uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Masoud Mohammadi Kamel
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Systems Science)
GRADE / DEGREE

Systems Science
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

An Information Theoretic Market Index

TITRE DE LA THÈSE / TITLE OF THESIS

C. Charalambous
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Nasiruddin Ahmed

Claude D'Amours

Gary W. Slater
LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCORAL STUDIES

An Information Theoretic Market Index

Masoud M. Kamel

Systems Science
University of Ottawa
Ottawa, Ontario, Canada

A Thesis submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of
Master of Science

Your file Votre référence
ISBN: 0-494-11356-1
Our file Notre référence
ISBN: 0-494-11356-1

Contents

List of Figures

Abstract

Nowadays, common indices act mainly as indicators of stock market performance and development. Since almost all of these indices are somehow a weighted average, thus, they are not able to answer the main question of investors, which is the likelihood that the status of a market changes or remains the same in next sessions. Therefore, by applying the principles of information theory we will try to suggest a new defined index as a new approach to stock market indicators.

In this thesis, we suggest two different ways of market analysis; first, we consider a single stock option and second, we consider a family of stock options. In the first case, we mainly consider three type of information measures; entropy, entropy rate and relative entropy. In the second case, we consider two information measures; entropy rate and joint entropy.

We also suggest several stochastic models as an estimate for stock price such as the Geometric Brownian motion model, with constant drift and volatility or with drift as a function of time, and Markov chain models. In addition, the analysis is carried by using models for the option price and the "rate of change" of the option price.

Acknowledgments

I would like to express my gratitude to all those who gave me the possibility to complete this thesis. I am deeply indebted to my supervisor Dr. Charalambous whose help, stimulating suggestions and encouragement helped me in all the time of research for and writing of this thesis.

I would like to thank Dr. Ahmed whose advice and comments as a committee of defense helped me in completing this work. I would also like to thank the other member of defense Dr. D'Amours for the comments he made to the thesis.

I thank my fellow graduate students including Farzad Rezaei and Mohammad Pour and members of the group at the University of Ottawa for the many useful discussions that helped me with the research.

Last but not the least, I would like to thank Dr. Thizy and Dr. Yeap and appreciate their support and concern in guiding me as the former and present director of program.

Summary and Conclusions

In this thesis, we suggest two different ways of market analysis. The first is based on a single stock option and the second is based on a family of stock options. In the first case, we mainly consider three types of information measures; entropy, entropy rate and relative entropy. In the second case, we consider two information measures; entropy rate and joint entropy.

In the single option case, when the entropy of a single stock at two consecutive periods is the same, we conclude that the probability of all prices remains constant during the two periods. When the entropy increases it means uncertainty increases. Therefore, we may conclude that if the entropy graph of a specific stock shows an increase, then the stock is more volatile. Finally, when the entropy decreases, we conclude that the uncertainty decreases.
In the multiple stock options case, when the entropy remains unchanged, similar to the single option section, the probability of items included in calculating the entropy remains unchanged. Therefore, there are two directions for a portfolio in such a case: firstly, when the rate of change of each individual remains the same and secondly, when the rate of change of all items changes with the same ratio. When the entropy increases the probability of each sequence item becomes closer to $\frac{1}{n}$, where n is the number of options included in a portfolio. In this case, the "rate of change" of different stocks are similar. When the entropy shows a decrease, we can conclude that there are a few stock prices which had noticeable changes, compared to the previous period of study.
We also employ entropy rate and relative entropy in the analysis. For the relative entropy, we may apply the distance between a given stock price and the corresponding state which grows according to the interest rates. The greater distance between the two processes, measured by the relative entropy, the worst the performance of the stock.

Chapter 1

Introduction

The birthplace of information theory was statistical physics, when the concept of entropy was introduced by the second law of thermodynamics. The entropy of a source of information is defined as the weighted average of the self information of each message item of that source, where the amount of uncertainty of a particular message is called self information. Later, information theory was applied in other fields of science such as economics and financial mathematics.

The application of information theory in the stock market was introduced by log-optimal portfolios to optimize the return of assets which were invested in a basket of stock options[1].

As well, common indices act mainly as indicators of stock market performance and development. If there is an increase in an index value, it means the value of the market has increased. On the contrary, if an index value goes down, the value of the market has decreased. In another words, we apply such an index as a tool to compare the value of a stock market during different sessions.

One of the most important factors that investors like to know is the likelihood that the status of a market will not change in the next session, for example, an investor needs to know that when the value of a market index is increasing, it will continue to increase. Obviously, the existing indices don't give us such information. Therefore, by applying the principles of information theory, the goal for this thesis is to define a new index to get more information.

Hence, as the first step, in this chapter we will introduce mathematical tools and economical backgrounds to employ them in subsequent chapters.

1.1 Mathematical Background

1.1.1 Differential Entropy

Let $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a Random Variable (R.V.) which is a measurable function on the probability space $(\Omega, \mathcal{F}, P)$. The probability measure P_X induced by X on $(\mathbb{R}, B(\mathbb{R}))$ is defined by

$$P_X(A) \triangleq P(\{\omega; X(\omega) \in A\}) = P(X \in A) \quad A \in \mathcal{B}(\mathbb{R}),$$

The function $F_X(.)$, called the probability distribution of X, is defined by

$$F_X(x) \triangleq P(\{\omega; X(\omega) \leq x\}).$$

If we assume that $F_X(x)$ be absolutely continuous with respect to Lebesgue measure, then $f(x) \triangleq \frac{d}{dx} F_X(x)$ exists and is called the probability density function of random variable X. We define the expectation of X by

$$E(X) \triangleq \int_{\mathbb{R}} f(x) dP_X(x)$$

The differential entropy of X is defined by

$$h(X) = E\big[\log \frac{1}{f(x)}\big] = -\int_{\mathbb{R}} f(x) \log f(x) dx \tag{1.1.1}$$

If a discrete R.V. taking value in a countable set $S_X = \{x_1, x_2, \cdots, x_n, \cdots\}$, then

$$\sum_{x_i \in S_X} P(\{\omega; X(\omega) = x_i\}) = \sum_{x_i \in S_X} P_X(x_i)$$

The distribution of the discrete R.V. is

$$F_X(A) = \sum_{x_i \in A \cap S_X} P(\{\omega; X(\omega) = x_i\})$$

Therefore, the entropy of the discrete R.V. is defined by

$$H(X) = -\sum_{x_i \in S_X} P_X(x_i) \log P_X(x_i) \tag{1.1.2}$$

Let P be a probability measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. If $f_x(x), x \in \mathbb{R}^n$ is the joint x-probability density, then

$$P_X(A) = \int_A p_X(x_1, x_2, \cdots, x_n) dx_1 \cdots dx_n \quad A \in \mathcal{B}(\mathbb{R}^n)$$

Then joint probability distribution is defined as

$$F_X(x_1, x_2, \cdots, x_n) = P(X_1(\omega) \leq x_1, \cdots, X_n(\omega) \leq x_n)$$

This can be written as

$$F_X(x_1, x_2, \cdots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} p(x_1, x_2, \cdots, x_n) dx_1 \cdots dx_n$$

Thus

$$p_X(x_1, x_2, \cdots, x_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F_X(x_1, x_2, \cdots, x_n)$$

1.1.2 Joint-Entropy

Let $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^n, B(\mathbb{R}^n))$ and $Y : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^n, B(\mathbb{R}^n))$ be R.V.s defined on probability space $(\Omega, \mathcal{F}, P)$.

And let P_X , P_Y be the probability mass function, if X, Y are discrete R.V.s.

Mutual Information:

Consider X , Y taking values in S_X, S_Y with probability densities $P_X(x_1, \cdots, x_n)$, $P_Y(y_1, \cdots, y_n)$ respectively, and $P_{X,Y}(x_i, y_i)$ is the joint probability density, the mutual Information $I(X;Y)$ is defined as following:

$$\begin{aligned} I(X;Y) &\triangleq \sum_{S_X} \sum_{S_Y} P_{X,Y}(x_1, \cdots, x_n, y_1, \cdots, y_n) \log \frac{P_{X,Y}(x_1, \cdots, x_n, y_1, \cdots, y_n)}{P_X(x_1, \cdots, x_n) P(y_1, \cdots, y_n)} \\ &= E_{P_{X,Y}(X,Y)} \log \frac{P_{X,Y}(X,Y)}{P_X(X) P_Y(Y)} \end{aligned}$$

Joint-Entropy :

Consider X , Y taking value in S_X, S_Y respectively, with joint probability $P_{X,Y}(x_1, \cdots, x_n, y_1, \cdots, y_n)$ joint entropy is defined as:

$$H(X,Y) = -\sum_{S_X}\sum_{S_Y} P_{X,Y}(x_1, \cdots, x_n, y_1, \cdots, y_n) \log P_{X,Y}(x_1, \cdots, x_n, y_1, \cdots, y_n)$$

Upper and Lower Bound of the Entropy

It can be shown that for a discrete R.V. X when $X \in \{x_1, x_2, \cdots, x_n\}$ there is a lower and upper bound defined as follows[2].

$$0 \leq H(X) \leq \log(n)$$

Entropy of a continuous R.V. can be negative, however, a R.V. X with finite variance is upper bounded[2].

$$H(X) \leq \frac{1}{2}\log((2\pi e)\sigma^2)$$

1.2 Economical Background

In this section we will introduce some definitions which will be used in subsequent chapters. Subsequently, we will discuss market indices. Market indices play a great role in financial analysis; our objective is to present some indices comparative to those that are common like DOW Jones.

1.2.1 Economic Terminology

After-Hours Trading: is the trading done after the regular market close(4:00 PM).

Blue Chip: is a company on the London Stock Exchange with large market capitalization, stable earnings, a consistent dividend record and a reputation as a reliable investment.

In general it also refers to those companies with large market capitalization, stable earnings, a consistent dividend record and a reputation as a reliable investment.

Commodity: is a real asset, such as gold or oil.

Complete Market: is a market in which every claim can be hedged.

Divisor: is a number that the index value is divided by, to get the weighted average. In some indices a divisor is also calculated.

Market: is a place for the exchanging of price information. Commonly situated in electronic space.

Portfolio: is a collection of security holdings.

Pre-Market Trading: is the trading done before the regular market opens(9:30 AM).

Security: is a piece of paper representing a promise.

Share: is a stock or equity.

Split: is the dividing a stock price into a fraction of it and realizing new number of share with the same fraction.

Spin off: is the distribution of a stock to its stockholders especially of stock of another company; usually a new company created by such a distribution.

Stock: is a security representing partial ownership of a company.

Stock Market: is a place for trading stocks.

Trade Volume: is the number of traded shares in a transaction or in a period of time.

Transaction Cost: is the charge for buying or selling a security.

1.2.2 Market Indices

Nowadays, the most common tools that investors use to analyze the atmosphere of the stock market are market indices. Since the stock market was established, scientists and investors have tried to find tools by which they were able to analyze the economical situation of the market and also to be able to forecast the market. Almost all market indices that are presented by different stock markets are somehow a weighted average of the total market value of their outstanding shares. The methodologies for calculating an index by most investment companies are almost the same. To show that stock market indices are weighted average, we will try to elaborate on the two major indices , the DOW Jones average and the NASDAQ index. They are important because the first is one of the oldest in North America and covers the most valuable companies and the second is the largest U.S. electronic stock market, with approximately 3,300 companies.

Our first concern is the DOW Jones Industrial Average (DJIA) which is the oldest index in the US and was published in 1884 by DOW Jones and Company Inc. The DJIA is an index of 30 "blue-chip" US stocks[13]. It's called an average because it was originally computed by adding up stock prices and then dividing the sum by a divisor which is calculated as follows.

$$D_{t+1} = D_t \frac{\Sigma C_t^a}{\Sigma C_t} \tag{1.2.3}$$

where

D_{t+1} is the divisor to be effective on trading session $t+1$.

D_t is divisor on trading session t.

C_t^a is the index components' adjusted closing prices for stock dividends; splits, spin offs and other applicable corporate actions on trading session t.

C_t is the index components' closing prices on trading session t.

Example 1.2.1. *To understand the above formulation we give an example.*
Consider 3 companies C1,C2 and C3 with the closing price \$5, \$10 and \$15 at time period t respectively. According to the DOW Jones formulation the average is calculated as $\frac{5+10+15}{3} = 10$. *Now assume that at the end of the day company C3 decides to splits its shares three-for-one, so by tomorrow every share of C3 will be \$5. If we calculate the average the answer will be \$6.67, but we know that nothing has been changed regarding the assets. To avoid such a distortion the devisor changes as follows. Knowing that* $D_t = 3$, *we can calculate the new divisor for* $t + 1$ *session.*
$\sum C_t^a = 5 + 10 + 5 = 20$ *which the summation of every item's price after the split*
$\sum C_t = 5 + 10 + 15 = 30,$ *and*
$D_{t+1} = 3 * \frac{20}{30} = 2.$
So the divisor for the next session is $D_{t+1} = 2$.

The second index we will consider is the NASDAQ. Unlike the Dow Jones, which provides only one index, the DJIA, the NASDAQ provides few indices such as the NASDAQ100, the NASDAQ Europe Composition (NECI) and so on. All of them have the same methodology, which we will try to explain below.
The NASDAQ indices act as indicators of the market's performance and development and mathematically they can be expressed as follows[14].

$$I(t) = I(0)\frac{\Sigma_i P(i,t)N(i,t)}{D(t)} \tag{1.2.4}$$

Where $D(t)$ is the devisor at time t which is calculated as:

$$D(t) = \left(\Sigma_i P(i,0)N(i,0) + \Sigma_{j=1}^{t} M(j-1)\right)\frac{I(0)}{I(j-1)} \tag{1.2.5}$$

Here $I(t)$ is value of the index at time t;

$I(0)$ is base index which is the value of index at base date, (i.e. 1000);

$P(i,t)$ is price of security i at time t;

$N(i,t)$ is the number of shares in issue for security i at time t;

$M(j)$ is new net money raised at time j through the issue of new companies, new shares, a rights issue, a capital reorganization or even a capital repayments, which may be negative. In the above mentioned formula the summation over i includes all companies in the index at the specific time, and the summation over j refers to all sessions (i.e. days), starting from the base date till the current time;

Consequently, $\Sigma_i P(i,t)N(i,t)$ is the total market value at time t and;

$\Sigma_i P(i,0)N(i,0) + \Sigma_{j=1}^{t} M(j-1)$ is the total value of the market after any adjustment.

Below is an example of how to calculate the value of the index and how to revise the divisor.

Example 1.2.2. *Consider three companies A,B and C with option price \$4,\$20 and \$30 with outstanding shares 25,35 and 20 million, respectively. Assuming an index starting at* $I(0) = 1000$, *we calculate the divisor at the beginning of the day as following:*

Since there is no adjustment with the base period market value, therefore, $\Sigma_{j=1}^{t} M(j-1) = 0$. *So by applying 1.2.5 we have*

$D(0) = Base\ Index \times \frac{Total\ Market\ Value\ After\ Adjustment}{Index\ Value} =$

$1000 \times \frac{4\times25{,}000{,}000+20\times35{,}000{,}000+30\times20{,}000{,}000}{1000} = 1000 \times \frac{1{,}400{,}000{,}000}{1000} = 1,400,000,000$

Now suppose that at the end of first session the price of companies A,B and C change to \$5 , \$21 and \$29 respectively, then the index value calculated by applying 1.2.4 is

$I(1) = Base\ Index \times \frac{Total\ Market\ Value}{Index\ devisor} =$

$1000 \times \frac{4\times25{,}000{,}000+20\times35{,}000{,}000+30\times20{,}000{,}000}{1{,}400{,}000{,}000} = 1000 \times \frac{1{,}440{,}000{,}000}{1{,}400{,}000{,}000} = 1028.57$

In the next step assume that company A issued 10 million new shares according to its fiscal year profit. Therefore, the outstanding shares for company A changes to 35 million. So we can calculate $\Sigma_{j=1}^{t} M(j-1) = 5 \times 10,000,000 = 50,000,000$. *Now the new devisor is calculated as*

$D(1) = 1000 \times \frac{1{,}440{,}000{,}000+50{,}000{,}000}{1028.57} = 1000 \times \frac{1{,}490{,}000{,}000}{1028.57} = 1,448,613,123.1$

and for the next session we have to use the new divisor consequently.

