Degree in Mathematics

Title: Commodity Valoration and its Derivatives
Author: Bru Martinell Chicano
Advisor: Josep J. Masdemont Soler
Department: Matemàtica Aplicada I
Academic year: 2015/16
Commodity Valoration and its Derivatives

Author: Bru Martinell Chicano

Advisor: Josep J. Masdemont Soler

Grau en Matemàtiques
Facultat de Matemàtiques i Estadística

January 2016
# Contents

1  Fundamentals of Probability Theory and Binomial Tree Model ........ 1  
   1.1  Background on Probability Theory ................................. 1  
   1.2  The binomial tree ............................................. 5  

2  Stochastic Modeling of Commodity Price Processes ..................... 9  
   2.1  Geometric Brownian motion .................................... 9  
   2.2  Mean-reversion .................................................. 16  
   2.3  Seasonality, jumps and stochastic volatility ..................... 21  

3  Plain Vanilla Options on Commodity Spot and Forward Prices .......... 31  
   3.1  Forward pricing .................................................. 32  
   3.2  Martingales ...................................................... 38  
   3.3  Neutral risk valuation ........................................... 42  
   3.4  Derivatives on geometric Brownian motion model ................. 46  
   3.5  Derivatives on mean-reversion model ............................. 57  

4  Calibrating the Models. A Real Approach to the Markets ............... 71  
   4.1  Maximum-likelihood estimation .................................. 71  
   4.2  Inference on the models ....................................... 74  

Bibliography ........................................................................ 83
INTRODUCTION

Financial mathematics is a subject of mathematics that models the behaviour of stock’s movements and set a “fair price” for the related products.

The stock movement has been largely studied and discussed. The work on it was accelerated due the apparition of the Black-Scholes formulae in 1973 (although the Nobel Prize was given in 1997) which allowed the estimation of the European call premium. Other models have been studied, allowing stochastic volatility or variable interest rate.

When talking about commodity modelling, the field is less explored. The commodity market is much more volatile that the stock market. One important difference is that it is a “real” market, meaning that there must be a place to store the commodity, transportation costs, maintenance costs and taking in account that the commodity can be used either as an invest asset or as a consumption asset.

Although the main difference is due the model used. It does not make much sense to think that the value of a tone of wheat can increase indefinitely as a Telefonica share. The related costs to the production can increase in the same way as the inflation but not as a firm.

The models used in commodities are mean-reversion models introduced by Vasicek in 1977. Actually, this model was thought to describe the evolution of interest rates.

With the models in hand, to predict or estimate the price is not a rough task. The pricing of derivatives has been always a delicate work. Due the nature of the market, a “bad” or “unfair” price will lead to loss either is the product overpriced or underpriced.

A closed form for pricing derivatives and other related products (such as the Black-Scholes formulae) is a rare and valuable thing. In general, when is not possible to reach the analytical solution, other methods are used such as
the Montecarlo method or the discretisation of the time that will be seen.

The classical approach to the financial world is through partial differential equations. In this thesis, it will be tackled by the stochastic calculus.

We present here an introduction to the commodity market and how are priced its relative derivatives. This thesis consists of 4 chapters, which are summarized below.

**Chapter 1** This chapter is a brief introduction to the background of Probability Theory that will be used along the thesis and to the binomial tree model. We will see basic definitions and properties of the measure theory and how it works when introducing the stochastic differential equations (SDEs), the Radon-Nikodym derivative and the Riemann-Stieltjes integration method for random variables are essentials for this work. The binomial tree model is introduced as the simplest model for any kind of asset. It is important to revise this model since in Chapter 3 will be introduced the trinomial tree model which is much more complicated.

**Chapter 2** We give here basic notes of the stochastic modelling and its main models. Here is where the stochastic calculus takes an important part. The classical models, as the geometric Brownian motion, are revisited and the more specific commodity models are exposed and studied in a rigorous way.

**Chapter 3** Once the basis are set, in this chapter there are calculated the derivatives prices related to each model. In the case of geometric Brownian motion modelled assets, the approach of Black, Scholes and Merton will be used. When treating the mean-reverting modelled assets it will be shown that it is not as easy as before and two numerical methods (the trinomial tree model and the Montecarlo method) will be presented.

**Chapter 4** This last chapter contains an approach to the real world. A real market application of how the data is treated and how to calibrate the models that have been introduced. Here the maximum-likelihood estimation will guide the way.

Powerful results on probability theory and calculus are applied in financial mathematics. To go deep in the financial modelling and pricing, some proofs on theoretical results have been skipped. All this proofs can be found in any graduate course on Probability Theory.
1

Fundamentals of Probability Theory and Binomial Tree Model

This chapter goes over the main results to deal with financial mathematics. The essential properties of the conditional expectation, the Radon-Nikodym derivative, the Riemann-Stieltjes integration and the binomial tree method are the principal concepts. Some concepts of this chapter can be found in any graduate introductory probability course or in any financial mathematics course, and so some basic proof are skipped in this section. Most concepts, results and proofs from this chapter can be found in [1], [2].

1.1 Background on Probability Theory

Definition 1.1.1. Given a probability space \((\Omega, \mathcal{A}, P)\), a filtration is a partitions sequence \(\mathcal{F} = \{\mathcal{F}_i\}_{0 \leq i \leq N}\) such that

(i) \(\mathcal{F}_0 = \Omega\),

(ii) \(\mathcal{F}_N = \{\{\omega\} : \omega \in \Omega\}\),

(iii) for all \(i < j\) then \(\sigma(\mathcal{F}_i) \subseteq \sigma(\mathcal{F}_j)\).

With \(\sigma(\mathcal{F}_i)\) is the \(\sigma\)-algebra generated by \(\mathcal{F}_i\).

Definition 1.1.2. Given a probability space \((\Omega, \mathcal{A}, P)\), a random process (or a stochastic process) \(X\) is a collection of \(\mathbb{R}\)-valued random variables \(\{X_t\}_{t \in T}\),

where \(T\) is a totally ordered set.
In general, the set $T$ represents the lapse of time within $X$ is being observed.

**Definition 1.1.3.** Let $\mathcal{F} = \{U_i\}$ be a partition of $\Omega$, then, the process $X$ is $\mathcal{F}$-measurable if it takes constant values on every $U_i$.

**Definition 1.1.4.** Let $\mathcal{F}$ be a filtration, the process $X$ is adapted to $\mathcal{F}_i$ if it is $\mathcal{F}_i$-measurable.

**Definition 1.1.5.** Given a probability space $(\Omega, \mathcal{A}, P)$, the conditional expectation of $X$ on $\mathcal{F}$, with $\mathcal{F}$ a partition of $\Omega$, is the function

$$E(X|A) : \Omega \rightarrow \mathbb{R}$$

$$\omega \rightarrow E(X|\mathcal{F})(\omega) = \frac{\sum_{\alpha \in \mathcal{F}(\omega)} P(\alpha)X(\alpha)}{\sum_{\alpha \in \mathcal{F}(\omega)} P(\alpha)}$$

where $\mathcal{F}(\omega)$ is the unique set in $\mathcal{F}$ within $\omega$ belongs.

One important property on the conditional expectation is the Tower Law. The Tower Law states how to treat a conditional conditional expectation.

**Proposition 1.1.6 (Tower Law).** If $\mathcal{F}$ is a filtration and $X$ a random process, then for all $0 \leq i \leq j \leq T$,

$$E\left(E(X|\mathcal{F}_j)|\mathcal{F}_i\right) = E(X|\mathcal{F}_j).$$

**Definition 1.1.7.** The process $X$ is a Markov chain if satisfies the Markov condition

$$P(X_n = k|X_0 = x_0, X_1 = x_1, ..., X_{n-1} = x_{n-1}) = P(X_n = k| X_{n-1} = x_{n-1})$$

for all $n \in T$ and all $k, x_0, x_1, ..., x_{n-1} \in \mathbb{R}$.

**Proposition 1.1.8.** The Markov property is equivalent to:

(i) $P(X_{n+1} = k|X_{n1} = x_{n1}, X_{n2} = x_{n2}, ..., X_{nk} = x_{nk}) = P(X_{n+1} = k| X_{nk} = x_{nk})$ for all $n_1 < n_2 < ... < n_k \leq n$ all in $T$.

(ii) $P(X_{m+n} = k|X_0 = x_0, X_1 = x_1, ..., X_m = x_m) = P(X_{m+n} = k| X_m = x_m)$ for all $m,n$ and $m+n$ in $T$.

**Note.** There is an alternative notation for processes that is used equally. The notation is given by

$$X_t = X(t).$$

This notation is specially useful when the process is in a sequence indexed by $n$. In this case the process is written as $X_n(t)$ meaning index $n$ and time $t$. 
One of the most important random process that will base the models on is the random walk.

**Definition 1.1.9.** For $n$ positive integer, define the random walk process $W_n(t)$ satisfying

(i) $W_n(0) = 0,$

(ii) every tic of time is separated $\frac{1}{n},$

(iii) the jumps up and down are equal and of size $\frac{1}{\sqrt{n}},$

(iv) the probabilities of up and down at any time are always $\frac{1}{2}.$

**Proposition 1.1.10.** For any $n \geq 0$ the random walk $W_n(t)$ has the Markov property.

This kind of movement will drive all the random processes. For the continuous case will be taken $n$ tending to the infinity.

The filtrations are interpreted as structures of information, what is known and what is still unknown. The way this information is interpreted will be shown to be the probability measure under which is being treated the filtration. In financial world the information and the way is treated is a real sensible field.
Definition 1.1.11. Given \((\Omega, \mathcal{F})\) a measurable space and \(P, Q\) two measures. Then \(P\) is absolutely continuous with respect \(Q\) if
\[Q(A) = 0 \Rightarrow P(A) = 0,\]
for all \(A \in \Omega\).

Definition 1.1.12. Given \((\Omega, \mathcal{F})\) a measurable space and \(P, Q\) two measures. Then \(P\) is equivalent to \(Q\) if each is absolutely continuous with respect the other.

Theorem 1.1.13. Given \((\Omega, \mathcal{A})\) and \(P, Q\) two measures. If \(Q\) is absolutely continuous with respect \(P\), then exists a random probability \(dQ/dP : \Omega \to \mathbb{R}\) such that, for all \(A \in \mathcal{A}\),
\[q(A) = \int_A \frac{dQ}{dP} dP.\]
Moreover, for all random variable \(X : \Omega \to \mathbb{R}\) and all \(A \in \mathcal{A}\),
\[\int_A X dQ = \int_A X \frac{dQ}{dP} dP.\]

En particular, \(E_Q(X) = E_P(X \frac{dQ}{dP})\). Define \(\frac{dQ}{dP}\) as the Radon-Nikodym derivative of \(Q\) with respect \(P\) or the density function of \(Q\) with respect \(P\).

Corollary 1.1.14. If \(P\) and \(Q\) are equivalent measures, then
\[\frac{dQ}{dP} = \left(\frac{dP}{dQ}\right)^{-1}.\]

This result is a tool for changing the probability measure under which is evaluated a process. Then the Radon-Nikodym is a way to change de “point of view” under which a process is being watched.

The last important concept is the Riemann-Stieltjes integral. The Riemann-Stieltjes integral is a new integration method, resembling the Riemann integral, for random variables.

Definition 1.1.15. Given a deterministic function \(f : \mathbb{R} \to \mathbb{R}\), a stochastic process \(X\) and two real numbers \(a < b\), then define
\[\int_a^b f(t)dX_t = \lim_{n \to \infty} \sum_{i=1}^{n} f(t_{i-1})(X_{t_i} - X_{t_{i-1}}),\]
where \(a = t_0 < t_1 < ... < t_{n-1} < t_n = b\) is a partition of the interval \([a, b]\).

Note. Notice that the Riemann-Stieltjes definition make sense only for continuous stochastic processes.
With this new integration method, the integrals where a stochastic variable takes part can be interpreted as a new random variable. The main modelling of prices (the continuous case) is done through stochastic differential equations (SDE). The SDE, usually, can be separated into two parts

$$dX_t = d\left(f(X_t, t)\right) + g(X_t, t)dW_t,$$  \hspace{1cm} (1.2)

where $X_t$ is the process modelled, $f(X_t, t)$ and $g(X_t, t)$ are deterministic functions (with no random terms) and $W_t$ a random variable. The integral formulation of an SDE is given by

$$\int_t^{t+s} dX_t = \int_t^{t+s} f(X_u, u)du + \int_t^{t+s} g(X_u, u)dW_u.$$  \hspace{1cm} (1.3)

Notice that

$$\int_t^{t+s} dX_t = X_{t+s} - X_t.$$

### 1.2 The binomial tree

The binomial tree is the most simple model (this do not mean useless) used in modelling stock movements. It can be used to price derivatives. It consists in discretise the time into intervals, then at any point of the tree the stock price can go up or down. Usually, to make easier the implementation, the stock price is the same when taking different paths but with the same number of steps up and steps down. This kind of tree is known as a recombined tree.

Then, the main variables participating in the binomial tree model are:

- $S$: The stock price, which can be evaluated at any node of the tree.
- $r$: The free-risk interest rate. This is the rate at which the long or the short position can request or invest money at any time.
- $\Delta t$: The time distance between any tic of time.
- $u$: Factor which multiplies the stock price when the stock price goes up.
- $d$: Factor which multiplies the stock price when the stock price goes down.

Then, if the stock price at initial time is $S_0$, then at time $t = \Delta t$ the stock price would be either $S_0u$ or $S_0d$. And if instead of buying the stock at initial time this money would have been invested, at time $t = \Delta t$ the investor would have $S_0e^{r\Delta t}$. 
The principal assumption that is made in the binomial tree is that it must be a fair play, in the sense of investing in the tree at any moment would have the same expected gain as investing the same amount of money at the free-risk interest rate $r$. The implementation of this fact is done through changing the stock probability of going up or going down. This new probability measure under which will be seen the tree is the neutral-risk probability $q$. Notice that

(i) under this probability the risk is balanced meaning that the risk assumed is exactly the one the operation has implicit,

(ii) the initial probability the tree had is totally ignored.

With this, can be found the probability of any branch of the tree. To calculate the price of a concrete derivative only must be calculated the expected payoff at every node and the expected payoff at the initial node will be the derivative price.

**Example 1.2.1.** The price of an stock at initial time is 100$, the expected fluctuations of the price per month is $\pm 10\%$, the free-risk interest rate is $r = 0.1$, the probability of the stock to go up is $p_u = \frac{3}{4}$ and the probability of the
stock to go down is \( p_d = \frac{1}{4} \). Then, calculate the prime of a European call at two months with strike price \( X = 95 \$ \).

The initial tree is given by

![Binomial tree diagram]

Figure 1.3: Binomial tree

At any node if the value of the stock is \( S \), then the value if going up is \( Su \) and the value if going down is \( Sd \). Then the neutral-risk probability \( q \) is the one such

\[
Suq + Sd(1 - q) = Se^{r\Delta t},
\]

\[
q = \frac{Se^{r\Delta t} - Sd}{Su - Sd}.
\]

In this particular case,

\[
u = 1.1
\]
\[
d = 0.9
\]
\[
r = 0.1
\]
\[
\Delta t = \frac{1}{12}
\]
\[
q = 0.5418
\]

The payoff of the European call is given by Table 1.1. Since if the market
price is lower than the strike price then the European call will not be exercised, else if the market price is higher than the strike price then the European call will be exercised. Then the call price at every node is calculated through the assumption of fair play.

\[
\begin{array}{c|cccc}
\text{Node} & D & E & F \\
\hline
\text{Call value ($)} & 26 & 4 & 0 \\
\end{array}
\]

Table 1.1: Call payoff

Finally, the initial call price is \( c_0 = 9.460377 \).
2

STOCHASTIC MODELING OF COMMODITY PRICE PROCESSES

This chapter shows the basic stock and interest rates models of price processes and how are adapted to the commodity price modelling. The main models are the geometric Brownian motion (which it is used in the Black-Scholes model) and the mean-reversion model. It will be shown how through the stochastic volatility and jumps in price trajectories is possible to adapt this models to the commodity models.

2.1 Geometric Brownian motion

The Brownian motion is the basis of the most important model in finance. The Brownian motion must be understood as a continuous random walk (Definition 1.1.9). That particular random walk (with \( n \to \infty \)) is what is known as a Brownian motion.

Definition 2.1.1. The process \( W = (W_t \mid t \geq 0) \) is a \( P \)-Brownian motion if and only if

(i) \( W_t \) is continuous, with \( W_0 = 0 \),

(ii) the value of \( W_t \) is distributed, under \( P \), as a normal random variable \( N(0, t) \),

(iii) the increment \( W_{s+t} - W_s \) is distributed as a normal \( N(0, t) \), under \( P \), and is independent of \( \mathcal{F}_s \), the history of what the process did up to time \( s \),

where \( P \) is a probability measure.
Remark. The third statement of the Brownian Motion definition it is the named previously Markov property.

There are few properties of the Brownian motion that deserve to be named. These properties will help to take care about Brownian motion’s random nature.

Property 2.1.2. If $W$ is a $P$–Brownian motion, then

(i) although $W_t$ is continue it is no differentiable at any point (with probability 1),

(ii) the Brownian motion $W$ will take any real value and will return to the 0 at some time,

(iii) when $W$ touches one value, immediately touches it an infinity of times, and eventually in the future,

(iv) at any scale $W$ will look similar, it has fractal nature.
Proposition 2.1.3. If $W_t$ and $\tilde{W}_t$ are two independent Brownian motions and $\rho$ is a constant between $-1$ and $1$, then the process $X_t = \rho W_t + \sqrt{1-\rho^2} \tilde{W}_t$ is a Brownian motion.

Proof. It is obvious that $X_t$ is a continuous function and that, since $W_0 = \tilde{W}_0 = 0$, $X_0 = 0$. Hence, if $W_t$ and $\tilde{W}_t$ are normally distributed

$$X_t \sim \rho N(0, t) + \sqrt{1-\rho^2} N(0, t) \sim N(0, \rho^2 t + (1 - \rho^2)t) = N(0, t).$$

Similarly, it is possible to prove that $X_{s+t} - X_s \sim N(0, t)$. The increment is certainly independent of both histories $(W_u | u \leq s)$ and $(\tilde{W}_u | u \leq s)$, and hence also independent of the history $(X_u | u \leq s)$. □

Arithmetic Brownian motion

The most basic model involving the Brownian motion is the arithmetic Brownian motion or “Brownian motion with drift”. It is a plain model that will help to understand more the Brownian motion and the basis of all the stochastic modelling. This model was proposed by the mathematician Louis Bachelier in 1900. Even so, the name is due Robert Brown who noted some properties exhibited by random movements of pollen particles at the surface of a liquid in 1827.

