STOCHASTIC ANALYSIS OF A EUROPEAN MARKET USING R

Capstone Design
Spring 2019
Sabrine Bentaleb
Supervised by Dr. L. Laayouni
STOCHASTIC ANALYSIS OF A EUROPEAN MARKET USING R

Capstone Final Report

Student Statement:
I affirm applying ethics to the design process and in the selection of the final proposed design. I held the safety of the public to be paramount and has addressed this in the presented design wherever may be applicable.”

Sabrine Bentaleb

Approved by the Supervisor

Dr. Lahcen Laayouni
ACKNOWLEDGMENTS

“Finance is not merely about making money. It’s about achieving our deep goals and protecting the fruits of our labor. It’s about stewardship and, therefore, about achieving the good society”.

Robert J. Shiller [1]

I was first introduced to the world of finance in AUI when I decided to specialize in this field. I would like to thank my finance Professors for the interest they helped me grow for this sphere of study. I also had the opportunity to relate theory to practice in internships and believe that finance is not only about profits but rather the way it is generated. I tried my best to apply ethics in my capstone project and perform extensive research. It has been a successful experience thanks to the precious help of my supervisor Dr. Lahcen Laayouni. I would like to thank him for his continuous support and attention to details. He provided me with guidance and advices during my project but also during the courses he previously taught me. I am forever grateful for the opportunity he gave me.

I finally express my gratitude to my family and friends for their support and encouragements throughout my undergraduate studies at Al Akhawayn University.
ABSTRACT
This report will consist of analysing a European market, CAC 40, and model it following a stochastic process using R. The random process relies on probability theory which provides mathematical definitions to these processes. Indeed, stochastic calculus applications to finance are endless such as the ‘Black-Scholes Option Pricing Formula’ and ‘Brownian Motion’.

The data is retrieved from the CAC 40 stock prices for a one-year period, 2018, and analysed using Excel. The findings match the EMH or efficient market hypothesis and the factors affecting the market trends. This confirms that Paris stock exchange is an efficient market and that all available information is reflected on stock prices. The index declined tremendously in 2018 due to a slowdown in global economic growth and Brexit political consequences on the European union.

The method chosen to model the CAC 40 predictions for 2019 is the simple random walk. Statisticians and mathematicians believe that random walk is the most accurate way to estimate the possible outcomes as it relies on the EMH. The model intends to predict these different outcomes using R and may be optimized to reduce generated statistical errors. The last step will be to compare the results of the model to reality. The aim of the model remains to predict the trends of the market and thus to capture profits.

Key words: European market, CAC 40, stochastic, market trends, forecast, R...
# TABLE OF CONTENTS

ACKNOWLEDGMENTS .................................................................................................................. 3

ABSTRACT ....................................................................................................................................... 4

LIST OF FIGURES ............................................................................................................................. 1

LIST OF TABLES ................................................................................................................................. 1

Chapter I: INTRODUCTION .............................................................................................................. 2

1 Efficient Market Hypothesis ....................................................................................................... 2
2 Stochastic Modelling ...................................................................................................................... 2
3 Stochastic Calculus Applications to Finance ................................................................................ 2
4 CAC 40............................................................................................................................................... 3
4.1 Presentation ................................................................................................................................... 3
4.2 Evolution ....................................................................................................................................... 4
4.3 Challenges .................................................................................................................................... 5
5 STEEPLE ANALYSIS ....................................................................................................................... 6
5.1 Societal Factor .............................................................................................................................. 6
5.2 Technological Factor ................................................................................................................... 6
5.3 Environmental Factor .................................................................................................................. 6
5.4 Ethical Factor .............................................................................................................................. 7
5.5 Political Factor ............................................................................................................................ 7
5.6 Legal Factor ............................................................................................................................... 7
5.7 Economic Factor ......................................................................................................................... 7

Chapter 2: DATA AND METHODOLOGY ......................................................................................... 9

1 Data................................................................................................................................................. 9
1.1 Definition of Data ....................................................................................................................... 9
1.2 Descriptive Statistics .................................................................................................................. 9
1.3 Stock Prices vs. Volatility.......................................................................................................... 11
2 Methodology ................................................................................................................................... 13
2.1 Stochastic Modeling .................................................................................................................. 13
2.2 Stochastic Applications in R ...................................................................................................... 14
2.3 Random Walk in Stock Markets ............................................................................................... 15
2.4 Simple Random Walk ............................................................................................................... 16
2.5 Advantages and Disadvantages ............................................................................................... 18

Chapter 3: RESULTS AND INTERPRETATION ............................................................................. 19

1 CAC 40 Predictions’ Model .......................................................................................................... 19
1.1 Simple Random Walk ............................................................................................................... 19
1.2 Normal Distribution Assumption ............................................................................................ 21
1.3 Data Analysis using Difference Equation ................................................................................ 24
1.3.1 Arithmetic and Geometric Random Walk ........................................................................ 25
1.3.2 EMH Assumptions .............................................................................................................. 26
1.4 Model Implementation .............................................................................................................. 27
2 Application of the Model using IQ Option .................................................................................. 29
2.1 Binary Option ........................................................................................................................... 29
2.2 Random Walk Index ............................................................................................................... 29
2.3 Investments’ Scenarios ............................................................................................................ 29

Chapter 4: FUTURE WORK AND CONCLUSION ........................................................................... 31
LIST OF FIGURES

Figure 1- The CAC 40 Evolution over the Years.......................................................... 5
Figure 2- The CAC 40 Stock Prices in 2018 ............................................................ 11
Figure 3- The CAC 40 Volatility in 2018................................................................. 12
Figure 4- Simple Random Walk Trials .................................................................... 17
Figure 5- Prediction of CAC 40 using Simple Random Walk – 10 Trials .................. 20
Figure 6- Prediction of CAC 40 using Simple Random Walk – 50 Trials .................. 20
Figure 7- CAC 40 Evolution for 2019...................................................................... 21
Figure 8- CAC 40 Prediction using Normal Distribution Assumption-Trial 1 ............. 23
Figure 9- CAC 40 Prediction using Normal Distribution Assumption-Trial 2 ............. 23
Figure 10- Daily Returns of CAC 40 in 2018............................................................ 24
Figure 11- Arithmetic Random Walk with Drift......................................................... 25
Figure 12- CAC 40 Predictions for 2019-Trial 1 ...................................................... 28
Figure 13- CAC 40 Predictions for 2019-Trial 2 ...................................................... 28
Figure 14- Investment Scenario 1 – Down Selection.................................................. 30
Figure 15- Investment Scenario 2 – Up Selection...................................................... 30

LIST OF TABLES

Table 1- Stock Prices’ Descriptive Statistics............................................................... 10
Table 2- Volatility Descriptive Statistics.................................................................... 12
Chapter I: INTRODUCTION

1 Efficient Market Hypothesis
Modern theory states that financial markets rely on the EMH or efficient market hypothesis which assumes that the stock price reflects all available information in the market. The efficient markets are characterized by the rapid correction of deviation in stock prices from true value. Indeed, each stock price is a representation of the discounted present values of the sum of future earnings [2]. Random variables are predominant in determining the stock price which continually change and adapt to internal or external factors. These variables range from daily return to volatility. The major factors can be political, economic, environmental or related to market psychology. They are all related as the economy of any country is affected by its political environment and market psychology dictates the ups and downs of the market. Indeed, each government shapes specific monetary and fiscal policy based on political and environmental factors. These policies affect the market trends as they shape international transactions which characterise the economic strength of the country. Market psychology regroups the speculation and expectation made which drive prices based on what to expect, and changes in supply and demand which create trends [3].