1.3 Literature Review

A large body of literature of financial economics and statistics is devoted to modeling of security of price behavior. The birthplace of financial mathematics is the work that Louise Bachelier introduced in his thesis in 1900. He had derived the price of an option where the share price movement is modeled by weiner stochastic process[1].

Most of existing models consider the value[2] of stock as the main variable to be modeled. It is obvious that stock value is partially observable via stock price. Therefore, from this point of view, models can be classified into two categories: macro- and micro-movement. Macro-movement models refer to daily, weekly and monthly closing price behavior, and micro-movement refer to transaction(trade-by-trade) price behavior. In early studies daily, weekly and monthly stock closing price were explored and micro models were developed in late 1980s when transaction data -interaday- became available and price micro models were emerged [9].

To claim that a specific model precisely represent behavior of an underlying asset is not wise and such a claim has not been made yet. Works on mathematical finance can be precise and they can be comprehensible. Sadly, It is believed the ones which are precise are not necessarily comprehensible, and those are comprehensible are not necessarily precise[3]. Therefore, from this point of view, we may classify models into two categories: the models express the general behavior of stock value and those try to explain a specific phenomenon like a big jump in market (for example, market clash in 1987 or 2000).

In this chapter, we will introduce some popular models and will explain the one we suggest for our work and its reason. According to my study, most developed models are as follows: geometric Brownian motion; jump diffusion process; and chaos theory for modeling the value

[1]Weiner process or Brownian motion will introduce later in modeling chapter

[2]Generally, we should differentiate between stock value and stock price. we may consider stock price as an observation for the value of the stock. Most existing models are expressing the behavior of the value, therefore, in this thesis we mainly mention stock price instead of stock value, however, there are some models that try to modeling the price itself.

of stock and conditional Poisson process for the stock price.

Geometric Brownian motion - Black-Scholes Model

In 1973 a very famous paper published by Black and Scholes in which a derivative pricing model introduced. They suggested models for stock price and Bond, by which they were able to stablish derivative pricing models. The suggested models are:

$$\begin{cases} dB_t = rB_t dt \\ dS_t = \mu S_t dt + \sigma S_t d\omega_t \end{cases}$$

where B_t is the price of Bond at time t, r is interest rate, and S_t is the stock price at time t, μ and σ are drift and volatility of the process respectively[3]. We previously mentioned the difference between macro- and micro movement models, and it is already stated that there is a strong relation between the macro and micro-movement if we assume the market follows GBM[9], therfore, we are able to apply the results from this study as if we are working on micro-movement model even though we do not have access to interaday data and this is one of the reasons that we propose such a model as basis of our study.

This model has been widely used in later studies for stock value behavior and subsequently for option pricing. Instead of assuming μ and σ to be constant which is implied by the model, other models developed for behavior of drift and volatility (for example μ follows a Markov model[4][5]). The reason of these studies on Black-Sholes model were two flaws emerged from many empirical investigations[10]. The first one called as "the asymmetric leptokurtic features", which can be interpreted that the asset return distribution is skewed to the left, and has a higher peak and two heavier tails than those of the normal distribution, the second one, called "the volatility smile", is suggesting that the implied volatility curve resembles a smile shape rather being constant. Many studies have been conducted to modify the BlackScholes model to explain the two empirical phenomena. For the first case a variety of models have

[3]We will discuss this model in more details in next chapters since we consider this model for stock value in this study

[4]We will introduce and apply this model in next chapters.

been proposed such as: (a) chaos theory, fractal Brownian motion, and stable processes; (b)time-changed Brownian motions; see[10] In a parallel development, different models are also proposed to incorporate the "volatility smile" in option pricing. Popular ones include: (a) stochastic volatility and ARCH models; (b) constant elasticity model (CEV) model; (c) normal jump models; (d) stochastic-volatility and jump-diffusion models; (e) models based on Levy processes [10].

Volatility Models - A Jump Diffusion Model

As explained, there are several type of models has been developed to justify the volatility of a price of an underlying asset (for example, a stock or stock index). There is a popular set of models, using the same method, which are called jump-diffusion. These are called jump models since all of them consist two different part, a continuous part of the model and a jump part. Many jump-diffusion models are using geometric Brownian motion for the continuous part and each specific one tries to suggest different model for the jump part. In this section, we will introduce two different jump models, the first one developed by Merton(1976) [9] called a standard jump-diffusion model.

$$dS_t = S_t\Big(\mu S_t dt + \sigma S_t d\omega_t + Y_{N_t} dN_t\Big)$$

where ω_t is standard Brownian motion, $N(t)$ a Poisson process with rate λ and the Y_i are i.i.d. R.V. with continuous density $p_Y(y)$. And its generator is defined as following[10]

$$A_\theta f(x) = A_1 f(x) + \lambda \int \big[f(x(1-z)) - f(x)\big] p_Y(z) dz$$

where θ includes μ, σ and parameters in $p_Y(y)$.

As another example of jump models, we introduce the following model. In this model, the continuous part is also modeled with geometric Brownian motion and the jump part, with the logarithm of jump size having a double exponential distribution and the jump times corresponding to the event times of a Poisson process. The following stochastic differential equation (SDE) is used to model asset price denoted by S_t [10],

$$dS_t = S_t\Big(\mu S_t dt + \sigma S_t d\omega_t + d\big(\sum_{i=1}^{N(t)} V_i - 1\big)\Big)$$

where ω_t is standard Brownian motion as in Black-Scholes, $N(t)$ a Poisson process with rate λ and $\{V_i\}$ is a sequence of independent identically distributed (i.i.d.) nonnegative R.V. such that $X = \log V$ has a double exponential distribution with the density

$$f_X(x) = \frac{1}{2\eta} e^{\frac{-|x-\kappa|}{\eta}}, \; 0 < \eta < 1$$

In other words

$$X - \kappa = \left\{ \begin{array}{ll} \xi & \text{, with probability } \frac{1}{2} \\ -\xi & \text{, with probability } \frac{1}{2} \end{array} \right\}$$

where ξ is an exponential R.V. with mean η and variance η^2. All source of randomness ω_t, $N(t)$ and X's are assumed to be independent (for more details and explanation, see [10]). Clearly, the jump part of this model is a SDE which has a solution, so it is a diffusion process and for this reason the model is called as jump-diffusion process[5].

Complex Models - Chaos Theory

As computers became more powerful in the past two decades, the usage of quantative models also became more popular. Therefore, instead of assuming a phenomenon follows a random process (stochastic), we may able to suggest a dynamic model which is deterministic or partially deterministic [6] . One of such a popular model which are applied in many branches of science is chaos theory. In essence, a chaotic system is a combination of a deterministic and a random process. The deterministic process can be characterized using regression fitting, while the random process can be characterized by statistical parameters of a distribution

[5] Diffusion process is a stochastic process which is the solution to a SDE

[6] Unlike the analytical models, which are random (stochastic) and mostly are too complicate to find a robust solution, in quantative models it is attempted to define the model into two parts, a deterministic part and a random part, in such a models the random part is a simple model which is much easier to have robustness.

function. Thus, using only deterministic or statistical techniques will not fully capture the nature of a chaotic system.

As empirical analysis shows, financial market is a chaotic system. There are few tests by which we may able to find if a phenomenon is chaotic, for a stock market, if we plot the stock market return on an hourly, daily, weekly and yearly basis and do not label them, we are not able to match the pattern to the appropriate time period. This self similarity, called geometric fractal, is a concept of a chaotic system. There are three main properties for a chaotic system which makes it unpredictable; (a) They are feedback systems, where outputs from a previous period are the input for the next period (for example in a stock market, stock price and volume of trade and other statistics of factors of the model); (b) An insignificant input may be compounded over time and greatly influence the behavior of the system, known as "butterfly effect"; (c)These systems are so sensitive according to initial conditions, as a small difference in initial condition produce great one in the final phenomenon [11].

It is shown[12] that even better results can be expected when we combine chaos theory with related theories to form hybrid systems (for example, a hybrid system of chaos theory and Artificial Neural Networks). Artificial Neural Networks are modeled on biological neural networks, or the human brain, and learn by themselves from patterns. This learning can then be applied in prediction or control. The major goal of Neural Network research these days is to develop models which can behave as adaptively as biological systems. To make neural networks behave with the flexibility of biological systems we must make them adaptive and chaotic behavior provides that sort of flexibility. Therefore, a chaotic neural network provides a means for prediction and control of a chaotic system such as an asset return.

Finally, as we stated, chaos theory leads us to a recursive quantative model. As the objective of this thesis is to apply information theory, by which we need an analytical model. Therefore, working on chaos theory will not lead us to our objective of the thesis.

Chapter 2

Criteria Definition

2.1 Introduction

It is well-known that entropy, and in general information theory, has been employed in many areas. Our objective is to apply certain concepts of information theory to economics, analyzing the behavior of the stock market. We propose application of entropy into two main category. The first is based on single stock option, where we will consider only information for one stock price, and the second one is based on multiple stock options, where we will consider a family of stock prices.

2.2 Single Stock

We will try to consider two variables which are 'Price' and 'Volume of trade' of a single option. Generally, the stock price is considered as the main variable to study the behavior of the stock market. In addition, outstanding shares are considered in calculating some indices, as explained in 1.2.2. Unlike the common indices, we will consider the total number of shares traded during the session for which the entropy measure is calculated.

2.2.1 Single Stock Price

Let $S(t)$ denote the price of a specific stock option at time $t \in [0, T]$.

Partition of $[0, T]$ into $0 = t_0 \leq t_1 \leq \cdots \leq t_N = T$ Define the probability mass function as

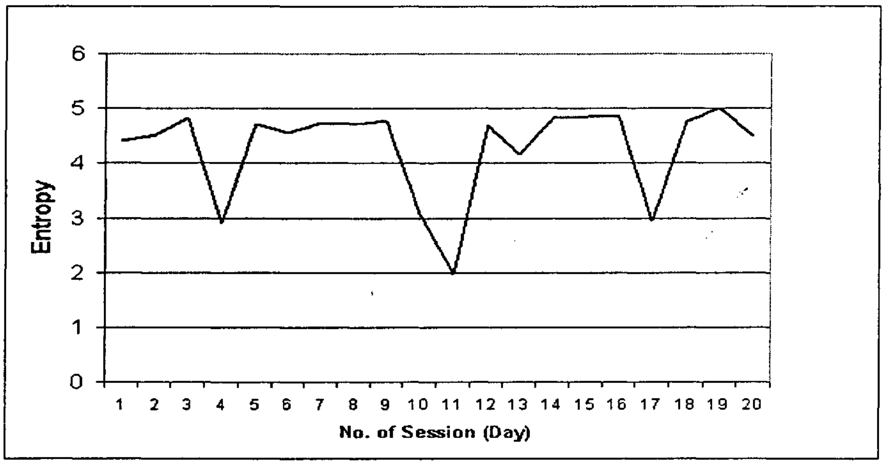

Figure 2.2.1: Entropy of Price for a Single Stock Option

follows.

$$q(t_k) \triangleq \frac{\left(\frac{S(t_{k+1})-S(t_k)}{S(t_k)}\right)^2}{\sum_{j=1}^{N-1}\left(\frac{S(t_{j+1})-S(t_j)}{S(t_j)}\right)^2}\ ,k \in [0, N-1]$$

clearly $\sum_{k=1}^{N-1} q(t_k) = 1$. Then the entropy associated with $q(.)$ is defined as

$$H(q(.)) = -\sum_{k=1}^{N} q(t_k) \log\ q(t_k) \tag{2.2.1}$$

Chart 2.2.1 represents a sample graph for entropy of a single option according to the above formula.

2.2.2 Single Stock Price and Volume of Trade

Let $v(t)$ denote the number of a specific stock share traded at time $t \in [0, T]$ and $S(t)$ denote the price of a the same option at time $t \in [0, T]$. Partition $[0, T]$ into $0 = t_0 \leq t_1 \leq \cdots \leq t_N = T$ we define probability mass function as following:

$$q(t_k) \triangleq \frac{v(t_k)\ S(t_k)}{\sum_{j=0}^{N-1} v(t_j)\ S(t_j)},\quad k \in [0, N-1]$$

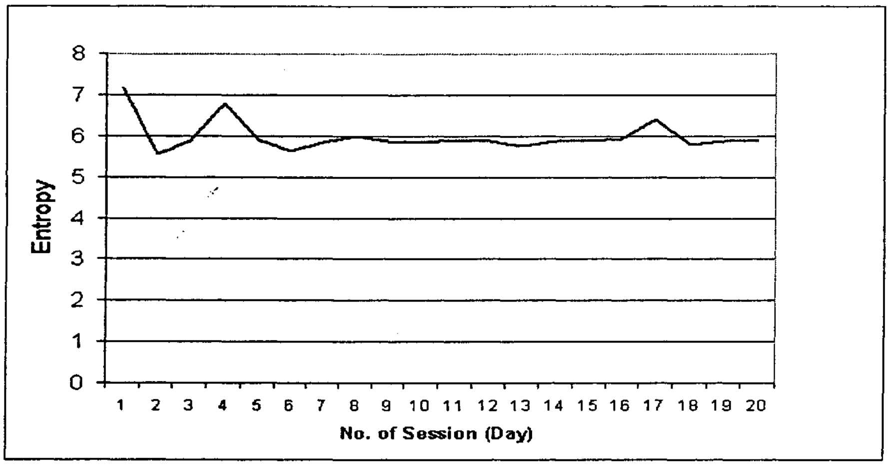

Figure 2.2.2: Entropy for a Single Stock Option with considering Price and Volume of Trade

Then the entropy associated with $q(.)$ can be defined as follows.

$$H(q(.)) = -\sum_{k=1}^{N} q(t_k) \log\ q(t_k) \tag{2.2.2}$$

Chart 2.2.2 represents a sample graph for entropy of a single option based on volume of trade, according to the above formula.

2.3 Multiple Stock Price

We have considered a single option in the market so far and this approach does not give us a proper approximate of the behavior of the market in many cases. In this section, we will consider more than one stock option to find out the possible corelation between different stock options.

Let $S_i(t)$ denote the price of the option $i \in [1, M]$ at time $t \in [0, T]$. Partition the period of time $[0, T]$ into $0 = t_0 \leq t_1 \leq \cdots \leq t_N = T$, we will define the probability mass function as follows.

$$q_i(t_k) = \frac{\left(\frac{S_i(t_{k+1})-S_i(t_k)}{S_i(t_k))}\right)^2}{\sum_{i=1}^{M}\left(\frac{S_i(t_{k+1})-S_i(t_k)}{S_i(t_k))}\right)^2} \quad ,k \in [0, N-1] \ \ ,i \in [1, M]$$

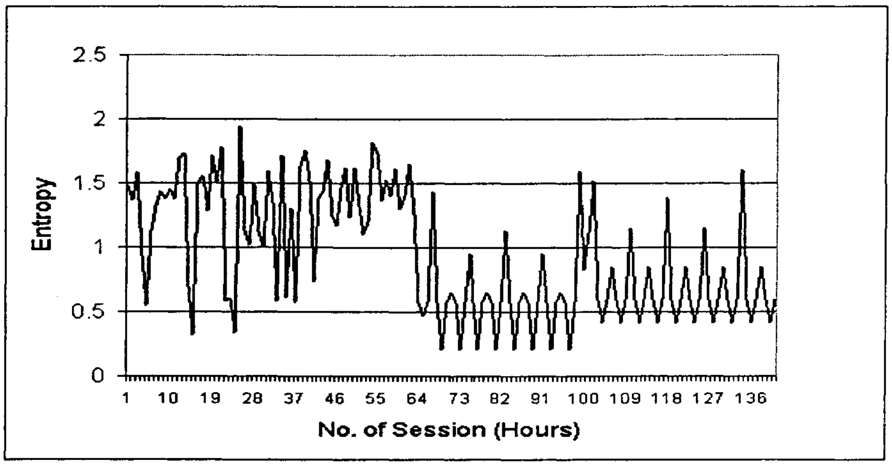

Figure 2.3.3: Entropy for Multiple Stock Options Including Companies Alcatel,Apple,AT&T Corp.,AT&T Canada,Ericsson,Intel,Microsoft,Sun,3M

Then the entropy associate with $q(.)$ can be defined as following:

$$H(t_k, q_1(t_k), ..., q_M(t_k)) = -\sum_{i=1}^{M} q_i(t_k) \log\ q_i(t_k) \tag{2.3.3}$$

2.3.1 Arbitrary Index

Every popular index on the market includes some stocks. For example the Dow Jones Industrial Average (DJIA) includes 30 wealthy and successful companies. It's believed the behavior of these companies, with respect to their stocks' price, can express the economic situation in general.

