DETERMINISTIC AND STOCHASTIC TOPICS
IN FINANCE

Ovidiu Calin
Department of Mathematics
Eastern Michigan University
Ypsilanti, MI 48197 USA
ocalin@emich.edu
## Contents

1 Determinism or Stochasticity? 3  
1.1 Determinism and Semi-determinism 3  
1.2 How to Measure 5  
1.3 Can Random be Deterministic? 6  

2 Regression 9  
2.1 Deterministic Regression 9  
2.2 Stochastic Regression 13  
2.3 Calibration 16  
2.4 Alternatives to the Method of Least Squares 21  
2.4.1 The Maximum Likelihood Method 21  
2.4.2 The Maximum Entropy Method 23  

3 Modeling Stochastic Rates 25  
3.1 Deterministic versus Stochastic Calculus 25  
3.2 Langevin’s Equation 27  
3.3 Equilibrium Models 28  
3.3.1 Rendleman and Bartter Model 29  
3.3.2 Vasicek Model 30  
3.3.3 Cox-Ingersoll-Ross Model 32  
3.4 No-arbitrage Models 34  
3.4.1 Ho and Lee Model 34  
3.4.2 Hull and White Model 34  
3.5 Nonstationary Models 36  
3.5.1 Black, Derman and Toy Model 36  

4 Bonds, Forward Rates and Yield Curves 39  
4.1 Bonds 39  
4.2 Yield 40  
4.3 Bootstrap Method 42  
4.4 Forward Rates 43
### 7 Martingale Measures

7.1 Martingale Measures ........................................... 121
    7.1.1 Is the stock price $S_t$ a martingale? ................. 121
    7.1.2 Risk-neutral World and Martingale Measure ............ 123
    7.1.3 Finding the Risk-Neutral Measure ....................... 124

7.2 Risk-neutral World Density Functions ....................... 125

7.3 Self-financing Portfolios ..................................... 127

7.4 The Sharpe Ratio .............................................. 127

7.5 Risk-neutral Valuation for Derivatives ..................... 128

### 8 Black-Scholes Analysis

8.1 Heat Equation .................................................. 131

8.2 What is a Portfolio? ............................................ 134

8.3 Risk-less Portfolios ........................................... 135

8.4 Black-Scholes Equation ....................................... 137

8.5 Delta Hedging ................................................... 139

8.6 Tradable securities ............................................. 139

8.7 Risk-less investment revised ................................. 141

8.8 Solving the Black-Scholes ..................................... 144

8.9 Black-Scholes and Risk-neutral Valuation .................. 147

8.10 Boundary Conditions .......................................... 148

8.11 Risk-less Portfolios for Rare Events ...................... 149

### 9 Black-Scholes for Asian Derivatives

9.1 Weighted averages .............................................. 153

9.2 Setting up the Black-Scholes Equation ...................... 155

9.3 Weighted Average Strike Call Option ......................... 156

9.4 Boundary Conditions .......................................... 158

9.5 Asian Forward Contracts on Weighted Averages ............. 162

### 10 American Options

10.1 Perpetual American Options ................................. 165
    10.1.1 Present Value of Barriers .............................. 165
    10.1.2 Perpetual American Calls ............................... 169
    10.1.3 Perpetual American Puts ............................... 170

10.2 Perpetual American Log Contract ............................ 173

10.3 Perpetual American Power Contract ......................... 175

10.4 Finitely Lived American Options ......................... 176
    10.4.1 American Call ........................................... 176
    10.4.2 American Put ........................................... 179
    10.4.3 MacMillan-Barone-Adesi-Whaley Approximation ....... 179
    10.4.4 Black’s Approximation ................................ 179
Chapter 1

Determinism or Stochasticity?

1.1 Determinism and Semi-determinism

Any system (physical, economic, financial, etc.) can be described by a certain number of parameters. In order to know and understand the system, one observes or measures the parameters. We shall assume we are able to measure all parameters with precise accuracy.\(^1\) They will describe the system as coordinates of a point in a parameter space. The number of parameters denotes the dimension of the space which models the configuration of the system.

For instance, a moving particle in the plane is characterized by its position and velocity, \((x, v)\), so the parameter space is 6-dimensional, the balance amount in a certificate of deposit account is characterized by a single number, so the system is 1-dimensional. If the system evolves in time, its parameters are functions of time, and the system can be represented by a curve in the aforementioned space. One of the the most important problems is to predict and describe the trajectories of the analyzed systems. These can be of three types, and they are discussed briefly in the following.

A system is called determinist if both future and past states of the system are uniquely determined by the present state of the system. Most systems studied by classical mechanics are determinist. For instance, if the present position, \(x_0\), and velocity, \(v_0\), of a car are known at time \(t_0\), then integrating the equations of motion, \(x''(t) = a\), with constant acceleration \(a\), one can find both future and past positions and velocities of the car, according to the

\(^1\)In Quantum Mechanics this is not possible because of the uncertainty principle.
formulas
\begin{align*}
v(t) &= v_0 + a(t - t_0) \\
x(t) &= x_0 + v_0(t - t_0) + \frac{1}{2}a(t - t_0)^2,
\end{align*}
where \( t \) is the time parameter. For \( t > t_0 \) we obtain the future states of the car, while for \( t < t_0 \) we obtain its past. The future and the past play symmetric roles for this deterministic problem.

A similar deterministic process occurs in finance when observing the value of money. An amount \( K \) of money at time \( t_0 \) worths \( Ke^{r(t-t_0)} \) at time \( t > t_0 \) and \( Ke^{-r(t-t_0)} \) at time \( t < t_0 \), provided the interest rate \( r \) is constant. This shows that the present state of the system, \( K \), determines uniquely the future and past states of the system.

The determinist systems are described by the theory of ordinary differential equations, see Arnold [1].

A system is called semi-determinist if only the future (or the past) states of the system are uniquely determined by the present state of the system, while the past (the future) is not. Heat propagation, smoke evolution, and ink diffusion are semi-determinist processes. They can be predicted in the future but cannot be traced back in time. The explanation of this phenomenon is given in Chapter 8 and it is based on the formula provided by Proposition 8.1.2, which solves a forward heat equation
\begin{align*}
\frac{\partial u}{\partial \tau} - \frac{\partial^2 u}{\partial x^2} &= 0 \\
\left. u(0, x) \right|_{\tau = 0} &= f(x),
\end{align*}
with the solution given by the convolution between the fundamental solution and the initial temperature
\[ u(\tau, x) = \int_{\mathbb{R}} G(\tau, y - x) f(y) \, dy, \quad \tau > 0, \]
where
\[ G(\tau, x) = \frac{1}{\sqrt{4\pi \tau}} e^{-\frac{x^2}{4\tau}}, \quad \tau > 0. \quad (1.1.1) \]
The expression \( p(y, x; \tau) = G(\tau, y - x) \) represents the transition probability of a particle starting at \( x \) at time \( t = 0 \) and arriving at \( y \) at time \( t = \tau \). Since the initial condition acts at any \( y \) as an initial density of heat, then the expression \( G(\tau, y - x) f(y) \Delta y \) represents the heat transferred from the interval \([y, y + \Delta y]\) to the point \( x \) within time \( \tau \). To find the heat at the point \( x \), we need to take into account the effects of all the above contributions, which is obtained by
the summation $\sum G(\tau, y - x)f(y)\Delta y$, which after taking the limit $\Delta y \rightarrow dy$, yields the aforementioned integral.

The semi-determinist processes are of utmost importance for finance. The unpredictable movements of stocks are modeled by Brownian motions, whose transition probabilities are of the form (1.1.1). Similar processes are used to model particles of ink or smoke, which diffuse in a liquid or atmosphere. The heat or diffusion equation in the case of finance takes the form of a backward heat-type equation, called the Black-Scholes equation. The past states of securities based on the underlying stocks can be predicted, while their future states cannot. The impossibility of foreseeing the future price of securities written on stocks lead to methods of computing the present price from future prices, which are provided in the contract at the expiration time. This is the reason why for pricing, let’s say an American option, the financial analyst has to work backwards through a tree, which models the multiple possible movements of the stock price.

The semi-determinist systems are mathematically described both by the theory of partial differential equations and stochastic calculus. This is why there are two main direct approaches to derivatives pricing in finance.

1.2 How to Measure

All exact sciences are based on observing and measuring some variables, and the more accurate this is done the more we know about the system. To measure a variable means to assign a real number to it at a certain time, when the measurement happens. In order to determine the state of a system one needs to measure all the parameters which defines the system. Then, using the measurements already done, one would like to predict future states of the system in the most accurate possible way.

In finance we measure interest rate, market volatility, stock prices, or different security prices (bonds, futures, options, etc.). When measuring the interest rate or volatility, we assign a percent, or a number between 0 and 1; when measuring a stock or a security price, we assign a certain number of dollars to it.

The process of measuring involves two ingredients:

(i) the time $t$ when the measurement occurs;

(ii) the information available to the one who measures at time $t$, denoted by $\mathcal{I}_t$.

If $S_t$ denotes the price of a stock at time $t$, and we would like to measure today the tomorrow’s price of the stock, $S_{t+1}$, then we cannot do it, because the information available today about the market, $\mathcal{I}_t$, is insufficient to determine
the price of the stock in the future. In this case, the variable $S_{t+1}$ is not measurable at time $t$.

Another example involves measuring the implied volatility, $\sigma_t$. This cannot be observed directly, like stock prices are. It needs to be computed from the prices of calls or puts, which need to be available at the measuring time $t$. Knowing these security prices means to have a grasp on the information $\mathcal{I}_t$.

The information $\mathcal{I}_t$ consists in the set of all possible events, which already happened until time $t$, or they look like might happen. The events which already happened are sure events in the market (they occurred with 100% probability); other events have certain probabilities of happening. They all influence the price of stocks. For instance, a possible strike for the workers of a certain airline has an influence in the stock market, even if it is not a sure event. The higher the odds of the strike, the larger the impact on the stock market is.

Hence, each element of the information set $\mathcal{I}_t$ has associated a probability. To set the notations, we write: if $A \in \mathcal{I}_t$ is an event, then $P(A)$ is the probability of $A$, i.e. a number between 0 and 1, describing the odds of the event $A$.

The complementary of an event, or the union and intersection of two events in $\mathcal{I}_t$ belongs also to $\mathcal{I}_t$; this provides a structure of σ-algebra, see Appendix.

### 1.3 Can Random be Deterministic?

Coin tossing is a very common example of random experiment. But how “random” is such a process?

Is it possible that some deterministic experiments to be actually disguised as random phenomena? This section will briefly cover this peculiar topic.

In probability theory we characterize the final state of the coin, not by a single possible outcome, but by two outcomes, with equal chances of occurrence, called a probability distribution. The only possible outcomes for the experiment are $H$ (heads) and $T$ (tails). If the coin is fair (flat, unweighed), we usually say that the odds of getting either $H$ or $T$ are $1/2$. Consequently, these approach implies that coin flipping is not a deterministic process.

Now, let’s take a closer look at the experiment. Naturally tossed coins obey the laws of classical mechanics, which are deterministic; hence the future states of the coin depend solely on the initial conditions. In order to model this, we consider a fixed frame in $\mathbb{R}^3$ with axes $Ox$, $Oy$ and $Oz$, which will be used to describe the motion of the coin. A coin moving in the $\mathbb{R}^3$ space can be considered as a plane region. It is known that a plane is completely determined by a given point and its normal direction. Using this fact, the coin can be completely described by the following 6 parameters:
the coordinates of the coin center, \((x, y, z)\); the first two are the coordinates of the projection onto the table, while \(z\) is the height the center;

- the angles made by the normal to the coin with the axis \(Ox, Oy\) and \(Oz\), denoted by \((\alpha, \beta, \gamma)\).

Hence, the state of the coin is a point in the following 6-dimensional space

\[(x, y, z, \alpha, \beta, \gamma) \in \mathbb{R} \times \mathbb{R} \times [0, \infty) \times [0, 2\pi) \times [0, 2\pi) \times [-\pi/2, \pi/2].\]

This is the parameter space where the coin dynamics evolves. The motion of the flipped coin corresponds to a continuous curve

\[t \rightarrow (x(t), y(t), z(t), \alpha(t), \beta(t), \gamma(t)), \quad 0 \leq t \leq \tau\]

in the above space, with \(\tau\) denoting the landing time. At time \(\tau\) the motion stops, and the final state can be observed. At landing we have \(z(\tau) = 0\) and \(\gamma(\tau) \in \{-\pi, \pi\}\). It is the value of \(\gamma(\tau)\) that produces the outcome for our experiment, since this defines the orientation of the final state of the coin. For instance, the final value \(\gamma(\tau) = \pi\) corresponds to \(H\) and \(\gamma(\tau) = -\pi\) to \(T\). Is the value \(\gamma(\tau)\) deterministic?

It certainly is. But the procedure to get the final outcome is not that straightforward. The coin dynamics can be described by a system of ordinary differential equations. The only acting force on the coin is gravity; the spinning coin has also kinetic energy of rotation and angular momentum of rotation relative to one of the diameters. Neglecting the friction with the air, texture of the surface, moisture, etc., and assuming that there is no precession (implied by the fact that the rotation axis is perpendicular on the normal to the coin), Keller [12] showed that the final outcome is completely determined by the initial upward velocity \(v_z\) and the rate of spin. A more complete analysis, involving also precession, can be found in Diaconis et al.[6].

For finding the landing time, \(\tau\), it suffices to find the positive solution of the equation \(z(\tau) = 0\), where the vertical height is given by

\[z(t) = z_0 + v_z t - \frac{1}{2} gt^2.\]

Even if the equations of motion of the flipping coin cannot be integrated in the general case, the best one can hope is to find conservation laws, such as the conservation of energy and momentum. The final state \(\gamma(\tau)\) is explicitly computable from the initial conditions, at least theoretically.

However, \(\gamma(\tau)\) is highly sensitive to small changes in the initial conditions, especially if the rotation rate is large. Since when throwing a coin, we do not have enough control or precision on our hand, this will drastically affect the final outcome. In fact, not the movement of the coin is unpredictable,
but our own hand control. The sensitivity of the solution has a lower order of magnitude than our biological sensitivity and locomotor control. If one can measure precisely the initial data, then the final state of the coin can be predicted with certainty.

If that much trouble is found in the analysis of such a simplistic experiment, we can imagine the difficulty faced in the case of interpreting stock markets, which are driven by millions of parameters. This is the main reason why stochastic calculus, which makes important simplifying assumptions, is more useful than a real deterministic analysis. However, in the day when the computation power will increase tremendously due to the use of quantum computers, the deterministic analysis will be used at its full capacity. This will change completely the way stochastic processes are analyzed, and real life simulations of stock markets will then become possible.
Chapter 2

Regression

Regression is used when calibrating financial models to the market, i.e. finding the values of the model parameters for which the model best fits the market observations. Regression can be either deterministic, or stochastic. The former looks for a best fit of the market data by a curve, while the latter approximates the data by a stochastic process, which depends on some parameters. Regression provides a procedure by which these parameters can be inferred such that the model approximates the true data, minimizing a certain error.