2 Stochastic Modelling
Several tools are used for predicting the possible outcomes of the market based on the factors that shape the stock prices’ trends. Stochastic models are one of these tools and are ways to quantify the probability of possible outcomes for random variables in a process over time. The analysis of stochastic processes starts with conditional probability under probability theory. Indeed, the theory provides mathematical definitions of random processes such as stochastic processes. These definitions are developed based on the study of experiments based on chance. The study is done at the beginning and the definitions are developed as probability develops through the experiments. The original interest behind probability theory was to give insights on how gambling games work. Since then, the scope extended beyond understanding games based on chance. By the early 2000’s, probability theory was not only related to number or ergodic theory, but also to “partial differential equations and differential geometry” [4].

3 Stochastic Calculus Applications to Finance
Stochastic modelling aims mainly to increase profitability by predicting future trends. Indeed, stochastic calculus applications in finance are countless such as the ‘Brownian Motion’, ‘Black-Scholes Option Pricing Formula’ or ‘Monte Carlo Simulation’. A stochastic process is also a
Brownian Motion when it starts from zero and the increments are independent and stationary on a continuous path. Black-Scholes equation is the output of stochastic modelling applied to finance based on economic reasoning. The equation is used to price European option by deriving volatility based on a particular differential equation [5]. A Markov process is a stochastic process with a Markov property, when the possible outcomes or stock prices’ predictions are based only on present information. Monte Carlo simulation is a random Markov process and generates samples of random variables that it evaluates using a performance function to extract probabilistic information [6]. The applications to stochastic calculus all intend to predict the random behavior the most accurately and increase profitability when applied to finance. The key to building accurate stochastic models remain to suit the intended purpose and not uniqueness as more than one model can represent a specific event but each model has a specific purpose [7].

4 CAC 40
An efficient market as the EMH or efficient market hypothesis mentions is a market where the stock price reflects all available information in the market. CAC 40 or the continuous assisted quotation is the index representing French economy. The index fluctuations reflect the country stability on the economic or political level. Therefore, CAC 40 stock price must go back quickly to its true value and the market as a whole is an efficient market. In addition, stock prices and most random variables affecting the market can be easily accessed online. Both conditions will facilitate the stochastic modelling aiming to increase profitability.

4.1 Presentation
The CAC 40 stands for Continuous Assisted Quotation and is a capitalization-weighted index of the top 40 companies out of 100 with the highest market caps on the Euronext Paris. Market capitalization quantifies the value of a company and is calculated by multiplying the share price by the number of shares outstanding. The index remains a good indicator of French economy as all the listed companies are French-domiciled and multinationals. In fact, almost half of shares is purchased by foreign investors.

The CAC 40 index needs continuous supervision. The index composition is reviewed quarterly by a committee following the companies ranking on the Euronext Paris. The ranking analyze the free float market capitalization, incorporated by the CAC 40 since 2003 by adjusting with a free float factor $F_{i,t}$, and the share turnover of the prior twelve months. The free float market capitalization excludes the shares held by the promoters and only accounts for the public share in the market capitalization computation, whereas the share turnover is a measure of stock
liquidity and is calculated by dividing the total number of shares traded by the average number of outstanding shares [8].

The 40 companies out of the 100 with the best ranking are chosen for the CAC 40 composition and represent a benchmark for good portfolio management. The index calculation requires first the market capitalization calculation of each company as the CAC 40 is a cap-weighted index. The index weightings maximum value allowed is 15%, a capping factor \( f_{i,t} \) is used to limit the weights to 15% when needed.

The overall index calculation computes the index value at day \( t \) for \( N \) share, usually 40. The market capitalization is calculated by multiplying the number of shares \( Q \) times the stock price \( C \). The final value of the market capitalization calculation for each company after adjustments is divided by the base date market capitalization of these companies. This base date market capitalization is adjusted by a coefficient \( K_t \) reflecting the transition from Franc to Euro in 1999, the base date. The overall index calculation is as follows [9]:

\[
I_t = 1000 \times \frac{\sum_{i=1}^N Q_{i,t} F_{i,t} f_{i,t} C_{i,t}}{K_t \sum_{i=1}^N Q_{i,0} C_{i,0}}
\]

4.2 Evolution

The index value is expressed in basis points (1 bps = 0.01%) and fluctuated along the years since its creation. The index was created after the Black Monday in 1987 where all the stock markets of the world crashed which led to a change in stock exchange transactions. The base value was first set to 1000 in 1987, 31st December, the currency used in France was still Francs. Financial markets are affected by external and internal factors, from economic factors to market psychology. In fact, the dot-com bubble that happened from 1995 to 2000 affected the CAC 40 index positively and its value increased up to 7000 bps in 2000. The dot-com bubble encouraged investors to purchase stocks from internet based companies as the stock prices were increasing and most world stock markets were impacted positively. Whereas, the financial crisis that occurred in Europe because of a threat of sovereign default led to a decrease in the CAC 40 value up to 3000 bps in 2011. The index value then recovered slowly and went up to 5000 in 2017 and 2018 [10].
4.3 Challenges

CAC 40 main challenge is to face fluctuations as seen in the evolution. These fluctuations occur according to external factors, whether economic, political or both. The index decreased up to 3000 bps in 2011 due to the financial crisis. Indeed, the index is influenced by international transactions as all the listed companies are French-domiciled but operate on the international level [10]. More than 45% of CAC 40 shares are purchased by foreign investors, theses investors have to be careful with the tax complications and any currency risks. Investors have the possibility to either invest directly in foreign stock on the Euronext Paris, have a direct exposure by investing in ETF or exchange traded funds, a marketable security that tracks the stock index on the NYSE Euronext Paris, or purchase individual components of the index using ADR’s or American Depository Receipts which entitles the investor to shares of foreign stocks being held at a depository bank. If an investor decides to go for the last option, liquidity risks have to be considered [8].

The CAC 40 index may face some challenges regarding the stock price fluctuations but the index always managed to recover the true value. The correction periods remained short over the years and the evolution shows a growth trend. Moreover, the index is an accurate representation of France economy as it takes into consideration the 40 companies with the highest market capitalization and most of them operate on the international level. French economy remains important for Morocco as the past colonialism and close geographical position makes France
economy affecting Morocco. The index has been chosen for an interest in predicting France economy and its reliability.