Investors choose a portfolio of stocks to increase the possibility of raising the value of the assets or decreasing the risk of loss according to their investment policy, which is either optimistic or pessimistic.

Therefore, we may apply the idea of multiple stock options to construct two different indices. Firstly, we will apply it to the same companies which are included in a common index, such as the Dow Jones. Subsequently, we are able to make a comparison between the DJIA index and the index created by us. Finally, we will choose an arbitrary set of stocks as a portfolio in order to create our own index.

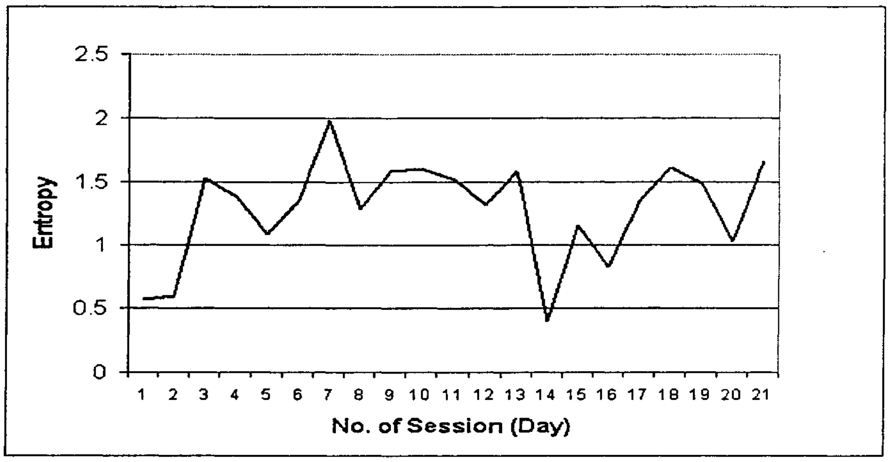

Figure 2.3.4: Entropy of closing price for Multiple Stock Options Including Companies Alcatel,Apple,AT&T Corp.,AT&T Canada,Ericsson,Intel,Microsoft,Sun,3M

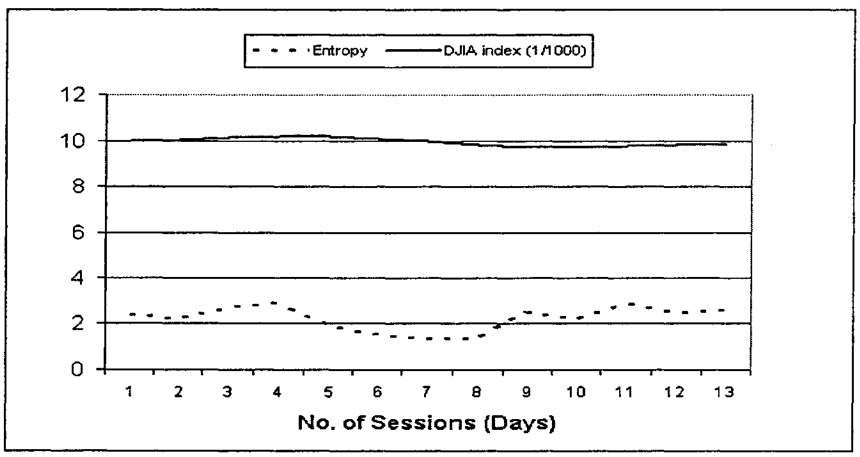

Figure 2.3.5: Entropy for Dow Jones Companies vs. DJIA index

For the first case, by using the formula (2.3.3) and considering the companies that are included in DJIA, we will calculate the index, as shown in graph 2.3.5. We believe that somehow this index gives more information than Dow Jones Industrial Average and we will try to support it later.

Similar to many common indices, we are able to consider option prices in two different ways. The first is closing price and the second is intra day which refers to every single transaction during a day. We considered closing price to calculate entropy so far, and we will continue calculating the entropy by considering intra day. Since we don't have access to the informations of every transaction we considered every minute as the time period for data gathering.

For the second one, ten IT and High-tech companies are selected as our defined index. The reason for picking such companies is that they were very bursty during the years 2001 & 2002, the period on which we gathered the data. Figure 2.3.5 shows entropy calculated for a set of companies where the closing price is considered in entropy calculation. Then,we will calculate the entropy, using the companies included in DJIA , to find out if the new index gives more information.

2.4 Discussion and Analysis

It is shown [2] that the largest value of the entropy is, for a specific period, the larger the uncertainty of the stock or a collection of stocks is.

We will discuss the entropy in three cases, firstly, when it increases; secondly, when it decreases; and finally, when it is constant or remains close to the previous period.

2.4.1 Single Stock

We will discuss the three above-mentioned cases, regarding the entropy of a single stock option. As stated in Section 2.2.1, each point on graph 2.2.1 is the entropy calculated for a specific period, i.e. a day or an hour. Therefore, when we talk about the entropies, we will refer to the points constructed on the graph.

The first case is when the entropy is constant or it is very close to the previous state. When the entropy of a single stock at two consecutive periods is the same, we conclude that the probability of all prices remains constant during the two periods. For example, if the probability of a stock's price being 10 dollars is $\frac{1}{3}$ during the first period, the probability remains the same during the second period. Consequently, we conclude that the expected value of the price remains the same regardless of how volatile the stock's price is during the two periods. In another words, we may conclude the capital assets of the company remain the same during the two periods.

There is an exception to the above conclusion, when the probability of two or more prices change among themselves without any change to the value of the probabilities. For example, suppose that probabilities for the price being 10 and 11 dolor, are $\frac{1}{3}$ and $\frac{1}{6}$ during the past period, respectively. If the probabilities switch so that we have $\frac{1}{3}$ probability for 11-dolor and $\frac{1}{6}$ probability for 10-dolor item, we will have the same entropy. By changing the scope of study to a micro, we are able to decrease such distortions. For example, if we consider every one hour as a period, we know that change of such probabilities are not likely to happen.

The second case is when the entropy shows an increase. When the entropy increases it means uncertainty increases and this can be interpreted as a convergence of probabilities of different state of prices. Ultimately, probabilities converge to a number which is called the maximum entropy of a finite sequence, when the sequence is uniformly distributed. Therfore, we can conclude that when the uncertainty increases the stock option becomes more volatile, since the probabilities converge. The more volatile a stock price is, the more risk we face, if we intend to invest in that stock. Therefore, we may conclude that if the entropy graph of a specific stock shows an increase, we are reluctant to invest in that stock.

The third case is when the entropy decreases. Therefore, we will conclude that the probability of a certain prices increase. This case may be the result of two different occurences, when the price of a stock either increases or decreases. Therfore, the decrease in entropy, which translates into a decrease of uncertainty, shows that price of the stock keeps going with the same trend. Therefore, if the price increases we like to invest in the stock and in case of decrease of price we don't. In another words, uncertainty decreases when the expected value of a stock price continues decreasing or increasing.

2.4.2 Multiple Stocks

In general, we need to consider the joint density of a portfolio and we will discuss it in the next chapters. The probability measure we considered is the possibility of occurrence of rate of change of a stock price, with regard to the price of other stock options in a portfolio, according to formula (2.3.3). We will discuss the three different cases of entropies, similar to what we did in a single option section. First we will discuss the most extreme cases of the entropy: the minimum number which is zero and the maximum one which is $\log n$, where n is the number of stock options included in a portfolio. The maximum case happens when the rate of change of all options are the same and the minimum one (zero) is when only one of the options shows change in the price and the rest remain unchanged.

When the entropy remains unchanged, similar to the single option section, the probability of items included in calculating the entropy remains unchanged. Unlike the single option, it cannot be interpreted as a change to the expected value of each option price. When the probability of each individual stock's price in a joint density remains the same, it means all the items have the same direction in their movement. Therefore, there are two directions for a portfolio in such a case: firstly, when the rate of change of each individual remains the same and secondly, when the rate of change of all items changes with the same ratio. For example, if the rate of change of option A were changed from 10 to 12 percent, the rate of change of option B would also be changed from 5 to 6 percent.

When the entropy increases, knowing the fact that the maximum entropy happens in a uniformly distributed sequence, we can conclude that when the entropy increases, the probability of each sequence item becomes closer to $\frac{1}{n}$, where n is the number of options included in a portfolio. In this case, the "rate of change" of different stocks are similar.

When the entropy shows a decrease, in which the entropy gets closer to the minimum state, we can conclude that there are few stock prices which had noticeable changes, compared to the previous period of study. In another words, there are more options with the rate of change close to zero.

In summary, since the extreme points in an entropy rarely happen in reality, our discussions were based on comparing the calculated entropy for a specific session with prior sessions. Therefore, to apply the set of mentioned interpretations, we should consider a moment of time called zero time as our origin, similar to other common indices. We need to analyze the status of companies in a portfolio at zero point. Then, we can analyze the later moments by using our index.

Our interpretation also has a vulnerable point. On the one hand, we did the discussion based on the most common probability measure used in reality of different prices of occuring, but the probability measure we've defined at the beginning of this chapter is different in concept. On the other hand, we can consider our defined probability measure similar to the probability of "rate of change" interpreted by a real probability measure. This may cause flaws in our interpretation. Therefore, we will consider a regular measure of probability in the following chapters.

Chapter 3

Mathematical Modeling

3.1 Introduction

A risky asset like a stock option is modeled by a random variable or a random process whose value changes over time in an uncertain way. To achieve the objective of the thesis, which is defining a new index, there are four main steps to be followed.

The first step in studying a stochastic process is to assume an appropriate model which explains the process. This assumption is not simple when a process is complicated. One example for such a complicate process is stock price. Generally, we are not able to find out what type of model exactly the stock price follows, although many studies have been done to introduce the stochastic models which capture behavior of a stock price. In this thesis we assume that the price follows two type of models, Geometric Brownian motion and a discrete time Markov Chain.

3.2 Preliminary Tools

3.2.1 Brownian Motion Process

A Random Process (R.P) $W \triangleq \{\omega_t; t \in [0,T]\}$ taking values in $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ is called a n-dimensional Brownian motion if and only if the following hold.

i) $P(\omega_0 = 0) = 1$

ii) Partition $t \in [0,T]$ into $N-1$ partitions such that

$0 \leq t_i \leq t_2 \leq \cdots \leq t_{n-1} \leq t_n \leq t_n, n \in \mathbb{N}_+$.

W has stationary independent increment, that is, for any disjoint intervals

$[t_i, t_{i+1}] \subset [0,T]$, The increments $\Delta\omega_{t_i} \triangleq \omega_{t_i+1} - \omega_{t_i}$ satisfy

$P(\Delta\omega_{t_1} \in A_1, \Delta\omega_{t_2} \in A_2, \cdots, \Delta\omega_{t_n} \in A_n)$

$= P(\Delta\omega_{t_1+s} \in A_1, \Delta\omega_{t_2+s} \in A_2, \cdots, \Delta\omega_{t_n+s} \in A_n)$

$= \Pi_{j=1}^{n} P(\Delta\omega_{t_j+s}) = \Pi_{j=1}^{n} P(\Delta\omega_{t_j}), \quad \forall A_j \in B(\mathbb{R}^n), [t_j+s, t_{j+1}+s] \subset [0,T], 1 \leq j \leq n$

iii) The increments $\omega_t - \omega_s$ are Gaussian with

$E[\omega_t - \omega_s] = 0$ and $Var(\omega_t - \omega_s) = \sigma^2(t-s)I_n, \ \forall s,t \in [0,T], \sigma^2 \in \mathbb{R}_+$

If $\sigma^2 = 1$ then W is called a standard Brownian motion.

3.2.2 Maximum Liklihood Estimate

Maximum likelihood (ML) estimation is a statistical result that has a good asymptotic properties when used in large sample statistical inference. Consider the R.V.

$Z_1^N \triangleq (z_1, \cdots, z_n) : \Omega \rightarrow \mathbb{R}^{N_P}$

defined on the family of probability spaces $\{(\Omega, \mathcal{F}, P_\theta); \theta \in \Theta\}$ where (Θ, A) is a Borel measurable space.

Let $f(Z_1^N, .) : \Theta \rightarrow \mathbb{R}+$ be a likelihood function of f_Θ at $Z_1^N \in \mathbb{R}^{N_P}$

Then $\hat{\theta}^N(.) : \mathbb{R}^{N_P} \rightarrow \Theta$, $\hat{\theta}^N \triangleq \hat{\theta}^N(Z_1^N)$

satisfying

$f(Z_1^N; \hat{\theta}^N(Z_1^N)) \geq f(Z_1^N; \theta)$, $\forall \theta \in \Theta$

is called the "ML estimate" of the true parameter θ^0

ML Estimation in i.i.d. Gaussian Variable

Let $Z = \{z_k; k \in \mathbb{N}+\}$ be a sequence of i.i.d. R,V's having distribution $N(\mu, \sigma^2)$.

$\theta = (\mu, \sigma^2) \in \Theta \triangleq \mathbb{R} \times (0, \infty)$. For a given sample $Z_1^N(\omega)$ which is generated by $\theta \in \Theta$ the "sample density" of Z_1^N is

$f(Z_1^N; \theta) = f(z_1, \cdots, z_N; \theta) = \prod_{i=1}^N f(z_i; \theta) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2}\frac{\sum_{i=1}^N (z_i - \mu)^2}{\sigma^2}}$

Hence

$$
\begin{aligned}
L(Z_1^N; \theta) &= -\log f(Z_1^N; \theta) \\
&= \frac{N}{2}\log 2\pi + \frac{N}{2}\log \sigma^2 + \frac{1}{2\sigma^2}\sum_{i=1}^N (z_i - \mu)^2
\end{aligned}
$$

Thus, $L(Z_1^N; \theta)$ is minimized over Θ at any point where $f(Z_1^N; \theta)$ is maximized over Θ. Therefore [1]

$$
\begin{aligned}
\frac{\partial}{\partial \mu} L(Z_1^N; \theta)\Big|_{\theta = \hat{\theta}} &= 0 \\
\frac{1}{\hat{\sigma}^2}\sum_{i=1}^N (z_i - \hat{\mu}).(-1) &= 0 \\
\frac{N}{\hat{\sigma}^2}\hat{\mu} - \frac{1}{\hat{\sigma}^2}\sum_{i=1}^N z_i &= 0 \\
\hat{\mu} = \frac{1}{N}\sum_{i=1}^N z_i &= \overline{z}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial \sigma^2} L(Z_1^N; \theta)\Big|_{\theta = \hat{\theta}} &= 0 \\
\frac{N}{\hat{\sigma}^2} &= \frac{1}{2(\hat{\sigma}^2)^2}\sum_{i=1}^N (z_i - \hat{\mu})^2 \\
\hat{\sigma}^2 &= \frac{1}{N}\sum_{i=1}^N (z_i - \hat{\mu})^2
\end{aligned}
$$

[1] $\hat{\mu}$,$\hat{\sigma}$ and $\hat{\theta}$ are estimate values for μ, σ and θ respectively.