Definition 2.1.4. A process $X$ that satisfies the stochastic differential equation

$$dX_t = \alpha dt + \sigma dW_t$$

(2.1)

is called an arithmetic Brownian motion. Where $\alpha, \sigma \in \mathbb{R}$ with $\sigma > 0$, $dX_t$ represents the increase (or decrease) of $X$ over an infinitesimal time interval $dt$ and $dW_t$ represents the change of a Brownian motion, $W_t$, over an interval $dt$.

Property 2.1.5. If $X$ is an arithmetic Brownian motion. The equation (2.1) implies that $E(dX_t) = \alpha dt$. Hence,

$$\alpha = \frac{1}{dt} E(dX_t)$$

represents the expected change of $X$ per unit of time. The constant $\alpha$ is also known as the drift of $X$. Also

$$\text{Var}(dX_t) = \sigma^2 dt,$$

thus, the dispersion range of $X$ increases with the standard deviation $\sigma$, also known as the volatility parameter.
Remark. Because of the Brownian motion’s Markov property, $X$ will also have this property. This will play a key role in the research of pricing models. Although in the practice of technical analysis for pricing it is important to use the whole history of prices, which is supposed to provide buying or selling signals.

![Arithmetic Brownian motion realization](image)

Figure 2.2: Arithmetic Brownian motion realization ($T = 3, \mu = 0.1, \sigma = 0.1$)

The equation (2.1) may also be written as

$$X(t + dt) = X(t) + \alpha dt + \sigma dW_t$$

this leads to the following proposition.

**Proposition 2.1.6.** At date $t$, $X_T$ follows the distribution law

$$\mathcal{L}(X_T|\mathcal{F}_t) = N(X_t + \alpha(T - t), \sigma^2(T - t))$$  \hspace{1cm} (2.2)

**Proof.** Integrating the equation (2.1) between dates $t$ and $T$ provides

$$X_T = X_t + \alpha(T - t) + \sigma(W_T - W_t).$$
Recalling the Markov property of the Brownian motion $W_T - W_t \sim N(0, (T - t))$, hence
\[ \mathcal{L}(X_T|\mathcal{F}_t) = N(X_t + \alpha(T - t), \sigma^2(T - t)). \]

\[ \square \]

**Corollary 2.1.7.** The variable $X_T$ may take negative values.

**Proof.** Since $X_T$ is a normal variable, $X_T$ can take negative values. \[ \square \]

**Example 2.1.8.** Let us observe a particular case of an arithmetic Brownian motion. Let be $X$ the arithmetic Brownian motion that satisfies the stochastic differential equation
\[ dX_t = \sigma dW_t. \]

Now $E(dX_t) = 0$, what means that $X_t$’s average does not change. This representation will be appropriate to represent the spread between two commodities whose price differences remain on average the same over time.

The possible negative values of $X$ is an important handicap since it is not possible to have negative prices. The possible negative values for the stock price would contradict the principle of limited liability attached to a stock. This fact represents the main limitation of this plain and accurate model. This leads to the geometric Brownian motion.

**Geometric Brownian motion**

To avoid the negative values handicap of the arithmetic Brownian motion model, Paul Samuelson introduced in 1965 a revisited version of the Bachelier’s model. This time, instead of modelling the change of $X$ as an arithmetic Brownian motion, the return of the stock will be modelled as an arithmetic Brownian motion. Hence, the stochastic differential equation of the stock price $S_t$ would be
\[ \frac{dS_t}{S_t} = \mu dt + \sigma dW_t \quad (2.3) \]

**Definition 2.1.9.** A process $S$ that satisfies the equation (2.3) is called a geometric Brownian motion.

This model has been used as fundamental assumption in many works. For example, this is the basis formula of the Black-Scholes-Merton pricing formula for options on stocks that was worthy of a Nobel prize in 1997.

The ratio $\frac{dS_t}{S_t}$ must be interpreted as the return obtained by investing in the stock for the period $(t, t + dt)$ with no dividend payment.
2. STOCHASTIC MODELING OF COMMODITY PRICE PROCESSES

Figure 2.3: Geometric Brownian motion realization \((T = 3, \mu = 0.1, \sigma = 0.1)\)

**Remark.** Note that due to the Markov property of the Brownian motion, \(S\) will also have this property. This may be easy to see if it is taken \(dS_t = S(t + dt) - S(t)\) and the equation (2.3) is written as

\[
S(t + dt) = S(t)(1 + \mu dt + \sigma dW_t)
\]

**Property 2.1.10.** The right-hand side of the equation (2.3) is a normal variable. Hence returns are normally distributed. Moreover, in the Black-Scholes model returns are normally distributed.

**Proposition 2.1.11.** The process \(U\) defined by

\[
U_t = \ln(S_t) \quad \forall t \geq 0
\]

is an arithmetic Brownian motion, where \(S\) is a geometric Brownian motion.
Proof. Using Itô’s lemma
\[
dU_t = d\left( \ln(S_t) \right) \\
= \frac{1}{S_t} dS_t - \frac{1}{S_t^2} (dS_t)_{\sigma(dt)} \\
= \left( \mu - \frac{\sigma^2}{2} \right) dt + \sigma dW_t
\]
It is obtained that the process \( U \) satisfies the equation
\[
dU_t = \left( \mu - \frac{\sigma^2}{2} \right) dt + \sigma dW_t \tag{2.4}
\]
Thus, \( U \) is an arithmetic Brownian motion. \( \square \)

**Corollary 2.1.12.** If \( U \) is an arithmetic Brownian motion, then, \( S_t = e^{U_t} \) is a geometric Brownian motion.

**Corollary 2.1.13.** If \( U \) is a process that satisfies \( (2.4) \) then
\[
\mathcal{L}(U_T | \mathcal{F}_t) = N\left( U_t + \left( \mu - \frac{\sigma^2}{2} \right)(T-t), \sigma^2(T-t) \right) \tag{2.5}
\]
and
\[
\mathcal{L}(S_T | \mathcal{F}_t) = \ln N\left( \ln(S_t) + \left( \mu - \frac{\sigma^2}{2} \right)(T-t), \sigma^2(T-t) \right). \tag{2.6}
\]

**Property 2.1.14.** If \( S \) is a geometric Brownian motion, the equation \( (2.3) \) implies that
\[
\mu = \frac{1}{dt} E\left( \frac{dS_t}{S_t} \right)
\]
and the constant parameter \( \mu \) represents the expected return per unit of time. If it is done the same process with the variance, we obtain
\[
\text{Var}\left( \frac{dS_t}{S_t} \right) = \sigma^2 dt.
\]

Remark. Note that the sign of \( \mu \) it is usually positive since no investor would buy a stock offering a negative expected return. Moreover, under some equilibrium conditions the expected return will be the risk-free rate \( r \) plus a risk premium. In resume, the drift \( \mu \) is usually higher than \( r \) and positive.

Remark. In the Black-Scholes model, it is supposed that the stock price grows on average over time. This assumption it is not always true for commodity prices. The Black-Scholes formula in its original form should not be used to price options on the spot prices of commodities. Furthermore, it is untrue the constant assumption of the volatility \( \sigma \). This volatility may be changed to a deterministic function of time at a low mathematical cost. It may be made stochastic either through the introduction of stochastic volatility.
Example 2.1.15. The main handicap of the arithmetic Brownian motion is the fact that the price process may take negative values. To see that with geometric Brownian motion it is not possible note that to take negative values the price $S_t$ would be 0 at some time and since we are dividing $dS_t$ by $S_t$ it is not possible. In an extreme case (large negative drift), the price will tend to 0.

![Geometric Brownian motion realization](image)

Figure 2.4: Geometric Brownian motion realization ($T = 3, \mu = -1, \sigma = 0.5$)

2.2 Mean-reversion

In commodity pricing, the mean reversion processes are common. Commodity prices neither grow nor decline on average over time. They tend to mean-revert to a level which may be viewed as the marginal cost of production.

Example 2.2.1. In agricultural prices; as prices decline, farmers find it less attractive to produce the commodity and supply decreases creating upward pressure on the price. Similarly, as the price of an agricultural commodity increases, farmers are more likely to devote resources to producing the commodity creating downward pressure on the price.

The mean-reversion process in commodity prices may be understood as the mean-reversion process on the price when an arbitrage occasion appears. This
was first introduced in finance by Vasicek in 1977. He introduced the mean-reversion phenomenon through the Ornstein-Uhlenbeck process to describe the short-term rate dynamics.

**Definition 2.2.2.** An Ornstein-Uhlenbeck process, $X$, satisfies the stochastic differential equation

$$dX_t = a(b - X_t)dt + \sigma dW_t$$  \hspace{1cm} (2.7)

where $a, b, \sigma > 0$ and $W_t$ is a Brownian motion.

**Proposition 2.2.3.** If $X$ is an Ornstein-Uhlenbeck process, it has the mean-reversion property.

*Proof.* The expected change of $X$, from date $t$, is given by

$$E(dX_t|\mathcal{F}_t) = a(b - X_t)dt.$$  

Thus if $X_t$ is smaller than $b$ the expected change is positive; if $X_t$ is greater than $b$, the expected change is negative. This is the named *mean-reversion property.* \hfill \Box

![Figure 2.5: Ornstein-Uhlenbeck process realization ($T = 3$, $a = 5$, $b = 10$, $\sigma = 0.1$)](image)

To calculate the distribution law of the Ornstein-Uhlenbeck processes it is necessary the following lemma.
Lemma 2.2.4. Let \( W_t \) be a Brownian motion, \( T, t, a \) integers such that \( T > t \). Then

\[
\int_t^T e^{at} dW_t = N\left(0, \frac{1}{2a} (e^{2aT} - e^{2at})\right) \tag{2.8}
\]

Proof. Using the Riemann-Stieltjes integral definition it is given that

\[
\int_t^T e^{at} dW_t = \lim_{n \to \infty} \sum_{i=1}^n e^{kt_i} (W_{t_i} - W_{t_{i-1}})
\]

where \( \{t_i\}_{0 \leq i \leq n} \) is a partition of the interval \( (t, T) \). By the Brownian motion’s definition \( W_{t_i} - W_{t_{i-1}} \sim N(0, \Delta t) \) with \( \Delta t = t_i - t_{i-1} \) \( \forall i \). Note that all this normal distributions are independent due the Markov property of the Brownian motion. Thus

\[
\lim_{n \to \infty} \sum_{i=1}^n e^{kt_i} (W_{t_i} - W_{t_{i-1}}) = \lim_{n \to \infty} \sum_{i=1}^n e^{kt_i} N(0, \Delta t) = N\left(0, \lim_{n \to \infty} \sum_{i=1}^n e^{2at_{i-1}} \Delta t\right).
\]

Since \( \Delta t = \frac{T-t}{n} \), thus

\[
\lim_{n \to \infty} \sum_{i=1}^n e^{2at_{i-1}} \Delta t = \int_t^T e^{2as} ds.
\]

Hence

\[
N\left(0, \lim_{n \to \infty} \sum_{i=1}^n e^{2at_{i-1}} \Delta t\right) = N\left(0, \int_t^T e^{2as} ds\right) = N\left(0, \frac{1}{2a} (e^{2aT} - e^{2at})\right)
\]

\( \square \)

Proposition 2.2.5. If \( X \) is a process satisfying (2.7), then at date \( t \)

\[
\mathcal{L}(X_T \mid \mathcal{F}_t) = N\left(e^{-a(T-t)}X_t + b(1 - e^{-a(T-t)}), \sigma^2 \frac{1 - e^{-2a(T-t)}}{2a}\right). \tag{2.9}
\]
Proof. Let $Z$ be a process such that

$$ Z_t = e^{at} X_t. $$

Hence using Itô’s lemma

$$ dZ_t = d(e^{at} X_t) = X_t a e^{at} dt + e^{at} dX_t = ab e^{at} dt + \sigma e^{at} dW_t $$

Integrating between the dates $t$ and $T$ provides

$$ Z_T - Z_t = b(e^{aT} - e^{at}) + \sigma \int_t^T e^{as} dW_s. $$

Using the Lemma 2.2.4,

$$ Z_T = Z_t + b(e^{aT} - e^{at}) + N \left( 0, \frac{1}{2a} (e^{2aT} - e^{2at}) \right) $$

Undoing the change of variable,

$$ X_T = e^{-a(T-t)} X_t + b(1 - e^{-a(T-t)}). $$

Thus

$$ \mathcal{L}(X_T | \mathcal{F}_t) = N \left( e^{-a(T-t)} X_t + b(1 - e^{-a(T-t)}), \sigma^2 \frac{1 - e^{-2a(T-t)}}{2a} \right). $$

This model is applied to the spot rate, $r(t)$. Hence, as seen previously, the equation (2.7) may be written as

$$ r(t + dt) = r(t) + a(b - r(t)) dt + \sigma dW_t $$

which, at date $t$, is an affine function on $dW_t$. Thus, as in the Brownian motion, the spot rate is, like $dW_t$, normally distributed and may take negative values. This, obviously a unkind property for prices. Resembling the geometric Brownian motion, the Ornstein-Uhlenbeck’s processes will be written as

$$ \frac{dS_t}{S_t} = k(\theta - \ln(S_t)) dt + \sigma dW_t \quad (2.10) $$

where $\theta$ is the drift term (which will be the equilibrium price around which the logarithm’s price will fluctuate) and $k$ is the mean-reversion force. As $k$ is higher the tendency to $\theta$ will increase.
Note. Differently of the Ornstein-Uhlenbeck process, the mean-reversion process (or exponential Vasicek process) (2.10) fluctuates around $e^\theta$. I.e. $\ln(S_t)$ fluctuates around $\theta$.

**Definition 2.2.6.** A process $S$ following the stochastic differential equation (2.10) is called a mean-reversion process.

**Proposition 2.2.7.** If $S$ is a mean-reversion process, then the process $\ln(S_t)$ is an Ornstein-Uhlenbeck process.

**Proof.** Let $Z_t$ be

$$Z_t = \ln(S_t).$$

Thus, with this change of variable and using the Itô’s

$$dZ_t = d\left(\ln(S_t)\right) = \frac{1}{S_t}dS_t - \frac{1}{S_t^2}(dS_t)^2 = k(\theta - Z_t)dt + \sigma dW_t - \frac{\sigma^2}{2}dt = k\left(\theta - \frac{\sigma^2}{2k} - Z_t\right)dt + \sigma dW_t.$$

I.e.,

$$dZ_t = k\left(\theta - \frac{\sigma^2}{2k} - Z_t\right)dt + \sigma dW_t. \quad (2.11)$$

That is a Ornstein-Uhlenbeck process with drift term $\theta - \frac{\sigma^2}{2k}$. □

**Corollary 2.2.8.** If $X$ is an Ornstein-Uhlenbeck process, then, $S_t = e^{X_t}$ is a mean-reversion process.

**Corollary 2.2.9.** If $Z$ is a process that satisfies (2.11) then

$$\mathcal{L}(Z_T | \mathcal{F}_t) = N\left(e^{-k(T-t)}Z_t + \left(\theta - \frac{\sigma^2}{2k}\right)\left(1 - e^{-k(T-t)}\right), \frac{\sigma^2}{2k}\left(1 - e^{-2k(T-t)}\right)\right), \quad (2.12)$$

and

$$\mathcal{L}(S_T | \mathcal{F}_t) = \ln N\left(e^{-k(T-t)}\ln(S_t) + \left(\theta - \frac{\sigma^2}{2k}\right)\left(1 - e^{-k(T-t)}\right), \frac{\sigma^2}{2k}\left(1 - e^{-2k(T-t)}\right)\right). \quad (2.13)$$

For contracts of long time maturity it is observed that the volatility goes to zero with this model and this is not observed in the market. To improve the model, jumps and stochastic volatility will be introduced.
2.3 Seasonality, jumps and stochastic volatility

The Mean-reversion model and the geometric Brownian motion are the most common models used when modelling the commodities price processes. In fact, there are some modifications that can be implemented in these models. In this section it will be introduced the concepts of seasonality, jumps and stochastic volatility.

Seasonality

Lots of commodities, such as agricultural or natural gas, exhibit seasonality in prices. This fact is due to harvest cycles or due to changes in consumption through a cycle. E.g. the electric consumption is much higher during the summer due the use of air conditioners. There are several models involving the seasonality feature. Geman, [3], uses an interesting model modifying the spot price $S_t$ by

$$\ln(S(t)) = f(t) + X(t)$$

(2.14)

where $f(t)$ is a deterministic periodic component that will provide the spot price with seasonality, and where $X$ is a process following either an arithmetic Brownian motion (2.1) or an Ornstein-Uhlenbeck process (2.7).
2. STOCHASTIC MODELING OF COMMODITY PRICE PROCESSES

The function \( f(t) \) is usually expressed as a sin or cos with annual or semi-annual periodicity and it is derived from a database of spot prices. 

*Note.* Note that due to (2.14), hence,

\[
S(t) = e^{f(t)} \cdot e^{X(t)}
\]

what means that the process \( S \) follows a process similar to \( e^{X(t)} \) that would be a geometric Brownian motion or a mean-reversion process (depending if the process \( X \) is an arithmetic Brownian motion or an Ornstein-Uhlenbeck process) modified by a seasonality ratio given by \( e^{f(t)} \) that will never cancel the price process \( S \) but will reduce its effect.

In order to understand the equation (2.14) in the same terms that previously it is interesting to have the associated stochastic differential equation.

**Proposition 2.3.1.** Given an Ornstein-Uhlenbeck process, \( X \), and a deterministic function, \( f_t \), the process of prices, \( S \), associated to (2.14) satisfy the stochastic differential equation

\[
\frac{dS_t}{S_t} = \left( a(b - X_t) + \frac{1}{2} \sigma^2 \right) dt + df_t + \sigma dW_t.
\]

(2.15)

*Proof.* The equation (2.14) is equivalent to

\[
S_t = e^{f_t + X_t}.
\]

Let \( Y \) be the process

\[
Y_t = f_t + X_t.
\]

By the Itô’s lemma,

\[
dS_t = e^{Y_t}dY_t + \frac{1}{2} e^{Y_t}(dY_t)^2|_{0}(dt)
\]

Note that \( S_t = e^{Y_t} \), hence

\[
dS_t = S_t(df_t + dX_t) + \frac{1}{2} S_t(a(b - X_t)dt + \sigma dW_t + df_t)^2|_{0}(dt).
\]

Since \( f_t \) is a deterministic function,

\[
df_t = f'_t dt
\]

hence,

\[
\frac{dS_t}{S_t} = \left( a(b - X_t) + \frac{1}{2} \sigma^2 \right) dt + df_t + \sigma dW_t
\]

\[\square\]
This model includes the seasonality feature at a low mathematical cost. For this motive, this is an interesting way for the modelling of commodity seasonal prices.

