open systems in which many subunits interact nonlinearly in the presence of feedback.

More than one hundred years ago, these concepts were applied successfully in fields far from the natural sciences. Concepts like power law distributions and random walk were first used in social sciences. Pareto [1] investigated the statistical character of the wealth of individuals in a stable economy modeling it through a power law of the type $y \sim x^{-1.5}$ (where y is the number of people having income x or greater than x). In 1900, a Ph.D. student of Poincare developed the first formalization of a random walk. In his thesis "Speculation Theory" Bachelier [2] suggested a probabilistic description of price fluctuations in the financial market and developed the mathematics of the Brownian motion to model the time evolution of asset prices. He determined the probability of price changes by writing down (what is now called) the Chapman-Kolmogorov equation and recognizing that (what is now called) a Wiener process satisfies a diffusion equation (rediscovered years latter by Einstein [3]). The first theoretical description of a random walk was performed in 1905 by Einstein [3] and in latter years its mathematics was made more rigorous by Wiener [4].

Bachelier work was not recognized by the scientific community at that time, perhaps because of its application to financial markets. His theory was actually forgotten until 1944 when Ito used it as motivation to introduce his calculus and a variant of the Brownian motion which is the geometric Brownian motion [5, 6]. Since the 1950s mathematicians began to show interest in the modeling of stock market prices. Bachelier original proposal that the price changes are Gaussian distributed was replaced to a model in which stock prices are log-normal distributed (i.e. the differences of the logarithms of prices are Gaussian distributed) [7]. In 1973, Black, Scholes and Merton [8, 9, 10] used the geometric Brownian motion to construct a theory for determining the price of stock options. This theory nowadays represents the milestone of mathematical finance although is clear that the model needs correction in its application because provide only a first approximation of what is observed in real data. For this reason several alternative models have been proposed [7]. One of them is the Maldebrot's hypothesis [11] which states that prices changes are Levy stable distributed.

Apart from the modeling of stock market prices, concepts like turbulence have been used in order to understand the behavior of financial markets. Mantegna and Stanley [12, 13] developed a parallel analysis of price dynamics and fluid velocity. Specifically, they analyzed the time evolution of the S&P 500 index and the velocity of a turbulent fluid at high Reynolds number. By other hand, Ghashghaie et al. [14] proposed a formal analogy between the velocity of a turbulent fluid and the currency exchange rate in the foreign exchange market. The results of these investigations are discussed in Appendix B: Financial Markets and Turbulence.

Another field of interest is the simultaneous investigation of several stock-price time series. The presence of correlations or anticorrelations between stocks has been long known and they play an important role in the determination of selecting the most efficient portfolio [15, 16]. Mantegna [17] studied how important are these correlations to detect the amount of synchronization present in the dynamics of a pair of stocks traded in a financial market. This approach is discussed in Appendix C: Correlations and Anticorrelations between Stocks.

Measuring and controlling financial risks is a major concern in the economic world. Several measures of risk have been introduced [10, 18] in order to reduce losses and allow these risks even to be traded. However, the classical definitions are considered "weak" to explain "rare events" where true financial risk resides. In Appendix D: Risk Measures, we review some classical ideas on financial risk to illustrate their weakness and then we will explore some theoretical ideas which could explain these "rare events".

Some of the areas that are undergoing investigation by physicists consists in the complete characterization of the stochastic process of price changes of a financial asset. For example: the shape of the distribution of price changes [18, 19, 20, 21], the temporal memory [12, 22, 23] and the higher order statistical properties [24, 25, 26]. Also there is a big interest in the development of a theoretical model that is able to encompass all the essential features of real financial markets [27, 28, 29, 30, 31, 32, 33, 34]. Other areas of investigation deals with rational pricing of a derivative product when some of the canonical assumptions of the Black-Scholes model are relaxed [18, 35, 36] and with optimal portfolio selection [37, 38, 39, 40].

In the present report we focus in the modeling of stock market prices. We start with a review of the statistical properties of financial time series and stochastic process, and finally, we discuss how this concepts have been used in order to analyze and model the financial markets (Black-Scholes theory). For an introduction in probability theory refer to Appendix A, and for a discussion about turbulence, correlation and risk refer to appendixes B, C, and D.

This project was developed with Latex and Matlab.

Chapter 2

Statistics of Real Prices

Nowadays is quite easy to obtain access to financial databases containing thousands of asset time series. This fact allow us to analyze their statistical features. The goal in this chapter is to present a brief review of the statistical properties of financial time series. For our analysis, we used Matlab and the package hist_stock_data [41]. This package uses the Matlab function urlread to access to the financial databases in Yahoo! Finance website and download and sort historical stock data for a user-specified time period. For example, the code bellow will provide us the price of the stock index S&P 500 (GSPC) during the time period January 1992 (01011992) to January 2002 (01012002)

$$\begin{aligned} price &= hist\_stock\_data('01011992','01012002','\hat{}GSPC'); \\ y &= price.Close(end : -1 : 1); \end{aligned}$$

The chosen stock index for our analysis is the Standard and Poor's 500 (S&P 500) US stock index during the time period January 1992 to January 2002. In this period the index rose from 392 to 1134 points.

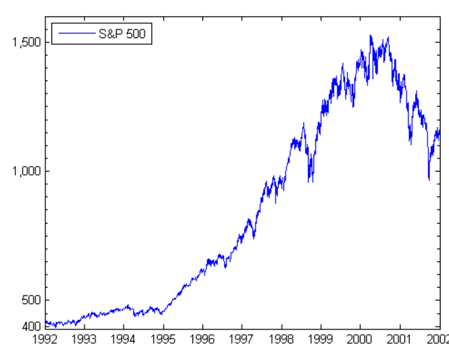

Figure 1: Standard and Poor's 500 (S&P 500) US stock index from January 1992 to January 2002. The data was extracted from the page finance.yahoo.com through the Matlab function urlread.

2.1 Price Increments and Returns

The absolute price increment δx between two instants separated by a time interval τ is defined as

$$\delta x_k = x_{k+1} - x_k = x(t+\tau) - x(t) \tag{2.1}$$

where $x_k = x(t)$ is the price of a specific asset at time t ($t \equiv k\tau$). The price return (or relative price increment) is defined as

$$\eta_k = \frac{\delta x_k}{x_k} \approx \log x_{k+1} - \log x_k \tag{2.2}$$

The last equation holds for τ sufficiently small so that the returns over that time scale are small.

The peaks in a return chart gives us information of the moments of high volatility of an asset price. The price returns for the studied asset are shown in figure 2

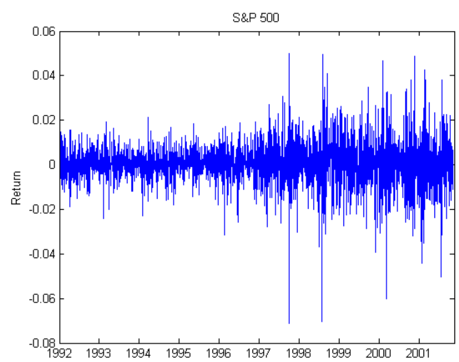

Figure 2: Price Returns for the S&P500 during the period analyzed. The peaks show moments of high volatility.

In the whole modern literature it is postulated that the relevant variable of study is the price return instead of the price increment [10]. As we will see, this is related with the fact that the prices are log-normal distributed instead of Gaussian distributed.

2.2 Autocorrelation

The correlations between price increments are measured by the temporal two-point correlation function $C^\tau(k,l)$ which is defined as

$$C^\tau(k,l) = \frac{< \eta_k \eta_{k+l} >_e}{< \eta_k^2 >_e} \tag{2.3}$$

where $< ... >_e$ is the empirical average over the whole time series. For uncorrelated increments, the correlation function should be equal (or around) zero for

$k \neq l$ (with an empirical RMS equal to $\sigma = 1/\sqrt{N}$, where N is the number of independent points used for the computation). The correlations between price increments for the S&P 500 are shown in figure 3.

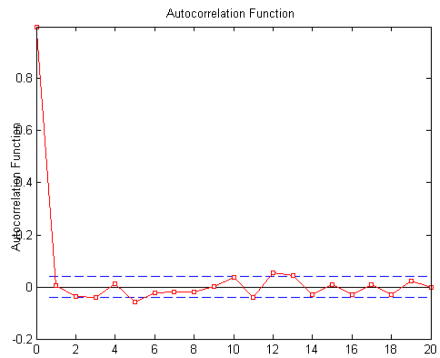

Figure 3: Autocorrelation function for the S&P500 during the period analyzed.

2.3 Price Variogram

Another way to observe the presence or absence of linear correlations is to study the variogram of log-change prices which is defined as

$$V(l) =< (\log x_k - \log x_{k+1})^2 >_e \tag{2.4}$$

A variogram measures how much in average the logarithm of price differs between two instant of times. If the returns η_k have zero mean and are uncorrelated, the variogram $V(l)$ grows linearly with the lag l, with a pre-factor equal to the volatility (as can seen in figure 5).

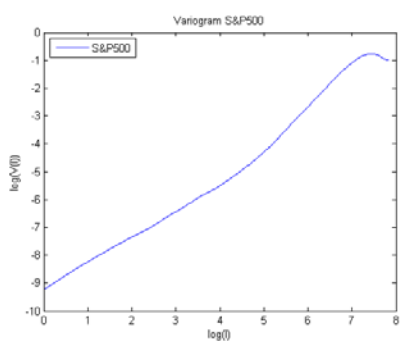

Figure 4: Variogram for the S&P500 index during the period analyzed.

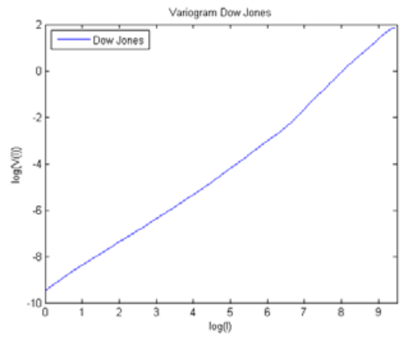

Figure 5: Variogram for the Dow Jones index for the period between the years 1950 and 2000. Except for the few last points the variogram shows no sign of saturation.

2.4 Power Spectrum

Another alternative way of presenting the same results is through the power spectrum which is defined as

$$S(w) = \frac{1}{N} \sum_{k,l} < \eta_k \eta_{k+l} > e^{iwl} \tag{2.5}$$

When the spectrum is flat is called "white noise", where all the frequencies are represented with equal weights. This corresponds to uncorrelated increments. The power spectrum for the S&P 500 is shown in figure 6.

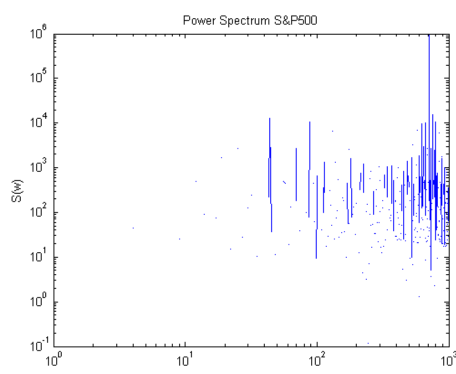

Figure 6: Power spectrum $S(w)$ for the S&P500.

2.5 Price Correlogram

The correlation function (or correlogram) of the variables η_k (the return) is defined as

$$C(l) =< \eta_{k+l}\eta_k > - < \eta >^2 \quad (2.6)$$

The correlogram for the returns of our case under analysis is shown in figure 7.

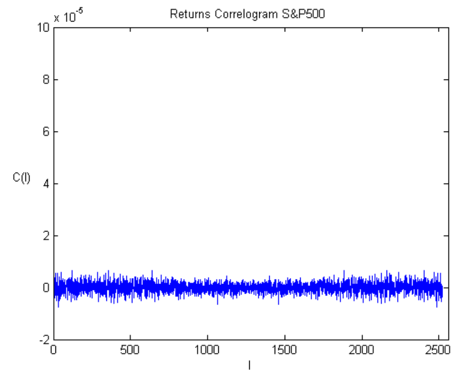

Figure 7: Correlogram for the returns of the S&P 500.