3.3 Stock Price as a Geometric Brownian Motion

It is often suggested [2] that stock price can be modeled using Geometric Brownian motion (GBM). In this section we employ such a models. Therefore a model for the market can be introduced as follows.

Let $S_t,\ t \in [0,T]$, be the price of the specific option at time t, then the Stochastic Differential Equation (SDE) of the price is given by:

$$dS_t = f(t)S_t dt + g(t)S_t dw_t \tag{3.3.1}$$

where $f(t)$ is drift and $g(t)$ is volatility of the process $\frac{dS_t}{S_t}$ at time t and dw_t is Brownian motion increment.

Defining "rate of change" of price as $y(t) = \frac{dS_t}{S_t}$, then

$$y(t) = \frac{dS_t}{S_t} = f(t)dt + g(t)dw_t \tag{3.3.2}$$

According to the above assumption, the price follows a GBM, the two equations stated in (3.3.1) and (3.3.2), will be our main focus, for the rest of the work.

3.3.1 Log-Normal Model - Constant Drift and Volatility

First we model the drift and the volatility as constant coefficient for the period of time $[0,T]$. Therefore we regard μ as the drift and σ as the volatility so the equation in (3.3.1) is given by.

$$dS_t = \mu S_t dt + \sigma S_t dw_t \tag{3.3.3}$$

Consequently, we have

$$dy(t) = \frac{dS}{S_t} = \mu dt + \sigma dw_t \tag{3.3.4}$$

[2] As previously explained one of main models applied for stock price is the Black-Scholes model in which stock price follows GBM

Using the Ito integral the SDE mentioned in (3.3.3), has the following solution.

$$S_t = S_0 \exp\{(\mu - \frac{1}{2}\sigma^2)t + \sigma w_t\} \tag{3.3.5}$$

The discrete form of the model in 3.3.2 is

$$y(t_k) = \mu(t_{k+1} - t_k) + \sigma \Delta W(t_k)$$

and according to the definition of B.M., ΔW has the Gaussian density which is defined as $\Delta W(t_k) \approx N(0; t_{k+1} - t_k)$

3.3.2 Log-Normal Model - Markov Chain Drift

In this part the volatility is assumed to be a constant coefficient, and μ is considered as a function of time. Thus (3.3.3) changes to

$$dS_t = \mu(t) S_t dt + \sigma S_t dw_t \tag{3.3.6}$$

To make it easier, we further assume that the drift μ_t follows a finite state Markov model. Since the process is already considered as GBM, it is proved that such an assumption is justified [5]. Suppose the drift $\{\mu_t\}$ is described by a countable state Markov model which takes value in a finite set ρ, where $\rho = \{e_1, e_2, \cdots, e_N\}$ and e_i is a unit vector with 1 in the ith element and 0 elsewhere [8]. We define a new process X_t which determines the state of μ_t at time t. the model for the stock price is given by

$$S_t = S_0 \exp\{\int_0^t < X_u, \rho > du + \sigma w_t\} \tag{3.3.7}$$

Later in the estimation section we shall use the above mentioned model to estimate the drift and to compute the transition probabilities of the Markov chain.

3.4 Option Price as a Discrete Markov Model

In this section we discretize the price model as a finite-state, homogeneous, discrete-time Markov chain $S_{t_k},\ k \in N$ and we suppose either S_0 is given or its distribution is known. Suppose that the state space of S_{t_k} has N elements which is identified as $S_{t_k} = \{e_1, \cdots, e_N\}$ where e_i are unit vectors in $\mathbb{R}^N$ with unity as the ith element and zeros elsewhere. Consider a matrix A as the transition matrix for the process, defined by

$$A = (a_{ji}) \in \mathbb{R}^{N\times N} \text{ and } a_{ji} = P\{S_{t_{k+1}} = e_j \mid S_{t_k} = e_i\}$$

Then a model for the price can be identified as follows.

$$S_{t_{k+1}} = AS_{t_k} + V_{t_{k+1}} \tag{3.4.8}$$

where, $P(S_{t_{k+1}}) = A^{tr}P(S_{t_k})^3$ and matrix A does not depend on time since we previously assumed that the model is homogeneous.

It also can be shown that $\{V_{t_k},\ k \in N\}$ is a sequence of martingale increments[5].

3.5 Multi-Dimension Log-Normal Model

It is often suggested that multiple stock option prices can be modeled by multi-dimension GBM process. Such a model for a multiple stock option price is described by.

$$dS_t^i = S_t^i(\mu_t^i dt + \sum_{j=0}^{n} \sigma_{ij}(t)dB_t^j)\ ,\ i = 1, 2, \cdots, n \tag{3.5.9}$$

Where B_t^j is a Brownian motion.

Clearly, by defining the rate of change of the option price, denoted by $dy_t^i \triangleq \frac{dS_t^i}{S_t^i}$,we consider a Gaussian process, that is easier to work with than the origin model. Following the same approach as in the single option model, we define

$$dy_t^i = \mu_t^i dt + \sum_j \sigma_{ij}(t)dB_t^j\ ,\ i = 1, 2, \cdots, n \tag{3.5.10}$$

[3] A^{tr} denotes the transpose of matrix A

3.6 Estimation and Filtering

In this section we estimate parameters of above proposed models, using maximum likelihood techniques.

3.6.1 Estimation of Drift and Volatility

First we consider the model with drift and volatility being parameters. Using maximum likelihood estimation for μ and σ by first computing the (3.3.1) joint probability density function.

$$P(y(t_1), y(t_2), ...y(t_N)) = \Pi_{i=1}^{N} P(y(t_i)) = \Pi_{i=1}^{N}\left(\frac{e^{-\frac{(y(t_i)-\mu(t_{i+1}-t_i))^2}{2\sigma^2(t_{i+1}-t_i)}}}{(2\pi\sigma^2(t_{i+1}-t_i))^{\frac{N}{2}}}\right)$$

We have the following maximum likelihood estimates.

$$L = \max_{(\mu,\sigma^2)} log(\Pi_{i=1}^{N}(\frac{e^{-\frac{(y(t_i)-\mu(t_{i+1}-t_i))^2}{2\sigma^2(t_{i+1}-t_i)}}}{(2\pi\sigma^2(t_{i+1}-t_i))^{\frac{1}{2}}}))$$

$$\begin{aligned}
&= \max_{(\mu,\sigma^2)} \sum_{i=1}^{N} \left[\log\left(e^{-\frac{(y(t_i)-\mu(t_{i+1}-t_i))^2}{2\sigma^2(t_{i+1}-t_i)}}\right) - \frac{1}{2}\log(2\pi\sigma^2(t_{i+1}-t_i))\right] \\
&= \max_{(\mu,\sigma^2)} \sum_{i=1}^{N} \left[-\frac{(y(t_i)-\mu(t_{i+1}-t_i))^2}{2\sigma^2(t_{i+1}-t_i)} - \frac{1}{2}\log(2\pi\sigma^2(t_{i+1}-t_i))\right]
\end{aligned}$$

Hence,

$$\begin{aligned}
\frac{\partial}{\partial\mu}L\Big|_{\mu=\hat{\mu},\sigma^2=\hat{\sigma}^2} &= \sum_{i=1}^{N} \frac{2}{2\hat{\sigma}^2(t_{i+1}-t_i)}(y(t_i)-\mu(t_{i+1}-t_i))(-(t_{i+1}-t_i)) \\
&= -\sum_{i=1}^{N}\frac{1}{\hat{\sigma}^2}(y(t_i)-\mu(t_{i+1}-t_i)) = 0
\end{aligned}$$

Therefore,

$$\begin{aligned}
\hat{\mu} &= \frac{\sum_{i=1}^{N} y(t_i)}{\sum_{i=1}^{N}(t_{i+1}-t_i)} \\
&= \frac{1}{t_N - t_1}\sum_{i=1}^{N} y(t_i)
\end{aligned}$$

We also have

$$\begin{aligned}\frac{\partial}{\partial\sigma^2}L\Big|_{\mu=\hat{\mu},\sigma^2=\hat{\sigma}^2} &= \sum_{i=1}^{N}\frac{(y(t_i)-\hat{\mu}(t_{i+1}-t_i))^2}{(t_{i+1}-t_i)}\frac{1}{(\hat{\sigma}^2)^2}-\frac{1}{2}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}^2} \\ &= \frac{1}{2(\hat{\sigma}^2)^2}\sum_{i=1}^{N}\frac{(y(t_i)-\hat{\mu}(t_{i+1}-t_i))^2}{(t_{i+1}-t_i)}-\frac{1}{2}\frac{N}{\hat{\sigma}^2}=0\end{aligned}$$

Thus,

$$\hat{\sigma}^2=\frac{1}{N}\sum_{i=1}^{N}\frac{(y(t_i)-\hat{\mu}(t_{i+1}-t_i))^2}{(t_{i+1}-t_i)}$$

Therefore, the estimates of μ and σ are the sample mean and variance.

3.6.2 Estimation of Drift as a Function of Time

In this section we consider a model with drift as a function of time, as stated in (3.3.6). Again by considering the "rate of change" model , we can perform the estimation easier. Since, we have no priori knowledge on $\mu(t)$, we can approximate it using a orthogonal basis $\{\phi_i(t)\}_{i=1}^{m}$

$$\mu(t)=\sum_{i=1}^{M}\theta_i\phi_i(t)$$

Where $\{\theta_i\}_{i=1}^{M}$ are the parameters to be estimated. The basis $\{\phi_i(t)\}_{i=1}^{M}$ may be regarded as Gaussian and can be considered as a function of $S(t)$. Therefore, one possibility is to assume

$$dy(t_k)=\sum_{i=1}^{M}\theta_i\phi_i(S_{t_k})dt+\sigma_i dw \tag{3.6.11}$$

The joint density of $y(t_k)$ is

$$f(y_1,...,y_n)=\Pi_{k=1}^{N}\frac{e^{\frac{-(y(t_k)-\sum_{i=1}^{M}\theta_i\phi_i(t_k)(t_{k+1}-t_k))^2}{2\sigma^2(t_{k+1}-t_k)}}}{(2\pi\sigma^2(t_{k+1}-t_k))^{\frac{1}{2}}}$$

The log likelihood function is

$$L=-\sum_{k=1}^{N}\frac{\{y(t_k)-\sum_{i=1}^{M}\theta_i\phi_i(t_k)(t_{k+1}-t_k)\}^2}{2\sigma^2(t_{k+1}-t_k)}-\frac{1}{2}\sum_{k=1}^{N}\log 2\pi\sigma^2(t_{k+1}-t_k)$$

Hence,

$$\begin{aligned}
\frac{\partial}{\partial\theta_l}L\Big|_{\theta_l=\hat{\theta}_l} &= -\sum_{k=1}^{N}\frac{1}{2\sigma^2(t_{k+1}-t_k)}2\{y(t_k)-\sum_{i=1}^{M}\hat{\theta}_i\phi_i(t_k)(t_{k+1}-t_k)\}(-\phi_l(t_k)(t_{k+1}-t_k)) \\
&= \sum_{k=1}^{N}\frac{\phi_l(t_k)(t_{k+1}-t_k)}{\sigma^2(t_{k+1}-t_k)}\{y(t_k)-\sum_{i=1}^{M}\hat{\theta}_i\phi_i(t_k)(t_{k+1}-t_k)\} \\
&= \frac{1}{\sigma^2}\sum_{k=1}^{N}\{y(t_k)\phi_l(t_k)-\phi_l\sum_{i=1}^{M}\hat{\theta}_i\phi_i(t_k)(t_{k+1}-t_k)\}=0, \quad l=1,2,\cdots,M
\end{aligned}$$

Thus, we have

$$\sum_{k=1}^{N}y(t_k)\phi_l(t_k)=\sum_{k=1}^{N}\sum_{i=1}^{M}\hat{\theta}_i\phi_l(t_k)\phi_i(t_k)(t_{k+1}-t_k), \quad l=1,2,\cdots,M$$

Finally, we have

$$\begin{pmatrix}
\sum_{k=1}^{N}\phi_1(t_k)\phi_1(t_k)(t_{k+1}-t_k) & \cdots & \sum_{k=1}^{N}\phi_M(t_k)\phi_1(t_k)(t_{k+1}-t_k) \\
\vdots & & \\
 & & \\
\sum_{k=1}^{N}\phi_1(t_k)\phi_M(t_k)(t_{k+1}-t_k) & \cdots & \sum_{k=1}^{N}\phi_M(t_k)\phi_M(t_k)(t_{k+1}-t_k)
\end{pmatrix}
\begin{pmatrix}\hat{\theta}_1\\ \vdots \\ \\ \hat{\theta}_M\end{pmatrix}
=\begin{pmatrix}\sum_{k=1}^{N}y(t_k)\phi_1(t_k)\\ \vdots \\ \\ \sum_{k=1}^{N}y(t_k)\phi_M(t_k)\end{pmatrix} \tag{3.6.12}$$

Calculating $\hat{\theta}_i$ from the equation (3.6.12) we are able to estimate $\mu(t)$ at different times.

3.6.3 Estimation of Multi-Dimension Log-Normal Model Parameters

As mentioned, the multi-dimension model is log-normally distributed. Since working with a Gaussian model is easier than a log-normal model we introduce the "rate of change" denoted by $dy_t = \frac{dS_t}{S_t}$. The joint density function of a multidimensional Gaussian is called joint Gaussian distributed and it is given by

$$P_y(y_t) = \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \exp\{\frac{-1}{2}(y_t - \mu)^{tr}\Sigma^{-1}(y_t - \mu)\}$$

where, the mean vector and the covariance matrix are denoted by $\mu = (\mu_1, \cdots, \mu_n)$ and $\Sigma = E\{(y_t - \mu)(y_t - \mu)^{tr}\}$. Using ML estimation, we are able to find the optimum estimation for mean vector and covariance matrix. We suppose that the estimate mean vector $\hat{\mu_t}$ and covariance matrix $\hat{\Sigma}$ are constant during a time period $t \in [0, T]$

$$P(y_{t_1}, \cdots, y_{t_N}) = \prod_{i=1}^{N} P_y(y_{t_i}) = \prod_{i=1}^{N} \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \exp\{\frac{-1}{2}(y_{t_i} - \mu)^{tr}\Sigma^{-1}(y_{t_i} - \mu)\}$$

Then

$$\begin{aligned}
L &= -\log\,[\prod_{i=1}^{N} \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \exp\{\frac{-1}{2}(y_{t_i} - \mu)^{tr}\Sigma^{-1}(y_{t_i} - \mu)\}] \\
&= -\sum_{i=1}^{N} \log\,[\frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \exp\{\frac{-1}{2}(y_{t_i} - \mu)^{tr}\Sigma^{-1}(y_{t_i} - \mu)\}] \\
&= -\sum_{i=1}^{N}\,[-\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma| - \frac{1}{2}\{(y_{t_i} - \mu)^{tr}\Sigma^{-1}(y_{t_i} - \mu)\}] \\
&= \sum_{i=1}^{N} \frac{n}{2}\log(2\pi) + \frac{1}{2}\sum_{i=1}^{N}\log|\Sigma| + \frac{1}{2}\sum_{i=1}^{N}\{(y_{t_i} - \mu)^{tr}\Sigma^{-1}(y_{t_i} - \mu)\}
\end{aligned}$$

To find the $\max_{(\mu,\Sigma)} L$ we need to calculate partial derivatives for μ and Σ. Therefore, we have

$$\frac{\delta}{\delta\mu}L\bigg|_{\mu=\hat{\mu},\Sigma=\hat{\Sigma}} = \sum_{i=1}^{N} 2(y_{t_i} - \hat{\mu})^{tr}\hat{\Sigma}^{-1}(-1)$$
$$= \sum_{i=1}^{N}(y_{t_i} - \hat{\mu})^{tr} = N\hat{\mu} - \sum_{i=1}^{N} y_{t_i} = 0$$

Therefore, the ML estimate for mean vector is

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} y_{t_i} \tag{3.6.13}$$

Where $y_{t_i} = (y_{t_i}^1, \cdots, y_{t_i}^n)$.