**Example 2.3.2.** If \( f_t = \frac{\sin(2\pi t)}{25} \) thus,

\[
\frac{dS_t}{S_t} = \left( a(b - X_t) + \frac{2\pi}{25} \cos(2\pi t) \right) dt + \frac{1}{2} \sigma^2 dt + \sigma dW_t
\]

which may be written as

\[
\frac{dS_t}{S_t} = a(\hat{b}_t - X_t)dt + \frac{1}{2} \sigma^2 dt + \sigma dW_t
\]

with \( \hat{b}_t = b + \frac{2\pi}{25} \cos(2\pi t) \). This may be interpreted as a process that mean-revert to \( \hat{b}_t \), whose value changes with time. This is what gives \( S_t \) the seasonality feature. Moreover appears a drift term (as in geometric Brownian motion), \( \frac{1}{2} \sigma^2 dt \).

---

Figure 2.7: Realization of \( S_t \) with \( T = 3, \sigma = 0.1, a = 5, b = \ln(10) \)

One of the limitations of this model is the assumption of deterministic seasonality, which is a rough approximation taking in account that the processes always had an unpredictable part. The model used presented in [4] is a
two-state variable model. They consider that the spot price \( S_t \) is described by

\[
dS_t = (\alpha + \beta X_t)S_t dt + \sigma S_t dW_{1t}
\]

(2.16)

where \( X_t \) is a stochastic process

\[
dx_t = (c_1 \sin(2\pi t + c_2) - aX_t)dt + bdW_{2t}
\]

(2.17)

and in the most general case they allow for correlation between the two Brownian increments \( dW_{1t} \) and \( dW_{2t} \) with the coefficient

\[
corr(dW_{1t}, dW_{2t}) = \rho dt
\]

(2.18)

In essence, this model treats the spot price \( S_t \) as a geometric Brownian motion with a stochastic drift term, \( (\alpha + \beta X_t) \), that has one constant part and the other one follows a mean-reversion seasonal drift term. Note that the correlation (2.18) links the \( S \) drift term to its noise part.

This model is much more complex than the Geman one. In this work it will be used the Geman model, for its simplicity and low mathematical cost. The Barbu and Burrage model can be studied deeply in [4].

### Jumps in price trajectories

All the models seen assume that the world has a kind of solidity faculty that assures continuity of the trajectories they produce. This is, in fact, untrue. Important events in history have affected the trend of the market; the crash of 1987 or the Lehman crisis in 2008 are clear examples. This breaking news are traduced as price jumps in the market. This jumps were introduced in 1976 by Merton through a jump component added to the diffusion term. This diffusion term is usually modelled as a Poisson process.

**Definition 2.3.3.** A Poisson process with intensity \( \lambda \) is a process \( N \) taking values in \( \mathbb{N} \) such that:

(i) if \( s < t \) then \( N_s \leq N_t \) and \( N_0 = 0 \),

(ii) \( P(N(t + h) = n + m \mid N(t) = n) = \begin{cases} \lambda h + o(h) & \text{if } m = 1, \\ o(h) & \text{if } m > 1, \\ 1 - \lambda h + o(h) & \text{if } m = 0, \end{cases} \)

(iii) if \( s < t \), the number \( N(t) - N(s) \) of emissions in the interval \((s, t]\) is independent of the times of emissions during \([0, s] \).
The process $N$ may be interpreted as a counting process. Every “occurrence” or “emission” of $N$ will be a jump in the price process. For this, it would be interesting to know the distribution of $N(t)$.

**Theorem 2.3.4.** Let $N$ be a Poisson process with intensity $\lambda$, hence, $N$ has the Poisson distribution with parameter $\lambda t$. I.e.

$$P(N(t) = j) = \frac{(\lambda t)^j e^{-\lambda t}}{j!}, \quad \forall j \in \mathbb{N}.$$ 

**Proof.** Conditioning $N(t+h)$ to $N(t)$

$$P(N(t+h) = j) = \sum_i P(N(t) = i)P(N(t+h) = j \mid N(t) = i)$$

$$= P(N(t) = j - 1)(\lambda h + o(h)) + P(N(t) = j)(1 - \lambda h + o(h)) + o(h)$$

$$= P(N(t) = j - 1)\lambda h + P(N(t) = j)(1 - \lambda h) + o(h).$$

Let $p_j(t) = P(N(t) = j)$. Thus, $p_j(t)$ satisfies

$$p_j(t+h) = \lambda h p_{j-1}(t) + (1 - \lambda h)p_j(t) + o(h) \quad \text{if } j \neq 0,$$

$$p_0(t+h) = (1 - \lambda h)p_0(t) + o(h).$$

Subtract $p_j(t)$ from each side of the first equation, divide by $h$, and let $h \downarrow 0$ to obtain

$$p'_j(t) = \lambda(p_{j-1}(t) - p_j(t)) \quad \text{if } j \neq 0; \quad (2.19)$$

likewise

$$p'_0(t) = -\lambda p_0(t). \quad (2.20)$$

The boundary condition is

$$p_j(0) = \begin{cases} 
  1 & \text{if } j = 0, \\
  0 & \text{otherwise.} 
\end{cases}$$

Solving (2.20) with the condition $p_0(0) = 1$ it is obtained

$$p_0(t) = e^{-\lambda t}.$$ 

Substituting this into (2.19) with $j = 1$ it is obtained

$$p_1(t) = \lambda te^{-\lambda t}.$$ 

And by induction

$$p_j(t) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}.$$
2. STOCHASTIC MODELING OF COMMODITY PRICE PROCESSES

Figure 2.8: Probabilities of a Poisson distribution with different $\lambda$ parameters

Note. The Poisson distribution is commonly used for counting uncommon events. E.g. the counting of printing errors per page in a book or the number of $\alpha$ particles emitted by a radioactive substance. The parameter $\lambda$ may be interpreted as the expected number of events or occurrences per unit of time.

A Poisson process can be understood through a different formulation which provides much insight into its behaviour. The process can be characterized giving the \emph{interarrival times} between the jumps an exponential distribution.

\textbf{Definition 2.3.5.} Given $N_t$ a Poisson process, the time of the $n$th arrival $T_n$ is defined by

$$T_n = \inf\{t \mid N(t) = n\}.$$

Note that $T_0 = 0$.

\textbf{Definition 2.3.6.} The \emph{interrival times} are the random variables $\{X_i\}_{i \geq 1}$ given by

$$X_n = T_n - T_{n-1}.$$

Note. From knowledge of $N_t$ it is possible to find the values of $\{X_i\}_{i \geq 1}$ by the previous definitions. Conversely, it is possible to construct $N_t$ from $\{X_i\}_{i \geq 1}$ by

$$T_n = \sum_{i} X_i, \quad N(t) = \max\{n \mid T_n \leq t\}.$$
Theorem 2.3.7. The random variables \( \{X_i\}_{i \geq 1} \) are independent, each having the exponential distribution with parameter \( \lambda \).

Proof. First of all consider \( X_1 \):

\[
P(X_1 > t) = P(N(t) = 0) = e^{-\lambda t}
\]
since \( N_t \) has the Poisson distribution with parameter \( \lambda \) (Theorem 2.3.4). And so \( X_1 \) is exponentially distributed. Now, conditioning on \( X_1 \),

\[
P(X_2 > t \mid X_1 = t_1) = P(\text{no arrival in } (t_1, t_1 + t] \mid X_1 = t_1).
\]
The event \( \{X_1 = t_1\} \) is referred to arrivals during the time interval \([0, t_1]\), whereas the event \( \{\text{no arrival in } (t_1, t_1 + t]\} \) is referred to arrivals after time \( t_1 \). From the Definition 2.3.3, these events are independent, hence

\[
P(X_2 > t \mid X_1 = t_1) = P(\text{no arrival in } (t_1, t_1 + t]) = e^{-\lambda t_1}.
\]
Thus \( X_2 \) is independent of \( X_1 \), and has the same distribution. Similarly,

\[
P(X_{n+1} > t \mid X_1 = t_1, \ldots, X_n = t_n) = P(\text{no arrival in } (T, T + t])
\]
where \( T = t_1 + t_2 + \cdots + t_n \), and the claim of the theorem follows by induction on \( n \). \qed

This theorem allow to understand and forecast better the new arrivals of informations that will be translated as jumps in the price trajectories.
Proposition 2.3.8. If $N$ is a Poisson process with intensity $\lambda$ hence

$$dN_t = \begin{cases} 
0 & \text{with probability } 1 - \lambda dt \\
1 & \text{with probability } \lambda dt
\end{cases}$$

Proof. It comes directly from the second statement of the Definition 2.3.3. □

The jump component is formed by the Poisson process and a real valued random variable $U_t$ that will determine the magnitude of the jump. This leads to the geometric Brownian motion with jumps or the mean-reversion processes with jumps:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t + U_t dN_t$$  \hspace{1cm} (2.21)

$$\frac{dS_t}{S_t} = k(\theta - \ln(S_t))dt + \sigma dW_t + U_t dN_t$$  \hspace{1cm} (2.22)

Note. The usual and practical assumption is that the Brownian motion $dW_t$ and the Poisson process $dN_t$ are not correlated.

The common choice for the distribution of $U_t$ is a normal distribution. The two big problems with jump processes are: the impossibility to build a riskless portfolio; and the difficulty with parameters estimative. I.e., although the implementation of the jump component has a low mathematical cost, the pricing and calibrating of assets, derivatives and models become a really complex problem. The jump model will not be discussed later.

Stochastic volatility models

All the seen models assume that volatility is constant. In practice volatility varies through time. The daily fluctuations of the return of asset prices typically exhibit volatility clustering where large moves follow large moves and small moves follow small moves. Also, the distribution of asset price returns is peaked and fat-tailed, this indicates a mixture of heterogeneous distribution with different variances. The first approximation to this problem would be supposing that the volatility parameter is a known deterministic function. The price process is then

$$\frac{dS_t}{S_t} = \mu dt + \sigma_t dW_t$$

where $S_t$ is a geometric Brownian motion. This assumption implies that the instantaneous volatility of an asset is predictable. Actually, that is not the best approximation. The volatility parameter can be approximated through
a stochastic model. This leads to a two-state variable model: the stock price and its volatility.

The modelling of stochastic volatility is due to Heston in 1993, [5]. This model takes the price standard deviation, $\sqrt{v_t}$, as a mean-reversion process to zero correlated with the asset price process. This lead to the model

$$\begin{align*}
\text{d}S_t &= \mu S_t \text{d}t + \sqrt{v_t} S_t \text{d}Z_{1t}, \\
\text{d}\sqrt{v_t} &= -\beta \sqrt{v_t} \text{d}t + \delta \text{d}Z_{2t}, \\
\text{d}Z_{1t}\text{d}Z_{2t} &= \rho \text{d}t,
\end{align*}$$

(2.23)

where $Z_{1t}$ and $Z_{2t}$ are two Brownian motions correlated with $\rho$. It is clear that $\sqrt{v_t}$ mean-revert to 0. The point of the model is to express the second equation of (2.23) depending of $\text{d}v_t$ the differential of the price’s variance.

**Proposition 2.3.9.** The stochastic differential equation

$$\text{d}\sqrt{v_t} = -\beta \sqrt{v_t} \text{d}t + \delta \text{d}Z_{2t}$$

is equivalent to

$$\text{d}v_t = k(\overline{v} - v_t) \text{d}t + \eta \sqrt{v_t} \text{d}Z_{2t},$$

(2.24)

where $k = 2\beta$, $\overline{v} = \frac{\delta^2}{2\beta}$, $\eta = 2\delta$.

**Proof.** Since $\text{d}v_t = \text{d}(\sqrt{v_t})^2$ using the Itô’s lemma it is given

$$\begin{align*}
\text{d}v_t &= \text{d}(\sqrt{v_t})^2 \\
&= 2\sqrt{v_t} \text{d}\sqrt{v_t} + (\text{d}(\sqrt{v_t})^2)_{|\text{d}t} \\
&= 2\sqrt{v_t}(-\beta \sqrt{v_t} \text{d}t + \delta \text{d}Z_{2t}) + \delta^2 \text{d}t \\
&= (\delta^2 - 2\beta v_t) \text{d}t + 2\delta \sqrt{v_t} \text{d}Z_{2t} \\
&= k(\overline{v} - v_t) \text{d}t + \eta \sqrt{v_t} \text{d}Z_{2t}
\end{align*}$$

with $k = 2\beta$, $\overline{v} = \frac{\delta^2}{2\beta}$, $\eta = 2\delta$. \qed

Summarizing, the Heston model is given by

$$\begin{align*}
\text{d}S_t &= \mu S_t \text{d}t + \sqrt{v_t} S_t \text{d}Z_{1t}, \\
\text{d}v_t &= k(\overline{v} - v_t) \text{d}t + \eta \sqrt{v_t} \text{d}Z_{2t}, \\
\text{d}Z_{1t}\text{d}Z_{2t} &= \rho \text{d}t,
\end{align*}$$

(2.25)

All model parameters are assumed to be constant. Privault, [6], holds that evidences based on financial market data shows that the variations in volatility tend to be negatively correlated with the variations of underlying asset prices.
In this way, the volatility clustering it is better explained. Note that the parameter $\overline{\tau}$ it is often estimated as the average variance rate. Moreover, Hull and White propose a slightly different model, using the Heston’s model notation:

\[
\begin{align*}
    dS_t &= \mu S_t dt + \sqrt{v_t} S_t dZ_{1t}, \\
    dv_t &= k(\overline{\tau} - v_t) dt + \eta \sqrt{v_t} dZ_{2t}, \\
    dZ_{1t} dZ_{2t} &= 0.
\end{align*}
\] (2.26)

Basically there are two differences:

(i) The Brownian motion that drive the price process and the price’s variance process are not correlated. This is a simplification of the Heston model. Assuming that the correlation $\rho$ it is different of zero is more complicated.

(ii) Hull and White’s model gives a freedom degree including the parameter $\alpha$. The case $\alpha = 0.5$ leads to the Heston $\rho = 0$ model.

Working on this type of model, Hull and White calculate the European call price. And for the case $\alpha = 0.5$ Hull and White provide a series expansion and Heston, [5], provides an analytic result.

The stochastic volatility impact on pricing is fairly small in absolute term for options that last less than a year (although in percentage terms it can be quite large for deep-out-of-the-money options). It becomes progressively larger as the life of the option increases.
Plain Vanilla Options on Commodity Spot and Forward Prices

The aim of this chapter is to assign a price to the main derivatives on commodities. To achieve this goal various techniques are used. The neutral risk valuation and the martingales theory are essential to price this kind of products. To price the options related to the mean-reversion models the neutral risk valuation becomes a tough path. This products will be priced through the Monte Carlo method, a simulation method.

Before going inside this chapter it is important to recall some items seen in the introduction:

\( S_t \), underlying asset price at time \( t \).

\( T \), maturity date.

\( X \), strike price.

\( F_t \), strike agreed at time \( t \) for a forward contract price.

\( f_t \), forward contract price at time \( t \).

\( c_t \), European call price at time \( t \).

\( p_t \), European put price at time \( t \).

\( c_0 \), European call premium.

\( p_0 \), European put premium.

Position: buyer (long) or seller (writer).

The first derivative to study are the forwards involving commodities.
3.1 Forward pricing

The forward has a singular property, the forward contract price \( f_t \) and the forward price \( F_t \) do not depend on the model attributed to the underlying asset \( S_t \). The first step is to deduce \( f_t \) and \( F_t \) for non-dividend paying stocks. All this propositions are argued in the same way. The prices are set making the following assumptions:

(i) No taxes, no transaction costs: “frictionless markets”.

(ii) Interest rates are constant: \( r \) denotes the continuously compound rate.

(iii) There are no arbitrage opportunities: with a zero initial wealth and taking no risk at initial date, the final wealth will be surely zero at date \( T \).

**Proposition 3.1.1.** For a non-dividend paying stock \( S \), the forward price set at time \( t \) is given by

\[
F_t = S_t e^{r(T-t)}. \tag{3.1}
\]

*Proof.* To prove this equality, it will be constructed an arbitrage strategy for the inequalities

\[
F_t < S_t e^{r(T-t)}, \quad F_t > S_t e^{r(T-t)}.
\]

For the first one, the strategy is given by Strategy 1 (Table 3.1).

<table>
<thead>
<tr>
<th>At time ( t )</th>
<th>At time ( T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Take long position in a forward starting at ( t )</td>
<td>Receive ( S_t e^{r(T-t)} ) from investment</td>
</tr>
<tr>
<td>Short one unit of asset for ( S_t )</td>
<td>Buy asset for ( F_t ) due the forward obligation</td>
</tr>
<tr>
<td>Invest ( S_t ) at ( r ) for ( T-t )</td>
<td>Close short sell</td>
</tr>
<tr>
<td>Profit realized: ( S_t e^{r(T-t)} - F_t &gt; 0 )</td>
<td></td>
</tr>
</tbody>
</table>

*Table 3.1: Strategy 1*

Which is a free risk operation with benefits. The strategy for the second inequality is given by Strategy 2 (Table 3.2).

Which is also a free risk operation with benefits. This means that, a priori, the unique price arbitrage-free is

\[
F_t = S_t e^{r(T-t)}. \tag{3.2}
\]

\( \square \)
3.1. FORWARD PRICING

At time $t$
- Take short position in a forward starting at $t$
- Buy one unit of asset for $S_t$
- Borrow $S_t$ at $r$ for $T-t$

At time $T$
- Sell asset for $F_t$ due the forward obligation
- Pay $S_te^{r(T-t)}$ for the loan with interest

Profit realized: $F_t - S_te^{r(T-t)} > 0$

Table 3.2: Strategy 2

*Note.* The forward price at maturity is exactly the asset value. *I.e.* $F_T = S_T$.

*Note.* Most of the proves using the construction of arbitrage strategies assure that the price could be that one. In the case of the forward it has been proved that $F_t$ cannot be less or more than $S_te^{r(T-t)}$. But what is not proved is that $S_te^{r(T-t)}$ is an *arbitrage-free* price. It has not been proved if $S_te^{r(T-t)}$ is the fair price or not, it has been proved that if $F_t \neq S_te^{r(T-t)}$ then exists an arbitrage strategy.

*Note.* It is important to note that no model, related to the behaviour of $S_t$, has been used. Although, as will be seen later, this is not the forward price that will be used in commodities.

The next step is to calculate the forward contract price $f_t$.