2.6 Hurst Exponent

The Hurst exponent measures the average span of the fluctuations. We can define the average distance between the high and low in a window of size $t = l\tau$ as

$$\widetilde{H}(l) =< \max(x_k)_{k=n+1,n+l} - \min(x_k)_{k=n+1,n+l} >_n \quad (2.7)$$

The Hurst exponent H is defined from $\widetilde{H}(l) \propto l^H$. When the increments are Gaussian $H = 1/2$. For TLD distribution, one finds $\widetilde{H}(l) \propto Al^{1/u}$ (for $l \ll N$) and $\widetilde{H}(l) \propto \sqrt{D\tau l}$ (for $l \gg N$) with $1 < u < 2$. The exponent evolves from $H = 1/u$ to $H = 1/2$. For fractional Brownian motion $H = 1 - \nu/2$.

2.7 Distribution of Returns over Different Time Scales

The previous results are compatible with the simplest scenario where the price returns are beyond a certain correlation time of a few tens of minutes on liquid markets. A much finer test of this assumption consists in studying directly the probability distributions of the log-price differences $\log(x/x_0) = \sum \eta_k$ on different time scales $N = T/\tau$. If the increments were iid, then the distributions on different time scales can be obtained from the one pertaining to the elementary time scale τ. More precisely $P(\log x/x_0, N) = [P_1]^N$, where P_1 is the distribution of elementary returns.

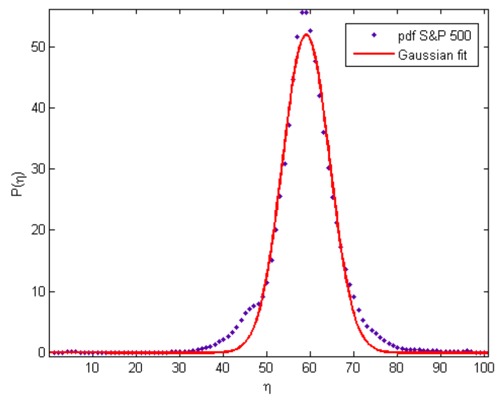

Figure 8: Pdf for S&P 500 and a Gaussian fit.

From the analysis of the distribution of the returns over different time scales have been observed the next facts [7, 10]: The price returns are uncorrelated beyond a time scale of a few tens of minutes in liquid markets. On shorter time scales strong correlation effects are observed. For time scales of the order of days, weak residual correlations can be detected in long time series of stocks and index returns. Price returns have strongly non Gaussian distributions (as can be seen in figure 8) but can be fitted by truncated Levy or student distribution. The tails are much fatter than those of a Gaussian distribution and can be fitted by a Pareto (power law) tail with and exponent μ from 3 to 5 in liquid markets. This exponent can be smaller in less mature markets (such as emerging country markets) where pure Levy truncated can be observed.

Chapter 3

Stochastic Processes

In the last chapter we studied how to access to the information available in the web about some specific asset and how to analyze their statistical properties. A widely accepted belief in financial theory is that these time series are unpredictable. This constitutes the cornerstone of the description of price dynamics as stochastic processes [7, 10]. Chaos theory has shown that these unpredictable time series can arise from deterministic nonlinear systems [42]. The results in physical and biological systems have motivated studies to determine if the time evolution of asset prices in financial markets might be due to underlying nonlinear deterministic dynamics of a number of variables [7]. In general, it is chosen to work within a paradigm that asserts price dynamics are stochastic processes, although it cannot be ruled out that financial markets follow chaotic dynamics [6]. This is motivated by the observation that the time evolution of asset prices depends on all the information affecting the investigated asset.

Since Bachelier our models for stock market prices have changed from the initial idea that price changes are Gaussian distributed, to log-normal distributed and then to Levy distributed (as we mention in the introduction). Asserting that price changes are Gaussian distributed is equivalent to say that prices are performing a Brownian motion and log-normal distributed, a geometrical Brownian motion. The last one is also the main assumption in the Black - Scholes theory for option pricing. In the present chapter, we start to accomplish our objective of modeling financial markets. In order to do that, we will introduce the concepts of random walk, Brownian and geometric Brownian motion and how these concepts have been used to model the stock market price changes.

3.1 Introduction

One fundamental characteristic of a stochastic (or random) process is the uncertainty in its future evolution. In stochastic processes, probability can evolve in time. Formally speaking, a stochastic process is a family $(X_t)_{t\in I}$ of random variables, where I can be discrete or continuos. The distribution function of X

at time t is defined as

$$p(x,t) = \int_{\Omega} \delta(x - X_t(\omega)) d\mu(\omega) \tag{3.1}$$

and the n-point distribution function is defined as

$$p_n(x_1, t_1; ...; x_n, t_n) = \int_{\Omega} \delta(x_1 - X_{t_1}(\omega))...\delta(x_n - X_{t_n}(\omega)) d\mu(\omega) \tag{3.2}$$

With help of equation (3.2), we can define the time dependent moments

$$\langle X(t_1)...X(t_n) \rangle = \int_{R^n} dx_1...dx_n x_1...x_n p_n(x_1, t_1; ...; x_n, t_n) \tag{3.3}$$

and the covariance matrix for two stochastic processes

$$Cov\left[X_i(t_1), X_j(t_2)\right] = \langle X_i(t_1) X_j(t_2) \rangle - \langle X_i(t_1) \rangle \langle X_j(t_2) \rangle \tag{3.4}$$

The diagonal elements in equation (3.4) are called autocorrelation functions and those off-diagonal are the cross-correlation functions.

A stochastic process is called stationary iff for all n

$$p_n(x_1, t_1 + \Delta t; x_2, t_2 + \Delta t; ...; x_n, t_n + \Delta t) = p_n(x_1, t_1; x_2, t_2; ...; x_n, t_n) \tag{3.5}$$

3.1.1 Martingale Processes

In a "fair game" (whatever was the history of the game), the probability of the outcomes is always the same (e.g. in a coin toss). A sequence $(X_n)_{n \in N}$ is called absolutely fair when for all $n = 1, 2...$ we have $E[X_1] = 0$ and $E[X_{n+1}|X_1, ..., X_n] = 0$. If we define another sequence of random variables $(Y_n)_{n \in N}$ by $Y_n = E[Y_1] + X_1 + ... + X_n$ we have $E[Y_{n+1}|Y_1, ..., Y_n] = E[Y_{n+1}|X_1, ..., X_n] = Y_n$.

A sequence $(Y_n)_{n \in N}$ is a martingale iff $E[Y_{n+1}|Y_1, ..., Y_n] = Y_n$, submartingale iff $E[Y_{n+1}|Y_1, ..., Y_n] \geq Y_n$ and supermartingale $E[Y_{n+1}|Y_1, ..., Y_n] \leq Y_n$. This means that the conditional expected value of an observation at some time $n + 1$ (given all the observations up to some earlier time n) is equal to the observation at that earlier time n. In other words, the best estimator for the next value given all the information obtained from the past values, is the actual value. If the process is a martingale, the expectation value of the increments is zero (its changes are a fair game).

3.1.2 Markov Processes

A Markov process is a random process whose future probabilities are determined by its most recent values. Formally speaking, for all n and $t_1 < t_2 < ... < t_n$

$$p(x_n, t_n | x_{n-1}, t_{n-1}, ..., x_1, t_1) = p(x_n, t_n | x_{n-1}, t_{n-1}) \tag{3.6}$$

This means that, in order to know the probability of occurrence of (x_n, t_n), we only need to know the actual state (x_{n-1}, t_{n-1}) of the system and not the whole history.

Chapman-Kolmogorov, Master and Fokker-Planck equations

The consistency equation for the conditional probabilities of a Markov process is know as the Chapman-Kolmogorov equation which is given by (for $n = 3$, $t_3 \geq t_2 \geq t_1$)

$$p(x_3, t_3|x_1, t_1) = \int_{-\infty}^{\infty} p(x_3, t_3|x_2, t_2)p(x_2, t_2|x_1, t_1)dx_2 \tag{3.7}$$

The integro-differential version of the Chapman-Kolmogorov equation is the master equation

$$\frac{\partial}{\partial t}p(x,t) = \int p(x',t)\omega(x|x')dx' - \int p(x,t)\omega(x'|x)dx' \tag{3.8}$$

which is a consistency equation for the transition probabilities of a stationary Markov process. A truncation of second order of the Kramers-Moyal expansion of the master equation [6] leads to the Fokker-Planck equation

$$\begin{aligned}\frac{\partial}{\partial t}p(x,t) &= -\frac{\partial}{\partial x}[a_1(x)p(x,t)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[a_2(x)p(x,t)] \qquad (3.9)\\ \frac{\partial}{\partial t}p(x,t|x_0,t_0) &= -\frac{\partial}{\partial x}[a_1(x)p(x,t|x_0,t_0)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[a_2(x)p(x,t|x_0,t_0)]\end{aligned}$$

where a_1 is the drift and a_2 is the diffusion coefficients.

If we are modeling a diffusion process we can use either the Fokker-Planck equation to describe the time evolution of its probability density or the Langevin equation (eq. 3.10) approach [6] to describe the time evolution of its sample paths.

$$\dot{x} = v(x(t)) + b(x(t))\eta(t) \tag{3.10}$$

3.2 Stochastic Processes and Stock Markets

Bachelier was the first who proposed that financial markets follow a "random walk" and they can be simply modeled by standard probability calculus [6]. However, the first two approaches that were commonly adopted by market professionals to predicting stock prices were the "chartist" (or "technical") theories and the theory of fundamental or intrinsic value analysis. Chartist techniques attempt to use knowledge of the past behavior of a price series to predict the probable future behavior of the series. The basic assumption in this approach is that history tends to repeat itself, i.e. past patterns of price behavior will tend to recur in the future (price changes are dependent). By other hand, the assumption of the fundamental analysis approach is that at any point in time an asset has an intrinsic value (an equilibrium price) which depends on the earning potential of the asset. Through a careful study of the factors which can affect the asset, the analyst should, in principle, be able to determine whether the actual price is above or below its intrinsic value and "predict" its future value

[2, 47, 48]. Malkiel proved that following any of these methods will produce inferior results over passive strategies [49].

Bachelier ideas were considered some years latter in order to attempt predict the stock prices changes from a more scientific approach. The random walk hypothesis is a financial theory stating that stock market prices evolve according to a random walk and thus the prices of the stock market cannot be predicted. This is consistent with the efficient-market hypothesis [2, 47, 48] which states that one cannot consistently achieve returns in excess of average market returns on a risk-adjusted basis, given the information available at the time the investment is made [43]. Often, random walks are assumed to be Markov chains or Markov processes. Empirical studies have demonstrated that prices do not completely follow random walks. Low correlations exist in the short term, and slightly stronger correlations over the longer term [6, 7, 10].

A random walk is a mathematical formalisation of a trajectory which consists of taking successive random steps of size $\pm s$. In figures 9 and 10 we show a simulation of 10 random walks with step length equal to 1 and $n = 100$ and 500 steps, respectively.

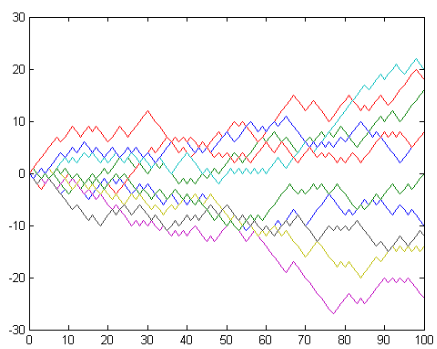

Figure 9: Simulation of 10 random walks with step length equal to 1 and $n = 100$.