2.1 Deterministic Regression

Assume $n$ measurements have been made at times $t_1, \ldots, t_n$, with the results $y_1, \ldots, y_n$. The deterministic regression consists in finding a curve which best fits the data, see Fig. 2.1 a. In most cases the curve has to be chosen from a family of parameterized curves $\psi_\xi(t)$, and the problem becomes the one of finding the value of the parameter $\xi = (\xi_1, \ldots, \xi_k) \in \mathbb{R}^k$ for which a certain goodness of fit is minimized.

The procedure of finding the best fit parameter has the following geometric explanation. We start by noting that the assignment

$$\Sigma : \xi \mapsto (\psi_\xi(t_1), \ldots, \psi_\xi(t_n)) \in \mathbb{R}^n$$

is a $k$-hypersurface in $\mathbb{R}^n$. The observations vector $(y_1, \ldots, y_n)$ can be considered as the coordinates of a point $P$ in $\mathbb{R}^n$. If $Q$ is an arbitrary point on the hypersurface $\Sigma$ given by coordinates $\psi_\xi(t_j)$, the Euclidean distance between the points $P$ and $Q$ is given by

$$\text{dist}(P, Q) = \sqrt{\sum_{j=1}^n (\psi_\xi(t_j) - y_j)^2}.$$
The minimum of this distance is achieved when $Q$ is the orthogonal projection of $P$ onto $\Sigma$, see Fig. 2.1 b.

This is equivalent with minimizing the sum of the squared errors

$$F(\xi) = \sum_{j=1}^{n} (\psi_{\xi}(t_j) - y_j)^2.$$  

Note that the partial derivatives can be written as

$$\partial_{\xi_k} F(\xi) = 2 \sum_{j=1}^{n} (\psi_{\xi}(t_j) - y_j) \partial_{\xi_k} \psi_{\xi}(t_j) = 2 \langle \overrightarrow{PQ}, \partial_{\xi_k} \psi_{\xi} \rangle.$$  

Using that $\partial_{\xi_k} \psi_{\xi}$ is a tangent vector to the hypersurface $\Sigma$ at $Q$, it follows that the normal equation

$$\partial_{\xi_k} F(\xi) = 0, \quad k = 1, \ldots, n$$

is equivalent with the fact that $PQ$ is normal to the tangent plane $T_Q \Sigma$.

To conclude, the best fitting curve (regression curve), $\psi_{\xi}$, is the one for which the parameter $\xi$ satisfies the equations

$$\sum_{j=1}^{n} (\psi_{\xi}(t_j) - y_j) \partial_{\xi_k} \psi_{\xi}(t_j) = 0, \quad k = 1, \ldots, n, \quad (2.1.1)$$
Sometimes these equation might not have solution, fact that translates as saying that the model $\psi_{\xi}$ is not a good candidate for fitting the observations. Other times the solution exists, and might be even unique, but there is no explicit solution available for it. In these cases the use of numerical methods is the only way to find the parameter.

If the observations point $P$ is close enough to the regression surface $\psi_{\xi}$, then standard geometric properties state the existence and uniqueness of the orthogonal projection $Q$, which corresponds to the existence of the best fitting model. Here, an important role is played by choosing the correct model $\psi_{\xi}$, which might be done such that it avoids the existence of a secondary orthogonal projection.

If $P$ is not close enough to the regression surface $\psi_{\xi}$, in order for $PQ$ to be of minimum length, some convexity condition needs to be satisfied. To see this, we compute the Hessian of $F$

$$\partial_{\xi_j} \partial_{\xi_k} F(\xi) = 2 \langle P\overrightarrow{Q}, \partial_{\xi_j} \partial_{\xi_k} \psi_{\xi} \rangle + 2 \langle \partial_{\xi_j} \psi_{\xi}, \partial_{\xi_k} \psi_{\xi} \rangle$$

$$= 2 \langle P\overrightarrow{Q}, \partial_{\xi_j} \partial_{\xi_k} \psi_{\xi} \rangle + 2g_{jk}(\xi),$$

where $g_{jk}(\xi)$ denotes the Riemannian metric on the hypersurface $\Sigma$. It is well known that $g_{jk}(\xi)$ is positive definite. If the surface is convex, then $\partial_{\xi_j} \partial_{\xi_k} \psi_{\xi}$ is also positive definite, and if assume that $P$ is on the correct side of the surface, then the first term is positive definite. This sums up to a positive definite behavior for the Hessian of $F$.

In the following we shall present a few classical examples of regression that have explicit solutions.

**Example 2.1.1 (Linear regression)** In this case the family of curves $\psi_{\xi}(t) = mt + b$ depends on two parameters $\xi = (m, b)$. This forms a 2-dimensional surface $\Sigma$, which actually is a plane. The sum of squared errors becomes in this case

$$F(m, b) = \sum_{j=1}^{n} (mt_j + b - y_j)^2.$$

Since $\partial_m \psi_{\xi} = (t_1, \cdots, t_n)$ and $\partial_b \psi_{\xi} = (1, \cdots, 1)$, then the regression equations (2.1.1) become the following linear system in $m$ and $b$

$$\sum_{j=1}^{n} (mt_j + b - y_j)t_j = 0$$

$$\sum_{j=1}^{n} (mt_j + b - y_j) = 0.$$
The unique solution is given by

\[ m^* = \frac{n \sum t_j y_j - \sum t_j \sum y_j}{n \sum t_j^2 - (\sum t_j)^2}, \quad b^* = \frac{\sum t_j^2 \sum y_j - \sum t_j \sum t_j y_j}{n \sum t_j^2 - (\sum t_j)^2}, \]

and the regression line has the equation \( \psi(t) = m^* t + b^* \).

**Exercise 2.1.2 (Quadratic regression)** Consider \( n \) observations \((t_1, y_1), \ldots, (t_n, y_n)\). Use the method of least squares to find formulas for the regression parameters \( a, b \) and \( c \) for the quadratic model \( \psi_{a,b,c}(t) = at^2 + bt + c \) which best fits the data.

**Exercise 2.1.3** Consider the observations \((t_j, y_j)\) given by the following table

<table>
<thead>
<tr>
<th>( t_j )</th>
<th>( y_j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.14</td>
<td>2.44</td>
</tr>
<tr>
<td>0.60</td>
<td>1.80</td>
</tr>
<tr>
<td>1.30</td>
<td>1.30</td>
</tr>
<tr>
<td>2.16</td>
<td>1.22</td>
</tr>
<tr>
<td>2.60</td>
<td>1.30</td>
</tr>
<tr>
<td>3.00</td>
<td>1.42</td>
</tr>
<tr>
<td>3.60</td>
<td>2.00</td>
</tr>
<tr>
<td>4.00</td>
<td>2.60</td>
</tr>
<tr>
<td>4.37</td>
<td>3.80</td>
</tr>
<tr>
<td>4.42</td>
<td>4.05</td>
</tr>
<tr>
<td>4.58</td>
<td>4.80</td>
</tr>
</tbody>
</table>

(i) Use either the closed formulas developed in Exercise 2.1.2 or the Excel Solver feature to find the regression quadratic model \( \psi_{a,b,c}(t) = at^2 + bt + c \).

(ii) Plot the table data and the regression quadratic model.

**Exercise 2.1.4 (Exponential regression)** Consider the 2-parameters family of curves \( \psi_\xi(t) = be^{rt} \), with \( \xi = (b, r) \). Find the regression parameters formulas for this model.

**Exercise 2.1.5 (Logarithmic regression)** Find the regression parameters formulas for the logarithmic model

\[ \psi_{a,b}(t) = a + b \ln t. \]
Let $X^{(\xi)}_t$ be a family of stochastic processes depending on the parameter vector $\xi = (\xi_1, \cdots, \xi_k)$, and consider $n$ observations $(t_j, y_j)$, $1 \leq j \leq n$. The stochastic regression consists in finding the value of the parameter $\xi$ for which the process $X^{(\xi)}_t$ best fits the data, see Fig. 2.2 a. This means to find $\xi$ such that the expectation of the sum of errors

$$F(\xi) = \sum_{j=1}^{n} \mathbb{E}[(X^{(\xi)}_{t_j} - y_j)^2]$$

reaches its minimum value. The estimated parameter $\xi$ minimizes the error between the values predicted by the model $X^{(\xi)}_{t_j}$ and data observations, $y_j$. In order to further express the variational equations, we denote by $p_\xi(x, t)$ the probability density function of $X^{(\xi)}_t$, and briefly recall its main properties:

1) $p_\xi(x, t) \geq 0$

2) $\int p_\xi(x, t) \, dx = 1$

3) $P(a < X^{(\xi)}_t < b) = \int_a^b p_\xi(x, t) \, dx.$

Using that

$$\mathbb{E}[X^{(\xi)}_{t_j}] = \int x p_\xi(x, t_j) \, dx,$$
we can write the goodness of fit in terms of \( p_\xi \)

\[
F(\xi) = \sum_{j=1}^{n} E[(X_{t_j}^{(\xi)} - y_j)^2] \\
= \sum_{j=1}^{n} E[(X_{t_j}^{(\xi)})^2] - 2 \sum_{j=1}^{n} y_j E[X_{t_j}^{(\xi)}] + \|y\|^2 \\
= \sum_{j=1}^{n} \int x^2 p_\xi(x, t_j) \, dx - 2 \sum_{j=1}^{n} y_j \int x p_\xi(x, t_j) \, dx + \|y\|^2.
\]

It is not hard to see that the variational equations \( \partial_\xi F(\xi) = 0 \) can be written as

\[
\sum_{j=1}^{n} \int (x^2 - 2 y_j x) \partial_\xi p_\xi(x, t_j) \, dx = 0, \quad r = 1, \cdots, k.
\]

(2.2.3)

There is no general explicit formula for the solution \( \xi \). In the following we shall consider the regression for a few particular families of stochastic processes, and try to obtain explicit formulas for the best fit parameter \( \xi \) in these cases.

**Example 2.2.1** Let \( \varphi_\xi : \mathbb{R} \to \mathbb{R} \), with \( \xi = (\xi_1, \cdots, \xi_k) \), be a family of continuous functions, and consider the family of stochastic processes \( X_{t_j}^{(\xi)} = \varphi_\xi(W_t) \).

Denote by

\[
\phi_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}, \quad t > 0
\]

the probability density of the Brownian motion \( W_t \). Then the expectation of the sum of squares errors can be written in terms of \( \phi_t \) as

\[
F(\xi) = \sum_{j=1}^{n} E[(X_{t_j}^{(\xi)})^2] - 2 \sum_{j=1}^{n} E[X_{t_j}^{(\xi)}] + \|y\|^2 \\
= \sum_{j=1}^{n} \int \varphi_\xi^2(x) \phi_{t_j}(x) \, dx - 2 \sum_{j=1}^{n} y_j \int \varphi_\xi(x) \phi_{t_j}(x) \, dx + \|y\|^2.
\]

Hence, the variational equations \( \partial_\xi F(\xi) = 0 \) take the following form

\[
\sum_{j=1}^{n} \int (\varphi_\xi(x) - y_j) \partial_\xi \varphi_\xi(x) \phi_{t_j}(x) \, dx = 0, \quad r = 1, \cdots, k.
\]

(2.2.4)

Even this system is still too complex to be solved explicitly in \( \xi \). The next examples deal with some particular forms of the function \( \varphi_\xi \).
Example 2.2.2 Consider the 1-parameter family of functions $\varphi_\xi(x) = \sigma|x|$, with $\xi = \sigma > 0$. The corresponding stochastic process is the reflected Brownian motion $X_t = \sigma|W_t|$. This process would be a good candidate if $y_j \geq 0$ and $y_j$ does not display any upward trend, see Fig. 2.2 b.

The expectation of the sum of squares errors can be evaluated as

$$F(\sigma) = \sum_{j=1}^{n} \mathbb{E}[(\sigma|W_{t_j}| - y_j)^2]$$

$$= \sum_{j=1}^{n} \{ \sigma^2 \mathbb{E}[W_{t_j}^2] - 2\sigma y_j \mathbb{E}[|W_{t_j}|] \} + ||y||^2$$

$$= \sigma^2 \sum_{j=1}^{n} t_j - 2\sigma \sum_{j=1}^{n} y_j \sqrt{\frac{2t_j}{\pi}} + ||y||^2,$$

which is quadratic in the parameter $\sigma$. The minimum of $F(\sigma)$ is obtained for the value

$$\sigma^* = \sqrt{\frac{2}{\pi} \sum \sqrt{t_j} y_j}.$$

Exercise 2.2.3 Find the regression parameters formulas for the stochastic model

$$X_t = \sigma|W_t| + b.$$  

Example 2.2.4 Consider the 2-parameter family stochastic process $X_t = e^{\sigma W_t + \mu}$. This would be a good choice in the case when the data $y_j$ is positive and also shows an upward trend. The associated goodness of fit function is

$$F(\mu, \sigma) = \sum_{j=1}^{n} \mathbb{E}[(e^{\sigma W_{t_j} + \mu} - y_j)^2]$$

$$= \sum_{j=1}^{n} \mathbb{E}[e^{2\sigma W_{t_j} + 2\mu} - 2y_j e^{\sigma W_{t_j} + \mu} + y_j^2]$$

$$= e^{2\mu} \sum_{j=1}^{n} e^{2\sigma^2 t_j} - 2e^\mu \sum_{j=1}^{n} y_j e^{\frac{1}{2} \sigma^2 t_j} + ||y||^2$$

$$= A(\sigma)x^2 + B(\sigma)x + C,$$

where

$$A(\sigma) = \sum_{j=1}^{n} e^{2\sigma^2 t_j}, \quad B(\sigma) = -2 \sum_{j=1}^{n} y_j e^{\frac{1}{2} \sigma^2 t_j},$$

$$C = ||y||^2, \quad x = e^\mu > 0.$$
Since $F$ is quadratic in $x$, its minimum is reached for

$$x = -\frac{B(\sigma)}{2A(\sigma)}, \quad (2.2.5)$$

which can be written equivalently as

$$\mu = \ln \left[ \sum_{y_j} y_j e^{\frac{1}{2} \sigma^2 t_j} \right]. \quad (2.2.6)$$

The variational equation $\partial_{\sigma} F(\mu, \sigma) = 0$ is equivalent to

$$x = -\frac{B'(\sigma)}{A'(\sigma)}. \quad (2.2.7)$$

Equating relations (2.2.5) and (2.2.7) imply the following equation satisfied by the critical value of $\sigma$

$$\frac{B(\sigma)}{2A(\sigma)} = \frac{B'(\sigma)}{A'(\sigma)}.$$ 

This can be written explicitly as

$$\sum_{t_j} e^{2\sigma^2 t_j} \sum_{t_j} e^{2\sigma^2 t_j} = 2 \sum_{t_j} y_j e^{\frac{1}{2} \sigma^2 t_j} \sum_{t_j} y_j e^{\frac{1}{2} \sigma^2 t_j}. \quad (2.2.8)$$

The critical point $(\mu^*, \sigma^*)$ of $F(\mu, \sigma)$ is the solution of the system of equations (2.2.6)-(2.2.8). The regression model is given by $X_t = e^{\sigma^* W_t + \mu^*}$.