5 STEEPLE ANALYSIS
CAC 40 index is a fair indicator of French economy which is among the leading ones in Europe. Financial markets are affected by internal and external factors which dictates the ups and downs of the index. STEEPLE analysis relate financial markets fluctuations and thus profitability opportunities to the following factors:

5.1 Societal Factor
The societal factor influences the economy in many aspects as social media are more widely used nowadays than before. Social media allowed the “democratization of information” which facilitates advertising for companies and lead to higher profitability. Social media also set the rules and define financial policies as major platforms like Facebook and Google control more than half of information online. Indeed, a financial crisis would for example happen due to a chaotic social background where information is not filtered and advertising is disrupted, whereas financial markets can expand tremendously when the country is socially stable. Financial markets’ expansion sets the ground for investing in portfolios and increase profitability. The major platforms build new ecosystems for ambitious entrepreneurs which develops the economy as a whole [11].

5.2 Technological Factor
Technological advances grew exponentially over the last decades and changes businesses’ operations and deliveries of products. Indeed, these technological advances facilitate price or volatility computations and more and more technological research and development is done to enhance financial products. Indeed, tech-related companies are the most profitable and active on the stock market in developed countries. Technology is thus affecting the financial and banking sectors and implies that the country has a good economic growth. Fin-Tech is what describes the merging of technology with finance, it affects financial institutions and banks through “customer service” which is crucial for any fin-tech company as it defines the level of profitability. These companies should make sure of clients’ satisfaction as their products can be intangible [12].

5.3 Environmental Factor
Finance is developing an environmental sense as more and more companies are becoming eco-friendly. These companies care as much for their development growth than for their
environmental growth. Moreover, finance can find solutions to the main environmental issues. Companies get specialized in producing environmental goods and delivering services like energy audit. These companies’ profitability can then be reinvested in clean energy and eco-friendly industries as whole [13].

5.4 Ethical Factor
The fall of Wall Street was mainly due the lack of ethics in finance which led to the collapse of the united states. Any customer seeking for financial services would examine closely the ethics of the institution and its popularity, small institutions would need to double efforts compared to a huge corporation. ‘Ethical finance’ is spreading and takes into consideration environmental and societal factors in addition to the financial returns. Investors attribute more value to the environmental and societal impact of their investments and choose wisely when building their portfolios [14].

5.5 Political Factor
During the last decades, economic power correlation to political power has been demonstrated as politics are affected by money. Finance scholars are more interested in studying political economy and vice versa. The correlation revolves around money as financial means can influence elections positively or negatively. Regulating this influence and political economy as a whole is crucial for financial markets’ equilibrium. Indeed, political stability has an impact on the market fluctuations and thus affect the stock prices. The relationship is established on both ways as markets’ stability lead to political stability and accurate prediction of the market trends ensure stability. Therefore, precise financial predictions can benefit politics and political stability affect the market trends positively as profitability opportunity increases [15].

5.6 Legal Factor
Legislative regulations in financial markets are established to ensure political stability and a healthy economy overall. Regulations ensure that products and services are well marketed and customer service reinforced. The legislative body is set to be independent from private financial institutions as the autonomy incurred will guarantee efficient operations in the interest of society and the economy as a whole. Each country relies on a governmental body for regulating markets. In the US, the SEM or ‘Securities Exchange Commission’ investigates all income statements while in Morocco, it’s the AMMC [16].

5.7 Economic Factor
Economic indicators are reflecting in financial markets and stock prices. Indeed, macroeconomic factors remain good predictors of the markets potential risk.
The first indicator is GDP or gross domestic product which is related to economic growth as a higher GDP means more job prospects and a powerful economy.

The second indicator is inflation and is directly related to risk, as inflation increase, volatility or variability of stock returns increase, which leads to economic growth slow down. The long-run economic growth can thus be predicted by estimating volatility of returns in efficient stock markets.

The third indicator is monetary policy which impacts the interest rate on the national level and the exchange rate on the international level. It is crucial to consider all the important aspects and goals of the economy before setting these policies.

The last component to take into consideration is stock liquidity. The higher the liquidity the higher the investments in the stock market, as lower default risk means lower volatility and market psychology is influenced positively [17].
Chapter 2: DATA AND METHODOLOGY

1 Data

Stock prices of the CAC 40 are retrieved for a one-year period, 2018. The data of 2019 will also be retrieved to test the model built. The descriptive statistics of the data is done on Excel and the model implementation on R. R is an accessible software used for statistical computing and graphics as it is free and runs on a wide variety of platforms. The final step will be to compare the results of the model with reality.

1.1 Definition of Data

The CAC 40 stock prices fluctuate following market trends affected by internal and external factors. The internal factors are related to the organizations’ operations and structure. The external ones are related to economy, politics and environment. Economic stability leads to political stability and vice-versa. The financial strategies followed by companies among the market should be ethical and have no or a positive impact on the environment. The resulting fluctuations quantify profitability opportunity and high stock prices remain a sign of profitable companies and thus a healthy economy.

The CAC 40 index remains reliable as its composition is continuously assessed (quarterly) by a committee and the capping factor is evaluated once a year during the first review meeting. Nevertheless, the index may encounter ups and downs in the stock prices which is measured by volatility, the standard deviation of an asset returns or a key determinant of asset’s profitability as a whole [18].

In fact, investing in stock markets means encountering risks. In addition to considering volatility, investors seek to increase their wealth by speculation and gambling. Speculation means studying the potential risks and hoping to get a bigger gain than risk, whereas gambling means investing money with an uncertain outcome following the game of chance. Gambling investment has to follow regulations in order not to become unsafe, especially the compulsory gambling which can be excessive and ruin investors’ life. Gambling causes and consequences have to be defined and eliminated [19].

1.2 Descriptive Statistics

The stock prices of the CAC 40 have been retrieved from yahoo finance for the period of one year, 2018. The descriptive statistics have been performed using excel.
The descriptive statistics summary of each stock price is tabulated. From the table, it can be observed that:

- The mean and median for all stock prices are almost equal (around 5300 bps) which shows that the normal distribution is symmetric and skewness is zero.

- The difference between the highest and lowest prices’ range is not too important (|993,24 – 1072,94|=79,7) which indicates that the fluctuations in the stock prices are not too erratic. The range or price volatility is a good indicator of risk. The investors can consider this parameter to decide when to invest. The price range over 2018 is around 1000 bps.

- The standard deviation which is a measure of volatility is around 200 when the average stock price is 5000 bps which confirms that that the stock prices are not too erratic.

The findings confirm that the CAC 40 in an efficient market as it fulfills the characteristics of the EMH or efficient market hypothesis. The deviations of the stock prices from their true value are small (around 200) as all available information is reflected on the stock prices. The overall results show a decline pattern in 2018 especially towards the end, as shown in the figure below:
Indeed, the CAC 40 has known the lowest decline in 2018 (10.95%) since 2011. The index started at 5300 bps which is above the symbolic threshold and increased up to 5657 bps in May after the tax reform executed by Donald Trump. The CAC 40 decreased until reaching around 4650 bps end 2018. The time plot shows that the lowest decline occurs during the last three months of 2018 mainly due to the slowing global economic growth. The slowdown can be explained by trade frictions that occurred between the United States and China. Moreover, environmental policies applied on Germany impacted negatively the car industry and sovereign risk affected supply and demand in Italy. In France, the street manifestations impacted negatively the tourism industry and thus the economy overall. In addition to the decrease in the global economic growth, ‘Brexit’ or UK divorce from EU led to political disagreements and a sense of denying the power of the European union [20]. These two factors impacted the CAC 40 negatively and experts predict a correction or reverse trend for 2019.