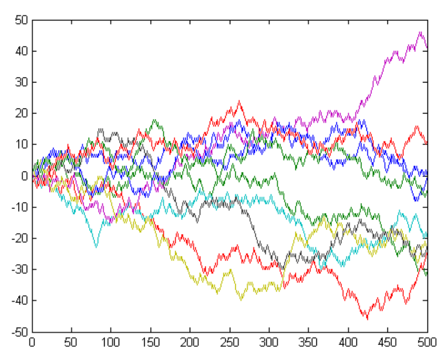

Figure 10: Simulation of 10 random walks with step length equal to 1 and $n = 500$.

The first and second moments for this process are $E\{x_i\} = 0$, $E\{x_i^2\} = s^2$, $E\{x(n\Delta t)\} = 0$, $E\{x^2(n\Delta t)\} = ns^2$.

As can be seen in figures 9 and 10, the variance of the process grows linearly with the number of steps n. Predicting stock prices is generally accepted to be a very difficult task. The stock prices behave very much like a random walk at least when the hit rates generated by most prediction methods in use are viewed. But this task is impossible if the time series to predict is an absolute random walk [44]. In that case, any algorithm for the prediction of the sign of a martingale of that process produces a 50% hit rate in a long run. When the series have a "near" to random walk behavior results of 54% hit rate in the predictions have been reported [45, 46].

Bachelier [2] suggested a probabilistic description of price fluctuations in the financial market and developed the mathematics of the Brownian motion to model the time evolution of asset prices. He determined the probability of price changes by writing down (what is now called) the Chapman-Kolmogorov equation (eq. 3.7) and recognizing that (what is now called) a Wiener process satisfies a diffusion equation. The first theoretical description of the Brownian motion was performed in 1905 by Einstein [3] and in latter years its mathematics was made more rigorous by Wiener [4].

Brownian motion is the assumable random movement of particles suspended in a fluid (a liquid or a gas). If $\rho(x,t)$ is the density of Brownian particles at point x at time t then it will satisfy the diffusion equation defined as

$$\frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2} \tag{3.11}$$

where D is the diffusion coefficient. Assuming all the particles were located at the origin at the initial time $t = 0$, the diffusion equation has as solution

$$\rho(x,t) = \frac{1}{(4\pi Dt)^{1/2}} e^{-x^2/4Dt} \tag{3.12}$$

In figure 11 we show a simulation of 10^4 steps, two dimensional Brownian motion. In this case we took the diffusion constant equal to 1.

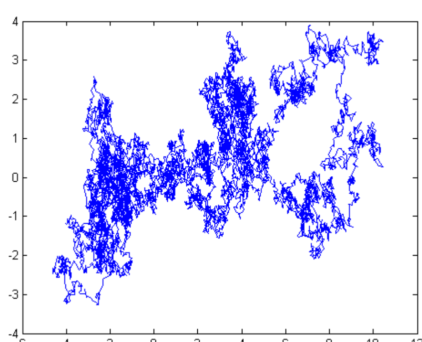

Figure 11: Simulation of 10^4 steps two dimensional Brownian motion. $D = 1$.

The mathematics of the Brownian motion is described by the Wiener process and it has several real-world applications. An often quoted example is stock market fluctuations. A Wiener process can be constructed as the scaling limit of a random walk, or other discrete-time stochastic processes with stationary independent increments. The continuous limit of a random walk may be achieved by considering the limit $n \to \infty$ and $\Delta t \to 0$ such that $t = n\Delta t$ is finite. Then $E\{x^2(t)\} = ns^2 = \frac{s^2 t}{\Delta t}$. To have consistency in the limits with $s^2 = D\Delta t$ it follows that $E\{x^2(t)\} = Dt$. The linear dependence of the variance $x^2(t)$ on t is characteristic of a diffusive process, and D is the diffusion constant. Usually it is implicitly assumed that for $n \to \infty$ and $\Delta t \to 0$ the stochastic process $x(t)$ is a Gaussian.

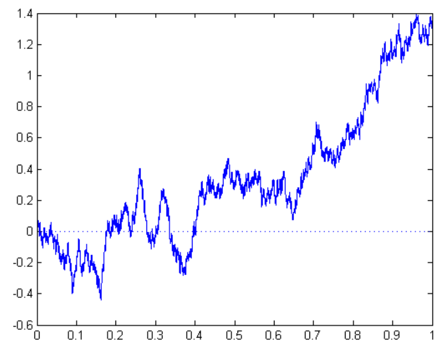

Figure 14: The continuous limit of a random walk, a "Gaussian walk".

It is important to note that Brownian motion is a diffusion process, a Gaussian process, a Lévy process, a Markov process and a martingale. On the one hand this makes it a very strong condition (and therefore the least realistic), on the other hand it makes it a very important generic stochastic process and is therefore used extensively for modeling financial markets.

Osborne [50] showed that the logarithms of common-stock prices, and the value of money, can be regarded as an ensemble of decisions in statistical equilibrium, and that this ensemble of logarithms of prices, each varying with time, has a close analogy with the ensemble of coordinates of a large number of molecules. Using a probability distribution function and the prices of the same random stock choice at random times, he was able to derive a steady state distribution function, which is precisely the probability distribution for a particle in Brownian motion.

Geometric Brownian motion (GBM) is the most commonly used time series for modeling in finance [6, 7, 10]. A stochastic process S_t is said to follow a GBM if it satisfies the following stochastic differential equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{3.13}$$

where W_t is a Wiener process (or Brownian motion) and μ (drift) and σ (volatility) are constants [5, 6]. A GBM is defined as a stochastic process in which the logarithm of the randomly varying quantity follows a Brownian motion [43]. Modeling stock price changes through a GBM is equivalent to say that the logarithm of prices changes are Gaussian distributed or that they are performing a Brownian motion. GBM is the main assumption to model stock prices in Black–Scholes theory which is the cornerstone of mathematical finance. In the next section we will analyze this theory in detail due to its importance and because it constitutes a good example of how the behavior of the financial markets can be modeled

Chapter 4

Modeling the Financial Market

Bachelier work was forgotten until 1944 when Ito used it as motivation to introduce his calculus and a variant of the Brownian motion which is the geometric Brownian motion [5, 6]. Since the 1950s mathematicians began to show interest in the modeling of stock market prices. Bachelier original proposal that the price changes are Gaussian distributed was replaced to a model in which stock prices are log-normal distributed [7]. In 1973, Black, Scholes and Merton [8, 9, 10] used the geometric Brownian motion to construct a theory for determining the price of stock options. This theory constitutes nowadays the cornerstone of mathematical finance.

4.1 From Brownian to Geometric Brownian Motion

The efficient market hypothesis [51] suggests that the stochastic process behind the price chances in the stock market is Markovian. A market is called efficient if the participants quickly and comprehensively obtain all information relevant to trading, if it is liquid and if there is low market friction. The efficient market hypothesis states that a market with these properties "digests" the new information so efficiently that all the current information about the market development is at all times completely contained in the present prices. That means that no advantage is gained by taking into account all or part of the previous price evolution [6, 7, 47, 51] (Markov assumption). This was the same conclusion that Bachelier made in his original work [2]. Specifically, he suggested a model for $S(t)$ as a Wiener process. However, this model suffer from the unrealistic property that it allows negative asset prices. This fact led to a refined version of a random walk model (the geometric Brownian motion [5]).

Lets define a model for $S(t)$. Lets suppose that at time $t = 0$ we deposit a sum $S(0)$ in a bank. The bank will grant a risk-free interest rate r for the deposit. If the interest is paid once at the end of time t, the initial sum has growth by

$$S(t) = S(0) + rtS(0) = S(0)(1 + rt) \tag{4.1}$$

and for n payments in time t, we obtain

$$S(t) = S(0)(1 + \frac{rt}{n})^n \tag{4.2}$$

which yields in the limit of continuously compounded interest

$$S(t) = S(0)(1 + \frac{rt}{n})^n|_{n\to\infty} \to S(0)\exp(rt) \tag{4.3}$$

Lets consider the asset price similar to a bank deposit but perturbed by stochastic fluctuations. Therefore, the price chance dS in the small time interval dt should consist of two contributions, a deterministic and a random one. In comparison with eq. (4.3) a reasonable ansatz for the deterministic part is

$$dS = \mu S(t)dt \tag{4.4}$$

where μ is called drift and measures the average growth rate of the asset price. According to the efficient market hypothesis, the stochastic nature of the price evolution should be Markovian and a possible choice for this contribution could be of the form (symmetric to (4.4))

$$dS = \sigma S(t)dW(t) \tag{4.5}$$

where σ is the volatility and measures the strength of the statistical price fluctuations. Combining both contributions (equations (4.4) and (4.5)) we get

$$dS = \mu S(t)dt + \sigma S(t)dW(t) \tag{4.6}$$

This equation defines a variant of Brownian motion which is called the geometric Brownian motion which a specific case of an Ito process defined by [5]

$$dS = a(S,t)dt + b(S,t)dW(t) \tag{4.7}$$

GBM introduces the parameters μ and σ. The drift μ describes the expected gain. The volatility σ quantifies the fluctuations around the average behavior, if the fluctuations are large, the asset is very volatile and an investment is considered risky.

Equation (4.6) suggests that actually the return dS/S is the relevant variable in our model and not the absolute change $dS = S(t + dt) - S(t)$. Using Ito formula we can rewrite equation (4.6) in terms of $\ln S(t)$. Ito formula is given by [5, 6]

$$df(S,t) = (\frac{\partial f}{\partial t} + a(S,t)\frac{\partial f}{\partial S} + \frac{1}{2}b(S,t)^2\frac{\partial^2 f}{\partial S^2})dt + b(S,t)\frac{\partial f}{\partial S}dW(t) \tag{4.8}$$

For $f(S,t) = \ln S$, $a(S,t) = \mu S$ and $b(S,t) = \sigma S$ we get

$$d \ln S = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW(t) \tag{4.9}$$

Therefore is the logarithm of the asset price and not the price itself (as Bachelier assumed) which performs a Wiener process with a constant drift. The variable $d \ln S$ is normally distributed with mean and variance given by [5, 6]

$$\begin{aligned} E[d \ln S] &= (\mu - \frac{1}{2}\sigma^2)dt \\ Var[d \ln S] &= \sigma^2 dt \end{aligned} \tag{4.10}$$

and the transition probability from (S,t) to (S',t') is given by [6]

$$p(S',t'|S,t) = \frac{1}{\sqrt{2\pi(\sigma S')^2(t'-t)}} \exp(-\frac{[\ln(S'/S) - (\mu - \sigma^2/2)(t'-t)]^2}{2\sigma^2(t'-t)}) \tag{4.11}$$

The price $S(t)$ has a log-normal distribution.

A good example of how financial markets have been modeled under the assumptions above is the Black - Scholes Theory for option pricing [8] which we analyze in the next section.

4.2 Option Pricing

An European option is a contract between two parties in which the seller of the option (writer) grants the buyer (holder) the right to purchase (call option) from the writer or to sell (put option) to him an underlying with current spot price $S(t)$ for a prescribed price K (strike price) at the expiry date T in the future. The key property of an option is that only the writer has an obligation. He must sell or buy the underlying asset for the strike price at time T. On the other hand, the holder has the possibility to exercise his option or not depending if he gains a profit i.e. if $S(T) > K$ for a call option, otherwise he can buy the underlying for a cheaper price $S(t) < K$ on the market. The question is: how much an option should cost and the answer is given by the Black-Scholes theory.

4.2.1 The Black - Scholes Theory

The Black - Scholes model for option pricing is based on the following assumptions [8]:

1. There is no credit risk, only market risk. The price of the option only depends on the markets fluctuations.

2. The market is maximally efficient, infinitely liquid and does not exhibit any friction. All relevant information is instantaneously and comprehensively available and at all times fully reflected by the current prices (Efficient Market Hypothesis).

3. Continuous trading is possible. The time interval Δt between successive quotations of the price of the underlying tends to zero.