**Exercise 2.2.5** Find the stochastic regression formulas for the following models

1. $X_t = \sigma W_t^2$
2. $X_t = \mu t + \sigma W_t$
3. $X_t = \mu t + \sigma W_t^2$
4. $X_t = \mu t + \sigma |W_t|$.

### 2.3 Calibration

Financial models, either deterministic or stochastic, depend on one or more parameters $\xi$. These parameters have to be determined from the available market data, such as stock prices or actively traded securities prices $y_j$ observed at time instances $t_j$. The procedure by which one infers the parameters values from market prices is called calibration of the model to the market.
Example 2.3.1 Consider the constant yield deterministic model for the price of a zero-coupon bond at time $t$, paying $1 at the expiration time $T$

$$P(t, T) = e^{-r(T-t)}.$$ 

In order to estimate the value of the rate $r$ it suffices to observe the market prices of $n$ bonds with time to expiration $\tau_j = T - t_j$

$$y_j = P(t_j, T), \quad j = 1, \ldots, n.$$ 

The rate $r$ has to be estimated such that the error between the observed and predicted bond prices is the smallest possible. In order to do this, form the goodness of fit function

$$F(r) = \sum_{j=1}^{n} \left( P(t_j, T) - y_j \right)^2 = \sum_{j=1}^{n} \left( e^{-r\tau_j} - y_j \right)^2$$

and set the normal equation $F'(r) = 0$, which does not have an explicit solution. In order to further simplifications, substitute $u = e^{-r}$, and consider the value of $u$ for which

$$G(u) = \sum_{j=1}^{n} \left( u^{\tau_j} - y_j \right)^2$$

is optimized. The equation $G'(u) = 0$ is equivalent to

$$\sum_{j=1}^{n} \tau_j u^{\tau_j - 1} (u^{\tau_j} - y_j) = 0. \quad (2.3.9)$$

This equation still does not have an explicit solution and consequently, numerical methods have to be employed for finding an estimation of the solution. Most numerical methods ask for an initial guess, which is of foremost importance for the convergence of the numerical procedure. One way to do this is to consider the solution of each of the equations $e^{-r(T-t_j)} = y_j$, given by $r_j = -\frac{\ln y_j}{T-t_j}$, and then consider their average

$$\bar{r} = \frac{1}{n} \sum_{j=1}^{n} r_j$$

as an initial guess. This will correspond to $\bar{u} = e^{-\bar{r}}$ as an initial start for the equation $(2.3.9)$.

In the following we shall describe its geometrical significance. Relation $(2.3.9)$ can be written equivalently as

$$\langle c(u), \mathbf{\bar{\rho}}(u) \rangle = 0,$$
Figure 2.3: a. PQ is normal to the curve c(u); b. The iterative method providing a sequence \((Q_n)\) of approximations of the orthogonal projection \(Q\).

where \(c'(u) = (\tau_1 u^{\tau_1-1}, \cdots, \tau_n u^{\tau_n-1})\) is the velocity vector to the \(n\)-dimensional curve \(c(u) = (u^{\tau_1}, \cdots, u^{\tau_n})\), which traces the prices of \(n\) bonds as a function of \(u = e^{-r}\). The vector \(\vec{P}(u)\) denotes the position vector of the curve \(c(u)\) with respect to the observations point \(P = (y_1, \cdots, y_n)\). Condition (2.3.9) represents the orthogonality relation between the vector \(\vec{P}(u)\) and the curve \(c(u)\). It is equivalent with stating that the normal plane to the curve at the point \(c(u)\) passes through the point \(P = (y_j)\), see Fig. 2.3 a.

The normal projection onto the curve \(c(u)\) can be obtained by an iterative procedure, which is an alteration of Newton's tangents method. This will be presented in the following. Let \(Q_0 = c(\bar{u})\) be the initial guess. Consider the tangent line to \(c(u)\) at \(Q_0\), called \(\ell_0\). Let \(Q_1\) be the point where the normal line from \(P\) to \(\ell_0\) intersects the curve \(c(u)\). Substitute \(Q_1\) in the role of \(Q_0\) and construct the point \(Q_2\). Inductively, we construct a sequence of points \(Q_n\) on the curve \(c(u)\). Its limit, \(Q = \lim_{n \to \infty} Q_n\), has the property that \(PQ\) is normal on the curve \(c(u)\). We skip the proof, but the reader can see that this seems clear at least for the case depicted in Fig. 2.3 b.

**Example 2.3.2** The price of a bond which pays $1 at maturity \(T = 10\) years is given by \(P(t,T) = e^{-r(T-t)}\). Use the method of least squares to estimate the parameter \(r\) from the following market data:

<table>
<thead>
<tr>
<th>(t_j)</th>
<th>(P(t_j, T))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.04</td>
</tr>
<tr>
<td>2</td>
<td>0.07</td>
</tr>
<tr>
<td>4</td>
<td>0.28</td>
</tr>
<tr>
<td>6</td>
<td>0.64</td>
</tr>
<tr>
<td>8</td>
<td>0.92</td>
</tr>
</tbody>
</table>
The goodness of fit function in this case is given by

\[ F(r) = (e^{-9r} - 0.04)^2 + (e^{-8r} - 0.07)^2 + (e^{-6r} - 0.280)^2 \\
+ (e^{-4r} - 0.64)^2 + (e^{-2r} - 0.92)^2, \quad 0 < r < 1. \]

Its graph is convex and is given by Fig.2.4 a. The minimum is realized for the value \( r^* = 0.18561 \), so the best fitting curve has the expression \( P(t,10) = e^{-0.18561(10-t)} \), see Fig.2.4 b.

**Example 2.3.3** The cash value in the case of stochastic spot rates

\[ r_t = r + \sigma W_t \]

is given by the model

\[ M_t = M_0 e^{rt + \sigma Z_t}, \]

where \( Z_t = \int_0^t W_s \, ds \) is the integrated Brownian motion. The parameters \( r \) and \( \sigma \) can be obtained by calibration. Consider the market observations \( y_j = M_{t_j}, j = 1, \ldots, n \). The goodness of fit function can be computed as

\[ F(r, \sigma) = \sum_{j=1}^n \mathbb{E}[(M_{t_j} - y_j)^2] \]

\[ = M_0^2 \sum_{j=1}^n e^{2rt_j} e^{\frac{4}{\pi} \sigma^2 t_j^3} - 2M_0 \sum_{j=1}^n y_j e^{rt_j + \frac{4}{\pi} \sigma^2 t_j^3} + \|y\|^2. \]

It is not hard to see that the normal equations

\[ \partial_r F = 0, \quad \partial_\sigma F = 0 \]
can be written explicitly in the form
\[
\sum_{j=1}^{n} e^{2rt_j} e^{2\sigma^2 t_j^3} = \sum_{j=1}^{n} e^{rt_j} e^{\frac{1}{2}\sigma^2 t_j^3} y_j \quad t_j
\]
\[
2 \sum_{j=1}^{n} e^{2rt_j} e^{2\sigma^2 t_j^3} = \sum_{j=1}^{n} e^{rt_j} e^{\sigma^2 t_j^3} y_j^3.
\]

Again, there is no explicit formula for the solution of the previous system. Numerical methods need to be used.

**Example 2.3.4** A popular model for the stock price \( S_t \) is given by the geometrical Brownian motion
\[
S_t = S_0 e^{(\mu - \sigma^2/2)t + \sigma W_t}.
\] (2.3.10)

This model depends on two parameters: \( \mu \), the return, and \( \sigma \), the stock volatility. The calibration to the market of the model consists in evaluating the parameters \( \mu \) and \( \sigma \) such that the model is the most consistent with the market data observations \( y_j = S_{t_j}, j = 1, \ldots, n \), in the sense of the least squares error.

Let \( r = \mu - \sigma^2/2 \). Then the goodness of fit function is
\[
F(r, \sigma) = \sum_{j=1}^{n} \mathbb{E}[(S_{t_j} - y_j)^2]
\]
\[
= S_0^2 \sum_{j=1}^{n} e^{2(r + \sigma^2)t_j} - 2S_0 \sum_{j=1}^{n} y_j e^{(r + \frac{1}{2}\sigma^2)t_j} + \|y\|^2,
\]

where we used the formula \( \mathbb{E}[e^{cW_t}] = e^{\frac{1}{2}c^2t}, c \in \mathbb{R} \). A straightforward computation shows that the normal equations
\[
\partial_r F = 0, \quad \partial_\sigma F = 0
\]
after some algebraic manipulations become
\[
S_0 \sum_{j=1}^{n} t_j e^{2(r + \sigma^2)t_j} = \sum_{j=1}^{n} y_j t_j e^{(r + \frac{1}{2}\sigma^2)t_j} \quad (2.3.11)
\]
\[
2\sigma S_0 \sum_{j=1}^{n} t_j e^{2(r + \sigma^2)t_j} = \sigma \sum_{j=1}^{n} y_j t_j e^{(r + \frac{1}{2}\sigma^2)t_j}. \quad (2.3.12)
\]

If \( \sigma \neq 0 \), then dividing the second equation by \( \sigma \) and subtracting from the first one yields
\[
S_0 \sum_{j=1}^{n} t_j e^{2(r + \sigma^2)t_j} = 0,
\]
which does not have any real solutions. Hence \( \sigma = 0 \). The parameter \( r \) can be found from equation (2.3.11). It turns out that the model (2.3.10) which best fits the data in the least squares sense is in fact deterministic, given by \( S_t = S_0 e^{r^* t} \), where \( r^* \) is the solution of equation (2.3.11).

2.4 Alternatives to the Method of Least Squares

The method of least squares uses the minimization of the Euclidean distance to estimate parameters. There are also other measures whose optimization leads to parameter estimations. We shall encounter some of the most convenient in the following.

2.4.1 The Maximum Likelihood Method

Let \( P(A \mid \xi) \) be the probability of the event \( A \), given the parameter \( \xi \). Then consider the observed data \( \{y_1, \cdots, y_n\} \) at time instances \( \{t_1, \cdots, t_n\} \) and try to match the stochastic process \( X_t^{(\xi)} \), which depends on the parameter \( \xi \), to the given data. Since we do not expect this to be a perfect fit, we need to determine the value of the model parameter \( \xi \) which makes most likely to observe \( X_t^{(\xi)} \) in a proximity of \( y_j \) at time \( t_j \). In order to do this, consider the probability

\[
P(X_t^{(\xi)} \in (y_j, y_j + dy) \mid \xi) = p_\xi(y_i, t_i) dy, \quad 1 \leq j \leq n,
\]

and define the likelihood function of parameter \( \xi \), given \( (t_j, y_j) \), by

\[
L(\xi; y_j, t_j) = \prod_{j=1}^{n} p_\xi(y_i, t_i).
\]

The maximum likelihood method consists in optimizing the likelihood function \( L(\xi; y_j, t_j) \). This way, all probabilities (2.4.13) are simultaneously maximized. The parameter \( \xi \) is obtained as a solution of the equation

\[
\frac{\partial L(\xi; y_j, t_j)}{\partial \xi} = 0.
\]

However, it is computationally more convenient to optimize the log-likelihood function

\[
G(\xi; y_j, t_j) = \ln L(\xi; y_j, t_j) = \sum_{j=1}^{n} \ln p_\xi(y_j, t_j).
\]

This is based on the fact that \( L \) and \( G \) have the same critical points \( \xi \). These points satisfy the equation

\[
\sum_{j=1}^{n} \frac{\partial p_\xi(y_j, t_j)}{p_\xi(y_j, t_j)} = 0.
\]
Example 2.4.1 Consider the observations \(((t_j, y_j))\) and the model \(X_t = \mu t + \sigma W_t\), where \(W_t\) is a Brownian motion. The model parameters in this case are \(\xi = (\mu, \sigma)\). Using that \(W_t \sim N(0, t)\), the associated likelihood function becomes

\[
L(\mu, \sigma; y_j, t_j) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2 t_j}} e^{-\frac{(y_j - \mu t_j)^2}{2\sigma^2 t_j}},
\]

and the log-likelihood function is

\[
G(\mu, \sigma; y_j, t_j) = -\ln(2\pi)^{n/2} - n \ln \sigma - \frac{1}{2} \sum_{j=1}^{n} \ln t_j - \frac{1}{2} \frac{1}{\sigma^2} \sum_{j=1}^{n} \frac{(y_j - \mu t_j)^2}{t_j}.
\]

Since

\[
\partial_\mu G(\mu, \sigma; y_j, t_j) = \frac{1}{\sigma^2} \left( \sum_{j=1}^{n} y_j - \mu \sum_{j=1}^{n} t_j \right)
\]

\[
\partial_\sigma G(\mu, \sigma; y_j, t_j) = \frac{1}{\sigma} \left( -n + \frac{1}{\sigma^2} \sum_{j=1}^{n} \frac{(y_j - \mu t_j)^2}{t_j} \right),
\]

the model parameters are given by

\[
\mu = \frac{\sum_{j=1}^{n} y_j}{\sum_{j=1}^{n} t_j}
\]

\[
\sigma^2 = \frac{1}{n} \sum_{j=1}^{n} \frac{(y_j - \mu t_j)^2}{t_j}.
\]

It is worth noting that the model parameters obtained by the least squares method are different than the ones obtained previously. They are estimated by the deterministic model

\[
\mu = \frac{\sum_{j=1}^{n} y_j}{\sum_{j=1}^{n} t_j^2}
\]

\[
\sigma^2 = 0.
\]

Exercise 2.4.2 Consider the stochastic process \(X_t = \sigma |W_t|\).

(i) Find the density function of \(X_t\);

(ii) Use the maximum likelihood method for the model \(X_t\) to estimate the parameter \(\sigma\) in terms of the given data \(((t_j, y_j))\).