**Proposition 3.1.2.** For a non-dividend paying stock $S$, the forward contract price set at time $t$ is given by

$$f_t = S_t - Xe^{-r(T-t)}.$$  \hfill (3.2)

**Proof.** To prove this equality, it will be constructed an arbitrage strategy for the inequalities

$$f_t < S_t - Xe^{-r(T-t)},$$
$$f_t > S_t - Xe^{-r(T-t)}.$$  

For the first one, the arbitrage strategy is given by Strategy 1 (Table 3.3). Which is a free risk operation with benefits. The strategy for the second inequality is given by Strategy 2 (Table 3.4).

Which is also a free risk operation with benefits. This means that, *a priori*, the unique *arbitrage-free* price is

$$f_t = S_t - Xe^{-r(T-t)}.$$

□
At time $t$

- Short one unit of asset for $S_t$.
- Buy a long position forward contract for $f_t$.
- Invest $X e^{-r(T-t)}$ at $r$ for $T-t$.

Profit realized: $S_t - f_t - X e^{-r(T-t)} > 0$

Table 3.3: Strategy 1

At time $T$

- Receive $X$ from investment.
- Buy asset for $X$ due the forward obligation.
- Close short sell.

Table 3.4: Strategy 2

At time $t$

- Borrow $X e^{-r(T-t)}$ at $r$ for $T-t$.
- Buy one unit of asset for $S_t$.
- Buy a short position in a forward contract for $-f_t$.

Profit realized: $f_t + X e^{-r(T-t)} - S_t > 0$

Corollary 3.1.3. The forward contract price is given by

$$f_t = (F_t - X)e^{-r(T-t)}. \quad (3.3)$$

Note. Using (3.3) thus

$$f_0 = f_T = 0. \quad (3.4)$$

This is the forward pricing basis. One of the most valuable result in finances is the put-call parity. The put-call parity is a relation between an European call price and its equivalent European put price.

Proposition 3.1.4. At date $t$, the European put $p$, which has the same strike $X$, maturity $T$, and same underlying asset $S$ as the European call $c$ has a price $p_t$ related to $c_t$ through the fundamental relationship

$$c_t - p_t = S_t - X e^{-r(T-t)} \quad (3.5)$$

where $t$ is any date prior or equal to the maturity $T$.

Proof. As seen previously, to prove (3.5) it will be seen that

$$c_t + X e^{-r(T-t)} < S_t + p_t,$$

$$c_t + X e^{-r(T-t)} > S_t + p_t.$$
At time $t$  
Short one unit of asset for $S_t$  
Sell one put for $p_t$  
Buy one call for $c_t$  
Invest $Xe^{-r(T-t)}$ at $r$ for $T-t$  
Profit realized: $S_t + p_t - c_t - Xe^{-r(T-t)} > 0$

Table 3.5: Strategy 1

At time $T$  
If $S_T > X$  
Receive $X$ from the investment.  
Exercise the call and buy the asset for $X$  
The put is not exercised  
Cancel short sell  
Profit realized: $S_t + p_t - c_t - Xe^{-r(T-t)} > 0$

Table 3.6: Strategy 1

At time $T$  
If $S_T < X$  
Receive $X$ from the investment  
Do not exercise the call  
The put is exercised, buy asset for $X$  
Cancel short sell

The arbitrage strategy built for the first inequality is Strategy 1 (table 3.5), which is a free risk operation with benefits. The strategy for the second inequality is given by Strategy 2 (Table 3.6).

Which is also a free risk operation with benefits. This means that, \( a \ priori \),

\[
c_t - p_t = S_t - Xe^{-r(T-t)}.\]

\[\square\]

\textit{Note.} The option pricing will be shown to be very complex. The put-call parity is a fast path to calculate the price if one of the call or put price has been calculated. It is important to note that it has not been used any model related to the behaviour of $S_t$.

\textit{Note.} The right-hand side of the equation (3.5) may be written as $f_t$. Thus, the put-call parity may be written as

\[
c_t - p_t = f_t.\]
There are special kinds of assets that report an income, \( I \), or a known yield, \( q \). This leads to a new forward pricing.

**Proposition 3.1.5.** The forward price and forward contract price for an underlying asset reporting a known income \( I \) or a known yield \( q \), in the period \([t, T]\), at time \( t \) are given by

\[
F_t = (S_t - I)e^{r(T-t)} \quad (3.6)
\]
\[
f_t = S_t - I - Xe^{-r(T-t)} \quad (3.7)
\]
\[
F_t = S_te^{(r-q)(T-t)} \quad (3.8)
\]
\[
f_t = S_te^{-q(T-t)} - Xe^{-r(T-t)} \quad (3.9)
\]

To incorporate the commodities properties this equations will be very useful. The main factors that have to be added to the model when dealing with commodities are:

(i) Storage costs

(ii) Consumption factor

(iii) Convenience yields

Most of the commodities must be stored. The main exception is the electricity. This storage has a constant cost that must be paid by the forward writer, since is who is holding the asset. The storage costs may be interpreted as a negative income. If \( U_t \) is the total cost of the storages in \( t \) thus

\[
F_t = (S_t + U_t)e^{r(T-t)}. \quad (3.10)
\]

This also can be added as a cost per annum proportion \( u \).

\[
F_t = S_te^{(r+u)(T-t)}. \quad (3.11)
\]

One of the most important factors when analysing commodity stock markets is to take in account the usage of the commodity. It is important to distinguish between consumption assets and investment assets. For this reason it is necessary to revisit the forward pricing method.

**Proposition 3.1.6.** If \( S \) is a consumption asset with storage costs \( U_t \), then, at date \( t \),

\[
F_t \leq (S_t + U_t)e^{r(T-t)}. \quad (3.12)
\]

If \( S \) is an investment asset with storage costs \( U_t \), at date \( t \) the forward price is given by (3.10).
Proof. To prove (3.10) and (3.12), the arbitrages strategies must be reviewed carefully. First of all, suppose that

\[ F_t > (S_t + U_t)e^{r(T-t)}. \]

In this case, the strategy in Table 3.2 can be used but instead of borrowing \( S_t \), it must be borrowed \( S_t + U_t \). There is no problem in implementing the strategy for any commodity.

Suppose next that

\[ F_t < (S_t + U_t)e^{r(T-t)}. \]

When the commodity is an investment asset, can be supposed that the investors hold the commodity only for investment. In this case, the investors will follow the strategy in Table 3.1 but instead of investing \( S_t \) it can be invested \( S_t + U_t \) since the storage cost has been saved. This argument cannot be used for a consumption commodity. Since it is a consumption asset, the commodity holders usually plan to use it in some way. They are reticent to sell the commodity in the market and buy forward contracts, essentially because a forward contract cannot be consumed or used in an industrial process. There is therefore nothing to stop the inequality from holding. Thus the only that can be asserted for a consumption commodity is

\[ F_t \leq (S_t + U_t)e^{r(T-t)}. \]

□

If storages costs are expressed as an annum proportion \( u \), thus the (3.12) equivalent result is

\[ F_t \leq S_t e^{(r+u)(T-t)}. \] (3.13)

Example 3.1.7. An oil refiner is unlikely to enter in a forward contract on crude oil in the same way as crude oil held in inventory. The held crude oil can be an input to the refining process, physical asset enables a manufacturer to keep a production process running and perhaps profit from temporary local shortages. A forward contract cannot be used in the same way.

Definition 3.1.8. The benefits from holding the physical asset is the convenience yield provided by the commodity. If the storage cost is \( U_t \), then the convenience yield \( y \) is defined such that

\[ F_t e^{y(T-t)} = (S_t + U_t)e^{r(T-t)} \] (3.14)

With the convenience yield, the forward price in consumption commodities may be written as

\[ F_t = S_t e^{(r+u-y)(T-t)} \] (3.15)
The convenience yield reflects the market’s expectation concerning the future availability of the commodity. As much big are the inventories, there is very little chance of shortages in the near future and the convenience yield will tend to be low. If there are much possibilities of shortages, the convenience yield will tend to be high.

Note. If the storage cost is written as an annum proportion $u$. Then, $u$ may be interpreted as a premium on the free-risk interest rate $r$. And the convenience yield may be interpreted as a known incoming yield, in the sense that $y$ measures the value of holding the commodity.

In general, all this modifications can be implemented through the cost of carry, $c$. This measures the storage cost plus $u$, the free-risk interest rate $r$, and the known yield $q$. In the case of commodities

$$c = r - q + u. \quad (3.16)$$

Then, the forward price may be written as

$$F_t = S_t^{c(T-t)}, \quad (3.17)$$

and in the case of consumption commodities

$$F_t = S_t^{(c-y)(T-t)} \quad (3.18)$$

where $y$ is the convenience yield.

Here has been assumed that the delivery location for spot and forward are the same. Including delivery costs may complicate more the equations.

### 3.2 Martingales

Up to now, it has been discussed which would be the best model to fit the commodity spot prices, always allowing a stochastic part based on the Brownian motion $W_t$. Recalling the definition of the Brownian motion, this Brownian motion it is seen under a probability measure $P$. This fact has not been treated or considered in detail. As seen in the introduction, a filtration may be interpreted as a history or as information. To understand better the role of this probability measure $P$, it is necessary to introduce the martingales.

**Definition 3.2.1.** A stochastic process $M$ is a martingale with respect to a measure $P$ if and only if

1. $E_P(|M_t|) < \infty \ \forall t$ and
3.2. MARTINGALES

(ii) $E_P(M_t|\mathcal{F}_s) = M_s \forall s \leq t$.

*Note.* The first condition is merely a technical fact. The second condition implies that, under one measure, the expected value of the process $M$ in the future it is exactly the same as now. It is not expected to drift upwards or downwards.

**Example 3.2.2.** The constant process $X = k$ is a martingale with respect to any measure since

$$E_P(X|\mathcal{F}_s) = k = X$$

for any measure.

**Proposition 3.2.3.** A $P$-Brownian motion is a $P$-martingale.

*Proof.* It make sense to think that the Brownian motion does not move consistently up or down, it is as likely to do either. To check this formally, it is necessary to see that $E_P(W_t|\mathcal{F}_s) = W_s$. By the definition, it is known that $W_t - W_s$ is independent of $\mathcal{F}_s$ and distributed as $N(0, t - s)$. Thus

$$E_P(W_t|\mathcal{F}_s) = E_P(W_s|\mathcal{F}_s) + E_P(W_t - W_s|\mathcal{F}_s)$$

$$= W_s + 0$$

$$= W_s.$$

□

**Proposition 3.2.4.** The process $X_t = W_t + \gamma t$ where $W_t$ is a $P$-Brownian motion is a $P$-martingale if and only if $\gamma = 0$.

*Proof.* It is given that

$$E(X_t|\mathcal{F}_s) = E(W_t + \gamma t|\mathcal{F}_s)$$

$$= E(W_t|\mathcal{F}_s) + \gamma t$$

$$= W_s + \gamma t$$

$$= X_s + \gamma(t - s).$$

By the martingale second condition, $X_s = X_s + \gamma(t - s)$. Thus $X_s$ is a $P$-martingale if and only if $\gamma = 0$. □

**Lemma 3.2.5** (Normality criterion). Let $X$ be a random variable. If $X$ is distributed as a $N(\mu, \sigma^2)$, then $E_P(e^X) = e^{\mu + \frac{1}{2}\sigma^2}$. In particular, $E_P(e^{kX}) = e^{k\mu + \frac{1}{2}k^2\sigma^2} \forall k \in \mathbb{R}$. 
Proof.

\[
E_P(e^X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{x} e^{-(x-\mu)^2/2\sigma^2} \, dx
\]

\[
= e^{\mu + \frac{1}{2} \sigma^2} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-(x-\mu-\sigma^2)^2/2\sigma^2} \, dx
\]

\[
= e^{\mu + \frac{1}{2} \sigma^2}.
\]

The last equality holds since the integral is the integral over all \( \mathbb{R} \) of the density function of a normal distribution. \(\square\)

The Cameron-Martin-Girsanov theorem (CMG) shows what is happening to a Brownian motion when the probability measure changes.

**Theorem 3.2.6** (CMG theorem). Let \((\Omega, \mathcal{A}, P)\) be a probability space, \(W_t\) a \(P\)-Brownian motion and \(\mathcal{F}_t\) a filtration. Let \(\gamma_t\) be \(\mathcal{F}_t\)-predictable satisfying

\[
E_P\left( \exp \left( \frac{1}{2} \int_0^T \gamma_t^2 \, dt \right) \right) < \infty
\]

then, exists a probability measure \(Q\) such that

(i) \(Q\) and \(P\) are equivalent.

(ii) \( \frac{dQ}{dP} = \exp \left( - \int_0^T \gamma_t \, dW_t - \frac{1}{2} \int_0^T \gamma_t^2 \, dt \right) \).

(iii) The process \(W_t = W_t + \int_0^T \gamma_s \, ds\) is a \(Q\)-Brownian motion.

Somehow this theorem is useful to control and correct the stochastic differential drifts. In fact, there exists too the CMG reciprocal theorem.

**Theorem 3.2.7** (CMG reciprocal theorem). Let \(W_t\) be a \(P\)-Brownian motion and \(Q\) a probability measure equivalent to \(P\). Then, exists a process \(\gamma_t\) that is \(\mathcal{F}_t\)-predictable such that the process

\[
W_t = W_t + \int_0^t \gamma_s \, ds
\]

is a \(Q\)-Brownian motion. Moreover, the Radon-Nikodym derivative is given by

\[
\frac{dQ}{dP} = \exp \left( - \int_0^T \gamma_t \, dW_t - \frac{1}{2} \int_0^T \gamma_t^2 \, dt \right).
\]
Example 3.2.8. Let $X$ be the random process that satisfies the SDE
\[ dX_t = \sigma_t dW_t + \mu_t dt \]
with $W_t$ a $P$-Brownian motion. The goal is to find the probability measure $Q$ under whom the process satisfies
\[ dX_t = \sigma_t d\hat{W}_t + \nu_t dt. \]

If the process
\[ \gamma_t = \frac{\mu_t - \nu_t}{\sigma_t} \]
satisfies the $L^2$ condition, hence by the CMG Theorem exists a probability measure $Q$ such that
\[ \hat{W}_t = W_t + \int_0^t \frac{\mu_s \nu_s}{\sigma_s} ds \]
is a $Q$-Brownian motion. Then the equivalent SDE is given by
\[ d\hat{W}_t = dW_t + \frac{\mu_t - \nu_t}{\sigma_t} dt. \]

Thus, under $Q$
\[ dX_t = \sigma_t d\hat{W}_t + \nu_t dt. \]

This leads to the main result in martingales. The martingale representation theorem relates two non-trivial $P$-martingales scaling one of them respect to the other one.

**Theorem 3.2.9** (Martingale representation theorem). Suppose that $M_t$ is a $Q$-martingale process. Then if $N_t$ is any other $Q$-martingale, there exists a $\mathcal{F}$-predictable process $\phi$ such that
\[ \int_0^T \phi_s^2 ds < \infty \]
satisfying that $\forall t \in [0, T]$
\[ N_t = N_0 + \int_0^t \phi_s dM_s. \]

Or in the SDE notation
\[ dN_t = \phi_t dM_t. \]

Essentially this result claims that if there is a measure $Q$ under which $M_t$ is a $Q$-martingale, then any other $Q$-martingale can be represented in terms of $M_t$. 
Corollary 3.2.10. If \( M_t \) is a \( P \)-martingale, then \( M_t \) has no drift.

Proof. Since \( M_t \) is a \( P \)-martingale, using the martingale representation theorem with a \( P \)-Brownian motion \( W_t \) is given that

\[
dM_t = \phi_t dW_t
\]

where \( \phi_t \) is a \( \mathcal{F} \)-predictable process. \( \square \)

Finally there are two more criterion to take care of when calculating prices: the martingale criterion and the exponential martingale criterion.

Lemma 3.2.11 (Martingale criterion). If \( X \) is a process satisfying

\[
dx_t = \sigma_t dW_t + \mu_t dt,
\]

and

\[
E_P \left( \int_0^T \sigma_s^2 ds \right) < \infty,
\]

then, \( X \) is a \( P \)-martingale if and only if \( \mu_t = 0 \) \( \forall t \in [0,T] \).

Lemma 3.2.12 (Exponential martingale criterion). If \( dX_t = \sigma_t X_t dW_t \), for some \( \mathcal{F} \)-predictable process \( \sigma_t \), then

\[
E_P \left( \exp \left( \frac{1}{2} \int_0^T \sigma_s^2 ds \right) \right) \iff X \text{ is a martingale}.
\]

3.3 Neutral risk valuation

All the martingale theory seen is important for the neutral risk valuation. To price the derivatives, it is assumed that the implicit risk in all the transactions is common. There are no risk preferences. In a world where investors are risk neutral, the expected return on all investment assets is the risk-free rate of interest \( r \). The reason is that risk-neutral investors do not require a premium to induce them to take risks. This assumption does, therefore, considerably simplify the analysis of derivatives.

Note. It is important to appreciate that neutral risk valuation (or the assumption that all investors are risk neutral) is merely an artificial device for obtaining the derivative’s price. This assumption provides the investor the certainty that this will be the fair price. In the way that if the price is different there will be an arbitrage opportunity.

To use this model, there are few assumptions that must be stated:
(i) There are not transactions costs.

(ii) It can be taken at any moment a long or a short position.

The aim of the valuation is to price a derivative.

**Definition 3.3.1.** The value of a derivative at time \( t \) is a random variable \( D_t : \Omega \to \mathbb{R} \). The value of the derivative at the maturity \( T \) is the payoff of the derivative.

**Definition 3.3.2.** The non-dividend paying bond is described by the process \( B = \{B_t\}_{0 \leq t \leq T} \) where

\[
B_t = B_0 e^{rt}
\]

is its value at time \( t \) and \( r \) the risk-free rate of interest.

When pricing, there are two main questions:

1. The value of \( D_0 \).
2. The hedging strategy.

**Definition 3.3.3.** The hedging strategy of a derivative \( D \) is the portfolio \( \Pi \) that must hold the short position at any time \( t \in [0, T] \). The portfolio \( \Pi_t = (\Phi_t, \Psi_t) \) is a pair of \( \Phi_t \) units of underlying asset and \( \Psi_t \) units of bonds. Satisfying:

(i) \( \Phi = \{\Phi_t\}_{0 \leq t \leq T} \) and \( \Psi = \{\Psi_t\}_{0 \leq t \leq T} \) are \( \mathcal{F}_t \)-predictable processes.

(ii) If \( V_t = \Phi_t S_t + \Psi_t B_t \) is the value of the portfolio \( \Pi_t \) at time \( t \), then the process \( V \) satisfies the SDE

\[
dV_t = \Phi_t dS_t + \Psi_t dB_t.
\]

(iii) The portfolio value at maturity is the derivative payoff. I.e.

\[
V_T = \Phi_T S_T + \Psi_T B_T = D_T.
\]

To deduce the existence of hedging strategies it will be used the martingale representation theorem (Theorem 3.2.9). Moreover the updated prices must be considered.