1.3 Stock Prices vs. Volatility
The descriptive statistics of the data showed that the stock prices do not fluctuate and their normal distribution is symmetrical which matches the EMH. CAC 40 is an efficient market so all available information is reflected on the stock price as the time plot of 2018 shows a decrease due to external factors. Standard deviation, a measure for volatility, was found to be small as the price goes back to true value quickly. The volatility should be examined more closely, especially its correlation to stock prices.
<table>
<thead>
<tr>
<th></th>
<th>Close</th>
<th>High</th>
<th>Low</th>
<th>Open</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>14,95</td>
<td>16,84</td>
<td>13,43</td>
<td>14,63</td>
</tr>
<tr>
<td>Median</td>
<td>14,06</td>
<td>15,15</td>
<td>12,77</td>
<td>13,80</td>
</tr>
<tr>
<td>Min</td>
<td>9,59</td>
<td>10,16</td>
<td>6,11</td>
<td>9,73</td>
</tr>
<tr>
<td>Max</td>
<td>30,45</td>
<td>96,27</td>
<td>23,17</td>
<td>27,84</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Close</th>
<th>High</th>
<th>Low</th>
<th>Open</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td>20,86</td>
<td>86,11</td>
<td>17,06</td>
<td>18,11</td>
</tr>
<tr>
<td>Variance</td>
<td>11,37</td>
<td>75,46</td>
<td>6,44</td>
<td>10,29</td>
</tr>
<tr>
<td>Std.</td>
<td>3,37</td>
<td>8,69</td>
<td>2,54</td>
<td>3,21</td>
</tr>
</tbody>
</table>

**Table 2: Volatility Descriptive Statistics**

The descriptive statistics summary for each stock volatility is tabulated, historical volatility over 2018. From the table, it can be observed that:

- The mean and median of the index volatility are almost equal which shows that the distribution is symmetrical and skewness is zero. The volatility distribution goes hand in hand with the stock price distribution.

- The difference between the high and low volatility range is not negligible, $86,11 - 17,06 = 69,05$. It can be explained by the spike in volatility the markets knew in February. Moreover, volatility quantifies the degree of variation in stock prices or the deviations from the mean value which is around 5000 bps.

The time plot of the index volatility of 2018 increases which confirms that as stock prices goes down, volatility goes up. Indeed, volatility or price range is a good indicator of risk which impacts the market psychology negatively and stock prices decreases.

**Figure 3: The CAC 40 Volatility in 2018**
The CAC 40 volatility increased tremendously in February as the fear of rise in US interest rates led to a spike in volatility which affected most markets. The index volatility attained 96.27 as the highest value the 8th of February after it opened at 30.45 and closed at 17.75. The volatility increase matches the CAC 40 decrease to 5100 bps in February. The stock price time plot shows that the price remained low until reversing the trend and increasing in April after the german sales became stable and the US monthly revenues for households increased [20]. Here again, the reverse trend in price led to a reverse trend in volatility as it started decreasing. The overall trend of the CAC 40 volatility in 2018 remains a growth trend whereas the CAC 40 stock prices in 2018 remains a decline trend.

2 Methodology

The methodology followed is stochastic modelling using simple random walk, a stochastic process. Random walk relies on the EMH or efficient market hypothesis as the stock price revolves around its true value in a random pattern. Data analysis of the stock prices of 2018 shows that the CAC 40 is an efficient market and follow the EMH as al available information is reflected on the price. The first step is to define stochastic modeling and its application in R and state the process advantages and disadvantages. The next step will be to define the variables to take into consideration and the external factors to consider. The model build will be optimized to match reality.

2.1 Stochastic Modeling

A mathematical model is the “quantitative description” of a natural event. The analysis of the model relies on a crucial parameter which is the usefulness. Stochastic is derived from the Greek and is related to any event happening by chance. ‘Stochastic’ model or random model is the antonym of ‘Deterministic’ model which means sure and certain. “A deterministic model predicts a single outcome from a given set of circumstances, whereas a stochastic model predicts a set of possible outcomes weighted by their likelihoods or probabilities”.

Events are not inherently stochastic or deterministic. The choice to model any event as stochastic or deterministic depends on the observer’s objective. The main criterion for deciding remains usefulness. Scientific modeling procedure relies on three main components:

(1) “a natural phenomenon under study,
(2) a logical system for deducing implications about the phenomenon,
(3) a connection linking the elements of the system under study to the logical system used to model it”.

13
There are relevant guidelines to follow when assigning probability values in a model and interpreting the results of the model. These guidelines are related to three broad principles which are often used to connect “the abstract elements of mathematical probability theory to a real or a natural phenomenon that is to be modeled”. These principles are the following:

1. “the principle of equally likely outcomes,
2. the principle of long run relative frequency,
3. the principle of odds making or subjective probabilities”.

All the relevant aspects of the event under study must be reflected in the stochastic model to be accurate. Moreover, the model must lead to calculation and allow the inference of crucial predictions about the event studied [5].

2.2 Stochastic Applications in R

Stochastic calculus applications in finance are endless. The model to be build will follow a stochastic process to predict the potential outcomes and thus increase profitability. Data analysis of the CAC 40 stock prices of 2018 showed that the date is normally distributed and do not fluctuate much and thus rely on the EMH or efficient market hypothesis. Forecasting can then be made easily on short-term period. The tool chosen to model the CAC 40 stock prices’ predictions for 2019 is R. R as a programming language is an efficient tool to represent stochastic processes. Packages in R are available to facilitate the analysis and lead to precise results when running the processes. These processes can be defined as sequences of random variables that is why the outcomes remain uncertain. The random variables in a stochastic process can take values comprised in a common space S which can be discrete or continuous. Therefore, the main parameters of a stochastic process are: “the index, state space, dependency relations among its random variables”. The analysis of stochastic processes starts with conditional probability which characterizes conditional distribution and conditional expectation. Conditional distribution is “the distribution of a random variable with either a probability mass function, if X is discrete, or a probability density function, if X is continuous” [21].

The joint probability mass function is a function of x and y, if x and y are discrete random variables, and is expressed as the following: $P(X = x, Y = y)$

The conditional PMF of Y given that X=x is expressed as:

$$P(Y = y | X = x) = \frac{P(X = x, Y = y)}{P(X = x)} \quad (2)$$

Where $P(X=x)>0$ and x is treated as fixed.
The joint density function is also a function of \( x \) and \( y \), if \( x \) and \( y \) are continuous random variables, and satisfies the following:

\[
P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s, t) \, dt \, ds \quad (3)
\]

The conditional DF of \( Y \) given that \( X=x \) is expressed as:

\[
f_Y|X(y|x) = \frac{f(x,y)}{f_X(x)}
\]

Where \( f_X \) is the marginal density function of \( x \) [20].