Similarly, for the covariance matrix we have

$$\frac{\delta}{\delta\Sigma}L\bigg|_{\mu=\hat{\mu},\Sigma=\hat{\Sigma}} = \frac{1}{2}\sum_{i=1}^{N}\hat{\Sigma}^{-1} - \frac{1}{2}\sum_{i=1}^{N}\hat{\Sigma}^{-2}(y_{t_i} - \hat{\mu})(y_{t_i} - \hat{\mu})^{tr}$$
$$= \frac{N}{2}\hat{\Sigma}^{-1} - \frac{1}{2}\sum_{i=1}^{N}\hat{\Sigma}^{-2}(y_{t_i} - \hat{\mu})(y_{t_i} - \hat{\mu})^{tr}$$
$$= \frac{N}{2}\hat{\Sigma} - \frac{1}{2}\sum_{i=1}^{N}(y_{t_i} - \hat{\mu})(y_{t_i} - \hat{\mu})^{tr} = 0$$

Therefore, the ML estimate for covariance matrix is

$$\hat{\Sigma} = \frac{1}{N}\sum_{i=1}^{N}(y_{t_i} - \hat{\mu})(y_{t_i} - \hat{\mu})^{tr} \tag{3.6.14}$$

Chapter 4

Information Measures for Stock Indices

4.1 Introduction

In this chapter we will define new indices by using information theory measures. We will consider two cases in analyzing stock options; a single stock option and multiple stock options. To study a single stock option, we mainly implement three types of information measures; entropy, entropy rate and relative entropy. We also consider two different random variable in our study; option price (S_t) and "rate of change" (y_t). Finally, in multiple stock options, relative entropy and joint entropy are the main information measures to be considered in our study.

4.2 Information Measures for Single stock price

Let S_t be the price of the specific option at time t, and let μ be the drift and σ be the volatility of the process. Then the Stochastic Differential Equation (SDE) satisfied by the price is

$$dS_t = \mu S_t dt + \sigma S_t dB_t \tag{4.2.1}$$

Using Ito integral, we have [3],

$$\begin{aligned} d\log S_t &= \frac{1}{S_t}dS_t + \frac{1}{2}(-\frac{1}{S_t^2})d < S_t, S_t > \\ & \frac{1}{S_t}[\mu S_t dt + \sigma S_t dB_t] - \frac{1}{2}\sigma^2 dt \\ \int d\log S_t &= \int \mu dt + \int \sigma dB_t - \int \frac{1}{2}\sigma^2 dt \\ \log S_t - \log S_0 &= \int (\mu - \frac{1}{2}\sigma^2)dt + \int \sigma dB_t \end{aligned}$$

Hence

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t} \tag{4.2.2}$$

Defining the "rate of change" of the price as $y(t) = \frac{dS}{S}$, we have

$$y(t) = \frac{dS_t}{S_t} = \mu dt + \sigma dB_t \tag{4.2.3}$$

Knowing the density of $y(t)$ and S_t, as Gaussian and Log-Normal, the objective of this chapter is to compute the relative entropy and entropy rate. For computing the above mentioned measures we need to calculate the entropy of rate of change and option price.

4.2.1 Entropy of the "Rate of Change" and Stock Price

Before computing the entropy of the "rate of change" of the price and the stock price, the density of them should be identified. It is known that the increment of a BM has a Gaussian distribution and because we assumed that μ and σ are constant, thus $y(t)$ has a Gaussian distribution as well. On the other hand, since $S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma(B_t - B_0)}$ at $B_0 = 0$, then we can conclude that $(\mu - \frac{1}{2}\sigma^2)t + \sigma(B_t - B_0)$ has Gaussian distribution.

Entropy of the "Rate of Change"

Based on definition for entropy we have

$$\begin{aligned} H(y_t) &= -E\big[\log P(y_t)\big] \\ &= -\int P(y_t)\log P(y_t) dy_t \end{aligned}$$

Hence

$$
\begin{aligned}
H(y_t) &= -\int_{-\infty}^{+\infty} \log(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{y_t-\mu}{\sigma})^2}) P(y_t) dy_t \\
&= -\int_{-\infty}^{+\infty} \big(\log \frac{1}{\sigma\sqrt{2\pi}} + \log e^{-\frac{1}{2}(\frac{y_t-\mu}{\sigma})^2}\big) P(y_t) dy_t \\
&= -\big(\log \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} P(y_t) dy_t - \frac{1}{2\sigma^2} \int_{-\infty}^{+\infty} (y_t-\mu)^2 P(y_t) dy_t\big)
\end{aligned}
$$

Since y_t is Gaussian, then

$\int_{-\infty}^{\infty} P(y_t)dy = 1$ and

$\int (y_t-\mu)^2 P(y_t)dy = Var(y_t) = \sigma^2$

Hence

$$
\begin{aligned}
H(y_t) &= -\log \frac{1}{\sigma\sqrt{2\pi}} + \frac{1}{2} \\
&= \frac{1}{2} + \frac{1}{2}\log(2\pi\sigma^2)
\end{aligned}
\tag{4.2.4}
$$

Entropy of the Stock Price

Based on definition for entropy we have

$$
\begin{aligned}
H(S_t) &= -E\big[\log P(S_t)\big] \\
&= -\int P(S_t) \log P(S_t) dS_t
\end{aligned}
$$

Since $S(t)$ is lognormal, then

$$
\begin{aligned}
H(S_t) &= -\int_0^{+\infty} \ln\big(\frac{1}{S_t\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\ln S_t-\mu}{\sigma})^2}\big) \frac{1}{S_t\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\ln S_t-\mu}{\sigma})^2} dS_t \\
&= -\int_0^{+\infty} \big[\ln 1 - \ln(S_t\sigma\sqrt{2\pi}) - \frac{1}{2}(\frac{\ln S_t-\mu}{\sigma})^2\big] \frac{1}{S_t\sigma\sqrt{2\pi}} e^{\frac{-1}{2}(\frac{\ln S_t-\mu}{\sigma})^2} dS_t
\end{aligned}
$$

By changing the variable as

$z_t = \frac{\ln S_t-\mu}{\sigma} \Rightarrow dz_t = \frac{dS_t}{S_t\sigma}$ then

$$
H(S_t) = -\int_{-\infty}^{+\infty} \big[\ln 1 - (\ln(\sigma\sqrt{2\pi}) + z_t\sigma + \mu) - \frac{1}{2}z_t^2\big] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_t^2} dz_t
$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\frac{1}{2}\ln(2\pi\sigma^2)e^{-\frac{1}{2}z_t^2}dz_t + \frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{1}{2}z_t^2}dz_t + \sigma\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}z_t e^{-\frac{1}{2}z_t^2}dz_t$$

$$+\frac{1}{2}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}z_t^2 e^{-\frac{1}{2}z_t^2}dz_t$$

Regarding $\mu_z = 0$ and $\sigma_z = 1$, where μ_z and σ_z are the drift and volatility of the process after the change of the variable, we can generate the first and second moment generator function of z as

$$E[z] = \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}ze^{-\frac{1}{2}z^2}dz = \mu_z = 0$$

$$E[z^2] = \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}z^2e^{-\frac{1}{2}z^2}dz = \sigma_z^2 + \mu_z^2 = 1$$

Hence

$$\begin{aligned} H(S_t) &= \frac{1}{\sqrt{2\pi}}\ln(2\pi\sigma^2)\sqrt{2\pi\sigma^2} + \frac{\mu}{\sqrt{2\pi}}\sqrt{2\pi\sigma^2} + \frac{1}{2} \\ &= \mu + \frac{1}{2}(\ln(2\pi\sigma^2) + 1) \\ &= \mu + \frac{1}{2}\ln(2\pi e\sigma^2) \end{aligned} \quad (4.2.5)$$

4.2.2 Entropy Rate of "Rate of Change"

Entropy Rate of the "Rate of change"

Based on the definition of the entropy rate we have

$$\mathcal{H}(y_t) = \lim_{n\to\infty} \frac{1}{n} H(y_n, \cdots, y_1) \tag{4.2.6}$$

in which $H(y_n, \cdots, y_1)$ is joint entropy of y_t and is defined as following:

$$H(y_n, \cdots, y_1) = -\int \log P(y_n, \cdots, y_1))\; dP(y_n, \cdots, y_1)$$

Since y_t is driven by BM increments, so it has independent increments. Therefore $P(y_n, \cdots, y_1)) = \prod_{i=1}^{n} P(y_i)$ then the joint entropy for countable n can be stated as

$$\begin{aligned} H(y_n, \cdots, y_1) &= -E \log P(y_n, \cdots, y_1)) \\ &= -\sum_{i=1}^{n} P(y_i) \log P(y_i) = \sum_{i=1}^{n} H(y_i) \end{aligned}$$

Hence, the entropy rate is

$$\begin{aligned} \mathcal{H}(y_t) &= \frac{1}{n} H(y_n, \cdots, y_1) \\ &= \frac{1}{n} \sum_{i=1}^{n} H(y_i) \end{aligned}$$

and by replacing the $H(y_i)$ with (4.2.4) we have

$$\mathcal{H}(y_t) = \frac{1}{n} \sum_{i=1}^{n} \big(\frac{1}{2} + \frac{1}{2} \log 2\pi\sigma^2\big) \tag{4.2.7}$$

4.2.3 Stock Price as a Discrete Time Markov Model

In this section we try to discretize the price model as a finite-state, homogeneous, discrete-time Markov chain S_k, $k \in N$ and we suppose X_0 is given or its distribution is known. Suppose that the state space of S_t has N elements it can be identified as

$$S_t = \{e_1, \cdots, e_N\}$$

where e_i are unit vectors in $\mathbb{R}^N$ with unity as the i^{th} element and zeros elsewhere. Considering a matrix A as the transition matrix for the process which is defined as follows.

$$A = (a_{ji}) \in \mathbb{R}^{N\times N} \text{ and } a_{ji} = P\{S_{t_{k+1}} = e_j \mid S_{t_k} = e_i\}$$

Hence, the model can be identified as follows [5],

$$S_{t_{k+1}} = AS_{t_k} + V_{t_{k+1}}$$

The probability distribution satisfies $(P_{t_{k+1}} = A^{tr}P_{t_k})$[1]. Next, we compute the entropy rate. For a stationary process the entropy rate is given by

$$\mathcal{H}'(S_t) = \lim_{n\to\infty} H(S_{t_n}|S_{t_{n-1}}, \cdots, S_{t_0}) \tag{4.2.8}$$

If the process is also Markovian, then

$$\mathcal{H}'(S_t) = \lim_{n\to\infty} H(S_{t_n}|S_{t_{n-1}}) = H(S_{t_1}|S_{t_0})$$

According to the definition of conditional entropy, we know that,

$$H(S_1 = e_j|S_0 = e_i) = -\sum_j\sum_i P(S_1 = e_j, S_0 = e_i)\log P(S_1 = e_j|S_0 = e_i)$$
$$= -\sum_j\sum_i P(S_0 = e_i)P(S_1 = e_j|S_0 = e_i)\log P(S_1 = e_j|S_0 = e_i)$$
$$= \sum_i P(S_0 = e_i)(\sum_j a_{ji}\log a_{ji})$$

Since $P(S_{t_0} = e_i)$ is known a priori, then

$$\mathcal{H}'(S_t) = -\sum_{i=0}^{N}\sum_{j=0}^{N} \mu_i a_{ji}\log a_{ji} \quad , \quad \mu_i = P(S_{t_0} = e_i) \tag{4.2.9}$$

[1] A^{tr} denotes the transpose of matrix A

4.2.4 Relative Entropy of Stock Price

Any suggested model for a stock price is an approximation that implies the stock may follows the underlying process. We suggested that the price of a stock can be modeled by a log-normal equation which has a specific drift. It is of interest to compare the price of a stock with an uncertain rate of return to the price of a stock that grows according to the interest rates or a nominal rate of return. To have such a comparison, we need a measure of distance between two models.

Relative entropy is a one possible measure, therefore, the objective of this section is to use it to determine how a stock behaves relative to its growth if it were to follow the interest rate[2].

The relative entropy or Kullback Leibler distance between two probability densities $P(x)$ and $Q(x)$ is defined by

$$D(P||Q) = E^{p}[\log \frac{dQ}{dP}] \tag{4.2.10}$$

Randon-Nikodym derivative

Next, in order to compute the relative entropy, we need to introduce Randon-Nikodym derivative (RND) between two measures denoted by P and $\tilde{P}$, defined on the same measurable space $(\Omega, \mathcal{F})$. According to RND theorem, between two measures P, and $\tilde{P}$ which is absolutely continuous with respect to P, shown as $\tilde{P} \ll P$[3], the derivative $\frac{d\tilde{P}}{dP}$ exists [3].

$$\Phi = \frac{d\tilde{P}}{dP} \rightarrow \tilde{P}(E) = \int_E \Phi dP;\ \forall E \in \mathcal{F} \tag{4.2.11}$$

Theorem 4.2.1. *Cameron-Martin-Girsanov theorem*

Let ω_t be P-Brownian motion. Suppose γ_t satisfies the boundedness condition $E_P \exp(\frac{1}{2}\int_0^t \gamma_t^2 dt) < \infty$. Then there is a new measure $\tilde{P}$ equivalent to P, such that $\tilde{\omega}_t = \omega_t + \int_0^t \gamma_s ds$ is $\tilde{P}$-Brownian motion. The RND of $\tilde{P}$ with respect to P is

[2] Assuming that the diffusion(volatility) for both models remain the same

[3] Means $\forall E \in \mathcal{F},\ if\ P(E) = 0\ then\ \tilde{P}(E) = 0$

$$\frac{d\tilde{P}}{dP} = \exp\big(- \int_0^T \gamma_t d\omega_t - \frac{1}{2}\int_0^T {\gamma_t}^2 dt\big)$$

Change of measure - Brownian motion plus constant drift

Suppose $dS_t = \mu dt + \sigma d\omega_t$ is the drifting B.M. process where $\{\omega_t\}$ is Brownian motion under measure P and μ and σ are constant. Using the C-M-G theorem 4.2.1 with $\gamma_t = \frac{\mu_t}{\sigma_t}$, there exists a new measure, called $\tilde{P}$, so that on every interval $[0, T]$,

$$\frac{d\tilde{P}}{dP}\Big|_{\mathcal{F}_T} = e^{-\gamma_t \omega_T - \frac{1}{2}\gamma_t^2 T}$$

Under the measure $\tilde{P}$ the process $\tilde{\omega}_t = \omega_t + \int_0^T \gamma_s ds$ is the Brownian motion. Moreover, the expected value of a process S_t under $\tilde{P}$ is related to that under P as follows.

$$E^{\tilde{P}}(S_t) = E^P(\frac{d\tilde{P}}{dP} S_t)$$

Change of measure - Geometric Brownian motion

Applying the C-M-G theorem for a process that follows Geometric Brownian motion under a probability measure P, It can be shown that there exist a new measure $\tilde{P}$, which is absolutely continuous with respect to P, then the process under the new measure also follows Geometric Brownian motion.

$$\begin{cases} P : dS_t = S_t(\mu dt + \sigma d\omega_t) \\ \tilde{P} : dS_t = S_t(r dt + \sigma d\tilde{\omega}_t) \end{cases}$$

Where ω_t is a P-BM and $\tilde{\omega}_t = \omega_t + \int_0^T \gamma_s ds$ is a $\tilde{P}$-BM. And the new measure $\tilde{P}$, equivalent to P, can be defined through

$$\frac{d\tilde{P}}{dP} = e^{-\int_0^t \gamma_t d\omega_t - \frac{1}{2}\int_0^t {\gamma_t}^2 dt}$$

Provided the condition $E^P(e^{\frac{1}{2}\int_0^T \gamma^2 dt}) < \infty$ holds. And $\gamma_t = \frac{\mu - r}{\sigma}$ where r is an arbitrary constant [3].