4. The time evolution of the asset price is stochastic and exhibits geometric Brownian motion (equation (4.6)).

5. The risk-free interest rate r and the volatility σ are constant. This assumption can be relaxed if r and σ are known functions of time. The Black-Scholes valuation formulas for European call and put options remain valid with r and σ replaced by [52]

$$r \to \frac{1}{T-t}\int_t^T r(t')dt' \tag{4.12}$$
$$\sigma^2 \to \frac{1}{T-t}\int_t^T \sigma^2(t')dt'$$

6. The underlying pays no dividends.

7. The underlying is arbitrarily divisible, i.e. the amount of underlying in the portfolio does not need to be an integer.

8. The market is arbitrage-free.

These assumptions can be used to derive the Black-Scholes pricing formulas.

The problem can be stated as follows: An European option confers to its holder the right to buy from a writer (call option C) or to sell to him (put option V) an asset for the strike price K at a future time T. This right has a value, the option price O, which must be a function of the current value S and the time t and also depends on K and T. At the expiry time, $C(S,T) = \max(S(T)-K, 0)$ and $P(S,T) = \max(K - S(T), 0)$ [8]. Given these limiting values, the problem consists in finding what are fair option prices when signing a contract at times $t < T$.

The option price $O(S,t)$ is a function of the stochastic variable S which performs an lto process, so its change in the time interval dt is given by the Ito's formula (equation (4.8))

$$dO = (\frac{\partial O}{\partial t} + \mu S(t)\frac{\partial O}{\partial S} + \frac{1}{2}(\sigma S(t))^2\frac{\partial^2 O}{\partial S^2})dt + \sigma S(t)\frac{\partial O}{\partial S}dW(t)$$
$$dO = (\frac{\partial O}{\partial t} + \frac{1}{2}(\sigma S(t))^2\frac{\partial^2 O}{\partial S^2})dt + \frac{\partial O}{\partial S}dS \tag{4.13}$$

in which equation (4.6) was inserted in order to obtain equation (4.13). Equation (4.13) represents the response of the option price O to the stochastic time evolution of S and it takes into account our six first assumptions [8].

In order to incorporate the last two assumptions, we consider that when signing the contract, the writer faces the risk that the odds are against him and he has to sell the underlying at the expiry date below the market price to the holder. Certainly, it is advisable to own a fraction $\Delta(t)$ $(0 < \Delta(t) < 1)$ of the underlying at any time t. This fraction should be adjusted depending on the changes of the asset price S. Imagine that the writer has sold a call option. If S rises, the option becomes more likely to be exercised, so $\Delta(t)$ should be increased, and vice versa. For these adjustments he needs money. Thus, it is also advisable to have a cash amount $\Pi(t)$ which he could in turn invest to increase its value. Certainly, this investment should be risk-less to avoid introducing further uncertainties. Taking advantage of both forms of advice, the writer should possess "wealth" $W(t)$ at time t, and we can define it as

$$W(t) = \Delta(t)S(t) + \Pi(t) \tag{4.14}$$

The writer should therefore require [6, 7, 8]

$$O(t) = W(t) = \Delta(t)S(t) + \Pi(t) \tag{4.15}$$

which must hold at any time, this means

$$\begin{aligned} dO &= \Delta(t)dS(t) + d\Pi \\ dO &= \Delta(t)dS(t) + r\Pi dt \end{aligned} \tag{4.16}$$

The first term of the right-hand side assumes that $\Delta(t)$ does not change in the time interval. A heuristic argument to justify this assertion is that a change of Δ should be interpreted as a reaction to a price fluctuation, so, the amount of underlying can only be adjusted after a price variation has occurred. The second term expresses the growth of the cash amount $\Pi(t)$. It reflects the requirement that the market is arbitrage-free. In such a market, there is no risk-less investment strategy which yields a better return than that granted by the bank. By comparing equations (4.13) and (4.16)

$$\Delta(t) = \frac{\partial O}{\partial S} \tag{4.17}$$

$$r\Pi = \frac{\partial O}{\partial t} + \frac{1}{2}(\sigma S(t))^2 \frac{\partial^2 O}{\partial S^2} \tag{4.18}$$

If the writer continuously adjusts the amount of underlying according to equation (4.17) during the lifetime of the option (called delta-hedge), he can eliminate the risk completely by calculating the option price from equation (4.18)

Inserting equations (4.15), (4.17) in equation (4.18) we obtain the **Black-Scholes equation** for European options [8]

$$\frac{\partial O}{\partial t} + \frac{1}{2}(\sigma S)^2 \frac{\partial^2 O}{\partial S^2} + rS\frac{\partial O}{\partial S} - rO = 0 \tag{4.19}$$

This equation is independent of the drift parameter μ and valid for any derivative which satisfies the assumptions compiled above, and in particular for call and put options.

The boundary conditions for a call option are [6, 8]

$$\begin{aligned} t &= T : C(S,T) = \max(S(T) - K, 0) \\ S &= 0 : C(0,t) = 0 \\ S &\to \infty : C(S,t)\tilde{}S \end{aligned} \tag{4.20}$$

and for a put option

$$\begin{aligned} t &= T : P(S,T) = \max(K - S(T), 0) \\ S &= 0 : P(0,t) = Ke^{-r(T-t)} \\ S &\to \infty : P(S,t)\tilde{}0 \end{aligned} \tag{4.21}$$

4.2.2 Solution and Interpretation

Now we present the solution of equation (4.19) subject to the boundary conditions (4.20) or (4.21) [8, 10, 52]. These solutions tell the writer which price he should charge for an option at time t, at which the contract is signed.

Lets consider a call option. Its price $C(S,t)$ is the solution of the Black-Scholes equation

$$\frac{\partial C}{\partial t} + \frac{1}{2}(\sigma S)^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0 \tag{4.22}$$

subject to

$$\begin{aligned} t &= T : C(S,T) = \max(S(T) - K, 0) \\ S &= 0 : C(0,t) = 0 \\ S &\to \infty : C(S,t)\tilde{}S \end{aligned} \tag{4.23}$$

By taking the following changes of variables

$$C = Kf(x,r),\ S = Ke^x,\ t = T - \frac{\tau}{\sigma^2/2} \tag{4.24}$$

we can rewrite equations (4.22) and (4.23) as

$$\frac{\partial f}{\partial \tau} = \frac{\partial^2 f}{\partial x^2} + (\kappa - 1)\frac{\partial f}{\partial x} - \kappa f \tag{4.25}$$

$$\begin{aligned} \tau &= 0 : f(x,0) = \max(e^x - 1, 0) \\ x &\to -\infty : f(x,\tau) \to 0 \\ x &\to +\infty : f(x,\tau) \sim e^x \end{aligned} \tag{4.26}$$

where $\kappa = 2r/\sigma^2$. This shows that the Black-Scholes theory depends on one single parameter κ.

Equation (4.25) would resemble to a diffusion equation if the last two terms were absent. In order to do this we take the following change of variables

$$f(x,\tau) = e^{ax+b\tau} g(x,\tau) \tag{4.27}$$

in (4.25), so

$$\frac{\partial g}{\partial \tau} = \frac{\partial^2 g}{\partial x^2} + [2a + (\kappa - 1)]\frac{\partial g}{\partial x} - [a^2 + (\kappa - 1)a - \kappa - b]g \tag{4.28}$$

By choosing a and b equal to $a = -1/2(\kappa - 1)$ and $b = -1/4(\kappa + 1)^2$ we obtain the diffusion equation

$$\frac{\partial g}{\partial \tau} = \frac{\partial^2 g}{\partial x^2} \tag{4.29}$$

with the boundary conditions

$$\begin{aligned}
\tau &= 0 : g(x,0) = \max(e^{(\kappa+1)x/2} - e^{(\kappa+1)x/2}, 0) \\
&\Rightarrow g(x,0)e^{-\alpha x^2}|_{|x|\to\infty} \to 0,\ \alpha > 0 \\
\tau &> 0 : g(x,\tau)|_{|x|\to\infty} \to e^{(\kappa+1)x/2} - e^{(\kappa+1)^2\tau/4} \\
&\Rightarrow g(x,\tau)e^{-\alpha x^2}|_{|x|\to\infty} \to 0,\ \alpha > 0
\end{aligned} \tag{4.30}$$

where $\alpha \in R^+$. As can be seen, g is well-behaved for $|x| \to \infty$, and an unique solution exists.

To find this solution, we use the Green's function method [52]. The function $g(x,\tau)$ has a meaning only for $\tau > 0$. However, if we introduce the variable

$$\overline{g}(x,\tau) = \Theta(\tau) g(x,\tau) \tag{4.31}$$

the time variable can be extended to $\tau < 0$. $\Theta(\tau)$ is the Heaviside function. This definition of $\overline{g}(x,\tau)$ turns the diffusion equation (4.29) into an inhomogeneous differential equation

$$(\frac{\partial}{\partial \tau} - \frac{\partial^2}{\partial x^2})\overline{g}(x,\tau) = g(x,\tau)\delta(\tau) = \overline{g}(x,0)\delta(\tau) \tag{4.32}$$

which is solved by

$$\overline{g}(x,\tau) = \int_{-\infty}^{\infty} \overline{g}(y,0) p(x,\tau|y,0) dy \tag{4.33}$$

if the integral kernel (Green's function) satisfies

$$(\frac{\partial}{\partial \tau} - \frac{\partial^2}{\partial x^2}) p(x,\tau|y,0) = \delta(x-y)\delta(\tau) \tag{4.34}$$

This is just the partial differential equation for the diffusion (Gaussian) propagator

$$p(x,\tau|y,0) = \frac{1}{\sqrt{4\pi\tau}} e^{\frac{-(x-y)^2}{4\tau}} \tag{4.35}$$

Combining these results, the solution of equation (4.29) can be written as

$$g(x,\tau) = \frac{1}{\sqrt{4\pi\tau}} \int_{-\infty}^{\infty} g(y,0) e^{\frac{-(x-y)^2}{4\tau}} dy \tag{4.36}$$

Inserting the boundary conditions for $g(y,0)$ in (4.35) we find

$$g(x,\tau) = \frac{1}{\sqrt{4\pi\tau}} \int_{-\infty}^{\infty} (e^{(\kappa+1)y/2} - e^{(\kappa-1)y/2}) e^{\frac{-(x-y)^2}{4\tau}} \tag{4.37}$$

and by introducing the new integration variable $z = (y-x)/(2\tau)^{1/2}$ this becomes

$$\begin{aligned} g(x,\tau) &= \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\pi}}^{\infty} (e^{(\kappa+1)(\sqrt{2\tau}z+x)/2} - e^{(\kappa-1)(\sqrt{2\tau}z+x)/2}) e^{\frac{-z^2}{2}} dz \\ &= \frac{1}{\sqrt{2\pi}} e^{(\kappa+1)x/2+(\kappa+1)^2\tau/4} \int_{-x/\sqrt{2\pi}}^{\infty} e^{-\frac{1}{2}(z-\sqrt{2\tau}(\kappa+1)/2)^2} dz \\ &\quad - \frac{1}{\sqrt{2\pi}} e^{(\kappa+1)x/2+(\kappa+1)^2\tau/4} \int_{-x/\sqrt{2\pi}}^{\infty} e^{-\frac{1}{2}(z-\sqrt{2\tau}(\kappa-1)/2)^2} dz \\ &= e^{(\kappa+1)x/2+(\kappa+1)^2\tau/4} N(d_1) - e^{(\kappa+1)x/2+(\kappa-1)^2\tau/4} N(d_2) \end{aligned} \tag{4.38}$$

where

$$\begin{aligned} d_1 &= \frac{\ln(S/K) + (r+\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \\ d_2 &= d_1 - \sigma\sqrt{T-t} \end{aligned} \tag{4.39}$$

and

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz \tag{4.40}$$

The function $N(x)$ is the probability that the normally distributed variable z adopts a value smaller than x. It is the cumulative probability distribution for the Gaussian distribution.