**Definition 3.3.4.** The updated price process \( Z \) is defined as

\[
Z_t = B_t^{-1} S_t.
\]

The updated payoff is also considered, \( Z_T = B_T^{-1} D_T \).
Proposition 3.3.5. Given a derivative $D$ over an underlying asset $S$, there exists a hedging strategy given by the portfolio $\Pi_t = (\Phi_t, \Psi_t)$.

Proof. The construction of the hedging strategy is:

(i) Applying the CMG theorem, is taken the probability measure $Q$ such that $Z$ is a $Q$-martingale.

(ii) The process $\{E_t\}_{0 \leq t \leq T}$ is considered where $E_t = E_Q(B_T^{-1}D_T | F_t)$. This process is adapted to the filtration $F_t$ and by the Tower’s Law (Proposition 1.1.6) it is also a $Q$-martingale.

(iii) Applying the martingale representation theorem (Theorem 3.2.9), it is taken the predictable process $\Phi$ such that $dE_t = \Phi_t dZ_t \quad \forall t \in [0, T]$.

(iv) It is considered the process $\Pi, \Pi_t = (\Phi_t, \Psi_t), \Phi_t, \Psi_t : \Omega \to \mathbb{R}$, defined by $\Phi_t$ units of $S$ and $\Phi_t = E_t - \Phi_t Z_t$ bonds.

Now must be seen that this hedging strategy satisfies the three conditions of the hedging strategy’s definition:

(i) $\Phi_t$ is a predictable process for the martingale representation theorem. And also is the $\Psi_t$ process.

(ii) It can be observed that the value of $\Pi_t$ at time $t$ is

$$V_t = \Phi_t S_t + \Psi_t B_t$$

$$= \Phi_t S_t + (E_t - \Phi_t Z_t)B_t$$

$$= B_t E_t.$$

Therefore,

$$dV_t = d(B_tE_t)$$

$$= (dB_t)E_t + B_t(dE_t)$$

$$= (dB_t)(\Phi_t Z_t + \Psi_t) + B_t(\Phi_t dZ_t)$$

$$= \Phi_t ((dB_t)Z_t + B_t dZ_t) + \Psi_t B_t$$

$$= \Phi_t d(B_t Z_t) + \Psi_t dB_t$$

$$= \Phi_t dS_t + \Psi_t dB_t.$$

(iii) The value of $\Pi$ at maturity $T$ is

$$V_T = \Phi_T S_T + (E_T - \Phi_T Z_T)B_T$$

$$= E_T B_T$$

$$= B_T E_Q(B_T^{-1}D_T | F_T)$$

$$= D_T.$$
At this point if the free-arbitrage assumption is imposed, it is given that

\[
D_t = V_t \\
= B_t E_t \\
= B_t E_Q(B_T^{-1}D_T | \mathcal{F}_t) \\
= e^{-r(T-t)} E_Q(D_T | F_t),
\]

where \( Q \) is the probability measure that makes \( Z_t \) a \( Q \)-martingale.

**Note.** It can be observed that at time \( t \) it can be constructed a process \( \Pi_t \) that allows to reach the maturity replicating the derivative payoff with no risk.

**Corollary 3.3.6.** The premium of the derivative at time 0 is the updated expected payoff,

\[
D_0 = e^{-rT} E_Q(D_T).
\]

**Note.** The premium of the derivative is the updated expected payoff under the neutral risk measure. This means simply that under the measure \( Q \) a derivative is a fair play, in the sense that the price is set to be fair. When the people speculates with derivatives is using its own probability measure in the way that its expectation will be bigger. This fact can be due that they have privileged information (what is nowadays under control) or a personal intuition.

In conclusion, given \( S \) the price process of the underlying asset, \( B \) the bond value and \( D_T \) the derivative’s payoff at the maturity \( T \). To calculate \( D_0 \),

(i) construct \( Q \) such that \( Z_t = B_t^{-1}S_t \) is a \( Q \)-martingale,

(ii) consider the \( Q \)-martingale \( E_t = E_Q(B_T^{-1}D_T | \mathcal{F}_t) \),

(iii) take the \( \mathcal{F} \)-predictable process \( \Phi \) such that \( dE_t = \Phi_t dZ_t \) (via the martingale representation theorem),

(iv) consider the portfolio \( \Pi_t = (\Phi_t, E_t - \Phi_t Z_t) \).

(v) The derivative value at time \( t \) is \( D_t = B_tE_t \). In particular, \( D_0 = e^{-rT} E_Q(D_T). \)
3.4 Derivatives on geometric Brownian motion model

The path to price derivatives is different depending on the model chosen. In this section the underlying asset of the derivative is the process $S$ following a geometric Brownian motion, i.e. the process $S$ satisfy the SDE

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

where $\mu$ is drift term, $\sigma$ the volatility and $W_t$ is a $P$-Brownian motion. This is the common stock model also known as the Black-Scholes model. In fact, the Black-Scholes model contain a bond $B_t$ and a stock $S_t$ with SDEs

$$\begin{align*}
    dB_t &= rB_t dt, \\
    dS_t &= S_t (\mu dt + \sigma dW_t),
\end{align*}$$

where $r$ is the constant interest rate.

**Note.** First of all will be studied the case were $S_t$ is a non-dividend-paying stock. Then, this feature will be considered.

Following the steps stated in the previous section, the first step to price a derivative $D$ is to make the discounted stock price $Z_t = B_t^{-1} S_t$ into a martingale. That is, if

$$\hat{W}_t = W_t + \int_0^t \gamma ds = W_t + t\gamma$$

is a $Q$-Brownian motion, then $Z_t$ has SDE

$$dZ_t = Z_t (\sigma d\hat{W}_t + (\mu - r - \sigma \gamma)dt).$$

The equality holds by the CMG theorem seen in previous sections. Then, $Z$ is a $Q$-martingale if

$$\gamma = \frac{\mu - r}{\sigma}$$

due the exponential martingale criterion. This leads to a condition on $S$ for the neutral risk valuation.

**Proposition 3.4.1.** If $S_t$ follows the SDE

$$dS_t = S_t (r dt + \sigma dW_t)$$  \hspace{1cm} (3.24)

where $r$ is the free-risk interest rate, $\sigma$ the asset’s volatility and $W_t$ a $P$-Brownian motion. Then, under $P$, the discounted price process

$$Z_t = B_t^{-1} S_t$$

is a $P$-martingale.
The Black-Scholes formula for pricing European calls and European puts

The derivatives that will be priced are the European call and the European put options. First of all will be calculated the call price through the neutral risk valuation method and then will be calculated the put price through the put-call parity (which easier and faster). The following result is one of the most important results in the financial modelling field.

Proposition 3.4.2 (European call price). If $c_t$ is the price of a European call at time $t$ of an non-dividend-paying underlying asset $S$ following the SDE (3.24) with strike $X$ at maturity time $T$, let be

$$d_1 = \frac{1}{\sigma \sqrt{T-t}} \left( \ln \left( \frac{S_t}{X} \right) + \left( r + \frac{1}{2} \sigma^2 \right) (T-t) \right),$$

(3.25)

$$d_2 = \frac{1}{\sigma \sqrt{T-t}} \left( \ln \left( \frac{S_t}{X} \right) + \left( r - \frac{1}{2} \sigma^2 \right) (T-t) \right),$$

(3.26)

and let $N(x)$ be the probability distribution function of a $N(0,1)$, i.e.

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-v^2/2} dv.$$

Thus

$$c_t = S_t N(d_1) - X e^{-r(T-t)} N(d_2).$$

(3.27)

Proof. First of all will be calculated the value of $c_0$. Then the $c_t$ value is the same that $c_0$ if instead of maturity $T$ the maturity taken is $T - t$.

The payoff of the European call is given by

$$D_T = c_T = \max(S_T - X, 0) = (S_T - X)^+.$$

At time $t$, the call value is

$$c_0 = V_0 = B_0 E_0 = B_0 E_Q(B_T^{-1} c_T | \mathcal{F}_0) = e^{-rT} E_Q((S_T - X)^+).$$

Under the probability measure $Q$ the value of $S_t$ is given by

$$S_t = S_0 \exp \left( \left( r - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \right).$$
with \( W_t \) a \( Q \)-Brownian motion. Thus,

\[
S_T = S_0 \exp \left( \left( r - \frac{1}{2} \sigma^2 \right) T + \sigma W_T \right) \\
= S_0 \exp(rT + Y),
\]
with \( Y = -\frac{1}{2} \sigma^2 T + \sigma W_t \sim N\left( -\frac{1}{2} \sigma^2 T, \sigma^2 T \right) \) under \( Q \). Hence

\[
c_0 = e^{-rT} E_Q \left( (S_0 \exp(rT + Y) - X)^+ \right) \\
= E_Q \left( (S_0 e^Y - X e^{-rT})^+ \right) \\
= E_Q \left( f(Y) \right)
\]
where \( f : \mathbb{R} \to \mathbb{R} \) is defined by

\[
f(y) = \left( S_0 e^y - X e^{-rT} \right)^+.
\]
Using that the probability distribution function of \( N(\mu, \sigma^2) \) is

\[
g(y) = \frac{1}{\sqrt{2\pi \sigma^2}} \exp \left( -\frac{(y - \mu)^2}{2 \sigma^2} \right)
\]
therefore

\[
c_0 = E_Q \left( f(Y) \right) \\
= \frac{1}{\sqrt{2\pi \sigma^2 T}} \int_{-\infty}^{\infty} f(y) \exp \left( -\frac{(y + \frac{1}{2} \sigma^2 T)^2}{2 \sigma^2 T} \right) dy.
\]
Then, since

\[
f(y) = 0 \iff S_0 e^y - X e^{-rT} \leq 0 \\
\iff e^y \leq \frac{X}{S_0} e^{-rT} \\
\iff y \leq \ln \left( \frac{X}{S_0} \right) - rT,
\]
then

\[
c_0 = \frac{1}{\sqrt{2\pi \sigma^2 T}} \int_{\ln \left( \frac{X}{S_0} \right) - rT}^{\infty} \left( S_0 e^y - X e^{-rT} \right) \exp \left( -\frac{(y + \frac{1}{2} \sigma^2 T)^2}{2 \sigma^2 T} \right) dy.
\]
Using the change of variable \( v = -\frac{1}{\sigma \sqrt{T}} \left( y + \frac{1}{2} \sigma^2 T \right) \), then, the changes are

(i) \( y = -\sigma \sqrt{T} v - \frac{1}{2} \sigma^2 T \),

(ii) \( \ln \left( \frac{X}{S_0} \right) - rT \leq y < \infty \iff \ln \left( \frac{X}{S_0} \right) - rT \leq -\sigma \sqrt{T} v - \frac{1}{2} \sigma^2 T < \infty \\
\iff \ln \left( \frac{X}{S_0} \right) - \left( r - \frac{1}{2} \sigma^2 \right) T \leq -\sigma \sqrt{T} v < \infty \\
\iff \frac{1}{\sigma \sqrt{T}} \left( \ln \left( \frac{S_0}{X} \right) + \left( r - \frac{1}{2} \sigma^2 \right) T \right) \geq v > -\infty, \)
(iii) $S_0 e^y - X e^{-rT} = S_0 \exp \left( -\sigma \sqrt{T} v - \frac{1}{2} \sigma^2 T \right) - X e^{-rT}$,

(iv) $-\frac{(y + \frac{1}{2} \sigma^2 T)^2}{2\sigma^2 T} = -\frac{1}{2} v^2$,

(v) $dy = -\sigma \sqrt{T} dv$.

Then,

$$c_0 = \frac{-\sigma \sqrt{T}}{\sqrt{2\pi}\sigma^2 Tr} \int_{-\infty}^{-d_2} e^{-\frac{1}{2}(\ln \left( \frac{S_0}{X} \right) + (r - \frac{1}{2} \sigma^2) T)} \left( S_0 e^{-\sigma \sqrt{T} v - \frac{1}{2} \sigma^2 T} - X e^{-rT} \right) e^{-\frac{1}{2}v^2} dv$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-d_2} e^{-\frac{1}{2}(\ln \left( \frac{S_0}{X} \right) + (r - \frac{1}{2} \sigma^2) T)} \left( S_0 e^{-\sigma \sqrt{T} v - \frac{1}{2} \sigma^2 T} - X e^{-rT} \right) e^{-\frac{1}{2}v^2} dv$$

$$= S_0 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-d_2} \exp \left( -\sigma \sqrt{T} v - \frac{1}{2} \sigma^2 T - \frac{1}{2} v^2 \right) dv - X e^{-rT} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_2} e^{-\frac{1}{2}v^2} dv$$

where $d_2 = \frac{1}{\sigma \sqrt{T}} \left( \ln \left( \frac{S_0}{X} \right) + (r - \frac{1}{2} \sigma^2) T \right)$. The first integral may be written as

$$\int_{-\infty}^{d_2} \exp \left( -\sigma \sqrt{T} v - \frac{1}{2} \sigma^2 T - \frac{1}{2} v^2 \right) dv = \int_{-\infty}^{d_2} \exp \left( -\frac{1}{2}(v + \sigma \sqrt{T})^2 \right) dv.$$

Substituting

$$c_0 = S_0 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-d_2} e^{-\frac{1}{2}(v + \sigma \sqrt{T})^2} dv - X e^{-rT} \int_{-\infty}^{d_2} e^{-\frac{1}{2}v^2} dv$$

$$= S_0 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_2+\sigma \sqrt{T}} e^{-\frac{1}{2}w^2} dw - X e^{-rT} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_2} e^{-\frac{1}{2}v^2} dv$$

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

where $d_1 = d_2 + \sigma \sqrt{T} = \frac{1}{\sigma \sqrt{T}} \left( \ln \left( \frac{S_0}{X} \right) + (r + \frac{1}{2} \sigma^2) T \right)$. $N(x) = \int_{-\infty}^{x} e^{-\frac{1}{2}v^2} dv$ is the distribution function of a normal distribution $N(0, 1)$ and $w = v + \sigma \sqrt{T}$. □

Once calculated the call price, it is simple to calculate the European put price.
Corollary 3.4.3. If \( p_t \) is the price of a European put at time \( t \) of an non-dividend-paying underlying asset \( S \) following the SDE (3.24) with strike \( X \) at maturity time \( T \), let be

\[
d_1 = \frac{1}{\sigma \sqrt{T-t}} \left( \ln \left( \frac{S_t}{X} \right) + \left( r + \frac{1}{2} \sigma^2 \right)(T-t) \right),
\]

\[
d_2 = \frac{1}{\sigma \sqrt{T-t}} \left( \ln \left( \frac{S_t}{X} \right) + \left( r - \frac{1}{2} \sigma^2 \right)(T-t) \right),
\]

and let \( N(x) \) be the probability distribution function of a \( N(0,1) \), i.e.

\[
N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{v^2}{2}} dv.
\]

Thus

\[
p_t = X e^{-r(T-t)} N(-d_2) - S_t N(-d_1).
\]

Proof. Since

\[
c_t = S_t N(d_1) - X e^{-r(T-t)},
\]

from put-call parity (3.5),

\[
c_t - p_t = S_t - X e^{-r(T-t)}.
\]

Thus,

\[
p_t = c_t - S_t + X e^{-r(T-t)}
\]

\[
= X e^{-r(T-t)} (1 - N(d_2)) - S_t (1 - N(d_1))
\]

\[
= X e^{-r(T-t)} N(-d_2) - S_t N(-d_1).
\]

□

Once the price of the European call and the European put has been calculated, the next step is to calculate the hedging strategy.

Proposition 3.4.4. The hedging strategy for an European call, \( c_t \), is given by the portfolio process

\[
\Pi_t \left( N(d_1), c_t - N(d_1)S_t \right).
\]

Proof. Recall that the underlying asset participation or shares for a derivative hedging can be calculated as

\[
\Pi_t = \frac{\partial D_t}{\partial S}.
\]
Then, in the case of the European call,

$$\Phi_t = \frac{\partial c_t}{\partial S} = N(d_1) + S N'(d_1) \frac{\partial d_1}{\partial S} - X e^{-r(T-t)} N'(d_2) \frac{\partial d_2}{\partial S},$$

where

$$N'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

$$d_1^2 = \frac{1}{\sigma^2(T-t)} \left( \ln^2 \left( \frac{S_t}{X} \right) + \left( r + \frac{1}{2} \sigma^2 \right)^2 (T-t)^2 + 2 \ln \left( \frac{S_t}{X} \right) \left( r + \frac{1}{2} \sigma^2 \right) (T-t) \right),$$

$$d_2^2 = \frac{1}{\sigma^2(T-t)} \left( \ln^2 \left( \frac{S_t}{X} \right) + \left( r - \frac{1}{2} \sigma^2 \right)^2 (T-t)^2 + 2 \ln \left( \frac{S_t}{X} \right) \left( r - \frac{1}{2} \sigma^2 \right) (T-t) \right),$$

Thus

$$\frac{N'(d_1)}{N'(d_2)} = \frac{\exp \left( -\frac{1}{2} d_1^2 \right)}{\exp \left( -\frac{1}{2} d_2^2 \right)} = \exp \left( -\frac{1}{2} (d_1^2 - d_2^2) \right) = \exp \left( -\frac{1}{2 \sigma^2 (T-t)} \frac{1}{2 r \sigma^2 (T-t)^2 + 2 \ln \left( \frac{S_t}{X} \right) \sigma^2 (T-t)} \right) = \exp \left( -r(T-t) - \ln \left( \frac{S_t}{X} \right) \right) = \exp \left( -r(T-t) \right) \frac{X}{S}.$$}

Hence,

$$SN'(d_1) - X e^{-r(T-t)} N'(d_2) = 0.$$

This can be used in the first expression since

$$\frac{\partial d_1}{\partial S} = \frac{\partial d_2}{\partial S} = \frac{1}{\sigma \sqrt{T-t}} \frac{1}{S},$$

then

$$SN'(d_1) \frac{\partial d_1}{\partial S} - X e^{-r(T-t)} N'(d_2) \frac{\partial d_2}{\partial S} = 0,$$

and finally

$$\Phi_t = N(d_1).$$

The bond part of the portfolio is proved through the definition. $\square$

Before introducing the dividend-paying options (and consequently the options on commodities), it is interesting to see what $N(d_2)$ represents.
Proposition 3.4.5. Under the neutral risk valuation probability measure \( Q \), the probability of exercising the European call option \( c_t \) is given by \( N(d_2) \).