Relative Entropy

As explained, relative entropy is a measure of distance between two probability densities or two distribution. and using the RND and G-M-C theorem it is shown how to calculate the derivative part of the relative entropy formula.

Consider a stock price S_t follows the SDE $dS_t = S_t(\mu dt + \sigma d\omega_t)$, where ω_t is a P^μ-BM. It is also shown that S_t can follow another SDE under a new probability measure denoted by $dS_t = S_t(rdt + \sigma d\tilde{\omega}_t)$, where $\tilde{\omega}_t$ is a P^r-BM. Suppose the P^r as a known density, we will state the relative entropy as follows.

$$D(P^r|P^\mu) = E^r[\log \frac{dP^\mu}{dP^r}]$$

Substituting the derivative in the above formula with what we shown in the previous section we have:

$$E^{P^r}\big(\log \frac{dP^r}{dP^\mu}\big) = E^{P^r}\big(\log e^{-\int_0^t \gamma_s d\omega_s - \frac{1}{2}\int_0^t \gamma_s^2 ds}\big)$$

Substitute ω_t with $\tilde{\omega}_t$ using formula $\tilde{\omega}_t = \omega_t + \int_0^T \gamma_s ds$ and $\gamma_t = \frac{\mu - r}{\sigma}$ given by C-M-G theorem, we have:

$$\begin{aligned} E^{P^r}\big(\log \frac{dP^r}{dP^\mu}\big) &= E^{P^r}\big(\log e^{-\int_0^t (\frac{\mu-r}{\sigma}) d\tilde{\omega}_s + \frac{1}{2}\int_0^t (\frac{\mu-r}{\sigma})^2 ds}\big) \\ &= E^{P^r}\big(-\int_0^t (\frac{\mu - r}{\sigma}) d\tilde{\omega}_s + \frac{1}{2}\int_0^t (\frac{\mu - r}{\sigma})^2 ds\big) \end{aligned}$$

$\tilde{\omega}$ is Brownian motion so $E^{P^r}\big(\int_0^t (\frac{\mu-r}{\sigma}) d\tilde{\omega}_s\big) = 0$, then

$$\begin{aligned} E^{P^r}\big(\log \frac{dP^r}{dP^\mu}\big) &= \frac{1}{2} E^{P^r}\big(\int_0^t (\frac{\mu - r}{\sigma})^2 ds\big) \\ &= \frac{1}{2}(\frac{\mu - r}{\sigma^2})^2 (t - 0) = \frac{(\mu - r)^2}{2\sigma^2} t \end{aligned} \tag{4.2.12}$$

We calculate the relative entropy, as in 4.2.12 [4] , now we need to justify the model in which we consider the probability density as known. We considered P^r as nominal and theoretically

[4] Although relative entropy is not a metric, in this special case it can be written as the Euclidean distance between μ and r.

r is an arbitrary constant but the problem is how interprete r to be able apply it in the real market. There are two suggestions for r, we may consider it as true interest rate or rate of return of the company's asset to be compared with the nominal rate of return. In the case that we considered r as interest rate, the SDE denoted by $dS_t = S_t r dt + S_t \sigma d\omega_t$ states that the first part $dS_t = S_t r dt$ is the Bond price, therefore, we try to formulate a stock price as the bond price plus a specific volatility modeled as Brownian motion. As the second case, we considered r as rate of return of the company that S_t represent the stock price of that company at time t. Rate of return which is function of time, can be considered constant at every interval $[0, T]$[5] and can be calculated according to sales, investments and so on.

[5] As we will discuss the issue in the next chapter, we will consider every hour as an interval to do estimations for μ and σ so T is constant for every interval and formula 4.2.12 can be writen as $E^{P^r}\left(\log \frac{dP^r}{dP^\mu}\right) = \frac{(\mu-r)^2}{2\sigma^2}T$

4.2.5 Drift as a Function of Time

Next, consider a price model given by

$$dS_t = S_t(\mu_t dt + \sigma d\omega_t)$$

in which μ_t changes according to a Markov model -this model called hidden Markov model (HMM)[6]-. Here μ_t takes value in $[r_1, \cdots, r_n]$ which are the states of a Markov chain.

We previously showed, using Ito integral, that

$$S_t = S_0(e^{\int_0^t (\mu_u - \frac{1}{2}\sigma^2)du + \sigma\omega_t})$$

Now, based on the assumption that μ_t follows HMM, we define $\rho_t = \mu_t - \frac{1}{2}\sigma^2$,therefore. ρ_t takes value in $[r_1 - \frac{1}{2}\sigma^2, \cdots, r_n - \frac{1}{2}\sigma^2]$

Defining $Y_t = \log S_t - \log S_0$ [8] then,

$$Y_t = \int_0^t (\mu_u - \frac{1}{2}\sigma^2)du + \sigma\omega_t$$

Let a process X_t be the state of the HMM which takes value in $[e_1, \cdots, e_n]$, where e_i is a unit vector with 1 in ith element and 0 elsewhere, then

$$Y_t = \int_0^t < X_u, \rho > du + \sigma\omega_t$$

where $< ., . >$ is the inner product.

Relative Entropy

The price of the stock is given by

$$S_t = S_0(e^{\int_0^t <X_u,\rho> du + \sigma\omega_t})$$

Using the C-M-G theorem for changing the measure, we have

$$\frac{dP^r}{dP^\mu} = e^{-\int_0^T \gamma_s d\omega_s - \frac{1}{2}\int_0^t \gamma_s^2 ds}$$

[6]Suppose $y(x) = f(g(x))$ where $g(x)$ follows a Markov process, since the states of the Markov process is not visible in y and hidden, therefore, $y(x)$ called a hidden Markov model, shown also by HMM [5]

where $\gamma_s = \frac{<X_s,\rho>-r}{\sigma}$, and it is already shown that according to C-M-G theorem that r is an arbitrary constant. Therefore,

$$D(P^r|P^\mu) = E^r[\log \frac{dP^r}{dP^\mu}]$$

then

$$\begin{aligned}
E^r\big(\log \frac{dP^r}{dP^\mu}\big) &= E^r\big(\log e^{-\int_0^t \gamma_s d\omega_s - \frac{1}{2}\int_0^t \gamma_s^2 ds}\big) \\
&= E^r\big(\log e^{-\int_0^t \gamma_s d\tilde{\omega}_s + \int_0^t \gamma^2 ds - \frac{1}{2}\int_0^t \gamma_s^2 ds}\big) \\
&= E^r\big(\log e^{-\int_0^t \gamma_s d\tilde{\omega}_s + \frac{1}{2}\int_0^t \gamma_s^2 ds}\big) \\
&= E^r\big(-\int_0^t \gamma_s d\tilde{\omega}_s + \frac{1}{2}\int_0^t \gamma^2 ds\big) \\
&= \frac{1}{2}E\big(\int_0^t (\frac{< X_s,\rho > -r}{\sigma})^2\big) ds \\
&= \frac{1}{2}\sum_{i=1}^{n} \int_0^t (\frac{< X_s,\rho > -r}{\sigma})^2 p_i(s) ds
\end{aligned} \tag{4.2.13}$$

where $p_i(t) = Prob(X(t) = e_i)$. Therefore, by estimating $p_i(t)$ and subsequently, the transition function of $X(t)$ we are able to calculate the relative entropy. Then we will apply it for the simulation section.

4.3 Information Measures for Multiple Stocks

Here, we will consider more than one stock option to determine the possible corelation between different stocks. We have elaborated our proposals based on an assumption that each stock's price is an individual independent process in previous sections. In many cases, this assumption does not give us a proper approximation of the behavior of a stock market. For example if we consider a stock option, like Microsoft, its movement is unaffected by the movement of other options. On the contrary, it is believed that a large movement on one stock is linked to the corresponding movement of other stocks. Specifically, if a stock option experiences a considerable movement, it is interpreted not only as being of the performance of the company but also as an effect of other companies in the market, mostly those that are in the same category as Microsoft. Therefore, it is suggested that the stocks are correlated. Consequently, a good model for several options should both describe each one individually and represent the interaction and dependency between them.

Multi-Dimension Brownian Motion

It is suggested that multiple stocks can be driven by multi-dimension Brownian motion. We introduce n-dimension Brownian motion, denoted by $B_t^1, \cdots, B_t^n$. B_t^i behaves as a one dimension Brownian motion.Therefore, the model for a single stock option can be stated as follows:

$$dS_t = S_t(\mu_t dt + \sum_i \sigma_t^i dB_t^i)$$

It is clear that the drift is not changed from the original definition (one factor), but the volatility is a vector. The variation of dS_t is $\sqrt{\sum_i (\sigma_t^i)^2}$, In other words, the total volatility of the process S_t is defined as $\sqrt{\sigma_1^2(t) + \cdots + \sigma_n^2(t)}$. Defining the volatility as a vector, we can interpret it as the interaction between different stock options. Let's consider the stock market that includes n different stocks. Therefore, we can conclude that the first source of variation for a specific option is itself, because of its own nature and that there are $n-1$ more sources of variation, considering the interaction of other options. Assuming the n-dimension BM we can state the following SDE for multiple stock options.

$$dS_t^i = S_t^i\ (\mu_t^i dt + \sum_{j=1}^{n} \sigma_{ij}(t) dB_t^j) \tag{4.3.14}$$

4.3.1 Joint Entropy of the "Rate of Change"

In the single option price chapter, the term 'rate of change of price' has been defined and it showed that it has Gaussian distribution. Since it is easier to work with Gaussian than log-normal, we define the "rate of change" for the multi-dimension Brownian motion to use it in this chapter as follows.

Let $dy_t \triangleq \frac{dS_t}{S_t}$ then,

$$dy_t = \mu_t dt + \sum_j \sigma_t^j dB_t^j$$

And for multiple stock option, we have,

$$dy_t^i = \mu_t^i dt + \sum_j \sigma_{ij}(t) dB_t^j$$

Hence, the joint entropy is defined as following :

$$H(dy_t^i) = -E\big[\log P(dy_t^1, \cdots, dy_t^n)\big]$$

It's already shown that dy_t^i follows Gaussian, therefore to be able to calculate the joint entropy we need calculate joint density function of dy_t^i.

Joint-Normal Distribution

The joint density function of a multi dimensional Gaussian distribution is called joint-normal distribution or multi-variate normal distribution and it is calculated as follows.

We denote the n-dimensional joint-normal distribution with mean vector $\mu = (\mu^1, \cdots, \mu^n)$ and the covariance matrix $\Sigma = E\{(dy_t - \mu)(dy_t - \mu)^{tr}\}$, the probability density function is

$$P_{y_t}(y_t) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\{\frac{-1}{2}(y_t - \mu)^{tr} \Sigma^{-1} (y_t - \mu)\}$$

Then

$$\begin{aligned}
H(dy_y^i) &= -E\big[\log P(dy)\big] \\
&= \frac{1}{2}E[(y_t-\mu)^{tr}(\Sigma^{-1})_{ij}(y_t-\mu)] + \frac{1}{2}\ln\big(2\pi|\Sigma|\big) \\
&= \frac{1}{2}E[\sum_{i,j}(dy_t^i-\mu^i)(\Sigma^{-1})_{ij}(dy_t^j-\mu^j)] + \frac{1}{2}\ln\big(2\pi|\Sigma|\big) \\
&= \frac{1}{2}\sum_{i,j}E[(dy_t^i-\mu^i)(dy_t^j-\mu^j)](\Sigma^{-1})_{ij} + \frac{1}{2}\ln\big(2\pi|\Sigma|\big) \\
&= \frac{1}{2}\sum_j\sum_i\Sigma_{ji}(\Sigma^{-1})_{ij} + \frac{1}{2}\ln\big(2\pi|\Sigma|\big) \\
&= \frac{1}{2}\sum_j I_{jj} + \frac{1}{2}\ln\big(2\pi|\Sigma|\big) + -
\end{aligned}$$

Considering that $\sum_{j=0}^{n} I_{jj} = n$, we have,

$$\begin{aligned}
H(dy_y^i) &= \frac{n}{2} + \frac{1}{2}\ln\big(2\pi|\Sigma|\big) \\
&= \frac{1}{2}\log\big(e^n 2\pi|\Sigma|\big)
\end{aligned} \tag{4.3.15}$$

Therefore, by calculating $|\Sigma|$, the determinant of the covariance matrix of the process, we can calculate the joint entropy of the "rate of change".

4.3.2 Relative Entropy of Multiple Stock model

As mentioned earlier, relative entropy is a measure of distance between two density functions. Since computing the density for multiple stock option is not an easy job, we will use relative entropy to find the distance between market density which is unknown and a known density such as interest rate. Hence, we will use C-M-G's theorem for n-dimensional geometric Brownian motion.

Theorem 4.3.1. *Cameron-Martin-Girsanov theorem n-factor*

Let $B_t = (B_t^1, \cdots, B_t^n)$ be an n-dimensional P-Brownian motion. Suppose $\gamma_t = (\gamma_t^1, \cdots, \gamma_t^n)$ satisfies the growth condition $E_P \exp(\frac{1}{2}\int_0^T \gamma_t^{tr}\gamma_t dt) < \infty$. Set $\tilde{B}_t^i = B_t^i + \int_0^t \gamma_s^i ds$. Then there *is a new measure $\tilde{P}$ equivalent to P, such $\tilde{B}_t \triangleq (\tilde{B}_t^1, \cdots, \tilde{B}_t^n)$ is n-dimensional $\tilde{P}$-Brownian motion. The RND of $\tilde{P}$ with respect to P is*

$$\frac{d\tilde{P}}{dP} = \exp\big(-\sum_{i=1}^{n}\int_0^T \gamma_t^i dB_t^i - \frac{1}{2}\int_0^T \gamma_t^{tr}\gamma_t dt\big)$$

Using the above theorem we are able to calculate relative entropy as follows. Let P^μ be the probability measure for the SDE (4.3.14). Consider another probability measure P^r which is absolutely continuous with respect to P^μ. Then

$$D(P^\mu|P^r) = E^{P^r}\log\frac{dP^r}{dP^\mu}$$

Let

$$\frac{dP^r}{dP^\mu} = \exp\big(-\sum_{j=1}^{n}\int_0^T \gamma_t^j dB_t^j - \frac{1}{2}\int_0^T \gamma_t^{tr}\gamma_t dt\big)$$

where $\gamma_t = (\sigma_t)^{-1}(\mu_t - r_t\mathbf{1})$, $\gamma_t = (\gamma_t^1, \cdots, \gamma_t^n)^{tr}$, $\mu_t = (\mu_t^1, \cdots, \mu_t^n)^{tr}$, σ_t is the matrix $(\sigma_{ij}(t))$, r_t is interest rate at time t and $\mathbf{1}$ is the constant vector $\mathbf{1} \triangleq (1, 1, \cdots, 1)$. then,

$$\begin{aligned} D(P^r|P^\mu) &= E^{P^r}\log\big(\exp\big(-\sum_{j=1}^{n}\int_0^T \gamma_t^j dB_t^j - \frac{1}{2}\int_0^T \gamma_t^{tr}\gamma_t dt\big)\big) \\ &= E^{P^r}\big(-\sum_{i=j}^{n}\int_0^T \gamma_t^j d\tilde{B}_t^j + \int_0^T \gamma_t^{tr}\gamma_t dt - \frac{1}{2}\int_0^T \gamma_t^{tr}\gamma_t dt\big) \\ &= -\sum_{j=1}^{n}\big(\int_0^T E^{\tilde{P}}(\gamma_t^j d\tilde{B}_t^j) + \frac{1}{2}E^{\tilde{P}}\big(\int_0^T \gamma_t^{tr}\gamma_t dt\big) \\ &= \frac{1}{2}E^{P^r}\big(\int_0^T \gamma_t^{tr}\gamma_t dt\big) = \frac{1}{2}\int_0^T E^{P^r}(\gamma_t^{tr}\gamma_t)dt \end{aligned}$$

In the single option chapter, it's stated that γ_t follows a HMM. Therefore, we are able to use the same logic to conclude

$$D(P^r|P^\mu) = \frac{1}{2}\sum_{j=1}^{n}\int_0^T (\gamma_t^{j^{tr}}\gamma_t^j)p_j(s)ds \qquad (4.3.16)$$

where $p_j(t) = Prob(S_t = e_j)$ is the probability of the states.