Exercise 2.4.3 Consider the observations \(((t_j, y_j))\) and the geometric Brownian motion model \(X_t = e^{\mu t + \sigma W_t}\). Find the maximum likelihood estimation of the parameters \(\mu\) and \(\sigma\).
2.4.2 The Maximum Entropy Method

Consider the model $X_t^{(\xi)}$ and the given data $(t_j, y_j)$. Another measure that can be optimized is the entropy

$$H(\xi; y_j, t_j) = -\sum_{j=1}^{n} p_\xi(y_j, t_j) \ln p_\xi(y_j, t_j).$$

The log-likelihood $\ln p_\xi(y_j, t_j)$ represents the information that the model $X_t^{(\xi)}$ is in the proximity of $y_j$ at time $t_j$ for the model parameter value $\xi$. This method intends to find the value of $\xi$ for which the average information associated with the data $(t_j, y_j)$ is maximum.
Chapter 3

Modeling Stochastic Rates

Elementary Calculus provides powerful methods of modeling phenomena from the real world. However, the real world is imperfect, due to multiple perturbation effects, and in order to study it, one needs to employ methods of Stochastic Calculus.

3.1 Deterministic versus Stochastic Calculus

In this section we shall consider a simple finance problem and solve it in both frames of deterministic and stochastic calculus and then compare the results.

The deterministic model
Consider the amount of money $M(t)$ at time $t$ invested in a bank account that pays interest at a constant rate $r$. The differential equation which models this problem is

$$dM(t) = rM(t)dt,$$  \hspace{1cm} (3.1.1)

i.e. the instantaneous relative rate of return $\frac{dM(t)}{M(t)}$ is proportional with the time interval $dt$. Given the initial investment $M(0) = M_0$, the account balance at time $t$ is given by $M(t) = M_0e^{rt}$, which is the solution of (3.1.1).

The stochastic model
In the real world the interest rate $r$ is not constant. It may be assumed constant only for a small amount of time, such as one day or one week. The interest rate changes unpredictably over time. This can be modeled in a few different ways. For instance, we may assume that the interest rate at time $t$ is given by the diffusion process $r_t = r + \sigma W_t$, where $\sigma > 0$ is a constant that

\hspace{1cm}

\footnotesize

\begin{itemize}
  \item At the time this book was written, the interest rates in US were frozen to a very low level for a few years, but in a normal market the rates fluctuate.
\end{itemize}
controls the volatility of the rate, and \( W_t \) is a Brownian motion process. The process \( r_t \) represents a diffusion that starts at \( r_0 = r \), with constant mean, \( \mathbb{E}[r_t] = r \), and variance proportional with the time elapsed, \( \text{Var}[r_t] = \sigma^2 t \). With this change in the model, the account balance at time \( t \) becomes a stochastic process \( M_t \) that satisfies the following stochastic differential equation

\[
\frac{dM_t}{M_t} = (r + \sigma W_t) dt, \quad t \geq 0.
\]  

(3.1.2)

Solving the equation

In order to solve this equation, we write it as \( dM_t - r_t M_t dt = 0 \) and multiply by the integrating factor \( e^{-\int_0^t r_s \, ds} \). We can check that

\[
d\left(e^{-\int_0^t r_s \, ds} \right) = -e^{-\int_0^t r_s \, ds} r_t dt
\]

\[
dM_t d\left(e^{-\int_0^t r_s \, ds} \right) = 0,
\]

since \( dt^2 = dt dW_t = 0 \). Using the product rule, the equation becomes exact

\[
d \left(M_t e^{-\int_0^t r_s \, ds} \right) = 0.
\]

Integrating yields the solution

\[
M_t = M_0 e^{\int_0^t r_s \, ds} = M_0 e^{\int_0^t (r + \sigma W_s) \, ds}
\]

\[
= M_0 e^{rt + \sigma Z_t},
\]

where \( Z_t = \int_0^t W_s \, ds \) is the integrated Brownian motion process. Since it is known that \( Z_t \sim N(0,t^3/3) \), then the moment generating function of \( Z_t \) is

\[
m(\sigma) = \mathbb{E}[e^{\sigma Z_t}] = e^{\sigma^2 t^3/6}.
\]

This implies

\[
\mathbb{E}[e^{\sigma W_s}] = e^{\sigma^2 t^3/6};
\]

\[
\text{Var}[e^{\sigma W_s}] = m(2\sigma) - m(\sigma) = e^{\sigma^2 t^3/3}(e^{\sigma^2 t^3/3} - 1).
\]

Then the mean and variance of the solution \( M_t = M_0 e^{rt + \sigma Z_t} \) are

\[
\mathbb{E}[M_t] = M_0 e^{rt} \mathbb{E}[e^{\sigma Z_t}] = M_0 e^{rt + \sigma^2 t^3/6};
\]

\[
\text{Var}[M_t] = M_0^2 e^{2rt} \text{Var}[e^{\sigma Z_t}] = M_0^2 e^{2rt + \sigma^2 t^3/3}(e^{\sigma^2 t^3/3} - 1).
\]

Conclusion

We shall make a few interesting remarks. If \( M(t) \) and \( M_t \) represent the balance at time \( t \) in the cases \( r \) constant and \( r_t \) stochastic, respectively, then

\[
\mathbb{E}[M_t] = M_0 e^{rt} e^{\sigma^2 t^3/6} > M_0 e^{rt} = M(t).
\]
This means that we expect to have more money in the account in the second case rather than in the first one. Similarly, a bank can expect to make more money when lending at a stochastic interest rate than at a constant interest rate. This inequality is due to the convexity of the exponential function. If \( X_t = rt + \sigma Z_t \), then Jensen’s inequality yields
\[
\mathbb{E}[e^{X_t}] \geq e^{\mathbb{E}[X_t]} = e^{rt}.
\]

### 3.2 Langevin’s Equation

We shall consider another stochastic extension of the equation (3.1.1). We denote now by \( M_t \) the wealth of a bank at time \( t \), which is invested at rate \( r \). At the same time, we shall allow for continuously random deposits and withdrawals which can be modeled by an unpredictable term, given by \( \alpha dW_t \), with \( \alpha \) constant. This makes sense if a large number of customers execute transactions at any time. The obtained equation
\[
dM_t = r M_t dt + \alpha dW_t, \quad t \geq 0 \tag{3.2.3}
\]
is called Langevin’s equation.

**Solving the equation**

We shall solve it as a linear stochastic equation. Multiplying by the integrating factor \( e^{-rt} \) yields
\[
d(e^{-rt} M_t) = \alpha e^{-rt} dW_t.
\]
Integrating we obtain
\[
e^{-rt} M_t = M_0 + \alpha \int_0^t e^{-rs} dW_s.
\]
Hence the solution is
\[
M_t = M_0 e^{rt} + \alpha \int_0^t e^{r(t-s)} dW_s. \tag{3.2.4}
\]

This is called the Ornstein-Uhlenbeck process. Since the last term is a Wiener integral, and hence it is normally distributed, we have that \( M_t \) is Gaussian with the mean
\[
\mathbb{E}[M_t] = M_0 e^{rt} + \mathbb{E}[\alpha \int_0^t e^{r(t-s)} dW_s] = M_0 e^{rt}
\]
and variance
\[
\text{Var}[M_t] = \text{Var} \left[ \alpha \int_0^t e^{r(t-s)} dW_s \right] = \frac{\alpha^2}{2r} (e^{2rt} - 1).
\]
It is worth noting that the expected balance is equal to $M_0 e^{rt}$, i.e. is the balance in a world where $r$ is constant. The variance for $t$ small is approximately equal to $\alpha^2 t$, which is the variance of $\alpha W_t$.

If the constant $\alpha$ is replaced by a random function $\alpha(t, W_t)$, the equation becomes

$$dM_t = rM_t dt + \alpha(t, W_t) dW_t, \quad t \geq 0.$$  

Using a similar argument we arrive at the following solution:

$$M_t = M_0 e^{rt} + \int_0^t e^{r(t-s)} \alpha(t, W_t) dW_s. \tag{3.2.5}$$

This process is not Gaussian. Its mean and variance are given by

$$\begin{align*}
\mathbb{E}[M_t] &= M_0 e^{rt} \\
\text{Var}[M_t] &= \int_0^t e^{2r(t-s)} \mathbb{E}[\alpha^2(t, W_t)] ds.
\end{align*}$$

The integral in equation (3.2.5) can be computed explicitly in just a few cases. For instance, if $\alpha(t, W_t) = e^{\sqrt{2r} W_t}$, using Application ?? with $\lambda = \sqrt{2r}$, we can work out for (3.2.5) an explicit form of the solution

$$M_t = M_0 e^{rt} + \int_0^t e^{r(t-s)} e^{\sqrt{2r} W_t} dW_s.$$  

The previous interest rate model is not realistic since it allows for negative rates. The Brownian motion hits $-r$ after a finite time with probability 1. However, it might be a good stochastic model for something such as the evolution of population, since in this case the rate can be negative.

### 3.3 Equilibrium Models

Let $r_t$ denote the spot rate at time $t$. This is the rate at which one can invest for the shortest period of time.\(^2\) For the sake of simplicity, we assume the interest

\(^2\)This is also called the short-time rate or the instantaneous rate.
rate $r_t$ satisfies an equation with one source of uncertainty.\footnote{This means the stochastic differential equation involves only one Brownian motion.} If the spot rates were differentiable, the drift term affected by the random fluctuations of the market would be written as

$$\frac{dr_t}{dt} = m(r_t) + \text{“noise”}.$$ 

This can be formalized as a stochastic differential equation

$$dr_t = m(r_t)dt + \sigma(r_t)dW_t.$$ \hspace{1cm} (3.3.6)

In this model the drift rate and volatility of the spot rate $r_t$ do not depend explicitly on the time $t$. There are several classical choices for $m(r_t)$ and $\sigma(r_t)$ that will be discussed in the following sections. It is worth noting that in the case of equilibrium models the spot rate $r_t$ determines the term structure and the bond price. We shall deal with this subject in more detail in Chapter 4.

3.3.1 Rendleman and Bartter Model

The model introduced in 1986 by Rendleman and Bartter [17] assumes that the short-time rate satisfies the process

$$dr_t = \mu r_t dt + \sigma r_t dW_t.$$ 

The growth rate $\mu$ and the volatility $\sigma$ are considered constants. This equation describes a geometric Brownian motion and its solution is given by

$$r_t = r_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}.$$ 

The distribution of $r_t$ is log-normal, with

$$\ln \frac{r_t}{r_0} \sim N\left(\mu - \frac{\sigma^2}{2}, \sigma^2 t\right).$$

Its mean and variance are given by

$$\mathbb{E}[r_t] = r_0 e^{(\mu - \frac{\sigma^2}{2})t} \mathbb{E}[e^{\sigma W_t}] = r_0 e^{(\mu - \frac{\sigma^2}{2})t} e^{\sigma^2 t/2} = r_0 e^{\mu t}.$$ 

$$\text{Var}[r_t] = r_0^2 e^{2(\mu - \frac{\sigma^2}{2})t} \text{Var}[e^{\sigma W_t}] = r_0^2 e^{2(\mu - \frac{\sigma^2}{2})t} e^{\sigma^2 t} (e^{\sigma^2 t} - 1) = r_0^2 e^{2\mu t} (e^{\sigma^2 t} - 1).$$

One of the disadvantages of this model is that in the long run $\mathbb{E}[r_t]$ becomes unboundedly large. This is the reason why the next two models incorporate the mean reverting phenomenon of interest rates. This means that in the long run the rate converges towards an average level. These models are more realistic and are based on economic arguments.
3.3.2 Vasicek Model

Vasicek’s assumption is that the short-term interest rates should satisfy the mean reverting stochastic differential equation

\[ dr_t = a(b - r_t)dt + \sigma dW_t, \]

with \( a, b, \sigma \) positive constants, see Vasicek [19].

Assuming the spot rates are deterministic, we take \( \sigma = 0 \) and obtain the ordinary differential equation

\[ dr_t = a(b - r_t)dt, \]

with the solution

\[ r_t = b + (r_0 - b)e^{-at}. \]

This implies that the rate \( r_t \) is pulled towards level \( b \) at the rate \( a \). This means that, if \( r_0 > b \), then \( r_t \) is decreasing towards \( b \), and if \( r_0 < b \), then \( r_t \) is increasing towards the horizontal asymptote \( b \). The term \( \sigma dW_t \) in Vasicek’s model adds some “white noise” to the process. In the following we shall find an explicit formula for the spot rate \( r_t \) in the stochastic case.

**Proposition 3.3.1** The solution of the equation (3.3.7) is given by

\[ r_t = b + (r_0 - b)e^{-at} + \sigma e^{-at} \int_0^t e^{as} dW_s. \]

The process \( r_t \) is Gaussian with mean and variance

\[ \mathbb{E}[r_t] = b + (r_0 - b)e^{-at}; \]
\[ Var[r_t] = \frac{\sigma^2}{2a} (1 - e^{-2at}). \]
Proof: Multiplying by the integrating factor $e^{at}$ yields
\[ d(e^{at} r_t) = a b e^{at} \, dt + \sigma e^{at} \, dW_t. \]
Integrating between 0 and $t$ and dividing by $e^{at}$ we get
\[ r_t = r_0 e^{-at} + b e^{-at}(e^{at} - 1) + \sigma e^{-at} \int_0^t e^{as} \, dW_s. \]
Since the spot rate $r_t$ is the sum between the predictable function $r_0 e^{-at} + b e^{-at}(e^{at} - 1)$ and a multiple of a Wiener integral, from Proposition (??) it follows that $r_t$ is Gaussian, with
\[
\mathbb{E}[r_t] = b + (r_0 - b)e^{-at},
\]
\[
\text{Var}(r_t) = \text{Var}\left[\sigma e^{-at} \int_0^t e^{as} \, dW_s\right] = \sigma^2 e^{-2at} \int_0^t e^{2as} \, ds
\]
\[ = \sigma^2 e^{-2at} \frac{(1 - e^{-2at})}{2a}. \]

The following consequence explains the name of mean reverting rate.

**Remark 3.3.2** Since $\lim_{t \to \infty} \mathbb{E}[r_t] = \lim_{t \to \infty} (b + (r_0 - b)e^{-at}) = b$, the process is mean reverting, i.e. the spot rate $r_t$ tends to $b$ as $t \to \infty$. This limit is approached exponentially fast. The variance tends in the long run to $\frac{\sigma^2}{2a}$, so the random market fluctuations interfering with the mean reversion are of magnitude $\frac{\sigma}{\sqrt{2a}}$, so weak mean reverting processes have large volatility.

Since $r_t$ is normally distributed, the Vasicek model has been criticized because it allows for negative interest rates and unbounded large rates. See Fig.3.1 for a simulation of the short-term interest rates for the Vasicek model.