Proof. From (2.5) it is given that

\[
\ln(S_T) = \ln(S_t) + \left( r - \frac{1}{2} \sigma^2 \right) (T - t) + \sigma(W_T - W_t)
\]

where \( W_t \) is a \( Q \)-Brownian motion. Note that \( \ln(S_T) = Y \sim N(\lambda, \rho^2) \) where \( \lambda = \ln(S_t) + \left( r - \frac{1}{2} \sigma^2 \right) (T - t) \) and \( \rho^2 = \sigma^2 (T - t) \). Then,

\[
P_Q(S_T > X) = P_Q(\ln(S_T) > \ln(X))
= P_Q(Y > \ln(X))
= \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} \int_{\ln(X)}^{\infty} \exp \left( -\frac{(x - \lambda)^2}{2\sigma^2(T-t)} \right) dx.
\]

Using the change of variable \( v = -\frac{1}{\sigma\sqrt{T-t}}(x - \lambda) \), then

(i) \( x = -\sigma\sqrt{T-t} \lambda + \lambda \),

(ii)

\[
\ln(X) \leq x \leq \infty \Leftrightarrow \ln(X) \leq -\sigma\sqrt{T-t} \lambda + \lambda < \infty
\Rightarrow \ln(X) - \lambda \leq -\sigma\sqrt{T-t} < \infty
\Rightarrow \frac{1}{\sigma\sqrt{T-t}}(\lambda - \ln(X)) \geq v \geq -\infty,
\]

(iii) \( -\frac{(x - \lambda)^2}{2\sigma^2(T-t)} = -\frac{1}{2} v^2 \),

(iv) \( dx = -\sigma\sqrt{T-t}dv \).

Substituting into the integral,

\[
P_Q(S_T > X) = \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} \int_{\ln(X)}^{\infty} \exp \left( -\frac{(x - \lambda)^2}{2\sigma^2(T-t)} \right) dx
= -\frac{\sigma\sqrt{T-t}}{\sqrt{2\pi\sigma^2(T-t)}} \int_{\frac{1}{\sigma\sqrt{T-t}}(\lambda - \ln(X))}^{\infty} e^{-\frac{1}{2} v^2} dv
= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{1}{\sigma\sqrt{T-t}}(\lambda - \ln(X))} e^{-\frac{1}{2} v^2} dv
= N \left( \frac{1}{\sigma\sqrt{T-t}}(\lambda - \ln(X)) \right).
\]
Returning to $d_2$ definition,

$$d_2 = \frac{1}{\sigma \sqrt{T-t}} \left( \ln(S_0) + \left( r - \frac{1}{2} \sigma^2 \right) (T-t) - \ln(X) \right)$$

$$= \frac{1}{\sigma \sqrt{T-t}} (\lambda - \ln(X)).$$

□

Once known the interpretation of $N(d_1)$ and $N(d_2)$, the call premium $c_t$ may be interpreted as the hedging price $S_t N(d_1)$ minus the strike price updated at time $t$ per the exercise probability $X e^{-r(T-t)} N(d_2)$, this term in fact is the expected benefit at time $t$. This second term may be interpreted as the proportional cost that will have the call at the maturity.

In the case of the European put, the terms $N(-d_1)$ and $N(-d_2)$ have a similar meaning.

**Corollary 3.4.6.** The hedging strategy for an European put, $p_t$, is given by the portfolio process

$$\Pi_t \{ N(-d_1), p_t - N(-d_1) S_t \}.$$

**Proof.** The proof is analogous to the hedging strategy for an European call. The main difference is the $S_t$ term, in this case instead of $N(d_1)$ is $N(-d_1)$. □

**Corollary 3.4.7.** Under the neutral risk valuation probability measure $Q$, the probability of exercising the European put option $p_t$ is given by $N(-d_2)$.

**Proof.** Since

$$N(-d_2) = 1 - N(d_2),$$

and

$$N(d_2) = P_Q(S_T > X),$$

recall that the probability of exercising an European put is given by $P_Q(S_T < X)$, then

$$P_Q(S_T < X) = 1 - P_Q(S_T > X)$$

$$= 1 - N(d_2)$$

$$= N(-d_2).$$

□

As in the European call, the put may be interpreted in the same terms. The put premium $p_t$ is the expected benefit updated at time $t$, that is given by $X e^{-r(T-t)} N(-d_2)$ minus the hedging cost (that in this case is afforded by the short position) $S_t N(-d_1)$. 


The Merton formula for dividend-paying stocks and commodities

In 1973 Merton developed a formula to price dividend-paying stocks. This was a big step for the pricing of commodities since the main difference between common stocks and commodities are the convenience yield and the storage costs. In this case, Merton formula allows to include the convenience yield as a particular case of dividend payment in a similar way as in forward pricing.

In this model, all assumptions of the Black-Scholes model are kept except that the underlying asset is supposed to report a continuous dividend payment at the rate $g$. I.e. the owner of a stock over the period $(t, t + dt)$ receives at date $t + dt$ the dividend $gS_tdt$.

**Proposition 3.4.8** (European call price with dividends). If $c_t$ is the price of a European call at time $t$ of an dividend-paying underlying asset $S$ at rate $g$ following the SDE (3.24) with strike $X$ at maturity time $T$, let be

$$d_1 = \frac{1}{\sigma \sqrt{T - t}} \left( \ln \left( \frac{S_t}{X} \right) + \left( r - g + \frac{1}{2} \sigma^2 \right) (T - t) \right),$$

$$d_2 = \frac{1}{\sigma \sqrt{T - t}} \left( \ln \left( \frac{S_t}{X} \right) + \left( r - g - \frac{1}{2} \sigma^2 \right) (T - t) \right),$$

and let $N(x)$ be the probability distribution function of a $N(0,1)$, i.e.

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{v^2}{2}} dv.$$

Thus

$$c_t = S_t e^{-g(T-t)} N(d_1) - X e^{-r(T-t)} N(d_2).$$

**Proof.** Since $X$ is fixed, it may be written that the call cost is

$$\begin{cases}
c_t = c(t, S_t), \\
\frac{dS_t}{S_t} = \mu dt + \sigma dW_t.
\end{cases}$$

Using the Itô’s lemma

$$dc = \left( \frac{\partial c}{\partial t} + \mu S_t \frac{\partial c}{\partial S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 c}{\partial S_t^2} \right) dt + \sigma S_t \frac{\partial c}{\partial S_t} dW_t,$$

and building the portfolio $\Pi$ at date $t$

$$V(t, S_t) = c(t, S_t) + nS_t.$$
\[ dV = \left( \frac{\partial c}{\partial t} + \mu S_t \frac{\partial c}{\partial S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 c}{\partial S_t^2} \right) dt + \sigma S_t \frac{\partial c}{\partial S_t} dW_t + ndS_t \]
\[ = \left( \frac{\partial c}{\partial t} + \mu S_t \frac{\partial c}{\partial S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 c}{\partial S_t^2} + n \mu S_t \right) dt + \left( \sigma S_t \frac{\partial c}{\partial S_t} + n \sigma S_t \right) dW_t. \]

Applying the neutral risk valuation method, the hedging portfolio has to be predictable. I.e. \( dV \) has to contain no random term. Then, the \( n \) chosen will be
\[ n = -\frac{\partial c}{\partial S_t}. \]

Then,
\[ dV = \left( \frac{\partial c}{\partial S_t} + \frac{1}{2} \sigma^2 S_t \frac{\partial^2 c}{\partial S_t^2} \right) dt. \]

The expected return of the portfolio \( V \) must be the free-risk rate of interest minus the dividend part, since the writer holds the underlying asset paying the known dividend rate \( g \). Then, by the neutral risk valuation imposition, it is given that
\[ dV = rV dt - gS_t dt. \]

Then yields,
\[ \frac{\partial c}{\partial S_t} + \frac{1}{2} \sigma^2 S_t \frac{\partial^2 c}{\partial S_t^2} = rV - gS_t, \]
\[ \frac{\partial c}{\partial S_t} + \frac{1}{2} \sigma^2 S_t \frac{\partial^2 c}{\partial S_t^2} - rV + gS_t = 0. \]

This PDE has the boundary condition
\[ C(T, S_T) = \max(S_T - X, 0). \]

The PDE is well defined and has only one solution. It can be checked then that the solution is
\[ c_t = S_t e^{-g(T-t)} N(d_1) - X e^{-r(T-t)} N(d_2). \]

\[ \square \]

Note. In fact, when including dividend payment the behaviour of the underlying asset has not changed at all. Then, it is coherent that the pricing formula changes slightly from the Black-Scholes formula. Actually, the only change is to discount the dividend part to the underlying asset price.

Note. With this change, it can be proved that the interpretations of \( N(d_1) \) and \( N(d_2) \) are still the same. Although might be token in account that the definition of \( d_1 \) and \( d_2 \) has changed.
Proposition 3.4.9 (European put price with dividends). If $p_t$ is the price of a European put at time $t$ of an dividend-paying underlying asset $S$ at rate $g$ following the SDE (3.24) with strike $X$ at maturity time $T$, let be

$$d_1 = \frac{1}{\sigma \sqrt{T-t}} \left( \ln \left( \frac{S_t}{X} \right) + \left( r - g + \frac{1}{2} \sigma^2 \right) (T-t) \right),$$

$$d_2 = \frac{1}{\sigma \sqrt{T-t}} \left( \ln \left( \frac{S_t}{X} \right) + \left( r - g - \frac{1}{2} \sigma^2 \right) (T-t) \right),$$

and let $N(x)$ be the probability distribution function of a $N(0,1)$, i.e.

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{v^2}{2}} dv.$$ 

Thus

$$-p_t = X e^{-r(T-t)} N(-d_2) - S_t e^{-g(T-t)} N(-d_1).$$

Proof. Adding the dividend-paying condition at $S_t$ and recalling the dividend paying forward equation (3.9), the put-call parity for dividend-paying assets is now

$$c_t - p_t = S_t e^{-g(T-t)} - X e^{-r(T-t)}.$$ 

Then, the put price can be calculated as

$$p_t = c_t - S_t e^{-g(T-t)} + X e^{-r(T-t)}$$

$$= X e^{-r(T-t)} (1 - N(d_2)) - S_t e^{-g(T-t)} (1 - N(d_1))$$

$$= X e^{-r(T-t)} N(-d_2) - S_t e^{-g(T-t)} N(-d_1).$$

□

A fundamental and immediate consequence of the Merton formula is that it provides the valuation of European calls written on commodity spots. It has been seen that the commodity price behaviour is the same as an underlying asset paying dividends where the dividend rate $g$ becomes the convenience yield $y$.

Note. One of the main handicaps for this model, especially when taking far maturities, is that it has been assumed constant interest rates and convenience yields. In distant maturities it becomes necessary to introduce a stochastic interest rate and convenience yield to represent the changing supply/demand conditions in the underlying commodity and the changes in the market. Normally, the stochastic convenience yield or interest rate is modelled as an Ornstein-Uhlenbeck process (as in Heston model).
3.5 Derivatives on mean-reversion model

In this section will be studied the case of the mean-reversion modelled assets. Recalling the mean-reversion model, the underlying asset $S_t$ follows the SDE

$$\frac{dS_t}{S_t} = k(\theta - \ln(S_t))dt + \sigma dW_t.$$ 

There exists an Ornstein-Uhlenbeck process $U_t$ such that

$$S_t = e^{U_t},$$

actually, the process $U_t$ satisfies the SDE

$$dU_t = k\left(\theta - \sigma^2 \frac{2}{k} - U_t\right)dt + \sigma dW_t.$$

As in the Black-Scholes model (or the Merton model) the bond price $B_t$ follows the same SDE (equation (3.22)). Also as before, the dividend-paying feature will be considered lately.

The following step, is to make the discounted price process $Z_t$ into a martingale. In this process will appear difficulties. This leads to two different methods to price this kind of options. As in the Black-Scholes case, to turn $Z_t$ into a martingale the drift term must be zero. The discounted price process it is given by

$$Z_t = B_t^{-1}S_t.$$ 

Then, applying Itô’s lemma

$$dZ_t = dB_t^{-1}S_t = dB_t^{-1}S_t + B_t^{-1}dS_t$$

$$= -rB_t^{-1}S_t dt + B_t^{-1}S_t k(\theta - \ln(S_t)) dt + S_t \sigma dW_t$$

$$= B_t^{-1}S_t \left(-r dt + S_t k(\theta - \ln(S_t)) dt + \sigma dW_t\right)$$

$$= Z_t \left(-r dt + S_t k(\theta - \ln(S_t)) dt + \sigma dW_t\right).$$

Since $Z_t$ is wanted to be a martingale, by the Corollary 3.2.10 thus

$$r = k(\theta - \ln(S_t)).$$

In the Black-Scholes model case, this condition was $r = \mu$. Impose $Z_t$ to be a martingale through this condition is a rough path.

The exact price calculation for mean-reversion models can be reached through tough calculus (see [7]) or through characteristic functions and the Fourier transform (tough calculus too, see [8]). To simplify the calculus and see other pricing methods, the methods that will be used are:

(i) The trinomial tree.

(ii) The Montecarlo method.
The trinomial tree

The binomial tree is a good method to price derivatives on geometric Brownian motion model in a discrete world. When using the mean-reversion model, it often proves to be convenient to use a trinomial rather than a binomial tree. The main advantage of a trinomial tree is that it provides an extra degree of freedom, this helps the tree to represent the mean-reverting property. There are several ways to implement the trinomial tree, in this work will be followed the approach of Hull (see [2, Section 30.7]). The construction of the trinomial tree is a complex method, therefore first will be shown how to construct a trinomial tree for an Ornstein-Uhlenbeck process $Z_t$ and then the method will be extended to exponential Ornstein-Uhlenbeck process $S_t$ (what was called before a mean-reversion process).

The construction of the binomial tree has two steps to follow. The first one is to construct a symmetric tree around the zero. The SDE followed by the process is

$$dX = -aXdt + \sigma dW_t.$$ 

This is an Ornstein-Uhlenbeck process that mean-reverts to zero. The spacing of $X$ in the tree is set, by Hull\(^1\), as

$$\Delta X = \sigma \sqrt{3 \Delta t}.$$ 

This election provides a faster convergence for the method.

Let $(i, j)$ be the node such that $t = i\Delta t$ and $X = j\Delta X$. Not every node will have the same probabilities for going up, down or taking the middle path. Because of the mean-reverting feature, the trinomial tree can be branched in three different ways.

This different branching methods are because of the mean-reverting property. In nodes far from the zero (which is the mean-reverting value), the attraction may be so high that one of the path probabilities may be negative. This may be understood that since there are “only” three movement option, and one is against the equilibrium, the mean-reverting force may be so big that the three most common sense paths are the ones that move towards the equilibrium. When $a > 0$, it is necessary to switch the branching method for sufficiently large (positive or negative) $j$. The branching method will be changed when one of the probabilities gets negative. Hull and White show that probabilities are always positive if $j_{\text{max}}$ is set equal to the smallest integer greater than

$$\frac{0.184}{a\Delta t}.$$ 

By the symmetric nature of the tree, \( j_{\text{min}} \) is set equal to \(-j_{\text{max}}\). The probabilities of transition of each node are given by \( p_u(i, j) \), \( p_m(i, j) \) and \( p_d(i, j) \) for taking the up, middle or down path respectively. To calculate these probabilities are set three conditions:

(i) They must sum the unity.

(ii) They must match the mean of \( X \).

(iii) They must match the variance of \( X \).

When using the branching method 3.1a, this may be written as

\[
\begin{align*}
    p_u + p_m + p_d &= 1, \\
    p_u \Delta X - p_d \Delta X &= -a_j \Delta X \Delta t, \\
    p_u (\Delta X)^2 + p_d (\Delta X)^2 &= \sigma^2 \Delta t + a^2_j^2 (\Delta X)^2 (\Delta t)^2.
\end{align*}
\]

for all node \((i, j)\). Using the expression for \( \Delta X \), the solutions are given by

\[
\begin{align*}
    p_u &= \frac{1}{6} + \frac{1}{2} (a^2_j^2 (\Delta t)^2 - a_j \Delta t), \\
    p_m &= \frac{2}{3} - a^2_j^2 (\Delta t)^2, \\
    p_d &= \frac{1}{6} + \frac{1}{2} (a^2_j^2 (\Delta t)^2 + a_j \Delta t).
\end{align*}
\]  

(3.37)

If the branching method is the given by 3.1b, the solutions are given thus by

\[
\begin{align*}
    p_u &= \frac{1}{6} + \frac{1}{2} (a^2_j^2 (\Delta t)^2 + a_j \Delta t), \\
    p_m &= -\frac{1}{3} - a^2_j^2 (\Delta t)^2 - 2a_j \Delta t, \\
    p_d &= \frac{7}{6} + \frac{1}{2} (a^2_j^2 (\Delta t)^2 + 3a_j \Delta t).
\end{align*}
\]  

(3.38)
If the branching method is given by 3.1c, the solution are given hence by

\[
\begin{align*}
    p_u &= \frac{7}{6} + \frac{1}{2} (a^2 j^2 (\Delta t)^2 - 3aj\Delta t), \\
    p_m &= -\frac{1}{3} - a^2 j^2 (\Delta t)^2 + 2aj\Delta t, \\
    p_d &= \frac{1}{6} + \frac{1}{2} (a^2 j^2 (\Delta t)^2 - aj\Delta t).
\end{align*}
\]

(3.39)

Note. The probabilities at one node only depend on \( j \). Furthermore, the tree is symmetrical.

**Example 3.5.1.** To illustrate this first step of the tree construction, suppose that \( \sigma = 0.01 \), \( a = 0.1 \), and \( \Delta t = 1 \) year. In this case,

\[
\Delta R = 0.0173, \\
j_{\text{min}} = -2, \\
j_{\text{max}} = 2.
\]

And the probabilities of the tree are given in the Table 3.7. The tree is set as the 3.2.