### 2.3 Random Walk in Stock Markets

Stock price behavior models are based on several theories. Market professionals tend to rely on two main approaches to build and develop models that predict stock prices. The two approaches are Chartist theory and Fundamental analysis.

#### 2.3.1 Chartist or Technical Theories

The first approach market professionals follow relies on history and thus past behavior in order to retrieve a pattern. This pattern will be the basis of the stock prices’ prediction. The approach assumes that history is reliable as it tends to “repeat itself”. Therefore, the sequence of price changes in securities is considered to be dependent [22].

#### 2.3.2 Fundamental or Intrinsic Value Analysis

The second approach market professionals follow assumes that securities have an intrinsic value or equilibrium price at any point in time. This equilibrium price is based on the security earning which depends on ‘fundamental factors’, such as “quality management, outlook for the industry and the economy, etc.” The actual price is then compared to the equilibrium price, above or below, and a prediction can be made about the future stock price [22].

#### 2.3.2 Random Walk

Statisticians and Mathematicians tend to rely on one main approach to build models that predict stock prices. This approach is *Random Walk*. As mentioned before, the random walk is a stochastic sequence of an independent random variable and is said to be simple when the variable takes either the value 1 or -1. Therefore, a random walk market is a market where the sequence of stock price changes in securities is independent. Random walk theory is also closely linked to EMH ‘efficient market hypothesis’. Random walk assumes that major securities’ transactions are efficient or in other words that all available information is reflected on the stock prices and that this information is accessible to all traders.
Overall, the theory of random walk in stock market prices presents challenges to both the technical and fundamental analysis. The technical analysis should show consistently that the pattern retrieved predicts the future prices accurately. Whereas, the fundamental analysis should be used only when the analyst has neither better insights nor new information as any new information will not be considered in current market prices on which depend the intrinsic value or equilibrium price estimation [22].

### 2.4 Simple Random Walk

Random walk predicts stock prices through several but equivalent methods. The simplest method of random walk theory remains “the drift less case with finite up or down steps at each time interval; at each decision point, the series either increases by one or decreases by one with equal probability”. The two other methods are the “one-dimensional random walk” with and without drift. Drift $\delta$ is a “nonrandom per-period change in the dependent variable” [2].

The simplest method or the simple random walk is defined as a stochastic process with $S_0 = cte$ as follows:

$$S_n = \sum_{k=1}^{n} X_k \quad (4)$$

- The variable $X$ is an independent distributed random variable. The current value of the variable is equal to “the past value plus an error term defined as white noise, a normal variable with zero mean and variance one” and it represented as: $X_t = X_{t-1} + \epsilon_t$.
- The change of $X$ is said to be random as the process predicts the variable outcomes based on the current value and does not take into consideration the change $(X_t - X_{t-1})$.
- The random walk is said to be simple when the variable takes either the value 1 or -1 with probabilities $p$ and $q = 1 - p$ as follows:

\[
\begin{align*}
    P(x = 1) &= p \\
    P(x = -1) &= q
\end{align*}
\]

- The mean of a random walk remains constant whereas the variance does not which makes the process nonstationary and variance increases with time [23].

The simple random walk model predicts different outcomes with a degree of error (figure 4). The model optimization would be to decrease this degree of error to match reality.
Several methods are used to analyze random walk, such as “difference equations which express the relationship between the elements of the sequence, the theory of Markov chains and branching processes or Martingales” [24].

Markov chains are “sequences of random variables” that take values in the state space, a finite space. These sequence have to satisfy the Markov property, when the predicted value $X_{n+1}$ depends on the past only through present value at time n. Markov chains have various applications, in stock predictions the possible outcomes are based only on present information.

If we suppose that in a finite set $S$ a probability space is given, a sequence of random variables is called ‘Markov chain’ when for all $n \in N$ and $s \in S$:

$$P(\xi_{n+1} = s | \xi_0, ..., \xi_n) = P(\xi_{n+1} = s | \xi_n) \quad (5)$$

The conditional probability of the predicted event $\xi_{n+1} = s$ “with respect to the random variable” is equal to the conditional probability of the predicted event “with respect to the $\sigma$-field generated by the random variables $\xi_0, ..., \xi_n$”. The $\sigma$-field regroups the variables and their complements and unions. The field can also contain an empty set.
Martingales is related to gambling or game of chance. Submartingale and supermartingale are related to the favourable or unfavourable chance winning the game. Martingales play an important role in mathematical finance, especially options. Indeed, some aspects of finance are connected to gambling. A sequence of random variables is defined as ‘Martingales’ with respect to a sequence of $\sigma$-field ordered on a increasing or decreasing pattern called filtrations if:

1. “$\xi_n$ is integrable for each $n=1,2,…$,”
2. $\xi_1, \xi_2, ...$ is adapted to $\mathcal{F}_1, \mathcal{F}_2, ...$,
3. $E(\xi_{n+1}|\mathcal{F}_n) =$” [25].

### 2.5 Advantages and Disadvantages

The simple random walk method remains widely used because it holds many advantages. The main one is its simplicity. The method can predict the stock market trends as it relies on the EMH. The simple process predicts the stock prices based on the current value and not the change as mentioned, successive stock price changes are believed to be independent and thus prediction is not based on past history.

However, the simple random walk method also holds some disadvantages. The random process revolves around its mean value, which is constant, as drift or volatility is not considered. The process is non-stationary and can affect the model accuracy. The computed outcomes fluctuate due to the statistical errors and can be referred to as ‘noisy’. These statistical errors can lead to asymmetric the normally symmetric outcomes erroneously. The model optimization requires to reduce the generated statistical errors; a possible solution would be to increase the number of variables [23].
Chapter 3: RESULTS AND INTERPRETATION

1 CAC 40 Predictions’ Model

The three required components of scientific modelling are fulfilled. The CAC 40 predictions model for 2019 aims to predict the market trends, the phenomenon under study, in order to capture profits. The logical system used is a stochastic process, simple random walk, which can lead to accurate conclusion about the phenomenon. Indeed, random walk theory relies on the EMH or efficient market hypothesis and presents challenges to both the technical and fundamental analysis that is why statisticians and mathematicians tend to rely on random walk to predict stock prices. Moreover, data analysis of the CAC 40 stock prices and volatility for 2018 shows that the data also follows the EMH and thus the CAC 40 is an efficient market. Simple random walk was chosen as a process due to its simplicity but the model may need to be optimized to reduce the generated errors. Finally, simple random walk is a stochastic sequence of a random variable $X_t$ which represents the stock price at time $t$ and thus the market trends on a defined period.