Finally, lets express the call price using equations (4.24), (4.27) and $a = -1/2(\kappa-1)$, $b = -1/4(k+1)^2$ as

$$C = Ke^{-1/2(\kappa-1)x-1/4(\kappa+1)^2\tau} g(x,\tau) \tag{4.41}$$

and using the original variables eq. (4.24) we get for a call option

$$C(S,t) = SN(d_1) - Ke^{-\tau(T-t)} N(d_2) \tag{4.42}$$

and for a put option (using the put-call parity [8])

$$P(S,t) = -S(1-N(d_1)) - Ke^{-\tau(T-t)}(1-N(d_2)) \tag{4.43}$$

The Black-Scholes equations (4.42) and (4.43) tell the writer which price he should charge for an (call or put) option at time t, at which the contract is signed. This price depends on the parameters K and T of the contract and on the market characteristics r and σ. The solutions do not depend on the drift the drift μ. This means that the option prices are identical for all assets evolving with the same volatility, no matter what μ is. The equations also provide the necessary information on how to eliminate the risk. The writer's portfolio only stays risk-less if he continuously adjusts the amount of underlying $\Delta(t)$ such that

$$\Delta(t) = \frac{\partial C}{\partial S} = N(d_1) \quad (0 \leq N(d_1) \leq 1) \tag{4.44}$$

$$\Delta(t) = \frac{\partial P}{\partial S} = -(1 - N(d_1)) \quad -1 \leq 1 - N(d_1) \leq 0 \tag{4.45}$$

Therefore, the first terms in the Black - Scholes equations can be interpreted as the fraction of underlying which the writer should buy (call option) or sell (put option) to maintain a risk-less position. Also the second terms can be identified with the cash amount $\Pi(t)$

$$\begin{aligned} C(S,t) &= SN(d_1) - Ke^{-\tau(T-t)}N(d_2) \\ C(S,t) &= S\Delta(t) + \Pi(t) \end{aligned} \tag{4.46}$$

so that

$$\Pi(t) = -Ke^{-\tau(T-t)}N(d_2) \tag{4.47}$$

The factor $-Ke^{-\tau(T-t)}$ is the strike price discounted to the present time, the so-called present value. For a call option, $\Pi(t)$ is given by the present value of the payment, due when the call is exercised, multiplied by the probability $N(d_2)$ that the call will be exercised. The minus sign indicates that this amount of money must be borrowed.

Finally, we can measure the sensitivity of the option price under a slight change of a single parameter while holding the other parameters fixed. These are called the "Greeks". Formally, they are partial derivatives of the option price with respect to the independent variables $\Delta = \frac{\partial C}{\partial S}$, $\Gamma = \frac{\partial^2 C}{\partial S^2}$, $V = \frac{\partial C}{\partial \sigma}$, $\Theta = \frac{\partial C}{\partial t}$, $\rho = \frac{\partial C}{\partial r}$ [8].

Extensions of the Model and main Disadvantages

The valuation formula can obtained using a risk neutrality argument [10] (above we used the method of arbitrage-free pricing). The method can also be extended for variable (but deterministic) rates and volatilities [6].

The Black–Scholes model disagrees with reality in a number of ways. Some of the disadvantages of the model are the underestimation of extreme moves, the assumptions of instant, stationary process, and specially, the continuous time and continuous trading which produce (liquidity, volatility) risk. However, Black–Scholes pricing is widely used in practice because it is a useful approximation when analyzing the directionality of prices when crossing critical points and also because the model can be adjusted to deal with some of its failures.

Alternative Model

In 1963, Benoit Mandelbrot suggested a stable distribution with a characteristic exponent, $\alpha < 2$ for the asset prices [11, 32]. The main disadvantage of his model is the fact that the variance of Levy distributions with $\alpha < 2$ diverges, whereas financial time series have a well-defined variance. However, a truncated Levy flight could give a description of the whole distribution of prices. The distribution of this process preserves the typical self-similar Levy scaling over an extended range before the truncation becomes effective and imposes a finite variance.

Chapter 5

Discussion

The first two approaches that were commonly adopted by market professionals to predicting stock prices were the "chartist" (or "technical") theories and the theory of fundamental or intrinsic value analysis. Chartist techniques attempt to use knowledge of the past behavior of a price series to predict their future probable behavior. The basic assumption in this approach is that price changes are dependent and past patterns will tend to recur in the future. By other hand, the fundamental analysis approach assumes that at any point in time an asset has an intrinsic value (an equilibrium price) which depends on the earning potential of the asset. Through a careful study of the factors that can affect the asset, the analyst should, in principle, be able to determine whether the actual price is above or below its intrinsic value and "predict" its future value [2, 47, 48]. However, it has been proved [49] that following any of these methods will produce inferior results over passive strategies.

Bachelier was the first who proposed that financial markets follow "random walks" and can be modeled by standard probability calculus [6]. In his thesis "Speculation Theory" Bachelier [2] suggested a probabilistic description of price fluctuations in the financial markets. He developed the mathematics of the Brownian motion to model the time evolution of asset prices and also the probability of price changes by writing down (what is now called) the Chapman-Kolmogorov equation and recognizing that (what is now called) a Wiener process satisfies a diffusion equation. However, Bachelier work was not recognized by the scientific community at that time, perhaps because of its application to financial markets and also because it suffers from the unrealistic property that it allows negative asset prices. This fact led to a refined version of a random walk model. In 1944 Ito used Bachelier theory as a motivation to introduce his calculus and a variant of the Brownian motion, the geometric Brownian motion [5, 6].

Geometric Brownian motion (GBM) is the most commonly used time series for modeling in finance [6, 7, 10]. A GBM is defined as a stochastic process in which the logarithm of the randomly varying quantity follows a Brownian motion [43]. Modeling stock price changes through a GBM is equivalent to say that the logarithm of prices changes are Gaussian distributed or that they are performing a Brownian motion. Bachelier original proposal that the price changes are Gaussian distributed was replaced to a model in which stock prices are log-normal distributed [7]. In 1973, Black, Scholes and Merton [8, 9, 10] used the geometric Brownian motion to construct a theory for determining the price of stock options. The Black-Scholes equations tell the writer which price he should charge for an (call or put) option at time t, at which the contract is signed. This theory constitutes nowadays the cornerstone of mathematical finance.

The Black–Scholes model disagrees with reality in a number of ways. Some of the disadvantages of the model are the underestimation of extreme moves, the assumptions of instant, stationary process, and specially, the assumption of continuous time and continuous trading, all of them producing some kind (liquidity, volatility,etc.) of risk. Actually, prices do not follow a strict stationary log-normal process, nor is the risk-free interest actually known (and is not constant over time) and also the variance has been observed to be non-constant [6] leading to models such as GARCH to model volatility changes.

However, Black–Scholes pricing is widely used in practice [53] because it is easy to calculate and explicitly models the relationship of all the variables. It is also a useful approximation when analyzing the directionality of prices when crossing critical points. It is used both as a quoting convention and a basis for more refined models [10]. Although volatility is not constant, results from the model are often useful in practice and helpful in setting up hedges in the correct proportions to minimize risk. Even when the results are not completely accurate, they serve as a first approximation to which adjustments can be made [8, 10, 43].

Acknowledgement 1 *The author would like to thank the Erasmus Mundus Master in Complex Systems Consortium for their support.*

Bibliography

[1] V. Pareto, Cours d'Economie Politique,Lausanne and Paris, 1897.

[2] L. Bachelier, Theorie de la speculation , Annales Scientifiques de I'Ecole Normale Superieure III-17, 21-86 (1900).

[3] A. Einstein, On the Movement of Small Particles Suspended in a Stationary Liquid Demanded by the Molecular-Kinetic Theory of Heat, Ann. Physik 17, 549-560 (1905).

[4] N. Wiener, Differential Space, J. Math. Phys. 2, 131-174 (1923)

[5] K. Ito, On Stochastic Differential Equations, Mem. Amer. Math. Soc. 4,1-51 (1951).

[6] P. Baschnagel, Stochastic Processes. From Physics to Finance, Springer, 1999.

[7] R.N. Mantegna-Stanley, An Introduction to Econophysics, Cambridge University Press, 2000.

[8] F. Black and M. Scholes, The Pricing of Options and Corporate Liabilities, J. Polit. Econ. 81, 637-654 (1973).

[9] R. C. Merton, Theory of Rational Option Pricing, Bell J. Econ. Management Sci. 4, 141-183 (1973).

[10] J.-P. Bouchad, Theory Of Financial Risks and Derivative Pricing: From Statistical Physics To Risk Management, Cambridge University Press, 2003.

[11] B. B. Mandelbrot, The Variation of Certain Speculative Prices, J. Business 36, 394-419 (1963).

[12] R. N. Mantegna and H. E. Stanley, Turbulence and Financial Markets, Nature 383, 587-588 (1996).

[13] R. N. Mantegna and H. E. Stanley, Stock Market Dynamics and Turbulence: Parallel Analysis of Fluctuation Phenomena, Physica A 239, 255-266 (1997).

[14] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y, Dodge, Turbulent Cascades in Foreign Exchange Markets, Nature 381, 767-770 (1996).

[15] E. J. Elton and M. J. Gruber, Modern Portfolio Theory and Investment Analysis, J. Wiley and Sons, New York, 1995.

[16] H. Markowitz, Portfolio Selection: Efficient Diversification of Investment, J.Wiley, New York, 1959.

[17] R. N. Mantegna, Degree of Correlation Inside a Financial Market, in Applied Nonlinear Dynamics and Stochastic Systems near the Millennium, edited by J. B. Kadtke and A. Bulsara AIP Press, New York, 1997.

[18] J.-P. Bouchaud and M. Potters, Theories des Risques Financiers, Eyrolles, Alea-Saclay, 1997.

[19] P. Gopikrishnan, M. Meyer, L. A. N. Amaral, and H. E. Stanley, Inverse Cubic Law for the Distribution of Stock Price Variations, Eur. Phys. J. B 3, 139-140 (1998).

[20] R. N. Mantegna, Levy Walks and Enhanced Diffusion in Milan Stock Exchange, Physica A 179, 232-242 (1991).

[21] R. N. Mantegna and H. E. Stanley, Scaling Behavior in the Dynamics of an Economic Index, Nature 376, 46-49 (1995).

[22] R. Cont, M. Potters, and J.-P. Bouchaud, Scaling in Stock Market Data: Stable Laws and Beyond, in Scale Invariance and Beyond, edited by B. Dubrulle, F. Graner, and D. Sornette Springer, Berlin, 1997.

[23] Y. Liu, P. Cizeau, M. Meyer, C.-K. Peng, and H. E. Stanley, Quantification of Correlations in Economic Time Series, Physica A 245, 437-440 (1997).

[24] A. Arneodo, J. F. Muzy, and D. Sornette, Direct Causal Cascade in the Stock Market, Eur. Phys. J. B 2, 277-282 (1998).

[25] P. Cizeau, Y. Liu, M. Meyer, C.-K. Peng, and H. E. Stanley, Volatility Distribution in the S&P 500 Stock Index, Physica A 245, 441-445 (1997).

[26] U. A. Muller, M. M. Dacorogna, R. B. Olsen, O. V. Pictet, M. Schwarz, and C. Morgenegg, Statistical Study of Foreign Exchange Rates, Empirical Evidence of a Price Change Scaling Law and Intraday Analysis, J. Banking and Finance 14, 1189-1208 (1995).

[27] P. Bak, M. Paczuski, and M. Shubik, Price Variations in a Stock Market with Many Agents, Physica A 246, 430-453 (1997).

[28] J.-P. Bouchaud and R. Cont, A Langevin Approach to Stock Market Fluctuations and Crashes, Eur. Phys. J. B 6, 543-550 (1998).

[29] G. Caldarelli, M. Marsili, and Y.-C. Zhang, A Prototype Model of Stock Exchange, Europhys. Lett. 40, 479-483 (1997).