Chapter 5

Implementation and Conclusion

5.1 Introduction

In previous chapters we introduced some mathematical models, by which to model the behavior of return value of an asset like a stock. Later in chapter 4, we derived some measure of information that we believe they will give us more information than popular indices such as DIJA. In this part of our study, we need some tools to verify if our proposed models and measure of information are leading us to the objective of this thesis.

In this chapter, first we will explain how data was gathered and will explain how we calculated different measure of informations. Later, we will discuss the general idea how these measures can be interpreted. Finally, in case that we find our models or measures are not precise enough to get us to the objectives, we need to define the steps that it is needed to be taken, as future work, to correct the vulnerable points of the study.

5.2 Data Gathering

We previously explored the difference between macro and micro-movement models for a stock market. Assuming that stock value follows a GBM model, we also stated that there is a strong relation between micro and macro movement models.

The more inquiry I made to access to intradya (tick-by-tick) data, the less successful I was. Therefore, I coded a program which downloaded data from the Internet on every minute basis.

Therefore, for every hour we had a time-series sequence of 60 data[1]. Applying normalization, we assumed that we have complete data and calculated the measured based on hourly basis. It means, each point on every graph represents the information measure calculated for an hour period. As the information measures calculated based on two categories: a single stock; and multiple stocks (a family of stocks). In this chapter, we also classify our discussions into same categories.

5.3 Single Stock

In previous chapters we have introduced three information measures for a single stock namely: entropy; entropy rate; and relative entropy. We also calculated this measures based on stock price and "rate of change" of a stock price. To interpret the measure of informations calculated for a single stock, first we need to repeat what we discussed in chapter 2. It is slightly different from that chapter, since we did the discussion based on an empirical model we suggested.

Entropy of Price and Rate of Change - General Discussion

In this section we will discuss the entropy of price and rate of change, based on the definition of entropy. Later we will discuss if our proposed models and calculations supports our general idea. The interpretation of the entropy for stock price is similar to the one explained in chapter 2. However, it is slightly different for the "rate of change" that we will discuss it as follows.

When the entropy of a single stock at two consecutive periods is the same, we conclude that the probability of all "rate of changes" remains constant during the two periods. For example, if the probability of a rate of change of a stock price being 0.01 unit is $\frac{1}{3}$ during

[1] As explained, entropy of a finite sequence is upper bounded by $\log N$, where N is the number of sequence. In case that we have different number of data for two different period (hour), the calculated entropies for these two period is different, even though the uncertainty stays the same in the two period. This phenomenon is one of the problem may cause distortion in the result of the calculated measures. So, we need to have consistent data, which was not possible, since we were related to the Internet and also to a third party web site[15].

the first period, the probability remains the same during the second period.

There is only one exception to the above conclusion, when the probability of two or more rate of change of prices change among themselves without any change to the value of the probabilities. For example, suppose that probabilities for the rate of change of price being 0.01 and 0.011 unit, are $\frac{1}{3}$ and $\frac{1}{6}$ during the past period, respectively. If the probabilities switch so that we have $\frac{1}{3}$ probability for 0.011 and $\frac{1}{6}$ probability for 0.01 item, we will have the same entropy. By changing the scope of study to a micro, we are able to decrease such distortions. For example, if we consider every one hour as a period, we know that change of such probabilities are not likely to happen and this is exactly what we did in the simulations. When the entropy increases, it means uncertainty increases and this can be interpreted as a convergence of probabilities of different state of rate of changes [2]. Ultimately, probabilities converge to a number which is called the maximum entropy of a finite sequence, when the sequence is uniformly distributed. Therfore, we can conclude that when the uncertainty increases the stock option becomes more volatile, since the probabilities converge. The more volatile a rate of change of a stock price is, the more risk we face, if we intend to invest in that stock. Therefore, we may conclude that if the entropy graph of a specific stock shows an increase, we are reluctant to invest in that stock.

When the entropy decreases, we will conclude that the probability of a certain rate of changes increase. This case may be the result of two different occurences, when the price of a stock either increases or decreases. Therfore, the decrease in entropy, which translates into a decrease of uncertainty, shows that price of the stock keeps going with the same trend. Therefore, if the price increases we like to invest in the stock and in case of decrease of price we don't. In another words, uncertainty decreases when the expected value of a stock price continues decreasing or increasing.

[2] As entropy increases, uncertainty is increases either. Therfore for a investor with high risk tolerance this is good, since the probability of a drastic increase in price to happens is high.

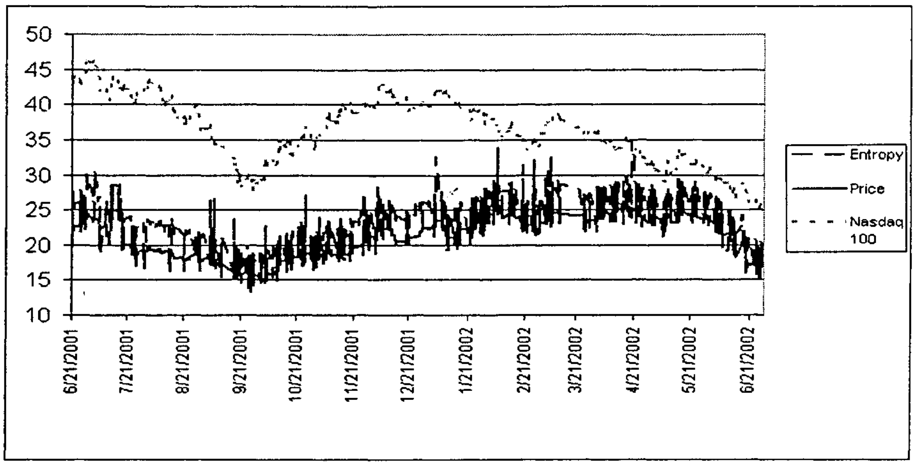

Figure 5.3.1: Entropy of Price for Apple - scaling for index is 1/40

Next, we need to support the above general conclusion for entropy of price and rate of change. Simulation is one popular tool that may help us to validate results of our study. For this reason we calculated the entropy for price and rate of change according to formula 4.2.5 and 4.2.4 respectively. Since, we consider every hour as the period of study, therefore, for each hour we have a number indicating the entropy of a specific stock (for example, Apple company) for the underlying period. Considering the price of that stock at the end of period -end of the hour- we have a one to one relation between entropy and stock price. Using these data we are able to draw a graph of entropy versus stock price. In figure 5.3.1, we consider stock price of Apple company and the calculated entropy and also the value of NASDAQ100. Since the index value is so large in comparison with stock price, we needed to rescale it, dividing it by 40, to show it in the graph. We can see that stock price for Apple company following the same trend as the index follows, and so does the entnropy. But it is obvious that entropy is so volatile in comparison with the price. It is not easy to conclude anything from this graph, because it shows a long period (one year)[3]. Therefore, we will show graphs with shorter duration. We will consider September and November of the year 2001. The reason of this choice is that the index shows decreasing in September and it shows

[3]We previously explained that we may consider a GBM model is stationary for very short term.

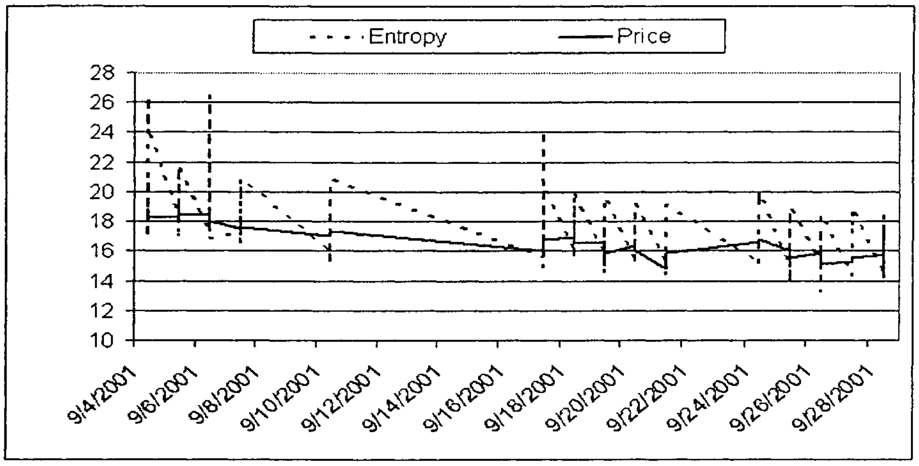

Figure 5.3.2: Entropy of Price for Apple for September 2001

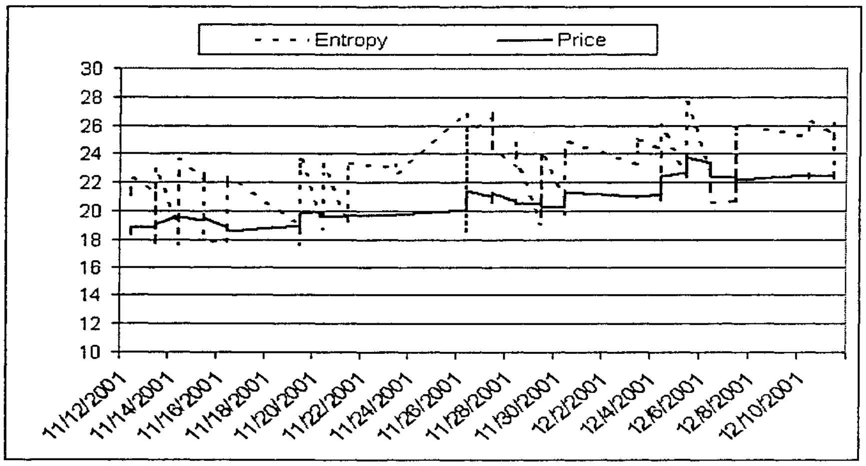

Figure 5.3.3: Entropy of Price for Apple for Mid-Nov to Mid-Dec

increasing in November. To see the graph in mote details, we will also choose one week in every period. As table 5.3.1 shows, for each day we have around 7 data, each one indicates for an hour, for example the first one is for 9:30 to 10:00 and the last one is for 15:00 to 16:00. To interpret the graphs we need to refer to an inconsistency in elements of the graph, where every entropy calculated with the data of the underlying period (an hour), the price is the last outcome of the market for the same period. In another word, the price is independent from the price fluctuation in that period, but they are all has been considered in calculating of the entropy.

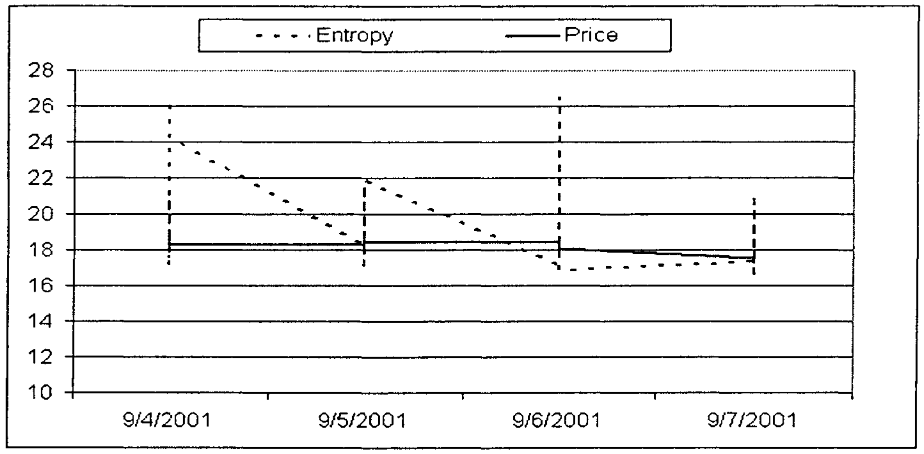

Figure 5.3.4: Entropy of Price for Apple for a week period in Sep-01

Although, entropy is much more volatile than the price, we are able to find some patterns. when ever entropy drops with a large amount, the price is not changing much, for example on Nov-14, entropy shows change for -4.5 unit and the price is only changing -0.009 dolor. On the contrary, in many cases when entropy rise with a large amount, the stock price changes considerebly. Therefore, when uncertainty decreases, as we expected, there sholdn't be a large change in the stock price and when the entropy increases we should expect large change in the price. For example , on Nov-14, entropy changes 6.2 unit and stock price shows 0.30 dolor. Finally, in many cases when the entropy is not changing considerably, price is not changing either. For example, on Nov-12 the 4th and 5th hour, where entropy changes 0.2 and 0.8 units prices changes 0.02 and 0.05. The graphs may arise a question why entropy is so close to price value. To answer this question, let's recall the formula 4.2.5, where entropy is calculated as following:

$$\mu + \frac{1}{2}\ln(2\pi e\sigma^2)$$

Clearly, the entropy is fluctuating around μ which is the sample mean of the stock price for every period (an hour).

Next, let's consider the graph for entropy of rate of change which is shown in figure 5.3.6. We previously discussed the entropy of rate of change in this section and also in previous

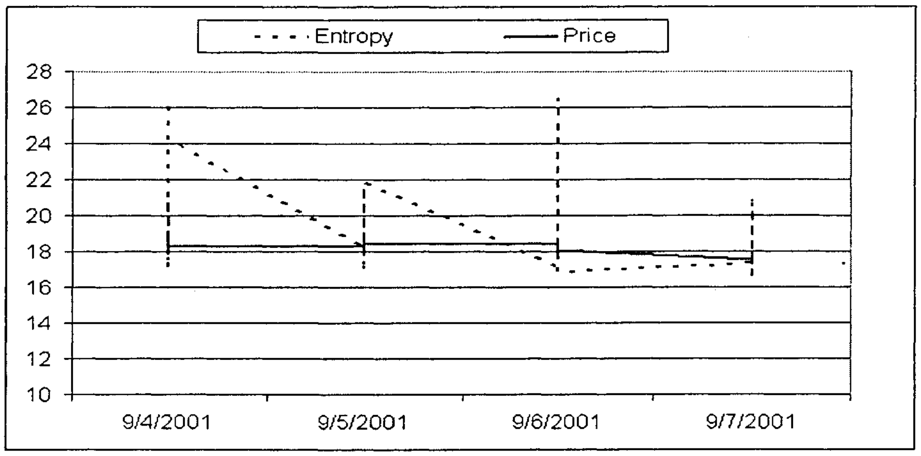

Figure 5.3.5: Entropy of Price for Apple for a week period in Nov-01

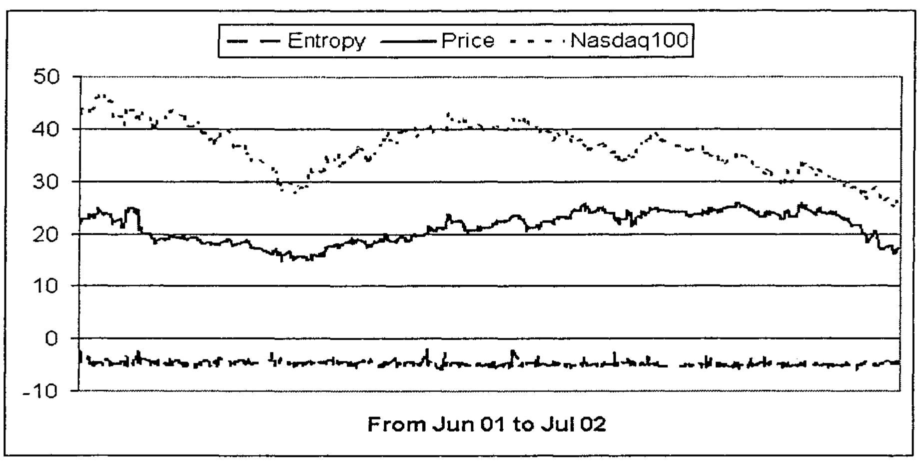

Figure 5.3.6: Entropy of Rate of Change for Apple - from Sep-01 to Jun-02

chapters. The interpretation of entropy of rate of change is similar to the entropy of price. Considering the formula for entropy of rate of change and entropy of price, it clears that the main difference is the drift for entropy of price.