1.1 Simple Random Walk

There are guidelines to assign probability values in a scientific model as mentioned in the methodology. The guidelines are three principles that relate the elements of mathematical theory to the phenomenon studied. Simple Random Walk is a stochastic or random sequence of independent distributed random variable that starts from $S_0 = cte$. The variable $X_t$ or the stock price at time $t$ goes up or down and thus takes either the value 1 or -1 with equal probability $P(x = 1) = 1/2$, $P(x = -1) = 1/2$ when symmetric. This fulfills the first principle of “equally likely outcomes”. Simple random walk process will aim to predict the stock market outcomes which goes on indefinitely and fulfills the second principle of “long run relative frequency interpretation”. Finally, simple random walk will predict several outcomes which fulfills the third principle of “odds making or subjective probability” as different probabilities can be assigned to the same event [5].

The probability mass function for each step related to the random variable $X$ is:

$$P(X = 1) = p, \quad P(X = -1) = 1 - p = q$$  \hspace{1cm} (6)

The density function for the $n$ first steps related to the random variable $X$ is:

$$P(X_n = k) = \binom{n}{n+k/2} p^{(n+k)/2} (1 - p)^{(n-k)/2}$$  \hspace{1cm} (7)
Figure 5: Prediction of CAC 40 using Simple Random Walk – 10 Trials

Figure 6: Prediction of CAC 40 using Simple Random Walk – 50 Trials [25]
Duration is equal to the number of trading days, there are 252 trading days in 2019. The CAC 40 opened with a value of 4642.18 bps on the 2\textsuperscript{nd} of January 2019 according to Yahoo finance. The simple random walk model have been iterated 10 times and 50 times. The second plot shows clearly that random walk revolves around a constant mean as the variance increases with time and is thus non-stationary as mentioned. Therefore, generated statistical errors can be made. The possible outcomes or predictions go up to 4680 bps which does not much reality or the first quarter of 2019, as shown in the graph below:

\[
X_t = X_{t-1} + \epsilon_t.
\]

The error term or white noise which generates statistical errors can be linked to expected return and volatility. The next simple random walk model will be based on normal distribution assumption for simplicity. The error term will be equated to the past price times $z$ normal distribution score to include expected return and volatility.

### 1.2 Normal Distribution Assumption

Data analysis of the CAC 40 stock prices for 2018 were found to have a symmetric normal distribution. The simple random walk model based on normal distribution will assume that the expected annual return and volatility are known and constant and that returns are normally
distributed [27]. These assumptions are not realistic as the financial market is unpredictable and the returns are non-stationary data. Nevertheless, the annual expected return and volatility can be estimated by considering the external factors and analyzing past expected return and volatility. The model remains an interesting starting point to investigate the impact of expected return and volatility on statistical errors. Therefore, the model is worth future development to match realistic assumptions. The simple random walk based on normal distribution is as follows:

$$X_t = X_{t-1} + X_{t-1}z \quad (8)$$

Where:

$$z = r_{daily} + \sigma_{daily}i$$

The estimates of the annual expected return and volatility can be made by analyzing the external factors. CAC 40 have been through the lowest decrease over the last 7 years in 2018, down to -10.95 % of annual return. The new year (2019) began with a correction as expected thanks to the more stable political environment after the Brexit and a boost to the global economic growth. The CAC 40 is increasing gradually, up to 15 % currently. Nevertheless, the growth perspective should be reevaluated and downgraded due to:

- Low macroeconomic values for France
- Commercial war which affects the exporting companies as China’s growth is not expected to rise as predicted.

Overall, the economic context provides expected benefits of 4-5% instead of 8.5% for the MSCI Europe, so the estimate for 2019 of expected annual return would be 4%.

The expected volatility is currently 5% but is expecting to increase due to the decreasing economic context. The yearly volatility of 2018 was 20% so the estimate for 2019 would be 12.5% [28].

The implementation for 2019 gave a plot that matches the actual stock prices for this year first four months, up to 5400 bps for the first trial and 5500 bps for the second. This proves that including expected annual return and volatility can reduce generated statistical errors and thus match reality. The application in R gave the following figures, the outputs may differ due to the random normal distribution:
Figure 8: CAC 40 Prediction using Normal Distribution Assumption - Trial 1

Figure 9: CAC 40 Prediction using Normal Distribution Assumption - Trial 2
1.3 Data Analysis using Difference Equation

The simple random walk model based on normal distribution assumption proved that expected return and volatility should be included to reduce statistical errors. The model is based on nonrealistic assumptions and thus should be adapted to reality. The first step in doing so is analyzing simple random walk using one of the methods previously stated. The model can be analyzed using the martingales derived or by analyzing the Markov chains derived. Difference equations will be used for simplicity. They show the relationship between successive values of the sequence by involving the difference between these values. The difference is the daily return of the CAC 40 and the data analyzed is the stock prices of 2018. Daily returns were computed based using the following formula: Daily return = (Closing Price of Today – Closing Price of Yesterday)/ Closing Price of Yesterday.

![Daily return 2018](image)

**Figure 10:** Daily Returns of CAC 40 in 2018

The daily return plot of 2018 was analyzed and found to match the pattern of arithmetic random walk with drift, nonrandom change in the dependent variable, after differentiation. The plot of the daily returns clearly shows that the prices keep going back to the mean value as the returns are proportional to the stock prices. This is a response to the EMH or efficient market hypothesis. Prices are either under or overvalued in random patterns and they eventually go back to their true values so the returns’ mean value remains constant. The results also confirm that volatility impacts the stock market as drift which can be derived from implied volatility affects the relationship between successive values of the simple random walk sequence.
1.3.1 Arithmetic and Geometric Random Walk

The main difference between arithmetic and geometric random walks is the way the random values are modeled. The Arithmetic process is modeled as the sum of random values and the geometric process is modeled as products of random values.

Arithmetic random walk with drift assumes that the stock prices can move up or down with a given probability and that the stock price movement follows a normal distribution $N(\mu, \sigma^2)$. The variable at time $t$ is equal to the variable at time $t-1$ plus the drift term $\delta$ and the error term or ‘white noise’ $\epsilon_t$ with mean value equal 0 and standard deviation $\sigma$ as follows [30]:

$$X_t = X_{t-1} + \delta + \epsilon_t \quad (9)$$

Geometric random walk is the logarithm transformations of data. Analogously to the simple random walk with steps or increments $\delta = X_t - X_{t-1}$, the relative increments for geometric random walk is $R_t = X_t/X_{t-1}$ which are assumed independent and identically distributed [31].

Brownian motion is a stochastic process considered as the limit of random walk. The arithmetic process links “discrete time random walks with their corresponding continuous time limits” as it starts from zero and the increments are independent and stationary on a continuous path. Whereas, geometric Brownian motion is considered as the stochastic equivalent of the exponential growth function and is an important application of stochastic calculus to finance.
Arithmetic Brownian motion is not as appropriate to predict stock prices as it does not take into consideration the negative values which can occur in the stock market. Moreover, a process able to predict the market trends should allow “the expected return to be proportional to the stock price” [32].