[30] D. Challet and Y. C. Zhang, On the Minority Game: Analytical and Numerical Studies, Physica A 256, 514-532 (1998).

[31] M. Levy and S. Solomon, Power Laws Are Logarithmic Boltzmann Laws, Intl J. Mod. Phys. C 7, 595-601 (1996).

[32] Mandelbrot, Fractals and Scaling in Finance, Springer-Verlag, New York, 1997.

[33] A. H. Sato and H. Takayasu, Dynamic Numerical Models of Stock Market Price: From Microscopic Determinism to Macroscopic Randomness, Physica A 250, 231-252 (1998).

[34] D. Sornette and A. Johansen, A Hierarchical Model of Financial Crashes, Physica A 261, 581-598 (1998).

[35] E. Aurell, R. Baviera, O. Hammarlid, M. Serva, and A. Vulpiani, A General Methodology to Price and Hedge Derivatives in Incomplete Markets, Int. J. Theor. Appl. Finance (in press).

[36] J.-P. Bouchaud and D. Sornette, The Black & Scholes Option Pricing Problem in Mathematical Finance: Generalization and Extensions for a Large Class of Stochastic Processes, J. Phys. I France 4, 863-881 (1994).

[37] S. Galluccio and Y. C. Zhang, Products of Random Matrices and Investment Strategies, Phys. Rev. E 54, R4516-R4519 (1996).

[38] S. Galluccio, J.-P. Bouchaud, and M. Potters, Rational Decisions, Random Matrices and Spin Glasses, Physica A 259, 449-456 (1998).

[39] M. Marsili, S. Maslov, and Y.-C. Zhang, Dynamical Optimization Theory of a Diversified Portfolio, Physica A 253, 403-418 (1998).

[40] D. Sornette, Large Deviations and Portfolio Optimization, Physica A 256, 251-283 (1998).

[41] http://www.mathworks.com/matlabcentral/fileexchange/.

[42] E. Ott, Dynamical Systems, Cambridge University Press, 1993.

[43] Wikipedia.

[44] Hellstrom, A Random Walk through the Stock Market, Licentiate Thesis, Umea University, 1998.

[45] Baestaens, Market inefficiencies, technical trading and neural networks, in Forecasting Financial Markets, Wiley, 1996.

[46] Tsibouris, Testing the efficient market hypothesis with gradient descent algorithms, in Neural Networks in the Capital Markets, Wiley, 1995.

[47] E. Fama, Random Walks In Stock Market Prices, Financial Analysts Journal 21 (5): 55–59 (1965).

[48] Cootner, The random character of stock market prices, MIT Press, 1964.

[49] Malkiel, A Random Walk Down Wall Street, W.W. Norton & Company, Inc., 1973.

[50] Osborne, Brownian Motion in Stock Markets, Operations Research, 1959.

[51] Fama, Eficient Capital Markets, J. Fin. 25, 383, 1970.

[52] Wilmott, Option Pricing, Mathematical Models and Computation. Oxford Financial Press, Oxford, 1993.

[53] Bodie, Investments, McGraw-Hill New York, 2008.

[54] Rychlik, Probability and Risk Analysis, An Introduction for Engineers, Springer, 2006.

[55] Panjer, Operational Risk Modeling Analytics, Wiley, 2006.

Chapter 6

Appendix A: Probability Theory

6.1 Long-term Frequencies

Lets consider the simple experiment of flipping a coin. The sample space is $S = \{0, 1\}$, which represents the physically observed outcomes $S = \{$"Heads", "Tails"$\}$. Lets repeat this experiment many times in independent manner. The i_{th} flip is denoted by X_i. Just if the coin is "fair" then $P(X_i = 1) = P(X_i = 0) = 1/2$ but in general, (a coin can be biased). There is a number p, $(0 \leq p \leq 1)$, such that $P(X_i = 1) = p$ and, $P(X_i = 0) = 1 - p$. Finding the exact value of p is not possible in practice. However, we can estimate its value through a procedure called the frequentist approach which is motivated by a fundamental result in theory of probability (the Law of Large Numbers (LLN)) [54] .

The law says the fraction of "tails" observed in the first n independent flips converges to p as n tends to infinity, i.e.

$$\overline{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n) \rightarrow p \tag{6.1}$$

as $n \rightarrow \infty$. Since $\sum_{i=1} X_i$ is equal to the number of times "tails" is shown in n flips we can interpret p as "long-term frequency" of "tails" in an infinite sequence of flips. Practically, one cannot flip a coin infinitely many times. We may expect that in practice $\overline{X} \neq p$. The error is defined as

$$E = p - \overline{X} \tag{6.2}$$

and the relative error

$$E_r = |p - \overline{X}|/p. \tag{6.3}$$

In Figure 15 we show a simulation of flipping a coin. As can be seen, for a relatively small number of flips ($n < 100$), the arithmetic mean (i.e. the fraction of "tails" observed in the first n independent flips) tends to p as n tends to infinity.

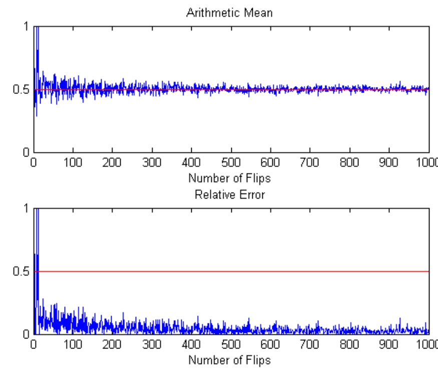

Figure 15: Simulation, tosses of a fair coin. Top: The arithmetic mean $\overline{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$. Bottom: The relative error $E_r = |p - \overline{X}|/p$, both as function of the number of flips. As can be seen, for a relatively small number of flips ($n < 100$), the arithmetic mean (i.e. the fraction of "tails" observed in the first n independent flips) tends to p as n tends to infinity, ($\overline{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n) \rightarrow p$, as $n \rightarrow \infty$). For this simulation, the number of flips was taken equal to 1000.

6.2 Probability Distributions

An experiment is defined as an observation of a given phenomenon under specified conditions. The result of an experiment is called an outcome. An event is a set of one or more possible outcomes. Probability is a measure of the likelihood of the occurrence of an event. A random variable is a function that assigns a numerical value to every possible outcome. A random variable is discrete if contains at most a countable number of values and continuous if the distribution function is continuous and differentiable everywhere with the possible exception of a countable number of values.

In order to describe a random process X for which the result is a real number, one uses a probability density $P(x)$, such that the probability that X is within a small interval of width dx around $X = x$ is equal to $P(x)dx$. The probability that X is between a and b is given by the integral of $P(x)$ between a and b,

$$\mathrm{Pr}(a < x < b) = \int_a^b P(x)dx \tag{6.4}$$

$P(x)dx$ is invariant upon any (monotonic) change of variable (i.e. $x \rightarrow y(x) = \gamma x$, $P(x)dx = P(y)dy$), non-negative ($P(x) \geq 0$ for all x) and must be normalized ($\int_{x_m}^{x_M} P(x)dx = 1$). The cumulative distribution function cdf of a random variable X is the probability that X is less than or equal to a given number x

$$\widetilde{P_<}(x) = \widetilde{P}_X(x) = \mathrm{Pr}(X < x) = \int_{-\infty}^{x} P(x')dx' \tag{6.5}$$

The distribution function must satisfy the following requirements: $0 \leq \widetilde{P}(x) \leq 1$ for all x and $\widetilde{P}(x)$ must be non decreasing. $\widetilde{P}(x)$ is right-continuous $\lim_{x\to-\infty}\widetilde{P}(x) = 0$ and $\lim_{x\to\infty}\widetilde{P}(x) = 1$. Similarly, one defines $\widetilde{P}_{>}(x) = 1 - \widetilde{P}_{<}(x)$. The survival function (also called the decumulative distribution function) $S(x)$ for a random variable X is the probability that X is greater than a given number. That is, $\Pr(X > x) = 1 - \widetilde{P}(x) = S(x)$ and also $0 \leq S(x) \leq 1$ for all x and $S(x)$ must be non increasing. $S(x)$ is right-continuous $\lim_{x\to-\infty} S(x) = 1$ and $\lim_{x\to\infty} S(x) = 0$. The probability density function is the first derivative (i.e., the slope) of the distribution function $P(x) = \widetilde{P}'(x) = -S'(x)$. The probability function also called the probability mass function $p(x)$ describes the probability at a distinct point x. The formal definition is $p(x) = \Pr(X = x)$. For discrete random variables, the distribution and survival functions can be recovered as $\widetilde{P}(x) = \sum_{y\leq x} p(y)$. The hazard rate or failure rate $h_X(x)$ or $h(x)$ is the ratio of the probability density function to the survival function at all points where the probability density function is defined that is $h(x) = P(x)/S(x)$.

6.2.1 Typical Values and Deviations

The typical values of X are: the most probable value x^*, the median x_{med} and the mean m. The most probable value x^* corresponds to the maximum of the function $P(x)$. The median x_{med} is the value of x such that $P_<(x_{med}) = P_>(x_{med}) = 1/2$. The mean m or $< x >$ is the average of all possible values of X, weighted by their corresponding probability

$$m =< x >= \int xP(x)dx \tag{6.6}$$

.

If one process is repeated several times, one expects the results to be scattered in a cloud of a certain "width" in the region of typical values of X. This fluctuations of X can be described by the mean absolute deviation (MAD) $E_{abs} = \int |x - x_{med}|P(x)dx$, by the roof mean square (RMS) (the variance) $\sigma^2 =< (x - m)^2 >= \int (x - m)^2 P(x)dx$ or by the "full width at half maximum" $w_{1/2}$.

6.2.2 Moments and Characteristic Function

One can define higher-order moments of the distribution $P(x)$ as the average of powers of X

$$m_n =< x^n >= \int x^n P(x)dx \tag{6.7}$$

The first moment of a distribution measures the expected rate of return. The second moment measures the spread of a distribution. The spread or width of a distribution measures the variability of a variable (the potential scenarios of outcomes). The third moment measures a distribution's skewness, that is, how the distribution is pulled to one side or the other. The fourth moment, or kurtosis, measures the peakedness of a distribution.

The characteristic function of $P(x)$ is defined as Fourier transform

$$\widehat{P}(z) = \int e^{izx} P(x) dx \tag{6.8}$$

also we have

$$P(x) = \frac{1}{2\pi} \int e^{-izx} \widehat{P}(z) dx \tag{6.9}$$

The moments of $P(x)$ can be obtained through successive derivatives of the characteristic function at $z = 0$,

$$m_n = (-i)^n \frac{d^n}{dz^n} \widehat{P}(z)|_{z=0} \tag{6.10}$$

The cumulants c_n of a distribution is a polynomial combination of the moments m_n. It can be obtained as the successive derivatives of the logarithm of its characteristic function

$$c_n = (-i)^n \frac{d^n}{dz^n} \log \widehat{P}(z)|_{z=0} \tag{6.11}$$

The third and fourth normalized cumulants are called the skewness $\lambda_3 = \frac{<(x-m)^3>}{\sigma^3}$ and kurtosis $\kappa = \lambda_4 = \frac{<(x-m)^4>}{\sigma^4}$. The kurtosis is often taken as a measure of the distance from a Gaussian distribution. When $\kappa > 0$ (leptokurtic distributions), the corresponding distribution density has a marked peak around the mean, and rather "thick" tails. When $\kappa < 0$, the distribution density has a flat top and very thin tails.

6.2.3 Gaussian and Log-Normal Distributions

A Gaussian of mean m and root mean square σ is defined as

$$P_G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}} \tag{6.12}$$

The kurtosis of a Gaussian variable is zero. A Gaussian variable is peculiar because "large deviations" are extremely rare. The quantity $e^{-\frac{(x-m)^2}{2\sigma^2}}$ decays so fast for large x that deviations of a few times σ are nearly impossible.