**Exercise 3.3.3** Let $0 \leq s < t$. Find the following expectations
(a) $\mathbb{E}[W_t \int_0^s W_u e^{au} \, du]$.  
(b) $\mathbb{E}[\int_0^t W_u e^{au} \, du \int_0^s W_v e^{av} \, dv]$. 

**Exercise 3.3.4** (a) Find the probability that $r_t$ is negative. 
(b) What happens with this probability when $t \to \infty$? 
(c) Find the rate of change of this probability. 
(d) Compute $\text{Cov}(r_s, r_t)$.

\[^4\]Starting with January 2015 the Swiss Bank start charging negative interest rates to banks and corporations.
3.3.3 Cox-Ingersoll-Ross Model

The Cox-Ingersoll-Ross (CIR) model assumes that the spot rates verify the stochastic equation

\[ dr_t = a(b - r_t)dt + \sigma \sqrt{r_t} dW_t, \]  

(3.3.8)

with a, b, \( \sigma \) positive constants, see [5]. Two main advantages of this model are:

- the process exhibits mean reversion.
- it is not possible for the interest rates to become negative.

A process that satisfies equation (3.3.8) is called a CIR process. In mathematics this is known under the name of Feller process, and its transition density is given by

\[
 p_t(y_0, y) = \frac{2a}{\sigma^2(e^{at} - 1)} \left( \frac{y}{y_0} \right)^{\nu/2} e^{a(1+\nu/2)t - \frac{2a(y_0 + e^{at}y)}{\sigma^2(e^{at} - 1)}} I_{\nu} \left( \frac{4a}{\sigma^2(e^{at} - 1)} \right),
\]

where \( \nu = \frac{2ab}{\sigma^2} - 1 \), and

\[
 I_{\nu}(z) = \sum_{n=0}^{\infty} \frac{1}{n!\Gamma(n + \nu + 1)} \left( \frac{z}{2} \right)^{\nu + 2n}
\]

is the Bessel function of index \( \nu \). The development of the aforementioned formula can be found in Gulisashvili [7].

In the following we shall compute its first two moments starting from the equation (3.3.8).
Multiplying by $e^{at}$ yields the exact equation
\[ d(e^{at}r_t) = abe^{at} dt + \sigma e^{at}\sqrt{r_t} dW_t. \]

Integrating between 0 and $t$ we obtain
\[ r_t = (r_0 - b)e^{-at} + b + \sigma e^{-at} \int_0^t e^{au} \sqrt{r_u} dW_u. \]

Using that an Ito integral has zero expectation, we get the following formula for the mean of the spot rate
\[ E[r_t] = \mu_t = (r_0 - b)e^{-at} + b. \tag{3.3.9} \]

Squaring the expression
\[ r_t = \mu_t + \sigma e^{-at} \int_0^t e^{au} \sqrt{r_u} dW_u, \]

and taking the expectation leads to
\[
E[r_t^2] - \mu_t^2 = \sigma^2 e^{-2at} E\left[ \left( \int_0^t e^{au} \sqrt{r_u} dW_u \right)^2 \right] \\
= \sigma^2 e^{-2at} \int_0^t e^{2au} E[r_u] du \\
= \sigma^2 e^{-2at} \int_0^t e^{2au} (b + (r_0 - b)e^{-au}) du \\
= \frac{\sigma^2}{a} e^{-2at} b (e^{at} - 1)(e^{at} + 1) + (r_0 - b)(e^{at} - 1) \\
= \frac{\sigma^2}{a} e^{-2at} (e^{at} - 1) \left( r_0 + \frac{b}{2} (e^{at} - 1) \right).
\]

Hence the variance is
\[ V ar(r_t) = \frac{\sigma^2}{a} e^{-2at} (e^{at} - 1) \left( r_0 + \frac{b}{2} (e^{at} - 1) \right). \]

A computation shows that
\[
\lim_{t \to \infty} E[r_t] = b, \quad \lim_{t \to \infty} V ar(r_t) = \frac{b \sigma^2}{2a},
\]

i.e. the process is mean reverting.

**Exercise 3.3.5** Let $\mu_k(t) = E[r_t^k]$ be the kth moment of $r_t$. Find a recursive formula for the moments $\mu_k(t)$, where $r_t$ satisfies the CIR process (3.3.8).
3.4 No-arbitrage Models

In the following models the drift rate is a function of time, which is chosen such that the model is consistent with the term structure. For these models the term structure is an input to the model.

3.4.1 Ho and Lee Model

The first no-arbitrage model was proposed in 1986 by Ho and Lee [9]. The model was presented initially in the form of a binomial tree. The continuous time-limit of this model is

\[ dr_t = \theta(t) dt + \sigma dW_t. \]

(3.4.10)

In this model \( \theta(t) \) is the average direction in which \( r_t \) moves and it is considered independent of \( r_t \), while \( \sigma \) is the standard deviation of the short rate. The solution process is Gaussian and is given by

\[ r_t = r_0 + \int_0^t \theta(s) ds + \sigma W_t. \]

If \( f(0, t) \) denotes the forward rate at time \( t \) as seen at time 0, it will be shown in section 4.5.1 that \( \theta(t) = \partial_t f(0, t) + \sigma^2 t \). Using that \( r_0 = f(0, 0) \) is the initial spot rate, the solution becomes

\[ r_t = f(0, t) + \frac{1}{2} \sigma^2 t^2 + \sigma W_t. \]

Exercise 3.4.1 Let \( r_t \) be the rate given by the Ho-Lee model.

(i) Prove that \( r_t = r_s + \int_s^t \theta(u) du + \sigma W_{t-s} \), for any \( 0 < s < t \);

(ii) Show the following integral average formula

\[ \frac{1}{t-s} \int_s^t \theta(u) du = E[r_t | \mathcal{F}_t] - r_s; \]

(iii) Prove that \( \theta(s) \) is the instantaneous average velocity of spot rates, i.e.

\[ \theta(s) = \lim_{t \to s} \frac{E[r_t | \mathcal{F}_s] - r_s}{t-s}. \]

3.4.2 Hull and White Model

The model proposed by Hull and White [10] is an extension of the Ho and Lee model model that incorporates mean reversion.
\[ dr_t = (\theta(t) - at)dt + \sigma dW_t, \]

(3.4.11)

with \( a \) and \( \sigma \) constants. We can solve the equation by multiplying by the integrating factor \( e^{at} \)

\[ d(e^{at} r_t) = \theta(t)e^{at}dt + \sigma e^{at}dW_t. \]

Integrating between 0 and \( t \) yields

\[ r_t = r_0 e^{-at} + e^{-at} \int_0^t \theta(s)e^{as} ds + \sigma e^{-at} \int_0^t e^{as} dW_s. \]

(3.4.12)

Since the first two terms are deterministic and the last is a Wiener integral, the process \( r_t \) is Gaussian.

The function \( \theta(t) \) can be calculated from the term structure (see section 4.5.2) as

\[ \theta(t) = \partial_t f(0,t) + af(0,t) + \frac{\sigma^2}{2a}(1 - e^{-2at}). \]

Then

\[
\int_0^t \theta(s)e^{as} ds = \int_0^t \partial_s f(0,s)e^{as} ds + a \int_0^t f(0,s)e^{as} ds \\
+ \frac{\sigma^2}{2a} \int_0^t (1 - e^{-2as})e^{as} ds \\
= f(0,t)e^{at} - r_0 + \frac{\sigma^2}{a^2} (\cosh(at) - 1),
\]

where we used that \( f(0,0) = r_0 \). The deterministic part of \( r_t \) becomes

\[ r_0 e^{-at} + e^{-at} \int_0^t \theta(s)e^{as} ds = f(0,t) + \frac{\sigma^2}{a^2} e^{-at} (\cosh(at) - 1). \]

An algebraic manipulation shows that

\[ e^{-at} (\cosh(at) - 1) = \frac{1}{2}(1 - e^{at})^2. \]

Substituting into (3.4.12) yields

\[ r_t = f(0,t) + \frac{\sigma^2}{2a^2}(1 - e^{at})^2 + \sigma e^{-at} \int_0^t e^{as} dW_s. \]
The mean and variance are:
\[
\begin{align*}
\mathbb{E}[r_t] &= f(0, t) + \frac{\sigma^2}{2a^2}(1 - e^{at})^2 \\
\text{Var}(r_t) &= \sigma^2 e^{-2at} \text{Var}\left[ \int_0^t e^{as} dW_s \right] = \sigma^2 e^{-2at} \int_0^t e^{2as} ds \\
&= \frac{\sigma^2}{2a}(1 - e^{-2at}).
\end{align*}
\]

### 3.5 Nonstationary Models

These models assume both $\theta$ and $\sigma$ as functions of time. In the following we shall discuss two models with this property.

#### 3.5.1 Black, Derman and Toy Model

The binomial tree of Black, Derman and Toy [3] is equivalent with the following continuous model of short-time rate

\[
\begin{equation}
\begin{aligned}
d(ln r_t) &= \left[ \theta(t) + \frac{\sigma'(t)}{\sigma(t)} \ln r_t \right] dt + \sigma(t) dW_t. \\
&= \theta(t) \sigma dt + \sigma(t) dW_t, \\
\end{aligned}
\end{equation}
\]

Making the substitution $u_t = \ln r_t$, we obtain a linear equation in $u_t$

\[
du_t = \left[ \theta(t) + \frac{\sigma'(t)}{\sigma(t)} u_t \right] dt + \sigma(t) dW_t.
\]

The equation can be written equivalently as

\[
\frac{\sigma(t) du_t - d\sigma(t) u_t}{\sigma^2(t)} = \frac{\theta(t)}{\sigma(t)} dt + dW_t,
\]

which after using the quotient rule becomes

\[
d\left[ \frac{u_t}{\sigma(t)} \right] = \frac{\theta(t)}{\sigma(t)} dt + dW_t.
\]

Integrating and solving for $u_t$ leads to

\[
u_t = \frac{u_0}{\sigma(0)} \sigma(t) + \sigma(t) \int_0^t \frac{\theta(s)}{\sigma(s)} ds + \sigma(t) W_t.
\]

This implies that $u_t$ is Gaussian and hence $r_t = e^{u_t}$ is log-normal for each $t$. Using $u_0 = \ln r_0$ and

\[
e^{\frac{u_0}{\sigma(0)} \sigma(t)} = e^\frac{\sigma(t)}{\sigma(0)} \ln r_0 = r_0^\frac{\sigma(t)}{\sigma(0)},
\]
we obtain the following explicit formula for the spot rate

$$r_t = r_0 e^{\sigma(t) \int_0^t \frac{\theta(s)}{\sigma^2(s)} ds} e^{\sigma(t)W_t}. \tag{3.5.14}$$

Since $\sigma(t)W_t$ is normally distributed with mean 0 and variance $\sigma^2(t)t$, the log-normal variable $e^{\sigma(t)W_t}$ has

$$\mathbb{E}[e^{\sigma(t)W_t}] = e^{\sigma^2(t)t/2}$$
$$\text{Var}[e^{\sigma(t)W_t}] = e^{\sigma(t)^2t}(e^{\sigma(t)^2t} - 1).$$

Hence

$$\mathbb{E}[r_t] = r_0 \int_0^t e^{\sigma(s) \int_0^s \frac{\theta(s)}{\sigma^2(s)} ds} e^{\sigma(s)t/2}$$
$$\text{Var}[r_t] = r_0^2 e^{2\sigma(t) \int_0^t \frac{\theta(s)}{\sigma^2(s)} ds} e^{\sigma(t)^2t}(e^{\sigma(t)^2t} - 1).$$

**Exercise 3.5.1** (a) Solve the Black, Derman and Toy model in the case when $\sigma$ is constant.
(b) Show that in this case the spot rate $r_t$ is log-normally distributed.
(c) Find the mean and the variance of $r_t$.

**Exercise 3.5.2** The following model was introduced by Black and Karasinski in 1991, see [4]:

$$d(\ln r_t) = \left[ \theta(t) - a(t) \ln r_t \right] dt + \sigma(t)dW_t.$$ 

Find the explicit formula for $r_t$. 

Chapter 4

Bonds, Forward Rates and Yield Curves

This chapter deals with the following financial equivalent instruments: bonds, forward rates and yield curves. Knowing any one of these provides full information on the other two. This makes possible to approach the interest rate market from multiple perspectives.

4.1 Bonds

A bond is a contract between a buyer and a financial institution (bank, government, etc) by which the financial institution agrees to pay a certain principal to the buyer at a determined time $T$ in the future, plus some periodic coupon payments done during the life time of the contract. If there are no coupons to be payed, the bond is called a zero coupon bond or a discount bond. Its price at any time $0 \leq t \leq T$ will be denoted by $P(t, T)$.

Consider first the case of deterministic rates:

(i) If spot rates are constant, $r_t = r$, then the price at time $t$ of a discount bond that pays off $1$ at time $T$ is given by

$$P(t, T) = e^{r(T-t)}.$$

(ii) If spot interest rates depend deterministic on time, $r_t = r(t)$, the bond formula becomes

$$P(t, T) = e^{-\int_t^T r(s) \, ds}.$$

However, in general, the spot rates $r_t$ are stochastic processes. In this case, the formula $e^{-\int_t^T r_s \, ds}$ is a random variable, and the price of the bond in this case can be calculated using a conditional expectation as of time $t$

$$P(t, T) = E[e^{-\int_t^T r_s \, ds} | F_t].$$
The bond price can be seen as the graph of a surface in $\mathbb{R}^3$ given by

$$(t, T) \to \Psi(t, T) = (t, T, P(t, T)),$$

where $0 \leq t \leq T$, see Fig. 4.1. The surface satisfies the boundary condition

$$P(t, T) \bigg|_{t=T} = 1, \quad \forall T \geq 0.$$

It is worth noting that the coordinate curves $T \to P(t, T)$ are deterministic while $t \to P(t, T)$ are stochastic. This makes possible to describe the bond surface by providing the stochastic model for the bond price, for instance, as

$$dP(t, T) = r_t P(t, T) dt + \nu(t, T) P(t, T) dW_t \tag{4.1.1}$$

$$P(T, T) = 1,$$

where $r_t$ is the spot rate and $\nu(t, T)$ is the bond volatility at time $t$ (and might also depend on the entire price history until time $t$). Since the bond pays off with certainty $\$1$ at time $T$, the volatility must decline to zero at maturity, i.e. $\nu(T, T) = 0$. The differential $dP(t, T)$ is taken with respect to $t$ and for avoiding confusions, it is denoted sometimes by $d_h P(t, T)$.

Equation (4.1.1) represents a one-factor model for the bond price, since the market uncertainty has only one noise source, $dW_t$. We shall stick with this simple model for the rest of the chapter.