1.3.2 EMH Assumptions

A random walk process is defined as a stochastic sequence of the independent distributed random variable X where the error term $\epsilon_t$ is defined as white or IID independent and identically distributed noise as follows:

$$\epsilon_t = X_t - X_{t-1} \quad (10)$$

White noise is defined widely as the time derivative of a Brownian motion. It can be defined as a stochastic process independent at unalike times and identically distributed or IID with a mean value of zero and an infinite variance with time $t$. Differencing a random walk leads to a stationary process. Therefore, the daily returns can be considered as a stationary data and thus more accurate to model [33].

Based on the EMH, statistical models commonly used to estimate financial returns consider $r_t = \mu + \epsilon_t$ where $\mu = E * r_t$ is the expected return and considered constant. We assume $var(\epsilon_t) = var(r_t) = \sigma^2$ is a positive and finite constant. “There are three different assumptions about the term $\epsilon_t$ from the weakest to the strongest:

- White noise: The term $\epsilon_t$ is defined as white noise and denoted as $\epsilon_t \sim N (\mu, \sigma^2)$, under this assumption, for all $t \neq s$ the correlation between $(\epsilon_t, \epsilon_s)$ is 0.

- Martingale difference: The term $\epsilon_t$ is defined as a martingale difference sequence as for any $t$ $E(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, ...) = E(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, ...)$. 

- IID or independent and identically distributed: The term $\epsilon_t$ is defined as IDD and denoted as $\epsilon_t \sim IDD (\mu, \sigma^2)$. It implies that the term is a martingale difference, and if it satisfies the defined sequence, for any $t > s$, $cov(\epsilon_t, \epsilon_s) = E(\epsilon_t \epsilon_s) = 0$. In other words, the term $\epsilon_t$ is a white noise series”.

The relationship between the three assumptions is that white noise is included in martingale differences which are both included in IID. IID is the strongest assumption but can be hard to apply in reality so martingale difference is widely used to predict return pattern under the EMH model. Moreover, white noise was also found to be a crucial variable to regress in addition to other variables impacting the market to predict the stock prices accurately [34].
1.4 Model Implementation

The analysis of the data using difference equation demonstrated that the daily returns were following an arithmetic random walk model with drift which can be derived from volatility. Moreover, the returns revolved around a constant mean which is a response to the EMH. The analysis confirmed the results of the simple random walk model based on normal distribution assumption model, expected return and volatility should be considered to reduce the generated statistical errors. Based on the EMH model, statisticians follow three main assumptions for the error term $\varepsilon_t$ in the model $r_t = \mu + \varepsilon_t$ to estimate financial returns. IDD or independent and identically distributed assumption is the strongest. “The log returns are assumed to be IDD normal, then the process is the exponential of a random walk” or geometric random walk [35]. Geometric random walks can be approximated as geometric Brownian motion which is a crucial application to stochastic calculus in finance as it considers the negative values and the returns to be proportional to stock prices.

The model final optimization will be to retrieve a trend based on the mean and standard deviation of the log returns of 2018 stock prices of CAC 40. The mean or expected return was found to be -0.0005 and the standard deviation 0.008688. The external factors should be considered in addition to historical data. The mean should be inverted and become positive 0.0005.

The recommendations that can implied from the model is to base investments on past volatility, expected return and white noise combined with the external factors. The CAC 40 correction in 2019 was proven to match reality for the first four months and the analysis of the external factors predict a growing expected return with growing volatility due to the low macroeconomic variables in France and the commercial war between China and the US. Moreover, stock prices have a martingale behavior, therefore it is crucial to retrieve the trend or log returns to back up the investment.
Figure 12: CAC 40 Predictions Model for 2019-Trial 1

Figure 13: CAC 40 Predictions Model for 2019-Trial 2 [36]
2 Application of the Model using IQ Option

2.1 Binary Option
Binary Option is advised for beginners in trading as it is based only on two selections, up or down. The selection is based on the prediction of the investor, rise or decline in price and is set with a specific deadline. Therefore, the gain or loss incurred is automatically added or subtracted from the investor’s account. A wide range of “currency pairs, commodities, corporate stocks, government bonds, and equity indices” is available for trading. Nevertheless, binary options do not allow traders to invest in the underlying security [37].

2.2 Random Walk Index
Technical indicators are used by investors to facilitate their predictions. The ‘Random Walk Index’ is one of these technical indicators. The index associates random movements to the stock price movements to determine its statistical trend and evaluate its significance. The index is computed for up and down trends as follows:

\[ RWIH = \frac{H - L(r)}{\text{ATR} \times \sqrt{n}} \] (11)

\[ RWIL = \frac{H(r) - L}{\text{ATR} \times \sqrt{n}} \] (12)

Where \( H \) and \( H(r) \) are the highest values of the current day and \( r \) days before, \( L \) and \( L(r) \) are the lowest values of the current day and \( r \) days before, and ATR is the average true range is a technical analysis and a volatility indicator. The average provides an indicator of the degree of price volatility.

The random walk index is used as indicator of stock price movements. An investor would take a long position or up selection when the long-term RWIH is greater than 1.0 and the short-term RWIL peaks above 1.0. Whereas, an investor would take a short position or down selection when the long-term RWIL is greater than 1.0 and the short-term RWIH peaks above 1.0 [38].

2.3 Investments’ Scenarios
The application of the model and random walk index are made using the simulator IQ Option. The two selections are investigated, the up and down selection. The CAC 40 predictions’ model shows a growth trend in April overall. Nevertheless, the trials of the model predictions show a decline trend towards the mid-April. The random walk index should be used to back up the index predictions’ model.

The first scenario is performed in mid-April; therefore, the selection made is down. The random
walk index confirms the prediction as the long-term RWIL is increasing up to 1.0 and the short-term RWIH is decreasing and peaks above 1.0.

The second scenario is towards the 20th of April; therefore, the selection made is up. The random walk index confirms the prediction as the long-term RWIH is increasing up to 1.0 and the short-term RWIL is decreasing and peaks above 1.0.

**Figure 14:** Investment Scenario 1 – Down Selection

**Figure 15:** Investment Scenario 2 – Up Selection
Chapter 4: FUTURE WORK AND CONCLUSION

1 Future Work
Mathematical finance has transformed the world of finance in the previous forty years thanks to the development of accurate computational methods and more precise mathematical models. These tools are considered the basis of computational finance which aims to calculate efficiently risk that any financial instrument produces. This field is considered interdisciplinary as it involves “financial mathematics, stochastics, statistics, numerics and scientific computing”. The mathematical tools have been developed and led to the extension of stochastic numerics to stochastic differential equations and Monte Carlo simulation to a multi-level method. Random walk applications to barriers in option-pricing has also been developed and simplified as the knowledge required to estimate probabilities is not primordial anymore [39].

Random walk remains an effective stochastic process to predict the stock prices and estimate the barriers in option-pricing. Moreover, random walk can be used effectively as a “network-based method for information mining”. Nevertheless, the process computational complexity may limit its efficiency on “larger scale datasets”. GPU or graphics processing units can be used to accelerate random walk computation on more heterogeneous networks [40]. Indeed, the report presents how random walk model was applied to predict the CAC 40 stock prices for 2019, which remains a small dataset. Therefore, computational complexity did not affect the random walk model efficiency. Future work and improvement would imply testing the model on larger scale datasets and test GPU efficiency in accelerating the random walk computation.