X is a log-normal random variable if $\log X$ is normal, or Gaussian. The log-normal distribution density is defined as

$$P_{LN}(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{\log^2(x/x_0)}{2\sigma^2}} \tag{6.13}$$

Its use in finance comes from the assumption that the rate of returns, rather than the absolute change of prices, are independent random variables. The increments of the logarithm of the price thus asymptotically sum to a Gaussian. Its moments are equal to $m_n = x_0^n e^{\frac{n^2\sigma^2}{2}}$.

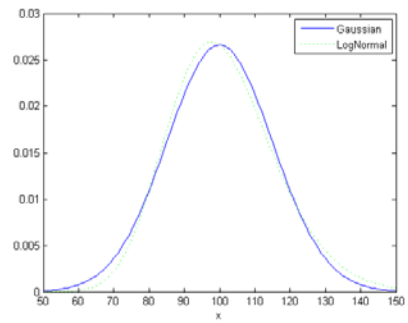

Figure 16: Comparison between a Gaussian (blue) and a log-normal (green) ddf. The difference between the two curves shows up in the tails. In this case: $m = x_0 = 100$, $\sigma_G = 15$, $\sigma_{LN} = 15\%$.

The Central Limit Theorem (CLT) states that a phenomenon resulting from a large number of small independent causes is Gaussian. There exists however a large number of cases where the distribution describing a complex phenomenon is not Gaussian: for example, the amplitude of earthquakes, the velocity differences in a turbulent fluid, the stresses in granular materials and the price fluctuations of most financial assets.

6.2.4 Levy, Discrete Poisson, Hyperbolic and Student Distributions

The tails of Levy distributions are much "fatter" than those of Gaussians, and are useful to describe multiscale phenomena. For large arguments these distributions exhibit power-law behavior, often called Pareto tails

$$L_\mu(x) \sim \frac{\mu A_\pm^\mu}{|x|^{1+\mu}} \text{ for } x \to \pm\infty \tag{6.14}$$

$A_\pm^\mu$ are called tail amplitudes, or scale parameters, $0 < \mu < 2$, gives the order of magnitude of the large (positive or negative) fluctuations of x. There is no a simple analytical expression for a symmetric Levy distribution but the full asymptotic series reads

$$L_\mu(x) = \sum_{n=1}^{\infty} \frac{(-)^{n+1}}{\pi n!} \frac{a_\mu^n}{x^{1+n\mu}} \Gamma(1+n\mu)\sin(\pi n\mu/2) \tag{6.15}$$

An interesting generalization which accounts for its exponential cut-off is given by the truncated Levy distributions.

Consider a set of points randomly scattered on the real axis, with a certain density ω (e.g. the times when the price of an asset changes). The number of points n in an arbitrary interval of length l is distributed according to the discrete Poisson distribution

$$P(n) = \frac{(\omega l)^n}{n!} e^{-\omega l} \tag{6.16}$$

The hyperbolic distribution interpolates between a Gaussian body and exponential tails

$$P_H(x) = \frac{1}{2x_0 K_1(\alpha x_0)} e^{-(\alpha\sqrt{x^2+x_0^2})} \tag{6.17}$$

where the normalization $K_1(\alpha x_0)$ is a modified Bessel function of the second kind.

The Student distribution is given by

$$P_S(x) = \frac{1}{\sqrt{\pi}} \frac{\Gamma((\mu+1)/2)}{\Gamma(u/2)} \frac{a^u}{(a^2+x^2)^{(1+u)/2}} \tag{6.18}$$

It also has power-law tails and coincides with the Cauchy distribution for $p = 1$.

6.2.5 Convolution

Lets consider X_1, X_2 independent random variables distributed according to $P_1(x_1)$ and $P_2(x_2)$ and also $X = X_1 + X_2$. The probability that $X = x$ within dx is given

$$P(x, N) = \int P_1(x')P_2(x - x')dx' \tag{6.19}$$

6.2.6 Correlations

If we assume that the correlation function $C_{i,j} = x_i x_j - m^2$ of X is non-zero for $i \neq j$ and the process is stationary,(i.e. that $C_{i,j}$ only depends on $|i - j|$: $C_{i,j} = C(|i-j|)$, with $C(\infty) = 0$). The variance of the sum of the independent random variables can be expressed in terms of the matrix C as

$$< x^2 >= \sum_{i,,j=1}^{N} C_{i,j} = N\sigma^2 + 2N \sum_{l=1}^{N} (1 - \frac{l}{N}) C(l) \tag{6.20}$$

where $\sigma^2 = C(0)$. If $C(l)$ decays faster than $1/l$ for large l, the sum over l tends to a constant for large N, and thus the variance of the sum still grows as N, as for the usual CLT.

Chapter 7

Appendix B: Financial Markets and Turbulence

On a qualitative level, turbulence and financial markets are similar. In turbulence one injects energy at large scale and the energy is transferred successively to smaller scales, as in financial markets with information. Also the price fluctuations in finance qualitatively resemble velocity fluctuations in turbulence.

7.1 Turbulence

Lets consider a fluid of kinetic viscosity v moving with velocity V in a pipe of diameter L. The complexity of the fluid is determined by the Reynolds number

$$R_e = \frac{LV}{v} \tag{7.1}$$

when the Reynolds number reaches a threshold value the fluid becomes turbulent. The Navier-Stokes equations describe the time evolution of an incompressible fluid

$$\begin{aligned} \frac{\partial}{\partial t}V(r,t) + (V(r,t)\cdot\nabla)V(r,t) &= -\nabla P + v\nabla^2 V(r,t) \\ \nabla\cdot V(r,t) &= 0 \end{aligned} \tag{7.2}$$

whose analytical solution has proved impossible and even the numerical are impossible for very large values of R_e. Kolmogorov showed that in the limit of infinite R_e the mean square velocity behaves approximately as $\left\langle [\Delta V(l)]^2 \right\rangle \sim l^{2/3}$. He also provides the exact relation for the third order $\left\langle [\Delta V(l)]^3 \right\rangle$ but fails to describe higher moments., and also fails to describe the intermittent behavior of velocity increments.

7.2 Parallel Analysis between Price Dynamics and Fluid Velocity

Mantegna and Stanley [12, 13] developed a parallel analysis between two systems, the S&P 500 index and the velocity of a turbulent fluid at high Reynolds number. Specifically, they analyzed the S&P 500 high-frequency time series during a six year period (1984-1989) and the wind velocity recorder in the atmospheric surface layer above the wheat canopy in the Connecticut Agricultural Research Station. By simple inspection, similarities and differences can be observed. But with a more deep analysis of their statistical properties both processes display intermittency and non-Gaussian features at short time interval and both are non stationary on short time scales but are asymptotically stationary.

Ghashghaie et all. [14] proposed a formal analogy between the velocity of a turbulent fluid and the currency exchange rate in the foreign exchange market. They supported their conclusion by observing that when measurements are made at different time horizons the shapes of the pdf's both change, both display leptokurtic profiles at short time horizons. However, the time correlation is completely different in the two systems and stochastic processes such as the TLF and the GARCH(1,1) also describe a temporal evolution of the pdf of the increments which evolves from a leptokurtic to a Gaussian shape, so such behavior is not specific to the velocity fluctuations of a fully turbulent fluid. The turbulence process becomes increasingly Gaussian as the time interval increases, but any scaling regime is observed.

The parallel analysis of velocity fluctuations in turbulence and index changes in financial markets shows that the same statistical methods can be used to investigate these systems. Both phenomena exhibit similarities like intermittency, non-Gaussian pdf, and gradual convergence to a Gaussian attractor in probability, and differences like the pdf's have different shapes in the two systems, and the probability of return to the origin shows different behavior. Also, for turbulence have not been observed a scaling regime whereas for index changes we observe a scaling regime spanning a time interval of more than three orders of magnitude. Moreover, velocity fluctuations are anticorrelated whereas index fluctuations are essentially uncorrelated.

Chapter 8

Appendix C: Correlations and Anticorrelations between Stocks

The presence of correlations or anticorrelations between stocks has been long known and they play an important role in the determination of selecting the most efficient portfolio [15, 16]. Mantegna [17] studied how important are these correlations to detect the amount of synchronization present in the dynamics of a pair of stocks traded in a financial market.

8.1 Simultaneous dynamics of Pairs of stocks

The log-changes in prices of stock i is defined as

$$S_i = \ln Y_i(t) - \ln Y_i(t-1) \tag{8.1}$$

where Y_i is the closure price of the stock i (at time t). The correlation coefficient coefficient the S_i and S_j of the stocks i and j

$$\rho_{ij} = \frac{\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle}{\sqrt{\left\langle S_i^2 - \langle S_i \rangle^2 \right\rangle \left\langle S_j^2 - \langle S_j \rangle^2 \right\rangle}} \tag{8.2}$$

If $\rho_{ij} = 1$ the changes in stock prices are completely correlated, $\rho_{ij} = -1$ completely anticorrelated and $\rho_{ij} = 0$ uncorrelated. Evidently, this values of ρ_{ij} vary in time. One way of determining how long is the characteristic time scale over which strongly correlated stocks maintain correlated is the deviation of ρ_{ij}

$$\delta_{ij} = \frac{\rho_{ij} - \langle \rho_{ij} \rangle}{\sigma} \tag{8.3}$$

From the time evolution of the prices of the stocks of Coca Cola and Procter & Gamble in 1990 is evident that the prices of the stocks are remarkably synchronized. By applying the approach was found a value of $\rho_{ij} = 0.73$ and during a period of five years (1990-1994) $\delta_{ij} > 1$.

Chapter 9

Appendix D: Risk Measures

Measuring and controlling financial risks is a major concern in the economic world. Several measures of risk have been introduced [10, 18] in order to reduce losses and allow these risks even to be traded. However, the classical definitions are considered "weak" to explain "rare events" where true financial risk resides.

Risks are categorized into five categories [55]: Credit, Market, Operational, Liquidity and Legal risk. The Basel Committee has developed a framework for the determination of minimum capital requirements for banks (Pillar I). However, such minimum capital requirements will need to be supported by a robust implementation related to the supervisory process and market conduct (Pillar II and III).

Pillar I: Minimum Capital Requirements: Definition of regulatory capital, risk-weighted assets, and the minimum ratio of capital to risk-weighted assets.

Pillar II: Supervisory Review Process: Prudential supervision by regulatory authorities of banks' capital adequacy as well as the banks' internal risk management systems.

Pillar III: Market Discipline: Market discipline by developing a set of disclosure requirements that will allow market participants to assess key pieces of information on the scope of application, capital, risk exposures, risk assessment processes, and hence the capital adequacy of the institution. These are especially useful when banks are given the authority to use bank-specific internal models in assessing their own risk profiles.

9.1 Risk and Volatility

Financial risk has been traditionally associated with the statistical uncertainty on the final outcome [10]. Its traditional measure is the RMS (the volatility). It is denoted by $R(T)$ the logarithmic return on the time interval T defined by

$$R(T) = \log[\frac{x(T)}{x_0}] \tag{9.1}$$

where $x(T)$ is the price of the asset X at time T, knowing that it is equal to x_0 today ($t = 0$). When $|x(T) - x_0| << x_0$, this definition is equivalent to $R(T) = x(T)/xo - 1$. If $P(x, T|xo, 0)dx$ is the conditional probability of finding $X(T) = x$ within dx, the volatility σ of the investment is the standard deviation of $R(T)$ defined by [10]

$$\sigma^2 = \frac{1}{T}[\int P(x,T|xo,0)R^2(T)dx - (\int P(x,T|xo,0)R(T)dx])^2 \tag{9.2}$$

The volatility is still now chosen as an adequate measure of risk associated to a given investment even if this definition includes in a symmetrical way both abnormal gains and abnormal losses. The theoretical foundations behind this particular definition of risk are numerous: First, operational; the computations involving the variance are relatively simple and can be generalized easily to multi-asset portfolios. Second. the Central Limit Theorem seems to provide a general and solid justification: by decomposing the motion from x_0 to $x(T)$ in $N = T/\tau$ increments, one can write

$$x(T) = x_0 \sum_{k=0}^{N-1} (1 + \eta_k) \tag{9.3}$$

where η_k is by definition the instantaneous return. Therefore we have

$$R(T) = \sum_{k=0}^{N-1} \log[1 + \eta_k] \tag{9.4}$$

Classically, we assume that η_k are independent variables but from the CLT we have that in the limit where $N \to \infty$, $R(T)$ becomes a Gaussian random variable centred on a given average return $\overline{m}T$, with $\overline{m} =< \log(1+\eta_k) > /\tau$, and whose standard deviation is given by $\sigma\sqrt{T}$. Therefore, in this limit, the entire probability distribution of $R(T)$ is parameterized by two quantities only $\overline{m}$ and σ: any reasonable measure of risk must therefore be based on σ. However, this is not true for finite N (which corresponds to the financial reality: there are only roughly $N \simeq 320$ half-hour intervals in a working month), especially in the tails of the distribution corresponding precisely to the extreme risks.