Entropy Rate

The second information measure that we calculated for a single stock is entropy rate of "rate of change". To interprete the entropy rate for the rate of change for a single stock, we will recall the definition of entropy rate . Entropy rate of a stochastic process is a time average of the conditional entropies of that process and it simply explains the change of the entropy,

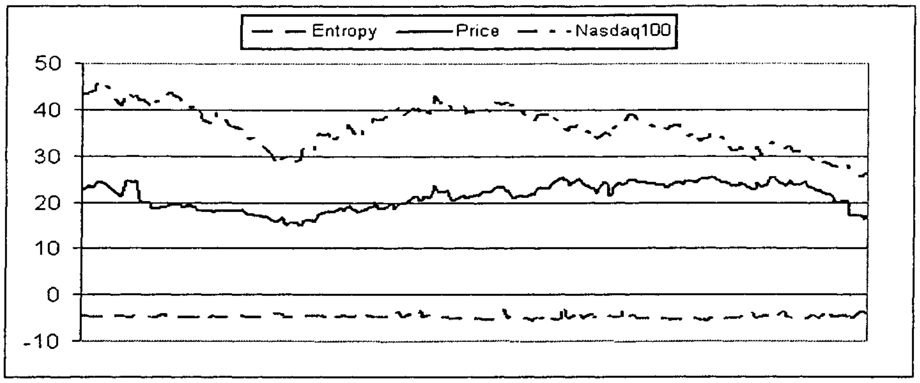

Figure 5.3.7: Entropy Rate of Rate of Change for Apple - from Jun-01 to Jul-02

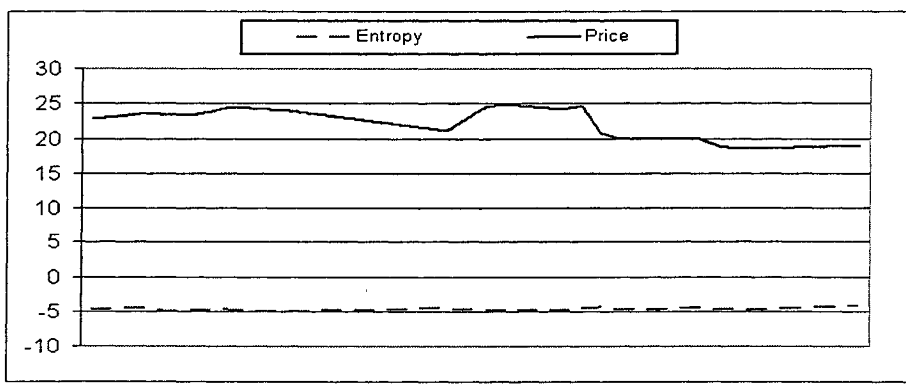

Figure 5.3.8: Entropy Rate of Rate of Change for Apple - from Jun-01 to Jul-02

uncertainty, by increasing the number of sequence of that process. In this case we took an average of entropy of rate of change and compared it with the closing price. Entropy rate shows how the entropy changes by time, and as shown in figure 5.3.7 and figure 5.3.9, it is less volatile than the entropy of rate of change. As we calculated in estimation section 3.6, the drift and volatility of the model we proposed(GBM) is simply sample mean and variance, therefore the average of entropies of one day, according to hourly calculations for entropy, is the same as the entropy of the same day by considering all data in the day. In another word, the entropy rate of rate of change for a specific day equals the entropy of

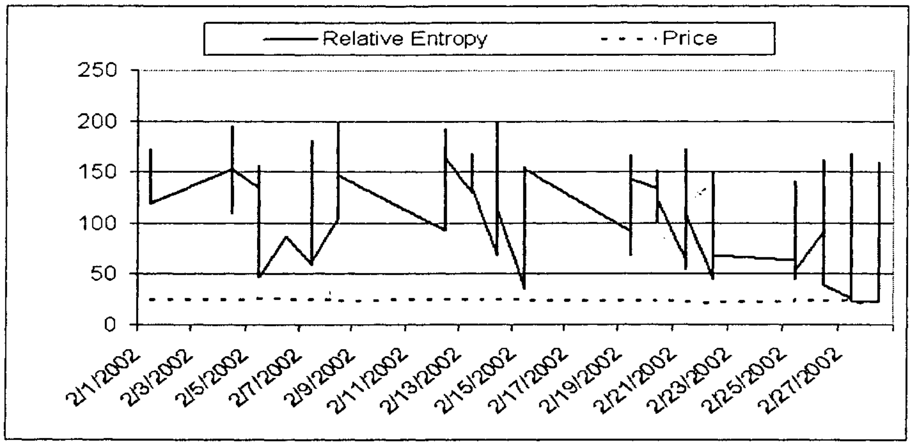

Figure 5.3.9: Relative Entropy of rate of change for Apple - from Jun-01 to Jul-02

rate of change for the same day. Consequently, we may interpret the entropy rate of rate of change as we did for the entropy of the rate of change. In case that we have access to the data according to every transaction (tick-by-tick), it is a very good tool to be calculated in a short period basis, for example, every 10-minute basis. Therefore, by having such a figure, we can interpret that by passing the time the rate of change of the price moves in a more certain way or not, and in case of it moves in a more certain way, we may like to invest on it if the price shows an increase.

Relative Entropy

Finally, let's recall the definition of relative entropy, where relative entropy is a measure of distance between two distributions, and it is applied to find the distance of an unknown distribution to a known one. In this thesis we consider interest rate as the known density to calculate the relative entropy of a stock price. The higher the relative entropy the greater is the distance between a stock price and interest rate. Since the interest rate mainly has a uniform distribution in a short period of time, the greater distance between a process with unknown distribution and a uniformly distributed process the more unpredictable the process is. When the relative entropy is constant or it is very close to the previous state at two consecutive periods, we may conclude that the volatility of the stock price doesn't change

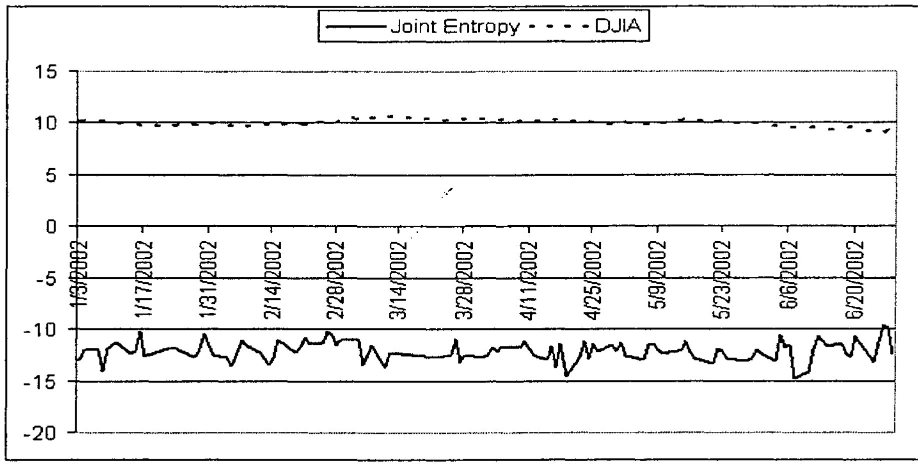

Figure 5.4.10: Joint Entropy of rate of change for Multiple Stock

during the two period. When the relative entropy shows decrease, it means that the distribution of the stock price becomes closer to a uniformly distributed process. Consequently, we can deduce that the stock price becomes more volatile so it takes more risk to invest on it. On the contrary, when the relative entropy increases, we can deduce that the stock price is less volatile.

5.4 Multiple Stock Options

We have also introduced two measures for multiple stock options; the relative entropy and the joint entropy. In calculation of relative entropy, we need to estimate the transition matrix of the process as shown in (4.3.16). The next measure was the joint entropy of the "rate of change", which was calculated as in (4.3.15). Figure 5.4.10 shows joint entropy for four companies such as Alcatel, Apple Intel and 3M.

The interpretation of the joint entropy for multiple stock option is similar to what we elaborated (2.4.2). Unlike the one we had, instead of the predefined probability measure in that chapter, we applied the joint density according the assumption that stock has a log-normal distribution. Therfore, we can state the conclusion as follows.

When the joint entropy remains unchanged, the probability of items included in calculating the entropy remains unchanged. Unlike the single option, it cannot be interpreted as a change to the expected value of each option price. When the probability of each individual stock's price in a joint density remains the same, it means all the items have the same direction in their movement. Therefore, there are two directions for a portfolio in such a case: firstly, when the rate of change of each individual remains the same and secondly, when the rate of change of all items changes with the same ratio.

When the joint entropy increases, knowing the fact that the maximum entropy happens in a uniformly distributed sequence, we can conclude that when the entropy increases, the probability of each sequence item becomes closer to $\frac{1}{n}$, where n is the number of options included in a portfolio. In this case, the "rate of change" of different stocks are similar.

When the joint entropy decreases, in which the entropy gets closer to the minimum state, we can conclude that there are few stock prices which had noticeable changes, compared to the previous period of study. In another word, there are more options with the rate of change close to zero.

5.5 Future Work

As we previously mentioned, there are some flaws that do not let us to match the interpretation we made based on the theory with what graphs showed us. The first flaw was that the entropy is not consistent with the price that we applied in graphs. The second one was the entropy, which is so fluctuated, that makes it difficult to decide what is happening and what will happen in next session. Hence, to be able to decide what is the trend of stock market, we either need some statistical study on the calculated entropy or we need to redefine our models or measures of information. We previously stated that the assumption of drift and volatility being constant, makes the two flaws called as empirical phenomena. Therefore, I suggest for next step we may conduct our study based on the model that was introduced in 3.3.2. We also calculated measure of information for this model in 4.2.13. Therefore, the study will focus on how we need to applying filtering for estimation of the parameters of model. Doing so, we are ready to step forward and calculate the measure of information according to the new model.

Table 5.3.1: Data used in figures 5.3.4 and 5.3.5

Date	Entropy	Price	Date	Entropy	Price
12-Nov-01	21.02063322	18.29	4-Sep-01	17.12705903	18.5
12-Nov-01	21.04068953	18.402	4-Sep-01	21.30260636	18.389
12-Nov-01	21.67182637	18.6	4-Sep-01	17.56637237	18.561
12-Nov-01	21.30472817	18.881	4-Sep-01	17.69989973	18.721
12-Nov-01	21.49249532	18.86	4-Sep-01	17.36721828	18.83
12-Nov-01	22.37565116	18.81	4-Sep-01	26.35669954	18.74
13-Nov-01	21.28756984	18.79	4-Sep-01	24.20959325	18.3
13-Nov-01	21.37150719	19.04	5-Sep-01	18.29943623	18.25
13-Nov-01	21.95866483	19.01	5-Sep-01	17.79072452	18.9
13-Nov-01	17.49316075	19.15	5-Sep-01	17.03894366	18.59
13-Nov-01	21.57927977	19.14	5-Sep-01	21.6084907	18.569
13-Nov-01	23.13240673	19.11	5-Sep-01	17.25765176	18.49
14-Nov-01	18.70564246	19.64	5-Sep-01	17.44409388	18.37
14-Nov-01	18.07398071	19.56	5-Sep-01	21.84107565	18.42
14-Nov-01	22.64690385	19.5	6-Sep-01	17.10436704	18.481
14-Nov-01	22.04057618	19.28	6-Sep-01	26.23932791	18.47
14-Nov-01	17.57931051	19.289	6-Sep-01	26.54997607	18.22
14-Nov-01	23.71324943	19.59	6-Sep-01	16.87063978	18.05
15-Nov-01	22.06118343	19.291	7-Sep-01	17.33960439	17.52
15-Nov-01	22.63187879	19.811	7-Sep-01	17.05982474	17.92
15-Nov-01	22.00355957	19.6	7-Sep-01	20.64591394	17.73
15-Nov-01	22.68735756	19.74	7-Sep-01	19.93442359	17.72
15-Nov-01	22.42371546	19.49	7-Sep-01	19.68005176	17.488
15-Nov-01	18.0896911	19.531	7-Sep-01	16.41675283	17.29
16-Nov-01	17.63721808	18.86	7-Sep-01	20.8863944	17.59
16-Nov-01	21.64731435	18.919			
16-Nov-01	22.32007089	18.951			
16-Nov-01	21.82753784	18.88			
16-Nov-01	21.48784386	18.61			
16-Nov-01	22.45390082	18.58			

Bibliography

[1] J.Kelly, "A new interpretation of information rate" Bell Sys. Tech. Inform. Theory 35:917-926, 1956.

[2] Thomas M.Cover and Joy A. Thomas,"*Elements of Information Theory*", Willy interscience publications, 1991.

[3] Martin Baxter and Andrew Rennie, "*Financial Calculus, An introduction to derivative pricing*", Cambridge University Press, Second Edition 2000, UK.

[4] F Black and M Scholes, "The Pricing of Options and Corporate Liabilities", Journal of political Economy, Vol. 81, 1973 pp. 637-659.

[5] Robert J. Elliott, Lakhdar Aggoun and John B. More, "*Hidden Markov Models, Estimation and Control*",, Springer-Verlog, 1995.

[6] Kenneth S. Miller , "*Multidimentional Gaussian Distribution*", The SLAM Series in Applied Mathematics ,John Wiley and Sons, Inc., 1964.

[7] Marvin K. Simon, "*Probability Distributions Involving Gaussian Random Variables* ",, KLUWER ACADEMIC PUBLISHERS, 2002.

[8] Robert J.Elliott and Raymond W. Rishel, "Estimating the implicit interest rate of a risky asset", Stochastic Processes and their Applications, Vol. 49, pp. 199-206.

[9] Yong Zeng, "A Partially-Observed Model for Micro-movement of Asset Prices with Bayes Estimation via Filtering",Mathematical Finance, Vol 13, 2003 pp. 411-444.

[10] Steven Kou, "A Jump Diffusion Model for Option Pricing with Three Properties: Leptokurtic Feature, Volatility Smile, and Analytical Tractability" , Management Science, Vol. 48, No. 8, August 2002 pp. 1086-1101.

[11] Ping Chen, "A Random-Walk or Color-Chaos on the Stock Market? Time-Frequency Analysis of S&P Indexes", Studies in Nonlinear Dynamics & Econometrics, Vol. 1, issue 2, 1996, pp. 89-103.

[12] Roman Lawrence, "Using Neural Networks to Forecast Stock Market Prices" , University of Manitoba Press, 1997.

[13] The official site of DOW Jones and Company, "*http://www.dowjones.com*"

[14] The official site of NASDAQ , "*http://dynamic.nasdaq.com/reference/IndexDescriptions.stm*"

[15] The official site of Globe and Mail , "*http://investdb.theglobeandmail.com*"