2 Conclusion
As addressed in my abstract, this report aims to analyze the CAC 40 and model it to predict the stock prices for 2019 and thus capture profits. The data retrieved is the stock prices and volatility of the index for a one-year period, 2018. The findings of the descriptive statistics using Excel confirm that the CAC 40 in an efficient market as it fulfills the characteristics of the EMH or efficient market hypothesis as all available information is reflected on the stock prices. The overall results show a decline pattern in 2018 especially towards the end and that as stock prices goes down, volatility goes up. Indeed, volatility or price range is a good indicator of risk which impacts the market psychology negatively and stock prices decreases.

The methodology followed is stochastic modelling using simple random walk, a stochastic process. Random walk relies on the EMH as the stock price revolves around its true value in a random pattern. Statisticians and mathematicians believe that random walk is the most accurate
way to estimate the possible outcomes as it presents challenges to both the technical and fundamental analysis. The random walk is said to be simple when the variable takes either the value 1 or -1 with probabilities p and q. The simple random walk implementation shows that the process revolves around a cte mean as the variance increases with time and is thus non-stationary. Therefore, the fact that the possible outcomes do not match reality can be explained by the generated statistical errors. The simple random walk model should be optimized to reduce statistical errors.

The first model optimization was to consider normal distribution assumption and relate the error term or white noise which generates statistical errors to expected return and volatility. The model assumes that the expected annual return and volatility are known and constant and that returns are normally distributed which is not realistic. The model remains crucial to investigate the impact of expected return and volatility on statistical errors. The simple random walk model based on normal distribution assumption proved that expected return and volatility should be included to reduce statistical errors.

The first step in adapting the non-realistic assumptions of the model is analyzing simple random walk using difference equations which show the relationship between successive values of the sequence by involving the difference between these values. The analysis of the daily returns of the stock prices of 2018 using difference equation demonstrated that the daily returns were following an arithmetic random walk model with drift which can be derived from volatility. Moreover, the returns revolved around a constant mean which is a response to the EMH. The analysis confirmed the results of the simple random walk model based on normal distribution assumption model, expected return and volatility should be considered to reduce the generated statistical errors.

Based on the EMH model, statisticians follow three main assumptions for the error term $\varepsilon_t$ to estimate financial returns. IDD is considered as the strongest assumption and are defined as log returns then the process is the exponential or geometric random walk. The model final optimization main goal was to retrieve a trend based on the mean and standard deviation of the log returns of 2018 stock prices of CAC 40. The mean or expected return was found to be 0.0005 and the standard deviation 0.008688 after considering the external effects in addition to historical data. The application of the model and random walk index is made using the simulator IQ Option. The two selections are investigated, the up and down selection. The random walk index should be used to back up the index predictions’ model.
References


APPENDICES

Appendix A: Simple Random Walk code on R

```r
> t <- 252
> n <- 10
> initial.value <- 4642.18
> GetRandomWalk <- function()
+ { initial.value + rnorm(t, MEAN(0, SDNRM(1)) ) }
> values <- replicate(n, GetRandomWalk())
> dev.new(height=8, width=12)
> plot(0:252, c(rnorm(10, t + 1)), main=sprintf(\"%s Random Walks\", n),
+ xlab=\"Time\", ylab=\"Price\",
+ ylim=4642.18 - 6.6 * c(-1, 1) + sqrt(t) )
+ mtext(sprintf(\"Across time \(0, 1, \ldots, n\)\",
+ \"Initial value \(\mu\)\",
+ \"Across time \(0, 1, \ldots, n\)\",
+ \"Initial value \(\mu\)\",
+ \"Across time \(0, 1, \ldots, n\)\")
+ for (i in 1:n) {
+ lines(0:252, values[,i], lwd=0.25)
+ }
+ }
```

Appendix B: Normal Distribution Model code on R

```r
> Z <- rnorm(252, 0, 1)
> mu <- 0.06
> sigma <- 0.125
> s <- 4642.18
> stockprice <- c(s)
> a <- 2
> t <- 1:253
> for (i in Z) {
+ S = s + s*mu/252 + sigma*sqrt (252)*z[i]
+ stockprice[i] <- s
+ a = a - 1
+ }
+ plot(t, stockprice, main="CAC 40 Predictions", xlab="2019", ylab="price", type="s")
+ ```
Appendix C: Arithmetic Model code on R

```r
> mu= 0.0006
> sigma=0.005669
> simulation<-exp(rnorm(252))
> stockprice=rep(NA,252)
> stockprice[1]<- 4442.18
> #start simulating prices
> for(i in 1:length(simulation))
> stockprice[i]=stockprice[i-1]*exp(rnorm(1,mu,sigma))
> 
> Price <- stockprice
> plot(Price, xlab="2019", col= "black")
> title(main="CAC 40 Predictions", col.main="black", font.main=1)
```

Appendix D: Log Returns Model code on R

```r
R version 3.6.3 (2019-03-11) -- "Great Truth"
Copyright (C) 2019 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)
R is a programming language and software environment for statistical computing and graphics. It is a GNU project
supported by the Free Software Foundation. R is partly written in C and Fortran.
Base R provides core statistical and graphical functionality which forms the heart of the system.
R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, etc.)
data visualization (plotting of both vectors and matrices), and a variety of uses for numerical linear algebra.
```

```r
> rwdrift <- arima.sim(model = list(order = c(0, 1, 0)), n = 252, mean = 1)
> px_drift <- ts(rwdrift, start = 3, frequency=1)
> px_drift diff <- diff(px_drift)
> plot(px_drift, diff, plot.type="single", xlab="2019", ylab="Return")
```
Appendix E: Initial Specifications

BENTALEB Sabrine
EMS
STOCHASTIC ANALYSIS OF A EUROPEAN MARKET USING R
LAAYOUNI L
Spring 2019

My capstone project will consist of analysing a European market (CAC 40) and model it following a stochastic process using R. The aim of my project will be to predict trends of the market and thus to capture profits.

The analysis of any market requires a detailed study of the factors affecting it, whether they are internal or external. The first step will be to define these factors and study how they affect the market trends on a specified period. The factors with less or short impact can be omitted. The second step will be to pick the random variables for specific financial products and estimate the possible outcomes within a forecast to overall predict the different situations. The information gathered will then be modelled following a stochastic process using R. The last step will be to compare the results of the model with reality.

The proposed timetable for my capstone will be literature review and research for next week. The following month will be dedicated to the capstone design from analysis of the market trends to the random variables to pick and estimation of the possible outcomes. The final two months will be dedicated to the model using R and implementation. The main objectives remain to match the results of the model with reality to overall predict the market trends.

The societal and ethical implication of this project remain to provide an effective model of the market trends, make sure to follow the code of ethics and cite any source of work.