One can give to σ the following intuitive meaning: after a long enough time T, the price of asset X is given by

$$x(T) = x_0 e^{\overline{m}T + \sigma\sqrt{T}\xi} \tag{9.5}$$

where ξ is a Gaussian random variable with zero mean and unit variance. The quantity σ gives us the order of magnitude of the deviation from the expected return. By comparing the two terms in the exponential, one finds that when $T \gg \widehat{T} \equiv \sigma^2/\overline{m}^2$ the expected return becomes more important than the fluctuations, which means that the probability that $x(T)$ is smaller than x_0 becomes small. The **security horizon** $\widehat{T}$ increases with σ. For a typical individual stock, one has $\overline{m} = 10\%$ per year and $\sigma = 20\%$ per year, which leads to a T as long as 4 years.

The quality of an investment is often measured by its **Sharp ratio** S that is, the signal-to-noise ratio of the mean return m^*T to the fluctuations $\sigma\sqrt{T}$

$$S = \frac{m^*\sqrt{T}}{\sigma} \tag{9.6}$$

where $m^* = \overline{m} - r$. The Sharpe ratio increases with the investment horizon and is equal to 1 precisely when $T = \widehat{T}$. The error in the sharpe radio reads

$$\Delta S = \frac{1}{\sqrt{N}}[\sqrt{T} + S\sqrt{2+\kappa}] \tag{9.7}$$

A purely additive description is actually more adequate at short time, so

$$x(T) = x_0 e^{\overline{m}T + \sigma\sqrt{T}\xi} \simeq x_0 + mT + \sqrt{DT}\xi \tag{9.8}$$

where $m \equiv \overline{m}x_0$ and $D \equiv \sigma^2 x_0^2$. The non-Gaussian nature of the random variable ξ is the most important factor determining the probability for extreme risks.

9.2 Value at Risk (VaR)

Behind the fact that generally financial risks are described through the volatility is the idea that the distribution of price changes is Gaussian. This definition have some disadvantages: The financial risk is based of losses and not in profits. A Gaussian model for the price fluctuations is never justified for extreme events which can lead to a very bad empirical determination of the variance. The solution of this problem consisted in remove from the analysis the contribution of this events. For a better control of the financial risk, another definition of risk is needed. One of them is the probability of extreme losses or the value at risk (VaR) which we will define bellow.

The probability to lose an amount $-\delta x$ larger than a certain threshold Λ on a given time horizon τ is defined as [10]

$$P[\delta x < -\Lambda] = P_<[-\Lambda] = \int_{-\infty}^{-\Lambda} P_\tau(\delta x) d\delta x \tag{9.9}$$

where $P_\tau(\delta x)$ is the probability density for a price change on the time scale τ. Alternatively, we can define the risk as a level of loss (VaR) Λ_{VaR} corresponding to a certain probability of loss P_{VaR} over the time interval τ

$$\int_{-\infty}^{-\Lambda_{VaR}} P_\tau(\delta x) d\delta x = P_{_{VaR}} \tag{9.10}$$

This definition means that a loss greater than Λ_{VaR} over a time interval $\tau = 1$ day happens only every 100 days on average for $P_{_{VaR}} = 1\%$. Lets note that this definition does not take into account the fact that losses can accumulate on consecutive time intervals τ, leading to an overall loss which might substantially exceed Λ_{VaR}. Also, this definition does not take into account the value of the maximal loss inside the period τ. In other words, only the closing price over the period $[k\tau, (k+1)\tau]$ is considered, and not the lowest point reached during this time interval.

More precisely, one can discuss the probability distribution $P(\Lambda, N)$ for the worst daily loss Λ (we choose $\tau = 1$ day to be specific) on a temporal horizon $T = N\tau$.

$$P(\Lambda, N) = N[P_>(-\Lambda)]^{N-1} P_\tau(-\Lambda) \tag{9.11}$$

For N large, this distribution takes a universal shape that only depends on the asymptotic behavior of $P_\tau(\delta x)$ for $\delta x \to -\infty$. In the important case for practical applications where $P_\tau(\delta x)$ decays faster than any power-law, one finds that $P(\Lambda, N)$ is given by $\frac{z}{\beta} e^{-z}$ where $z = e^{-\frac{x-u}{\beta}}$ which is represented in the next figure

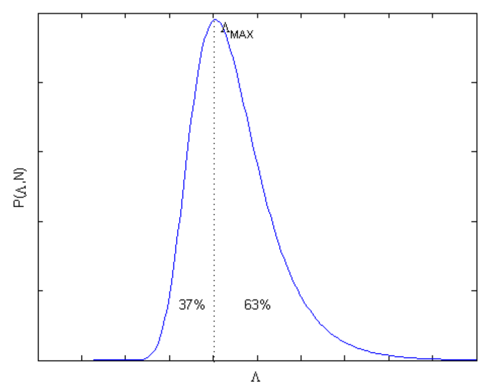

Figure 17: Gumbel distribution. $\Lambda_{\max}$ has a probability equal to 0.63 to be exceeded.

This distribution reaches a maximum precisely for $\Lambda = \Lambda_{vaR}$. The intuitive meaning of Λ_{vaR} is thus the value of the most probable worst day over a time interval T. Note that the probability for Λ to be even worse is equal to 63%. One could define Λ_{vaR} in such a way that this probability is smaller, by requiring a higher confidence level, for example 95%. This would mean that on a given horizon (for example 100 days), the probability that the worst day is found to be beyond Λ_{vaR} is equal to 5%. This Λ_{vaR} then corresponds to the most probable worst day on a time period equal to $100/0.05 = 2000$ days.

Chapter 10

Appendix E: Non Linear Correlations and Volatility Fluctuations

There is strong evidence supporting the conclusion that the volatility (of log price changes) is a time-dependent stochastic process [6, 7, 10]. One approach for describing stochastic processes characterized by a time dependent variance (volatility), is the ARCH processes introduced by Engle in 1982. These ARCH models have been applied to means and variances of inflation in the UK, stock returns, interest rates, and foreign exchange and can also be very attractive for describing physical systems. The difficulty in the determination of some coefficients in this approach led to the introduction of generalized ARCH processes called GARCH by Bollerslev in 1986.

10.1 ARCH(p)

An ARCH(p) process is a stochastic process with non constant variances conditional on the past (but constant unconditional variances) is defined by

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + ... + +\alpha_p x_{t-p}^2 \tag{10.1}$$

where α_i are positive variables and x_t is a random variable (with zero mean and σ_t^2 variance) characterized by a conditional pdf usually taken to be a Gaussian.

The ARCH(1) process is defined by

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 \tag{10.2}$$

and $S(t) = \sum_i^t x_{i.}$ The value of the variance for ARCH(1) with a Gaussian conditional pdf is [7]

$$\sigma^2 = \frac{\alpha_0}{1 - \alpha_1} \tag{10.3}$$

and the kurtosis is

$$\kappa = \frac{\langle x^4 \rangle}{\langle x^2 \rangle^2} = 3 + \frac{6\alpha_1^2}{1 - 3\alpha_1^2} \tag{10.4}$$

which is finite if $0 \leq \alpha_1 \leq 1/\sqrt{3}$. By varying α_0 and α_1 it is possible to obtain stochastic processes with the same unconditional variance but with different values of the kurtosis.

10.2 GARCH(1,1)

In this model, the volatility evolves through a simple feedback mechanism given by [7, 10]

$$\begin{aligned} \eta_i &= \epsilon_i \sigma_i \\ \sigma_i^2 &= \sigma_0^2 + \alpha(\sigma_{i-1}^2 - \sigma_0^2) + g(\eta_{i-1}^2 - \sigma_{i-1}^2) \end{aligned} \tag{10.5}$$

where ϵ_i are iid Gaussian random numbers of zero mean and unit variance. σ_0^2 is the unconditional average variance, α measures the strength with which the volatility reverts to its average value and g is a coupling parameter describing how the difference $\eta_{i-1}^2 - \sigma_{i-1}^2$ feedback in the present value of volatility.

In this model $<\eta_i>=0$ and $<\eta_i\eta_j>=0$ for $i \neq j$ and the unconditional variance is given by $<\eta_i^2>=<\sigma_i^2>=\sigma_0^2$. Setting $\delta_i = \eta_i^2 - \sigma_i^2$ and $\psi_i = g\eta_i^2 + (\alpha - g)\sigma_i^2 + \alpha\sigma_0^2$ transforms the system into

$$\begin{aligned} \delta_i &= (\psi_{i-1} + \sigma_0^2)\xi_i \\ \psi_i &= \alpha\psi_{i-1} + g(\psi_{i-1} + \sigma_0^2)\xi_i \end{aligned} \tag{10.6}$$

where $\xi_i = \epsilon_i^2 - 1$ is a zero mean non-Gaussian noise.

The volatilities are

$$<\sigma_i^2\sigma_j^2> - <\sigma_i^2><\sigma_j^2> = \frac{2g^2\sigma_0^4}{1-\alpha^2-2g^2}\alpha^{|i-j|} \tag{10.7}$$

this holds for $\alpha < 1$ and $2g^2 < 1 - \alpha^2$ (stationary model).

The square return correlation function is given by [7, 10]

$$C_2(i,j) = <\eta_i^2\eta_j^2> - <\eta_i^2><\eta_j^2> = \begin{cases} \frac{2\sigma_0^4(1-\alpha^2+g^2)}{1-\alpha^2-2g^2} & i = j \\ \frac{2g\sigma_0^4(1-\alpha^2+g\alpha)}{1-\alpha^2-2g^2}\alpha^{|i-j|-1} & i \neq j \end{cases} \tag{10.8}$$

and the unconditional kurtosis

$$\kappa = \frac{6g^2}{1-\alpha^2-2g^2} \tag{10.9}$$

which is zero when there is no feedback i.e. $g = 0$.

In Figure 18 we show a GARCH(1,1) process, consisting in a single path of 1000 observations for return series, innovations, and conditional standard deviation processes for the asset S&P500 (Figure 1). Innovations represents a single path of a unvaried time series. The first element of this time series contains the oldest observation, and the last element the most recent. Standard deviation is the time series of conditional standard deviations.

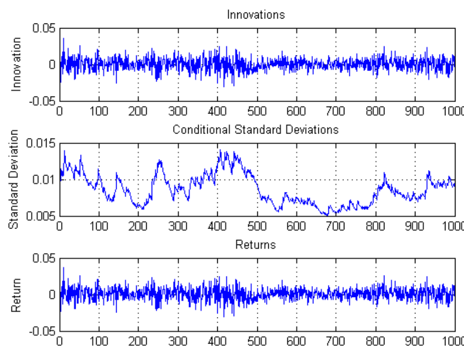

Figure 18: GARCH(1,1), single path of 1000 observations for return series, innovations, and conditional standard deviation processes (S&P500).