### 4.2 Yield

If the interest rates are constant, $r_t = r$, the bond price formula $P(t, T) = e^{-r(T-t)}$ enables to retrieve the rate as $r = -\ln P(t, T)/(T-t)$. If the interest
rates are deterministic, \( r_t = r(t) \), then the similar formula
\[
\bar{r} = \frac{1}{T-t} \int_t^T r(s) \, ds
\]
provides the mean interest rate \( \bar{r} = \frac{1}{T-t} \int_t^T r(s) \, ds \).

In the case of stochastic interest rates, a similar rate of distinguished importance can be defined as in the following. The *yield* of the bond is defined as
\[
R(t, T) = -\ln P(t, T)/(T - t), \quad t \leq T.
\]
This is equivalently with the inverted formula
\[
P(t, T) = e^{-R(t, T)(T-t)}.
\]

The assignment
\[
(t, T) \rightarrow R(t, T)
\]
defines the *term structure surface*. The coordinate curves
\[
T \rightarrow R(t, T)
\]
are deterministic smooth curves, while
\[
t \rightarrow R(t, T)
\]
are stochastic processes. In the following we shall compute the law of this process. Assuming that \( P(t, T) \) satisfies equation (4.1.1), Ito’s formula provides
\[
d \ln P(t, T) = \frac{1}{P(t, T)} dP(t, T) - \frac{1}{2P(t, T)^2} dP(t, T)^2 \quad (4.2.2)
\]
\[
= (r_t - \frac{1}{2} \nu(t, T)^2) \, dt + \nu(t, T) \, dW_t.
\]

Applying the product rule and using the previous formula we have
\[
dR(t, T) = d\left( \frac{\ln P(t, T)}{t - T} \right)
\]
\[
= d\left( \frac{1}{t - T} \right) \ln P(t, T) + \frac{1}{t - T} d \ln P(t, T)
\]
\[
= \frac{1}{T - t} (R(t, T) - (r_t - \nu(t, T)^2/2)) \, dt - \frac{1}{T - t} \nu(t, T) \, dW_t.
\]

Hence, the yield in terms of the spot rate and bond price volatility is given by
\[
dR(t, T) = \frac{1}{T - t} (R(t, T) - r_t + \nu(t, T)^2/2) \, dt - \frac{1}{T - t} \nu(t, T) \, dW_t. \quad (4.2.3)
\]
The spot rate, \( r_t \), can be obtained from the yield as in the following. Consider a small interval of time \( \Delta t \). Then a discount bond with maturity \( T = t + \Delta t \) has the price given by

\[
P(t, t + \Delta t) = e^{-R(t,t+\Delta t)\Delta t}.
\]

Comparing with the asymptotic relation

\[
P(t, t + \Delta t) = e^{-r_t \Delta t}, \quad \Delta t \to 0,
\]

we can retrieve the spot rate from the yield via the formula \( r_t = R(t, t) \). Moreover,

\[
r_t = \lim_{\Delta t \to 0} R(t, t + \Delta t) = -\lim_{\Delta t \to 0} \frac{\ln P(t, t + \Delta t)}{\Delta t} = -\lim_{\Delta t \to 0} \frac{\ln P(t, t + \Delta t) - \ln P(t, t)}{\Delta t} = -\frac{\partial}{\partial T} \ln P(t, T) \bigg|_{T=t},
\]

where we used the condition \( P(t, t) = 1 \). This shows how can the spot rate be retrieved from the bond price. However, in general, the spot rate \( r_t \) is not enough to recover \( P(t, T) \). In order to be able to do this, we need forward rates.

### 4.3 Bootstrap Method

The construction of the yield curve starting from bond prices is called the bootstrap method. Consider \( n \) discount bonds with time to maturity \( \tau_j = T_j - t \) and face values \( V_j \), \( 1 \leq j \leq n \). Their prices, \( P_j \), can be directly observed from the bond market, as in the next table:

<table>
<thead>
<tr>
<th>Bond price</th>
<th>Face value</th>
<th>Maturity</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_1 )</td>
<td>( V_1 )</td>
<td>( T_1 )</td>
</tr>
<tr>
<td>( P_2 )</td>
<td>( V_2 )</td>
<td>( T_2 )</td>
</tr>
<tr>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
</tr>
<tr>
<td>( P_n )</td>
<td>( V_n )</td>
<td>( T_n )</td>
</tr>
</tbody>
</table>

The bond price formula \( P_j = V_j e^{-R(t,T_j)(T_j-t)} \) implies

\[
R(t, T_j) = \frac{1}{T_j - t} \ln \frac{V_j}{P_j}, \quad 1 \leq j \leq n.
\]

The yield curve \( R(t, T) \) can be constructed by interpolation from the previous values, starting with the initial value \( R(t, t) = r_t \), which is the spot rate at time \( t \), see Fig. 4.2.
4.4 Forward Rates

Let $0 \leq t < T_1 < T_2$, and denote by $f(t, T_1, T_2)$ the forward rate as seen at time $t$ for the time period between $T_1$ and $T_2$. This means that if one enters a contract at time $t$ that pays $1$ at time $T_2$, then its price at time $T_1$ is obtained discounting by the forward rate between $T_1$ and $T_2$ as

$$V = e^{-f(t, T_1, T_2)(T_2 - T_1)}.$$  

Now, consider a discount bond that pays off the amount $V$ at time $T_1$. Its price at time $t$ is

$$P(t, T_1)V = P(t, T_1)e^{-f(t, T_1, T_2)(T_2 - T_1)}. \tag{4.4.4}$$

In order to avoid arbitrage opportunities, this price has to equal the value of a bond at time $t$ that pays of $1$ at maturity $T_2$, so

$$P(t, T_2) = P(t, T_1)V. \tag{4.4.5}$$

Equating relations (4.4.4)-(4.4.5) and solving for the forward rate, implies

$$f(t, T_1, T_2) = -\frac{\ln P(t, T_2) - \ln P(t, T_1)}{T_2 - T_1}. \tag{4.4.6}$$

The process followed by $f(t, T_1, T_2)$ can be computed using Ito’s formulas similar to (4.2.3)

$$d \ln P(t, T_1) = (r_t - \frac{1}{2} \nu(t, T_1)^2) dt + \nu(t, T_1) dW_t$$

$$d \ln P(t, T_2) = (r_t - \frac{1}{2} \nu(t, T_2)^2) dt + \nu(t, T_2) dW_t$$

and then taking the differential in (4.4.6) we obtain

$$df(t, T_1, T_2) = \frac{\nu(t, T_2)^2 - \nu(t, T_1)^2}{2(T_2 - T_1)} dt - \frac{\nu(t, T_2) - \nu(t, T_1)}{T_2 - T_1} dW_t. \tag{4.4.7}$$
It is worthy to note that $f(t, T_1, T_2)$ depends only on the bonds volatilities.

The \textit{instantaneous forward rate of borrowing} is defined by

\begin{equation}
    f(t, T) = \lim_{\Delta t \to 0} f(t, T, T + \Delta t) = -\frac{\partial}{\partial T} \ln P(t, T),
\end{equation}

where we made use of (4.4.6). $f(t, T)$ represents the rate at time $T$ as seen from time $t$. For $t = T$ this becomes the spot rate at time $t$

\begin{equation}
    r_t = f(t, t).
\end{equation}

\textbf{Exercise 4.4.1} Let $0 < t < T_1 < T_2 < T_3$. Show that the forward rate $f(t, T_1, T_3)$ can be written as the weighted average of partial forwards rates as follows:

\begin{equation}
    f(t, T_1, T_3) = \frac{T_3 - T_2}{T_3 - T_1} f(t, T_2, T_3) + \frac{T_2 - T_1}{T_3 - T_1} f(t, T_1, T_2).
\end{equation}

\textit{State and prove a generalization.}

The forward rate can be retrieved from instantaneous forward rates as an integral average as in the following.

\textbf{Proposition 4.4.2} Let $t < a < b$. Then

\begin{equation}
    f(t, a, b) = \frac{1}{b - a} \int_a^b f(t, \tau) \, d\tau.
\end{equation}

\textit{Proof:} Consider the equidistant division $a = T_0 < T_1 < \cdots < T_{n-1} < T_n = b$, with width $\Delta \tau$. A generalization of Exercise 4.4.1 to $n$ intervals writes as

\begin{equation*}
    f(t, a, b) = \frac{1}{b - a} \sum_{j=0}^{n-1} f(t, T_j, T_{j+1})(T_{j+1} - T_j) = \frac{1}{b - a} \sum_{j=0}^{n-1} f(t, T_j, T_j + \Delta \tau) \Delta \tau.
\end{equation*}

Taking $\Delta \tau \to 0$ yields

\begin{equation*}
    f(t, a, b) = \lim_{n \to \infty} \frac{1}{b - a} \sum_{j=0}^{n-1} f(t, T_j, T_j + \Delta \tau) \Delta \tau
\end{equation*}

\begin{equation*}
    = \frac{1}{b - a} \int_a^b f(t, \tau, \tau) \, d\tau = \frac{1}{b - a} \int_a^b f(t, \tau) \, d\tau.
\end{equation*}
Exercise 4.4.3 (Forward contract) Enter a contract at time $t$ to pay $K$ at time $T_1$ in order to receive $1$ at a later time $T_2$. Show the following equivalent formulas for the payment:

(i) $K = e^{-f(t,T_{1})(T_{2}-T_{1})}$;
(ii) $K = e^{-\int_{t}^{T_1} f(t,s) ds}$;
(iii) $K = P(t,T_2)/T(t,T_1)$.

Exercise 4.4.4 (Coupon bearing bond) Let $0 < t < T_1 < \cdots < T_n < T$. Consider a contract that pays off the cash amounts $c_1, \cdots, c_n$ at times $T_1, \cdots, T_n$, and a final payment of $1$ at time $T$. Show that the price of the contract at time $t$ is

$$C(t) = P(t,T) + \sum_{i=1}^{n} c_i P(t,T_i).$$

4.5 Single-Factor HJM Models

The process followed by the instantaneous forward rate $f(t,T)$ can be obtained from (4.4.7) making $T_1 = T$, $T_2 = T + \Delta t$ and taking $\Delta t \to 0$. The resulting process is

$$df(t,T) = \nu(t,T) \frac{\partial}{\partial T} \nu(t,T) dt - \frac{\partial}{\partial T} \nu(t,T) dW_t.$$  \quad (4.5.10)

This corresponds to a one-factor model, which is fully described by the bond volatility $\nu(t,T)$. Following Heath, Jarrow and Morton [8] we write

$$df(t,T) = m(t,T) dt - s(t,T) dW_t.$$  \quad (4.5.10)

Assuming $\nu(t,t) = 0$, the following relation between the drift and volatility holds

$$m(t,T) = s(t,T) \int_{t}^{T} s(t,u) du.$$  \quad (4.5.10)

The forward rate $f(t,T)$ can be written in terms of the initial forward rate as

$$f(t,T) = f(0,T) + \int_{0}^{t} m(\tau,T) d\tau - \int_{0}^{t} s(\tau,T) dW_\tau.$$  \quad (4.5.11)

The process for the spot rate The aforementioned one-factor (HJM) model for the forward rate provides a model for the spot rate. Making $T = t$ in (4.5.11) and using (4.4.9) yields

$$r_t = f(0,t) + \int_{0}^{t} m(\tau,t) d\tau - \int_{0}^{t} s(\tau,t) dW_\tau,$$  \quad (4.5.12)
where

\[ m(\tau, t) = \nu(\tau, t) \frac{\partial}{\partial t} \nu(\tau, t) \quad (4.5.13) \]

\[ s(\tau, t) = \frac{\partial}{\partial t} \nu(\tau, t). \quad (4.5.14) \]

In the case when the volatility \( \nu(t, T) \) is independent of both the interest rate \( r_t \) and its history, then the spot rate is normally distributed with the mean and variance given by

\[ E[r_t] = f(0, t) + \int_0^t m(\tau, t) \, d\tau \]

\[ Var(r_t) = \int_0^t s(\tau, t)^2 \, d\tau. \]

In the following we shall encounter a few familiar cases of the one-factor (HJM) model.

### 4.5.1 Ho-Lee Model

Let \( \nu(t, T) = -\sigma(T - t) \), with \( \sigma \) constant. Then (4.5.13)-(4.5.14) provide

\[ m(\tau, t) = \sigma^2(t - \tau) \]

\[ s(\tau, t) = -\sigma. \]

Substituting in (4.5.12) we obtain the following formula for the spot rate

\[ r_t = f(0, t) + \frac{1}{2} \sigma^2 t^2 + \sigma W_t. \quad (4.5.15) \]

The associated stochastic differential equation is given by

\[ dr_t = (\partial_t f(0, t) + \sigma^2 t) \, dt + \sigma \, dW_t. \]

Substituting \( \theta(t) = \partial_t f(0, t) + \sigma^2 t \) we arrive at the Ho-Lee model

\[ dr_t = \theta_t \, dt + \sigma \, dW_t. \]

Hence, the Ho-Lee model is a one-factor (HJM) model. The initial forward curve \( f(0, t) \) is an input to the model, in the sense that the function \( \theta(t) \) can be read from the forward curve available at time \( t = 0 \).
4.5.2 Hull and White Model

Assume the volatility is \( \nu(t, T) = -\frac{\sigma}{a}(1 - e^{-a(T-t)}) \). Then we can infer from the equations (4.5.13)-(4.5.14) that

\[
\begin{align*}
m(\tau, t) &= \frac{\sigma^2}{a} \left( e^{-a(t-\tau)} - e^{-2at-\tau} \right) \\
s(\tau, t) &= -\sigma e^{-a(t-\tau)}.
\end{align*}
\]

An elementary computation provides

\[
\int_0^t m(\tau, t) d\tau = \frac{\sigma^2}{a^2} \left( \frac{1}{2} (1 + e^{-2at}) - e^{-at} \right).
\]

Substituting in (4.5.12), we obtain

\[
rt = f(0, t) + \frac{\sigma^2}{a^2} \left( \frac{1}{2} (1 + e^{-2at}) - e^{-at} \right) + \sigma e^{-at} \int_0^t e^{at} dW. \quad (4.5.16)
\]

Comparing with the solution of the Hull-White model (3.4.11), which is given by

\[
rt = r_0 e^{-at} + e^{-at} \int_0^t \theta(s) e^{as} ds + \sigma e^{-at} \int_0^t e^{as} dW, \quad (4.5.17)
\]

we obtain the following integral equation in \( \theta(t) \)

\[
r_0 + \int_0^t \theta(s) e^{as} ds = e^{at} \left[ f(0, t) + \frac{\sigma^2}{a^2} \left( \frac{1}{2} (1 + e^{-2at}) - e^{-at} \right) \right].
\]

Solving for \( \theta(t) \) by differentiation, we obtain

\[
\theta(t) = \partial_t f(0, t) + af(0, t) + \frac{\sigma^2}{2a} (1 - e^{-2at}).
\]

This shows that the initial forward curve \( f(0, t) \) is an input for the Hull and White model. We also note that taking \( a \to 0 \) in the Hull and White model recovers Ho-Lee model.