![Tree for X, first step](image.png)

The second step is, starting as of the symmetric tree, to build a tree that represents the original process. For this, the nodes will be displaced by \( \alpha(t) \) where

\[
\alpha(t) = Z_t - X_t.
\]

(3.40)
3.5. DERIVATIVES ON MEAN-REVERSION MODEL

<table>
<thead>
<tr>
<th>Node</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>X (%)</td>
<td>0.000</td>
<td>1.732</td>
<td>0.000</td>
<td>-1.732</td>
<td>3.464</td>
<td>1.732</td>
<td>0.000</td>
<td>-1.732</td>
<td>-3.464</td>
</tr>
<tr>
<td>( p_u )</td>
<td>0.1667</td>
<td>0.1217</td>
<td>0.1667</td>
<td>0.2217</td>
<td>0.8867</td>
<td>0.1217</td>
<td>0.1667</td>
<td>0.2217</td>
<td>0.0867</td>
</tr>
<tr>
<td>( p_m )</td>
<td>0.6666</td>
<td>0.6566</td>
<td>0.6666</td>
<td>0.6566</td>
<td>0.0266</td>
<td>0.6566</td>
<td>0.6666</td>
<td>0.6566</td>
<td>0.0266</td>
</tr>
<tr>
<td>( p_d )</td>
<td>0.1667</td>
<td>0.2217</td>
<td>0.1667</td>
<td>0.1217</td>
<td>0.0867</td>
<td>0.2217</td>
<td>0.1667</td>
<td>0.1217</td>
<td>0.8867</td>
</tr>
</tbody>
</table>

Table 3.7: Probabilities on the tree

With

\[
dZ_t = (\theta - aZ_t)dt + \sigma dW_t.
\]

This difference \( \alpha \) adapts the tree for \( X_t \) to the tree for \( Z_t \). To make easier the notation, \( \alpha(i \Delta t) \) will be written as \( \alpha_i \). Then, it is necessary to define another variable. Define \( Q_{i,j} \) as the present value of a security that pays off 1 if node \((i,j)\) is reached and zero otherwise. The \( \alpha_i \) and \( Q_{i,j} \) are calculated using forward induction in such a way that the initial term structure is matched exactly. Although is less clear, these news variables can be calculated analytically by

\[
\alpha_m = \frac{\ln \left( \sum_{j=-n_m}^{n_m} Q_{m,j} \exp(-jX\Delta t) \right) - \ln(P_{m+1})}{\Delta t}, \tag{3.41}
\]

were \( n_m \) is the number of nodes on each side of the central node at time \( m\Delta t \) and with

\[
P_{m+1} = \sum_{j=-n_m}^{n_m} n_m Q_{m,j} \exp \left( - (\alpha_m + jX)\Delta t \right). \tag{3.42}
\]

And once \( \alpha_m \) is calculated, then

\[
Q_{m+1,j} = \sum_{k} Q_{m,k} q(k,j) \exp \left( - (\alpha_m + kX)\Delta t \right). \tag{3.43}
\]

Example 3.5.2. To illustrate this second step, and with the same tree that in the previous example, suppose that the continuously compounded zero rates are as shown in Table 3.8.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>3.430</td>
</tr>
<tr>
<td>1.0</td>
<td>3.824</td>
</tr>
<tr>
<td>1.5</td>
<td>4.183</td>
</tr>
<tr>
<td>2.0</td>
<td>4.512</td>
</tr>
<tr>
<td>2.5</td>
<td>4.812</td>
</tr>
<tr>
<td>3.0</td>
<td>5.086</td>
</tr>
</tbody>
</table>

Table 3.8: Zero rates

The value of \( Q_{0,0} \) is 1.0. The value of \( \alpha_0 \) is chosen to give the right price for a zero-coupon bond maturing at time \( \Delta t \). That is, \( \alpha_0 \) is set equal to the initial
\(\Delta t\)-period interest rate. Because \(\Delta t = 1\) in this example, \(\alpha_0 = 0.03824\). This defines the position of the initial node on the \(Z\)-tree in Figure 3.3. The next step is to calculate the values of \(Q_{1,1}\), \(Q_{1,0}\) and \(Q_{1,-1}\). There is a probability of 0.1667 that the \((1, 1)\) node is reached and the discount rate for the first time step is 3.82\%. The value of \(Q_{1,1}\) is therefore \(0.1667e^{-0.0382} = 0.1604\). Similarly, \(Q_{1,0} = 0.6417\) and \(Q_{1,-1} = 0.1604\).

Once \(Q_{1,1}\), \(Q_{1,0}\) and \(Q_{1,-1}\) have been calculated, \(\alpha_1\) must be determined. It is chosen to give the right price for a zero-coupon bond maturing at time \(2\Delta t\). Since \(\Delta X = 0.01732\) and \(\Delta t = 1\), the price of this bond as seen at node \(B\) is \(e^{-(\alpha_1+0.01732)}\), as seen at node \(C\) is \(e^{-\alpha_1}\) and as seen at node \(D\) is \(e^{-(\alpha_1-0.01732)}\). Thus, the price as seen at the initial node \(A\) is given by

\[
Q_{1,1}e^{-(\alpha_1+0.01732)} + Q_{1,0}e^{-\alpha_1} + Q_{1,-1}e^{-(\alpha_1-0.01732)}.
\]

From the initial term structure, this bond price should be \(e^{-0.04512/2} = 0.9137\). Substituting in the previous equation,

\[
0.1604e^{-(\alpha_1+0.01732)} + 0.6417e^{-\alpha_1} + 0.1604e^{-(\alpha_1-0.01732)}.
\]

This equation has solution

\[
\alpha_1 = 0.05205.
\]

This means that the central node at time \(\Delta t\) in the tree for \(Z\) is equal to 5.205\%. The following step would be to calculate \(Q_{2,1}\), \(Q_{2,0}\), \(Q_{2,-1}\), \(Q_{2,-2}\) and \(\alpha_2\) in the same way and so on.

---

**Figure 3.3: Tree for \(Z\), second step**
3.5. DERIVATIVES ON MEAN-REVERSION MODEL

Table 3.9: Probabilities on the tree

<table>
<thead>
<tr>
<th>Node</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z (%)</td>
<td>3.824</td>
<td>6.937</td>
<td>5.205</td>
<td>3.473</td>
<td>7.984</td>
<td>6.252</td>
<td>4.520</td>
<td>2.788</td>
<td></td>
</tr>
<tr>
<td>p_u</td>
<td>0.1667</td>
<td>0.1217</td>
<td>0.1667</td>
<td>0.2217</td>
<td>0.8867</td>
<td>0.1217</td>
<td>0.1667</td>
<td>0.2217</td>
<td>0.0867</td>
</tr>
<tr>
<td>p_m</td>
<td>0.6666</td>
<td>0.6566</td>
<td>0.6666</td>
<td>0.6566</td>
<td>0.0266</td>
<td>0.6566</td>
<td>0.6666</td>
<td>0.6566</td>
<td>0.0266</td>
</tr>
<tr>
<td>p_d</td>
<td>0.1667</td>
<td>0.2217</td>
<td>0.1667</td>
<td>0.1217</td>
<td>0.0867</td>
<td>0.2217</td>
<td>0.1667</td>
<td>0.1217</td>
<td>0.8867</td>
</tr>
</tbody>
</table>

Table 3.10: Probabilities on the tree

When extending the procedure to other models, such as the mean-reversion model that comes from

\[ d \ln(Z_t) = (\theta - a \ln(Z_t))\sigma dW_t, \]

appears when calculating the value of \( \alpha_m \) so that the tree correctly prices an \((m + 1)\Delta t\) zero-coupon bond. The \( \Delta t \)-period interest rate at the \( j \)th node at time \( m\Delta t \) is

\[ e^{\alpha_m + j\Delta \ln(Z_t)}. \]

Then, the price of a zero-coupon bond maturing at time \((m + 1)\Delta t\) is given by

\[ P_{m+1} = \sum_{j=-n_m}^{n_m} Q_{m,j} \exp \left( -\exp(\alpha_m + j\Delta \ln(Z_t))\Delta t \right). \]

This equation can be solved using a numerical procedure such as Newton-Raphson. Then

\[ \alpha_0 = \ln(Z_0). \]

Once \( \alpha_m \) has been calculated, then

\[ Q_{m+1,j} = \sum_k Q_{m,k} q(k, j) \exp \left( -\exp(\alpha_m + j\Delta \ln(Z_t))\Delta t \right). \]

Example 3.5.3. Let suppose that the continuously compounded zero rates are as shown in Table 3.8. Calculating the mean-reversion model associated with \( a = 0.22, \sigma = 0.25 \) and \( \Delta t = 0.5 \), the associated trinomial tree is the given by the Figure 3.4.

Table 3.11: Probabilities on the tree
Note. Using this method, it seems that the term $\theta$ it is not used during the calculation. In fact, this term is normally (to simplify) assumed constant, although it can be estimated using the forward curves. In the previous example this term is approximated by the Table 3.7.

Montecarlo method

The last method to calculate the price is the Montecarlo method. This is a widely used method in the estimation of random processes. This method is based in the Law of large numbers. To understand the essence, the weak form will be enough.

**Theorem 3.5.4** (Weak law of large numbers). Let $X_1, X_2, ..., X_n$ be a sequence of random variables with the same mean $\mu$ and variance $\sigma^2$, with $\sigma^2 < \infty$, and pairwise uncorrelated. Define

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

$\forall n > 0$. Then the sequence of random variables $(S_n)_{n \geq 1}$ converges in probability to the constant $\mu$.

The fundamental consequence of this theorem is that the expectation $\mu$ of a random process $X$ may be approximated by the arithmetic average of the
numbers obtained in a large sequence of independent draw of this random variable. With this theorem can be approximated too the probability of reaching a concrete range of values of $X$.

**Example 3.5.5.** Let $X$ be the random process given by rolling a six-sided dice. To calculate the probability of getting a result smaller than 3, the experience is repeated $n$ times with this results:

<table>
<thead>
<tr>
<th>$n$</th>
<th>Number of successes</th>
<th>Probability inferred</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>4</td>
<td>0.4000</td>
</tr>
<tr>
<td>100</td>
<td>26</td>
<td>0.2600</td>
</tr>
<tr>
<td>1000</td>
<td>314</td>
<td>0.3140</td>
</tr>
<tr>
<td>10000</td>
<td>3296</td>
<td>0.3296</td>
</tr>
<tr>
<td>100000</td>
<td>333084</td>
<td>0.3331</td>
</tr>
</tbody>
</table>

Table 3.11: Rolling results

Then it can be claimed that the probability of getting a result smaller than 3 is approximately $\mu = 0.3331$ with a standard deviation $\sigma = 0.0013$, which is really close to the real value $\frac{1}{3}$.

**Example 3.5.6.** To calculate the approximated value of $\pi$ consider a circle
inscribed within a square of side 1. Then, the ratio of areas is

\[ r = \frac{0.5^2 \pi}{1}. \]

Then

\[ \pi = 4r. \]

This ratio can be calculated through the Monte Carlo method. The values of \( \pi \) inferred are:

<table>
<thead>
<tr>
<th>( n )</th>
<th>Approximated ( \pi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>3.6000</td>
</tr>
<tr>
<td>100</td>
<td>3.3200</td>
</tr>
<tr>
<td>1000</td>
<td>3.1600</td>
</tr>
<tr>
<td>10000</td>
<td>3.1644</td>
</tr>
<tr>
<td>1000000</td>
<td>3.1433</td>
</tr>
</tbody>
</table>

Table 3.12: \( \pi \) approximation

Then, the value of \( \pi \) is approximately \( \mu = 3.1433 \) with a standard deviation \( \sigma = 0.0057 \).

Note. Note that the convergence of this method is slow.

Assuming that interest rates are constant, a derivative can be valued as follows:

(i) Sample a random path for \( S \) in a risk-neutral world.

(ii) Calculate the payoff from the derivative.

(iii) Repeat steps 1 and 2 to get many sample values of the payoff from the derivative in a risk-neutral world.

(iv) Calculate the mean of the sample payoffs to get an estimate of the expected payoff in a risk-neutral world.

(v) Discount this expected payoff at the risk-free rate to get an estimate of the value of the derivative.

As before, to apply this method must be rescaled the Brownian motion \( W_t \) to evaluate \( S \) in a risk-neutral world. The solution in the mean-reversion models proposed by [9] is to scale \( W_t \) by

\[ d\hat{W}_t = dW_t - qdt, \]
where $q$ represents the market price of risk. In [9] is proposed to take it equal to zero or constant (it is used $q = -0.1$).

The accuracy of the approximation through the Montecarlo method depends on the number of simulations made. The standard error of the estimate is given by

$$E = \frac{\sigma}{\sqrt{n}}.$$  \hfill (3.44)

To reach a 95% confidence interval for the price $D_0$ of the derivative is therefore given by

$$\mu - \frac{1.96\sigma}{\sqrt{n}} < D_0 < \mu + \frac{1.96\sigma}{\sqrt{n}}.$$

Note. The error is inversely proportional to the square root of the number of simulations. In particular, to increase the accuracy by a factor of 10, the number of trials must increase by a factor of 100.

**Example 3.5.7.** With this two methods, which would be the price of an European call option on the mean-reversion model of strike $X = 0.13$, maturity $T = 3$ years, $k = 0.2$, $\theta = 1$, volatility $\sigma = 0.26$ and starting at $S_0 = 0.2$?

To build the trinomial tree, first of all some data must be stated. It will be token time ticks of 1 year, therefore $\Delta t = 1$. Since the mean-reversion model is being used, there will be no changes in the calculus of the probabilities at each node. In this case, $\Delta X = \Delta \left( \ln(S_t) \right)$, then

$$\Delta \left( \ln(S_t) \right) = \sigma \sqrt{3\Delta t} \approx 0.4503.$$

The next important value is $j_{\max}$,

$$\frac{0.184}{k\Delta t} = 0.92 < 1 = j_{\max},$$

then

$$j_{\max} = 1, \quad j_{\min} = -1.$$

Thus, the tree shape will be something like Figure 3.6. Now, the path probabilities at each node can be calculated.

To calculate the path probabilities for the central nodes $A, C, F$ will be used (3.37), for the upper nodes $B, F$ (3.38) and for $D, G$ (3.39). Thus, the probabilities at each node are in Table 3.13.

To calculate the values of $S_t$ on the tree are necessary the variables $\alpha_m$ and $Q_{m,j}$. The procedure used is the same as in [2], to calculate the value of $\alpha_m$ it has been used the bisection method with an error under $10^{-6}$. The values of $S_t$ at each node are then in Table 3.14.
3. PLAIN VANILLA OPTIONS ON COMMODITY SPOT AND FORWARD PRICES

Figure 3.6: Initial tree

<table>
<thead>
<tr>
<th>Node</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_u$</td>
<td>0.1667</td>
<td>0.8867</td>
<td>0.1667</td>
<td>0.0867</td>
<td>0.8867</td>
<td>0.1667</td>
<td>0.0867</td>
</tr>
<tr>
<td>$p_m$</td>
<td>0.6666</td>
<td>0.0267</td>
<td>0.6666</td>
<td>0.0267</td>
<td>0.0266</td>
<td>0.6667</td>
<td>0.0267</td>
</tr>
<tr>
<td>$p_d$</td>
<td>0.1667</td>
<td>0.0867</td>
<td>0.1667</td>
<td>0.8867</td>
<td>0.0867</td>
<td>0.1667</td>
<td>0.8867</td>
</tr>
</tbody>
</table>

Table 3.13: Probabilities on the tree

<table>
<thead>
<tr>
<th>Node</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln(S)$</td>
<td>-1.6094</td>
<td>-1.1859</td>
<td>-1.6363</td>
<td>-2.0866</td>
<td>-1.1921</td>
<td>-1.6425</td>
<td>-2.0928</td>
</tr>
<tr>
<td>$S$</td>
<td>0.2000</td>
<td>0.3055</td>
<td>0.1947</td>
<td>0.1241</td>
<td>0.3036</td>
<td>0.1935</td>
<td>0.1233</td>
</tr>
</tbody>
</table>

Table 3.14: Values on the tree

<table>
<thead>
<tr>
<th>Node</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call value $c$</td>
<td>0.0440</td>
<td>0.0955</td>
<td>0.0516</td>
<td>0.0204</td>
<td>0.1284</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Node</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call value $c$</td>
<td>0.0590</td>
<td>0.0138</td>
<td>0.1749</td>
<td>0.0643</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table 3.15: Call value at each node

To find the cost of a call at the node $A$, only is need to calculate the payoff and to calculate the discounted expected value at every node. Thus, the value of a call with strike $X = 0.13$ at every node is in Table 3.15.

In conclusion, the price of the call $c$ with strike $X = 0.13$ at initial time will be $c_0 = 0.0440$.

When using the Montecarlo method, the SDE that is being estimated is

$$\frac{dS_t}{S_t} = k(\theta - \ln(S_t))dt + \sigma dW_t.$$
To produce the paths, this SDE may be written as

$$dS_t = S_t \left( k(\theta - \ln(S_t))dt + \sigma dW_t \right).$$

And by definition,

$$S_{t+dt} = S_t \left( 1 + k(\theta - \ln(S_t))dt + \sigma dW_t \right).$$

With $W_t \sim N(0, dt)$. Then, sampling 1,000,000 random paths for $S_t$ the mean of payoffs obtained is

$$c_0 = 0.0253,$$

with a standard deviation

$$s = 0.2253.$$

Both results are quiet different. This may be due to the fact that the increment of time $\Delta t = 1$ in the case of the trinomial tree is rather big. This fact affect the value of $c_0$ in two different ways:

(i) In first place, taking only three steps conditions the accuracy of the result.

(ii) In second place, the $j_{\text{max}}$ and $j_{\text{min}}$ associated to $\Delta t$ is only 1. This forces the tree to variate and contemplate other options.
The aim of this chapter is to go beyond the mathematical modelling and pricing of derivatives and analyse it in the real world. In the markets, the stocks are shown as prices not as processes satisfying a concrete SDE. In this chapter will be showed how to, from the data in the market, actually build the models seen before and determine the parameters associated to the underlying asset. The calibration is one of the most important part in finance since the expectations and the confidence intervals are determined from the parameters of the models.

To build this parameters from the market data, it will be used the maximum-likelihood estimation of the parameters and some basic algebra.

4.1 Maximum-likelihood estimation

There are several methods to estimate a parameter of a random variable $X$. The maximum-likelihood estimation method is the more intuitive one. In fact, the basis of the method is to choose the value such that the data observed has maximum probability. The problem is then reduced to an optimization problem.

In this section, it will be considered the random variable $X$ and its observations $\{X_n\}_n$. Define the parameter to estimate as the vector $\theta = (\theta_1, ..., \theta_k)$. And define the density function of $X$ with the parameters $\theta$ fixed as $f(x|\theta)$.

Definition 4.1.1. Given the observed data $X_1, ..., X_n$ with density function
4. CALIBRATING THE MODELS. A REAL APPROACH TO THE MARKETS

\( f(x, \theta) \), define the likelihood function for a fixed \( x \) as

\[
L_x : \Theta \to \mathbb{R}^+
\]

\( \theta \to L_x(\theta) = L(\theta|x) = \prod_{i=1}^{n} f(x_i|\theta) \)

**Example 4.1.2.** Let \( X_1, \ldots, X_n \) be a simple random sample of \( X \sim \text{exp}(\mu) \).

The density function is given by

\[
f(x_i|\mu) = \frac{1}{\mu} \exp\left(-\frac{x_i}{\mu}\right),
\]

\[
f(x_1, \ldots, x_n|\mu) = \frac{1}{\mu^n} \exp\left(-\frac{\sum_{i=1}^{n} x_i}{\mu}\right).
\]

Thus the likelihood function is

\[
L(\mu|x_1, \ldots, x_n) = \frac{1}{\mu^n} \exp\left(-\frac{\sum_{i=1}^{n} x_i}{\mu}\right).
\]

The likelihood function that will be used is the likelihood function of a \( N(\mu, \sigma^2) \).