**Exercise 4.5.1** Find the forward curve \( f(t, T) \) in the case of

(a) Ho-Lee model;

(b) Hull and White model.
4.5.3 Vasicek Model

Considering \( \nu(t, T) = \frac{\sigma}{a} e^{-a(T-t)} \), then equations (4.5.13)-(4.5.14) imply

\[
\begin{align*}
    m(\tau, t) &= -\frac{\sigma^2}{a} e^{-2a(t-\tau)} \\
    s(\tau, t) &= -\sigma e^{-a(t-\tau)}.
\end{align*}
\]

Substituting in (4.5.12) and performing the integration, we obtain

\[
    r_t = f(0, t) - \frac{\sigma^2}{4a} (1 - e^{-2at}) + \sigma e^{-at} \int_0^t e^{a \tau} dW_\tau.
\]

Comparing with the Ornstein-Uhlenbeck process

\[
    r_t = b + (r_0 - b) e^{-at} + \sigma e^{-at} \int_0^t e^{a \tau} dW_\tau,
\]

which is a solution of the Vasicek’s model (3.3.7), we infer

\[
    f(0, t) = b + (r_0 - b) e^{-at} + \frac{\sigma^2}{4a} (1 - e^{-2at}).
\]

This shows that in the case of the Vasicek’s model the initial forward curve is an output of the model.

**Exercise 4.5.2** Express the parameters \( a, b, \) and \( \sigma \) in terms of the initial derivatives of the initial forward curve \( \partial_t f(0, t) \big|_{t=0}, \partial^2_t f(0, t) \big|_{t=0}, \) and \( \partial^3_t f(0, t) \big|_{t=0} \). Why is this method limited in practical applications?

**Exercise 4.5.3** Let \( f(0, t) = 0.5 - 0.2 e^{-0.15t} \). Find the spot rate \( r_t \) and evaluate \( E[r_t | F_0] \) in the following cases:

(i) Ho-Lee model;

(ii) Hull and White model.

4.6 Relation Formulas

In this section we shall deal with the one-to-one relations among bond prices \( P(t, T) \), instantaneous forward rates \( f(t, T) \) and yield curves \( R(t, T) \).

**Proposition 4.6.1** The price of a discount bond paying off \( \$1 \) at time \( T \) can be evaluated as

\[
    P(t, T) = e^{-\int_t^T f(t, s) \, ds}.
\]

(4.6.18)
Proof: Integrating between \( t \) and \( T \) in the definition formula for instantaneously forward rates (4.4.8)

\[
f(t, \tau) = -\frac{\partial}{\partial \tau} \ln P(t, \tau),
\]
we obtain

\[
\int_t^T f(t, \tau) \, d\tau = \ln P(t, T) - \ln P(t, t) = \ln P(t, T),
\]
where we used \( P(t, t) = 1 \). Solving for \( P(t, T) \) we arrive at formula (4.6.18).

Proposition 4.6.2 The relations between yield curves and instantaneous forward rates are given by

\[
R(t, T) = \frac{1}{T - t} \int_t^T f(t, s) \, ds \tag{4.6.19}
\]

\[
f(t, T) = R(t, T) + (T - t) \frac{\partial}{\partial T} R(t, T). \tag{4.6.20}
\]

Proof: Writing the bond price in terms of the yield

\[
P(t, T) = e^{-R(t, T)(T - t)}
\]
and comparing with equation (4.6.18) implies

\[
R(t, T)(T - t) = \int_t^T f(t, s) \, ds, \tag{4.6.21}
\]
which implies the former relation. Then differentiating with respect to \( T \) in (4.6.21) and using product rule and Fundamental Theorem of Calculus, we obtain the latter relation.

It is worth noting that if \( R(t, T) \) increases with maturity \( T \), then the latter formula implies \( f(t, T) > R(t, T) \). Taking \( t = T \), we obtain the relation with the spot rate

\[
R(t, t) = r_t = f(t, t).
\]

Exercise 4.6.3 Find the stochastic differential equation of \( R(t, T) \) with respect to \( t \).
Exercise 4.6.4 Consider the forward rate satisfying the model
\[ df(t, T) = \alpha dt + \sigma dW_t, \quad t \leq T, \]
with \( \alpha \) and \( \sigma \) constants. Given the initial forward curve
\[ f(0, T) = 0.4 - 0.15e^{-0.2T}. \]

(i) Find the yield curve \( R(t, T) \);
(ii) Find the bond price \( P(t, T) \).

Exercise 4.6.5 Assume the spot rate is given by \( r_t = r + \sigma W_t \), with \( r \) and \( \sigma \) constants.

(i) Find the instantaneous forward rate \( f(t, T) \);
(ii) Find the yield curves \( R(t, T) \).

In the following we shall price the discount bond in a few model cases.

4.7 A Simple Spot Rate Model

We shall start by considering a simplistic example, which despite of the fact that does not have any market applicability, will serve our later computational purposes. Assume the spot rate \( r_t \) satisfies the stochastic equation
\[ dr_t = \sigma dW_t, \]
with \( \sigma > 0 \), constant. This corresponds to a spot rate which diffuses starting from an initial rate \( r \) as
\[ r_t = r + \sigma W_t, \quad t \geq 0. \]

We shall show that the price of the associated zero-coupon bond that pays off $1 at time \( T \) is given by the formula
\[ P(t, T) = e^{-r(t)(T-t)+\frac{1}{2}\sigma^2(T-t)^3}. \]

Let \( 0 < t < T \) be fixed. We start by writing future spot rates in terms of the present spot rates. This can be achieved by solving the stochastic differential equation, and obtaining for any \( t < s < T \)
\[ r_s = r_t + \sigma(W_s - W_t). \]
Integrating over the remaining life of the bond yields
\[ \int_t^T r_s \, ds = r_t (T - s) + \sigma \int_t^T (W_s - W_t) \, ds. \]

Then taking the exponential, we obtain the price of the bond at time \( t \)
\[ P(t, T) = E\left[ e^{-\int_t^T r_s \, ds} \mid \mathcal{F}_t \right] \]
\[ = e^{-r_t(T-t)} E\left[ e^{-\sigma \int_t^T (W_s - W_t) \, ds} \mid \mathcal{F}_t \right] \]
\[ = e^{-r_t(T-t)} E\left[ e^{-\sigma \int_t^T (W_s - W_t) \, ds} \right] \]
\[ = e^{-r_t(T-t)} e^{\frac{\sigma^2 (T-t)^3}{6}}. \]

In the second identity we took the \( \mathcal{F}_t \)-measurable part out of the expectation, while in the third identity we dropped the condition since \( W_s - W_t \) is independent of the information set \( \mathcal{F}_t \) for any \( t < s \). The fourth identity invoked the stationarity of the Brownian motion. The last identity follows from the fact that \(-\sigma \int_0^{T-t} W_s \, ds \) is normally distributed with mean 0 and variance \( \frac{\sigma^2 (T-t)^3}{6} \), and from the expression of the moment generating function of a normal distribution.

It is worth noting that for this model the price of the bond, \( P(t, T) \), depends only the spot rate at time \( t \), \( r_t \), and the time to maturity \( T - t \). This is not a realistic model since it states that if today’s spot rate is known, then this would suffice for computing the bond price; this would not take into consideration all the future market movements between today and maturity, which would definitely have an effect on the bond price.

**Term Structure** The yield curves are given by
\[ R(t, T) = -\frac{1}{T - t} \ln P(t, T) = r_t - \frac{1}{6} \sigma^2 (T - t)^2 \]
\[ = r - \frac{1}{6} \sigma^2 (T - t)^2 + \sigma W_t. \]

Then the instantaneous forward rate becomes
\[ f(t, T) = R(t, T) + (T - t) \frac{\partial}{\partial T} R(t, T) \]
\[ = r_t - \frac{1}{2} \sigma^2 (T - t)^2 = r - \frac{1}{2} \sigma^2 (T - t)^2 + \sigma W_t. \] (4.7.22)

The forward rates are below the yield curves, and the term structure is downward sloping, see Fig. 4.3. One major disadvantage of this model is that the rates become negative for \( T - t \) large.
Exercise 4.7.1 Assume the spot rates exhibit positive jumps of size $\sigma > 0$ and satisfy the stochastic equation

$$dr_t = \sigma dN_t,$$

where $N_s$ is a Poisson process with rate $\lambda$.

(i) Find the price of the associated zero-coupon bond that pays off $\$1$ at time $T$;

(ii) Compute the yield curve $R(t, T)$;

(iii) determine the formula for the instantaneous forward rates.

4.8 Bond Price for Ho-Lee Model

In the case of Ho-Lee model the initial term structure is an input to the model, and the explicit formula for the spot rate is given by relation (4.5.15)

$$r_t = f(0, t) + \frac{1}{2}\sigma^2 t^2 + \sigma W_t.$$

We shall show that the price of the discount bond paying $\$1$ at time $T$ is given by

$$P(t, T) = e^{\left(\frac{1}{2}\sigma^2 - f(0, t, T)\right)(T-t) - \frac{1}{6}\sigma^2 (T^3 - t^3)}.$$ 

This can be computed as in the following

$$P(t, T) = E\left[e^{-\int_t^T r_s \, ds} | \mathcal{F}_t \right]$$

$$= e^{-\left(T-t\right)f(0, t, T) - \frac{1}{6}\sigma^2 (T^3 - t^3)} E\left[e^{-\sigma W_T - W_t} | \mathcal{F}_t \right]$$

$$= e^{\left(\frac{1}{2}\sigma^2 - f(0, t, T)\right)(T-t) - \frac{1}{6}\sigma^2 (T^3 - t^3)}. \quad (4.8.23)$$
Interest Term Structure The bond formula (4.8.23) combined with relation $P(t, T) = e^{-R(t,T)(T-t)}$ imply the following formula for the yield curve

$$R(t, T) = f(0, t, T) - \frac{1}{2} \sigma^2 + \frac{1}{6} \sigma^2 (T^2 + Tt + t^2).$$

From here, the forward rates can be obtained using formula (4.6.20)

$$f(t, T) = R(t, T) + (T - t) \frac{\partial}{\partial T} R(t, T) = f(0, t, T) - \frac{1}{2} \sigma^2 + \frac{1}{6} \sigma^2 (T^2 + Tt + t^2) + (T - t) \left( \frac{\partial}{\partial T} f(0, t, T) + \frac{1}{3} \sigma^2 T + \frac{1}{6} \sigma^2 t \right) = f(0, t, T) + (T - t) \frac{\partial}{\partial T} f(0, t, T) + \frac{1}{2} \sigma^2 (T^2 - 1).$$

Exercise 4.8.1 Use relation (4.8.24) to find $f(0, t, T)$ in terms of $f(t, T)$. Consider the initial condition $f(0, t, t) = f(0, t)$.

4.9 Bond Price for Vasicek’s Model

The next result regarding bond pricing is due to Vasicek [19]. Its initial proof is based on partial differential equations, while here we provide an approach solely based on expectations.

Proposition 4.9.1 Assume the spot rate $r_t$ satisfies the model

$$dr_t = a(b - r_t)dt + \sigma dW_t,$$  

(4.9.25)

Then the price of the zero-coupon bond that pays $1 at time $T$ is given by

$$P(t, T) = A(t, T)e^{-B(t,T)r_t},$$  

(4.9.26)

where

$$B(t, T) = \frac{1 - e^{-a(T-t)}}{a},$$

$$A(t, T) = e^{\frac{(B(t,T)-Tt)(a^2 b - \sigma^2 / 2)}{a^2 - \sigma^2 B(t,T)^2} - \frac{a^2 B(t,T)^2}{4a^2}}. $$

Proof: The computation follows the idea of section 4.7. Integrating equation (4.9.25) between $t$ and $T$, and then taking the exponential, we arrive at

$$e^{-\int_t^T r_s ds} = e^{\frac{1}{a}(r_T-r_t)-\frac{\sigma}{a}(W_T-W_t)-b(T-t)}.$$  

(4.9.27)
The bond price is computed by taking the conditional expectation as of time $t$ in relation (4.9.26). In order to do this we need to express the terms in the exponent of the right side more conveniently.

We start with the computation of the difference $(r_T - r_t)$. Multiplying equation (4.9.25) by $e^{as}$ leads to the exact equation

$$d(e^{as}r_s) = abe^{as}ds + \sigma e^{as}dW_s,$$

which after integration between $t$ and $T$ yields

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

Solving for $r_T$, then subtracting $r_t$ and dividing by $a$ we obtain

$$\frac{1}{a}(r_T - r_t) = -r_t B(t, T) + bB(t, T) + \frac{\sigma}{a}e^{-aT} \int_t^T e^{as} dW_s. \quad (4.9.28)$$

The second exponent in the right side of (4.9.27) can be written as an integral

$$\frac{\sigma}{a}(W_T - W_t) = \frac{\sigma}{a} \int_t^T dW_s. \quad (4.9.29)$$

Substituting (4.9.28) and (4.9.29) into (4.9.27) we obtain

$$e^{-\int_t^T r_s ds} = e^{-r_t B(t, T)} e^{bB(t, T) - b(T-t)} e^{\frac{\sigma}{a} \int_t^T (e^{as} - 1) dW_s}. \quad (4.9.30)$$

Taking the measurable part out, then dropping the independent condition in the conditional expectation, the price of the discount bond at time $t$ becomes

$$P(t, T) = E[e^{-\int_t^T r_s ds} | F_t] = e^{-r_t B(t, T)} e^{bB(t, T) - b(T-t)} E\left[e^{\frac{\sigma}{a} \int_t^T (e^{as} - 1) dW_s}\right] = e^{-r_t B(t, T)} A(t, T),$$

where we denoted

$$A(t, T) = E\left[e^{\frac{\sigma}{a} \int_t^T (e^{as} - 1) dW_s}\right]. \quad (4.9.31)$$

As a Wiener integral, $\int_t^T (e^{as} - 1) dW_s$ is normally distributed with zero mean and variance given by

$$\int_t^T (e^{as} - 1)^2 ds = \frac{a}{2} B(t, T)^2 - B(t, T) + (T-t).$$
Then the log-normal variable \( e^{\frac{\sigma}{a} \int_t^T (e^{as} - a - 1) dW_s} \) has the mean

\[
E[e^{\frac{\sigma}{a} \int_t^T (e^{as} - a - 1) dW_s}] = e^{\frac{1}{2} \frac{\sigma^2}{a^2} \left[ -\frac{a}{2} B(t, T)^2 - B(t, T) + (T-t) \right]}
\]

Substituting in (4.9.31) and performing the following algebra

\[
bB(t, T) - b(T-t) + \frac{1}{2} \frac{\sigma^2}{a^2} \left[ -\frac{a}{2} B(t, T)^2 - B(t, T) + (T-t) \right]
= -\frac{\sigma^2}{4a} B(t, T)^2 + (T-t) \left( \frac{\sigma^2}{2a^2} - b \right) + \left( b - \frac{\sigma^2}{2a^2} \right) B(t, T)
= -\frac{\sigma^2}{4a} B(t, T)^2 + \frac{1}{a^2} (a^2 b - \sigma^2/2) (B(t, T) - T + t),
\]

we obtain

\[
A(t, T) = e^{\left( \frac{B(t, T) - T + t}{a^2} (a^2 b - \sigma^2/2) - \frac{\sigma^2 B(t, T)^2}{4a} \right)}.
\]

It is worth noting that \( P(t, T) \) depends on the time to maturity, \( T-t \), and the spot rate \( r_t \).