**Proposition 4.1.3.** Let \( X_1, \ldots, X_n \) be a simple random sample of \( X \sim N(\mu, \sigma^2) \). The likelihood function is

\[
L(\mu, \sigma^2|x_1, \ldots, x_n) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^2 \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\right), \quad (4.1)
\]

**Proof.** Since the density function is given by

\[
f(x_1, \ldots, x_n|\mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^2 \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\right),
\]

the likelihood function has the same expression but fixing the data instead of the parameters. \( \square \)

**Definition 4.1.4.** If \( \hat{\theta}(x) \) is a maximum of the likelihood function for a given sample \( x = (x_1, \ldots, x_n) \) of data, then define \( \hat{\theta}(x) \) as the maximum likelihood estimator (MLE) of \( \theta \).

When calculating the maximum of the likelihood there are some points to take care of:

(i) If \( L(\theta|x) \) is differentiable at \( \theta \), the solutions of \( \frac{\partial}{\partial \theta_i} L(\theta|x) = 0 \) with \( i = 1, \ldots, n \) are candidates to be the MLE. Note that the solutions should not be a maximum. Note that this method is only valid in \( \Theta \), the boundary of \( \Theta \) is not taken in account.
(ii) The maximum can be found directly by finding an upper bond for \( L(\theta|x) \) and then proving that exists one point that reaches that value.

(iii) Sometimes it is easier to find the maximum points of \( \ln(L(\theta|x)) \) instead of \( L(\theta|x) \). Since the logarithm is strictly increasing in \((0, \infty)\), the maximum points of \( \ln(L(\theta|x)) \) and \( L(\theta|x) \) will coincide.

**Proposition 4.1.5.** The MLE for a simple random sample \( X_1, ..., X_n \) such that \( X \sim N(\mu, \sigma^2) \) are given by

\[
\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i, \tag{4.2}
\]

\[
\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu})^2. \tag{4.3}
\]

**Proof.** In this case,

\[
L(\mu, \sigma^2|x) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left( \frac{-1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right).
\]

Taking the logarithm,

\[
\ln(L(\mu, \sigma^2|x)) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2
\]

Then, the candidates to be the MLE must satisfy

\[
\frac{\partial}{\partial \mu} \ln(L(\mu, \sigma^2|x)) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \theta) = 0,
\]

\[
\frac{\partial}{\partial (\sigma^2)} \ln(L(\mu, \sigma^2|x)) = -\frac{n}{2\sigma^2} - \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2 = 0.
\]

It can be verified that the Hessian matrix \( H \) is a negative matrix. This implies that the MLE are a maximum of the likelihood function.

The last result that will be needed to infer the parameters of the models is the invariance principle for estimators.

**Theorem 4.1.6 (Invariance principle for estimators).** If \( \hat{\theta} \) is the MLE of \( \theta \), then, for all function \( f(x) \) the MLE of \( f(\theta) \) is \( f(\hat{\theta}) \).

**Example 4.1.7.** The MLE for the standard deviation \( \sigma \) of \( X \), with \( X \sim N(\mu, \sigma^2) \), is given by

\[
\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu})^2}
\]
4.2 Inference on the models

Once the statistic basis are settled, the parameters of the models can be inferred. Recall the four basic models treated in the Chapter 2:

(i) Arithmetic Brownian motion: \(dX_t = \alpha dt + \sigma dW_t.\)

(ii) Geometric Brownian motion: \(\frac{dS_t}{S_t} = \mu dt + \sigma dW_t.\)

(iii) Ornstein-Uhlenbeck process: \(dX_t = a(b - X_t)dt + \sigma dW_t.\)

(iv) Mean-reversion process: \(\frac{dS_t}{S_t} = k(\theta - \ln(S_t))dt + \sigma dW_t.\)

Arithmetic and geometric Brownian motion

First of all will be studied the two firsts processes. The two processes can be related through \(Z_t = \ln(S_t),\) this change of variable leads to an arithmetic Brownian motion with a modified drift term.

Proposition 4.2.1. Given the observed data \(X_0, ..., X_n\) with \(X\) an arithmetic Brownian motion (satisfying the SDE (2.1)), then, the MLE estimators for \(\alpha\) and \(\sigma\) are given by

\[
\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} (X_i - X_{i-1}), \tag{4.4}
\]

\[
\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - X_{i-1} - \hat{\alpha} \Delta t)^2}. \tag{4.5}
\]

Proof. Since the difference \(\Delta X_t\) (which is \(dX_t\) when \(\Delta t \to 0\)) is distributed as

\[
\Delta X_t \sim N(\alpha \Delta t, \sigma^2 \Delta t),
\]

using the previous results on the normal distribution,

\[
\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} (X_i - X_{i-1}),
\]

\[
\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - X_{i-1} - \hat{\alpha} \Delta t)^2}.
\]
4.2. INFERENCE ON THE MODELS

In the case of geometric Brownian motion, what will be used is the change of variable \( U_t = \ln(S_t) \) that satisfies that \( U_t \) is an arithmetic Brownian motion, then, the same method will be used.

**Proposition 4.2.2.** Given the observed data \( S_0, ..., S_n \) with \( S \) a geometric Brownian motion (satisfying the SDE (2.3)), then, the MLE estimators for \( \mu \) and \( \sigma \) are given by

\[
\hat{\mu} = \hat{\alpha} + \frac{1}{2} \hat{\sigma}^2, \tag{4.6}
\]
\[
\hat{\sigma} = \sqrt{\frac{1}{n\Delta t} \sum_{i=1}^{n} (S_i - S_{i-1} - \hat{\alpha} \Delta t)^2}, \tag{4.7}
\]

with

\[
\hat{\alpha} = \frac{1}{n\Delta t} \sum_{i=1}^{n} (S_i - S_{i-1}).
\]

**Proof.** Define \( U_t = \ln(S_t) \), then, using Itô’s lemma and since \( S_t \) is a geometric Brownian motion

\[ dU_t = (\mu - \frac{\sigma^2}{2}) dt + \sigma dW_t. \]

Since \( U_t \) is an arithmetic Brownian motion, let \( \alpha = \mu - \frac{\sigma^2}{2} \), by the Proposition 4.2.1 it is known that

\[
\hat{\alpha} = \frac{1}{n\Delta t} \sum_{i=1}^{n} (S_i - S_{i-1}),
\]
\[
\hat{\sigma} = \sqrt{\frac{1}{n\Delta t} \sum_{i=1}^{n} (S_i - S_{i-1} - \hat{\alpha} \Delta t)^2}.
\]

Using the invariance principle for estimators (Theorem 4.1.6), it is given that

\[
\hat{\alpha} = \hat{\mu} - \frac{\hat{\sigma}^2}{2}.
\]

Thus

\[
\hat{\mu} = \hat{\alpha} + \frac{1}{2} \hat{\sigma}^2.
\]

\[ \square \]

**Example 4.2.3.** The price of the IBEX 35 is settled by the index IBEX. Between the dates August 8th of 2012 and August 8th 2014 the price behaved as a pure geometric Brownian motion (without jumps or strange performances).
4. CALIBRATING THE MODELS. A REAL APPROACH TO THE MARKETS

Applying the previous method, the resulting estimators are

\[ \hat{\mu} = 0.1868, \]
\[ \hat{\sigma} = 0.1857. \]

To contrast the accuracy of the estimator \( \hat{\mu} \), recall that \( \mu \) is the expected growth per year. In this case, the growth can be approximated as the growth between the mean of the first five values and the mean of the last five which in this case is 0.2018 per year which is quite close to \( \hat{\mu} \). There is no simple way to match if \( \hat{\mu} \) is or is not a good estimation of the parameter.

All the data has been extracted from the website www.investing.com.

Mean-reversion processes

The main difficulty when calibrating the models associated to Ornstein-Uhlenbeck processes and mean-reversion processes is that the differential of the process do not follow a normal distribution and unknowing the value of the parameters there is no change, a priori, to transform the process into an arithmetic Brownian motion (like in the geometric Brownian motion case). The MLE for this processes must be calculated from the definition of the method. There are no previous results (like the normal distribution MLE) that can be, a priori, used.

**Proposition 4.2.4.** Given the observed data \( X_0, \ldots, X_n \) with \( X \) an Ornstein-Uhlenbeck process (satisfying the SDE (2.7)), then, the MLE estimators for \( a, \)
4.2. INFERENCE ON THE MODELS

$b$ and $\sigma$ are given by

$$
\hat{a} = -\frac{1}{\Delta t} \ln \left( \frac{X_{xy} - \hat{b}(X_x + X_y) + n\hat{b}^2}{X_{xx} - 2\hat{b}X_x + n\hat{b}^2} \right), \quad (4.8)
$$

$$
\hat{b} = \frac{X_yX_{xx} - X_xX_{xy}}{n(X_{xx} - X_{xy}) - (X_x^2 - X_xX_y)}, \quad (4.9)
$$

$$
\hat{\sigma}^2 = \frac{2\hat{a}}{n(1 - \alpha^2)} \left( X_{yy} - 2\alpha X_{xy} + \alpha^2 X_{xx} - 2\hat{b}(1 - \alpha)(X_y - \alpha X_x) + n\hat{b}^2(1 - \alpha)^2 \right). \quad (4.10)
$$

With,

$$
\alpha = e^{-a\Delta t}
$$

$$
X_x = \sum_{i=1}^{n} S_{i-1},
$$

$$
X_y = \sum_{i=1}^{n} S_i,
$$

$$
X_{xx} = \sum_{i=1}^{n} S_i^2 - 1,
$$

$$
X_{xy} = \sum_{i=1}^{n} (S_{i-1} S_i),
$$

$$
X_{yy} = \sum_{i=1}^{n} S_i^2.
$$

Proof. Recall the probability density function $f(x)$ for a normal distribution $N(\mu, \sigma)$ is given by

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

Then, by the equation (2.9) the process $X_t$ has a conditioned probability density function

$$
f(X_t \mid X_{t-1}; a, b, \hat{\sigma}) = \frac{1}{\sqrt{2\pi \hat{\sigma}^2}} \exp \left( -\frac{(X_t - X_{t-1}e^{-a\Delta t} - b(1 - e^{-a\Delta t}))^2}{2\hat{\sigma}^2} \right),
$$

with

$$
\hat{\sigma}^2 = \sigma^2 \frac{1 - e^{-2a\Delta t}}{2a}.
$$
and $X_i = X_{i\Delta t}$. In this case, it will be much easier to use the logarithm of the likelihood function. Thus the logarithm likelihood function is

$$L(a, b, \tilde{\sigma}) = \sum_{i=1}^{n} \ln \left( f(X_i | X_{i-1}; a, b, \tilde{\sigma}) \right)$$

$$= -\frac{n}{2} \ln(2\pi) - n \ln(\tilde{\sigma}) - \frac{1}{2\tilde{\sigma}} \sum_{i=1}^{n} \left( X_i - X_{i-1}e^{-a\Delta t} - b(1 - e^{-a\Delta t}) \right)^2.$$ 

The next step is to find the values of $(a, b, \tilde{\sigma})$ that maximizes the likelihood function. These values must satisfy

$$\frac{\partial L}{\partial a} = 0,$$

$$\frac{\partial L}{\partial b} = 0,$$

$$\frac{\partial L}{\partial \tilde{\sigma}} = 0.$$ 

These equalities hold when

$$a = -\frac{1}{\Delta t} \ln \left( \frac{\sum_{i=1}^{n}(X_i - b)(X_{i-1} - b)}{\sum_{i=1}^{n}(X_{i-1} - b)^2} \right),$$

$$b = \frac{\sum_{i=1}^{n}(X_i - X_{i-1}e^{-a\Delta t})}{n(1 - e^{-a\Delta t})},$$

$$\tilde{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left( X_i - b - e^{-a\Delta t}(X_{i-1} - b) \right)^2.$$ 

The problem with these solutions is that the solutions depend on each other. In any case, both $a$ and $b$ are independent of $\tilde{\sigma}$. And $a$ or $b$ can be determined substituting one into the other condition. To work with them, the change of variable $X_x, X_y, X_{xx}, X_{xy}$ and $X_{yy}$ is taken. Then,

$$a = -\frac{1}{\Delta t} \ln \left( \frac{X_{xy} - b(X_x + X_y) + nb^2}{X_{xx} - 2bX_x + nb^2} \right),$$

$$b = \frac{X_y - e^{-a\Delta t}X_x}{n(1 - e^{-a\Delta t})}.$$ 

Substituting $a$ into the $b$ expression gives

$$nb = \frac{X_y - \left( \frac{X_{xy} - b(X_x + X_y) + nb^2}{X_{xx} - 2bX_x + nb^2} \right) X_x}{1 - \left( \frac{X_{xy} - b(X_x + X_y) + nb^2}{X_{xx} - 2bX_x + nb^2} \right)},$$

$$nb = \frac{X_y(X_{xx} - 2bX_x + nb^2) - (X_{xy} - bX_x - bX_y + nb^2)X_x}{(X_{xx} - 2bX_x + nb^2) - (X_{xy} - bX_x - bX_y + nb^2)},$$
4.2. INFERENCE ON THE MODELS

\[ nb = \frac{(X_y X_{xx} - X_x X_{xy}) + b(X_x^2 - X_x X_y) + b^2 n(X_y - X_x)}{(X_{xx} - X_{xy}) + b(X_y - X_x)}, \]

\[ nb(X_{xx} - X_{xy}) - b(X_x^2 - X_x X_y) = X_y X_{xx} - X_x X_{xy}. \]

And finally,

\[ \hat{a} = -\frac{1}{\Delta t} \ln \left( \frac{X_{xy} - \hat{b}(X_x + X_y) + n\hat{b}^2}{X_{xx} - 2\hat{b}X_x + n\hat{b}^2} \right), \]

\[ \hat{b} = \frac{X_y X_{xx} - X_x X_{xy}}{n(X_{xx} - X_{xy}) - (X_x^2 - X_x X_y)}, \]

\[ \hat{\sigma}^2 = \frac{2\hat{a}}{n(1 - \alpha^2)} (X_{yy} - 2\alpha X_{xy} + \alpha^2 X_{xx} - 2\hat{b}(1 - \alpha)(X_y - \alpha X_x) + n\hat{b}^2(1 - \alpha)^2). \]

\[ \square \]

As in the Brownian motion, in the case of mean-reversion will be used the change of variable \( Z_t = \ln(S_t) \) that satisfies that \( Z_t \) is an Ornstein-Uhlenbeck process, then, the same method will be used.

**Proposition 4.2.5.** Given the observed data \( S_0, ..., S_n \) with \( S \) a mean-reversion process (satisfying the SDE (2.10)), then, the MLE estimators for \( k \), \( \theta \) and \( \sigma \) are given by

\[ \hat{k} = -\frac{1}{\Delta t} \ln \left( \frac{X_{xy} - \hat{b}(X_x + X_y) + n\hat{b}^2}{X_{xx} - 2\hat{b}X_x + n\hat{b}^2} \right), \]

\[ \hat{\theta} = \hat{b} + \frac{\hat{\sigma}^2}{2\hat{k}}, \]

\[ \hat{\sigma}^2 = \frac{2\hat{a}}{n(1 - \alpha^2)} (X_{yy} - 2\alpha X_{xy} + \alpha^2 X_{xx} - 2\hat{b}(1 - \alpha)(X_y - \alpha X_x) + n\hat{b}^2(1 - \alpha)^2). \]
With,

\[ \alpha = e^{-a\Delta t}, \]
\[ \hat{b} = \frac{X_yX_{xx} - X_xX_{xy}}{n(X_{xx} - X_{xy}) - (X_x^2 - X_xX_y)}, \]
\[ X_x = \sum_{i=1}^{n} S_{i-1}, \]
\[ X_y = \sum_{i=1}^{n} S_i, \]
\[ X_{xx} = \sum_{i=1}^{n} S_{i-1}^2, \]
\[ X_{xy} = \sum_{i=1}^{n} S_{i-1}S_i, \]
\[ X_{yy} = \sum_{i=1}^{n} S_i^2. \]

**Proof.** Define \( Z_t = \ln(S_t) \), then, using Itô’s lemma and since \( S_t \) is a mean-reversion process

\[ dZ_t = k(\theta - \frac{\sigma^2}{2k} - Z_t)dt + \sigma dW_t. \]

Since \( Z_t \) is an Ornstein-Uhlenbeck process, let \( b = \theta - \frac{\sigma^2}{2} \), by the Proposition 4.2.4 it is known that

\[ \hat{\theta} = -\frac{1}{\Delta t} \ln \left( \frac{X_{xy} - \hat{b}(X_x + X_y) + n\hat{b}^2}{X_{xx} - 2\hat{b}X_x + n\hat{b}^2} \right), \]
\[ \hat{b} = \frac{X_yX_{xx} - X_xX_{xy}}{n(X_{xx} - X_{xy}) - (X_x^2 - X_xX_y)}, \]
\[ \hat{\sigma}^2 = \frac{2 \hat{\alpha}}{n(1 - \alpha^2)} (X_{yy} - 2\alpha X_{xy} + \alpha^2 X_{xx} - 2\hat{b}(1 - \alpha)(X_y - \alpha X_x) + n\hat{b}^2(1 - \alpha)^2). \]

Using the invariance principle for estimators (Theorem 4.1.6), it is given that

\[ \hat{b} = \hat{\theta} - \frac{\hat{\sigma}^2}{2k}. \]

Thus

\[ \hat{\theta} = \hat{b} + \frac{\hat{\sigma}^2}{2k}. \]
Example 4.2.6. The price of the crude oil is set by the index CLG6. Between the dates February 25th of 2011 and August 8th 2014 the price behaved as a pure mean-reversion process (without jumps or strange performances).

Applying the previous method, the resulting estimators are

\[ \hat{k} = 16.2384, \]
\[ \hat{\theta} = 4.5717, \]
\[ \hat{\sigma} = 0.2520. \]

The Crude Oil Volatility index (OVX) during this period has a mean of 0.1966 which is quite similar to \( \hat{\sigma} \). The mean of the logarithm of the price during this period is 4.5744 which is really close to the estimated mean-revert term \( \theta \). There is no simple way to match if \( \hat{k} \) is or is not a good estimation of the parameter.

As in the previous example, all the data has been extracted from the website www.investing.com.
Bibliography