**Exercise 4.9.2** Find the price of an infinitely lived bond in the case when spot rates satisfy Vasicek’s model.

**The term structure** Let \( R(t, T) \) be the continuously compounded interest rate at time \( t \) for a term of \( T-t \). Using the formula for the bond price \( P(t, T) = e^{-R(t,T)(T-t)} \) we get

\[
R(t, T) = -\frac{1}{T-t} \ln P(t, T).
\]

Using formula for the bond price (4.9.26) yields

\[
R(t, T) = -\frac{1}{T-t} \ln A(t, T) + \frac{1}{T-t} P(t, T) r_t.
\]

A few possible shapes of the term structure \( R(t, T) \) in the case of Vasicek’s model are given in Fig. 4.4.

### 4.10 Bond Price for CIR’s Model

In the case when \( r_t \) satisfy the Cox, Ingersoll, and Ross (CIR) model (3.3.8), the zero-coupon bond price has a similar form as in the case of Vasicek’s model

\[
P(t, T) = A(t, T) e^{-B(t,T)r_t}.
\]
The functions $A(t,T)$ and $B(t,T)$ are given in this case by

\[
B(t,T) = \frac{2(e^{\gamma(T-t)} - 1)}{(\gamma + a)(e^{\gamma(T-t)} - 1) + 2\gamma}
\]

\[
A(t,T) = \left( \frac{2\gamma e^{(a+\gamma)(T-t)/2}}{(\gamma + a)(e^{\gamma(T-t)} - 1) + 2\gamma} \right)^{2ab/\sigma^2},
\]

where $\gamma = (a^2 + 2\sigma^2)^{1/2}$. For details the reader can consult Cox, Ingersoll, and Ross [5].

**Exercise 4.10.1** Consider the spot rate $r_t$ satisfying the Black, Derman and Toy model (3.5.13).

(i) Find the price of the associated discount bond;

(ii) Match the function parameter $\theta(t)$ to the forward curve $f(0,t)$.

### 4.11 Mean Reverting Model with Jumps

We shall study a model for the short-term interest rate $r_t$, which is mean reverting and also incorporates jumps. Let $N_t$ be the Poisson process of constant
Consider the following model for the spot rate

\[ dr_t = a(b - r_t)dt + \sigma dM_t, \quad (4.11.32) \]

with \( a, b, \sigma \) positive constants. It is worth noting the similarity with the Vasicek’s model, which is obtained by replacing the process \( dM_t \) by \( dW_t \), where \( W_t \) is a one-dimensional Brownian motion.

Making abstraction of the uncertainty source \( \sigma dM_t \), this implies that the rate \( r_t \) is pulled towards level \( b \) at the rate \( a \). This means that, if for instance \( r_0 > b \), then \( r_t \) is decreasing towards \( b \). The term \( \sigma dM_t \) adds jumps of size \( \sigma \) to the process. The fact that these jumps have equal size is a limitation of the model, but will certainly help in finding a trackable formula. A few realizations of the process \( r_t \) are given in Fig. 4.5.

The stochastic differential equation (4.11.32) can be solved explicitly using the method of integrating factor, see Øksendal [16]. Multiplying by \( e^{at} \) we get an exact equation; integrating between 0 and \( t \) yields the following closed form formula for the spot rate

\[ r_t = b + (r_0 - b)e^{-at} + \sigma e^{-at} \int_0^t e^{as} dM_s. \quad (4.11.33) \]

Since the integral with respect to \( dM_t \) is a martingale, the expectation of the
The spot rate is
\[ E[r_t] = b + (r_0 - b)e^{-at}. \]
In the long run the mean \( E[r_t] \) tends to \( b \), which shows the mean reversion. Using the formula
\[ E \left[ \left( \int_0^t f(s) \, dM_s \right)^2 \right] = \lambda \int_0^t f(s)^2 \, ds \]
the variance of the spot rate can be computed as follows
\[
\Var(r_t) = \sigma^2 e^{-2at} \Var \left( \int_0^t e^{as} \, dM_s \right) = \sigma^2 e^{-2at} E \left[ \left( \int_0^t e^{as} \, dM_s \right)^2 \right] \\
= \frac{\lambda \sigma^2}{2a} (1 - e^{-2at}).
\]

It is worth noting that in the long run the variance tends to the constant \( \frac{\lambda \sigma^2}{2a} \), which is proportional with the frequency of jumps \( \lambda \).

In the following we shall find the value of a zero-coupon bond. Let \( \mathcal{F}_t \) be the \( \sigma \)-algebra generated by the random variables \( \{N_s; s \leq t\} \). This contains all information about jumps occurred until time \( t \). This information is used in the computation of the bond value as
\[
P(t,T) = E \left[ e^{-\int_t^T r_s \, ds \big| \mathcal{F}_t} \right].
\]

**Proposition 4.11.1** Assume the spot rate \( r_t \) satisfies the mean reverting model with jumps
\[ dr_t = a(b - r_t) \, dt + \sigma dM_t. \]
Then the price of the zero-coupon bond that pays off \$1 at time \( T \) is given by
\[ P(t,T) = A(t,T) e^{-B(t,T) r_t}, \]
where
\[
B(t,T) = \frac{1 - e^{-a(T-t)}}{a}, \\
A(t,T) = \exp \left( (b + \frac{\lambda \sigma}{a}) B(t,T) \right) \\
+ [\lambda (1 - \frac{\sigma}{a}) - b](T - t) - \lambda e^{-\sigma a} \int_0^{T-t} e^{\sigma x} e^{-ax} \, dx.
\]

**Proof:** The first part of this computation mimics closely the one done in the case of Vasicek’s model. Integrating the stochastic differential equation between \( t \) and \( T \) yields
\[ -a \int_t^T r_s \, ds = r_T - r_t - \sigma (M_T - M_t) - ab(T - t). \]
Taking the exponential we obtain
\[ e^{-\int_t^T r_s \, ds} = e^{\frac{1}{a} (r_T - r_t) - \frac{\sigma}{a} (M_T - M_t) - b(T - t)}. \] (4.11.34)

Multiplying in the stochastic differential equation by \( e^{as} \) yields the exact equation
\[ d(e^{as} r_s) = ab e^{as} \, ds + \sigma e^{as} \, dM_s. \]
Integrate between \( t \) and \( T \) to get
\[ e^{aT} r_T - e^{at} r_t = b(e^{aT} - e^{at}) + \sigma \int_t^T e^{as} \, dM_s. \]
Solving for \( r_T \) and subtracting \( r_t \) yields
\[ r_T - r_t = -r_t(1 - e^{a(T-t)}) + b(1 - e^{a(T-t)}) + \sigma e^{-aT} \int_t^T e^{as} \, dM_s. \]
Dividing by \( a \) and using the notation for \( B(t,T) \) yields
\[ \frac{1}{a} (r_T - r_t) = -r_t B(t,T) + b B(t,T) + \frac{\sigma}{a} e^{-aT} \int_t^T e^{as} \, dM_s. \]
Substituting into (4.11.34) and using that \( M_T - M_t = \int_t^T dM_s \), we get
\[ e^{-\int_t^T r_s \, ds} = e^{-r_t B(t,T) e^{bB(t,T)} - b(T-t) e^{\frac{\sigma}{a} \int_t^T (e^{as-aT} - 1) \, dM_s}}. \] (4.11.35)
Taking the predictable part out and dropping the independent condition, the price of the zero-coupon bond at time \( t \) becomes
\[
P(t,T) = \mathbb{E}[e^{-\int_t^T r_s \, ds} | \mathcal{F}_t] = e^{-r_t B(t,T) e^{bB(t,T)} - b(T-t) e^{\frac{\sigma}{a} \int_t^T (e^{as-aT} - 1) \, dM_s} | \mathcal{F}_t]
\]
\[ = e^{-r_t B(t,T) A(t,T)}, \]
where
\[ A(t,T) = e^{bB(t,T)} - b(T-t) e^{\frac{\sigma}{a} \int_t^T (e^{as-aT} - 1) \, dM_s} | \mathcal{F}_t]. \] (4.11.36)
We shall compute in the following the right side expectation. From Bertoin [2], p. 8, the exponential process
\[ e^{\int_0^t u(s) \, dN_s + \lambda \int_0^t (1 - e^{u(s)}) \, ds} \]
is an \( \mathcal{F}_t \)-martingale, with \( \mathcal{F}_t = \sigma(N_u; u \leq t) \). This implies
\[ \mathbb{E}[e^{\int_t^T u(s) \, dN_s} | \mathcal{F}_t] = e^{-\lambda \int_t^T (1 - e^{u(s)}) \, ds}. \]
Using \( dM_t = dN_t - \lambda dt \) yields

\[
E\left[e^{\int_t^T u(s) \, dM_s} \Big| \mathcal{F}_t\right] = E\left[e^{\int_t^T u(s) \, dN_s} \Big| \mathcal{F}_t\right] e^{-\lambda \int_t^T u(s) \, ds} = e^{-\lambda \int_t^T (1+u(s)e^{-u(s)}) \, ds}.
\]

Let \( u(s) = \frac{e}{a}(e^{a(s-T)} - 1) \) and substitute in (4.11.36); then after changing the variable of integration we obtain

\[
A(t, T) = e^{bB(t,T)-b(T-t)} E\left[e^{\frac{a}{\lambda} \int_t^T (e^{as} - aT - 1) \, dM_s} \Big| \mathcal{F}_t\right]
\]

\[
= e^{bB(t,T)-b(T-t)} e^{\lambda \int_0^{T-t} [1 + \frac{a}{\lambda} (e^{-ax} - 1) - e^{\frac{a}{\lambda} (e^{-ax} - 1)}] \, dx}
\]

\[
= e^{bB(t,T)-b(T-t)} e^{\lambda(T-t) + \frac{a}{\lambda} \left( B(t,T) - (T-t) \right) - \lambda \int_0^{T-t} e^{\frac{a}{\lambda} (e^{-ax} - 1)} \, dx}
\]

\[
= e^{(b+\frac{a}{\lambda})B(t,T)} e^{\lambda(1-\frac{a}{\lambda}) - b(T-t)} e^{\lambda \int_0^{T-t} e^{\frac{a}{\lambda} e^{-ax}} \, dx}
\]

\[
= \exp\{ (b + \frac{\lambda \sigma}{a}) B(t,T) \}
\]

\[
+ \left[ \lambda \left( 1 - \frac{\sigma}{\lambda} \right) - b \right] (T-t) - \lambda e^{-\sigma a} \int_0^{T-t} e^{\frac{a}{\lambda} e^{-ax}} dx \}.
\]  

(4.11.37)

\[\blacksquare\]

**Evaluation using Special Functions** The integral in the last term of formula (4.11.37) cannot be computed explicitly; for its evaluation we need to use some special functions.

The solution of the initial value problem

\[
f'(x) = \frac{e^x}{x}, \quad x > 0
\]

\[
\lim_{x \to 0} f(x) = -\infty
\]

is the **exponential integral function** \( f(x) = Ei(x), \ x > 0 \). This is a special function that can be evaluated numerically in MATHEMATICA by calling the function ExpIntegralEi[x]. For instance, for any \( 0 < \alpha < \beta \)

\[
\int_\alpha^\beta \frac{e^t}{t} \, dt = Ei(\beta) - Ei(\alpha).
\]

The reader can find more details regarding the exponential integral function in Abramovitz and Stegun [13]. The last integral in the expression of \( A(t, T) \) can be evaluated using this special function. Substituting \( t = \frac{a}{\lambda} e^{-ax} \) we have

\[
\int_0^{T-t} e^{\frac{a}{\lambda} e^{-ax}} \, dx = \frac{1}{a} \int e^{\frac{a}{\lambda} e^{-ax(T-t)}} \frac{e^t}{t} \, dt
\]

\[
= \frac{1}{a} \left[ Ei\left( \frac{\sigma}{a} \right) - Ei\left( \frac{\sigma}{a} e^{-(T-t)} \right) \right].
\]
4.12 A Model with pure Jumps

Consider the spot rate \( r_t \) satisfying the stochastic differential equation
\[
\frac{dr_t}{t} = \sigma dM_t,
\]  
(4.12.38)
where \( \sigma \) is a positive constant denoting the volatility of the rate. This model is obtained when the rate at which \( r_t \) is pulled toward \( b \) is \( a = 0 \), so there is no mean reverting effect. This type of behavior can be noticed during a short time in a highly volatile market; in this case the behavior of \( r_t \) is mostly influenced by jumps.

The solution of (4.12.38) is
\[
r_t = r_0 + \sigma M_t = r_0 - \lambda \sigma t + \sigma N_t,
\]
which is an \( \mathcal{F}_t \)-martingale. The rate \( r_t \) has jumps of size \( \sigma \) that occur at the arrival times \( t_k \), which are exponentially distributed.

Next we shall compute the value \( P(t, T) \) at time \( t \) of a zero-coupon bond that pays the amount of $1 at maturity \( T \). This is given by the conditional expectation
\[
P(t, T) = E[e^{-\int_t^T r_s \, ds} | \mathcal{F}_t].
\]
Integrating between \( t \) and \( s \) in equation (4.12.38) yields
\[
r_s = r_t + \sigma (M_s - M_t), \quad t < s < T.
\]
And then
\[
\int_t^T r_s \, ds = r_t (T - t) + \sigma \int_t^T (M_s - M_t) \, ds.
\]
Taking out the predictable part and dropping the independent condition yields
\[
P(t, T) = E[e^{-\int_t^T r_s \, ds} | \mathcal{F}_t] = e^{-r_t(T-t)}E[e^{-\sigma \int_t^T (M_s - M_t) \, ds} | \mathcal{F}_t] = e^{-r_t(T-t)}E[e^{-\sigma \int_0^{T-t} M_r \, dr}] = e^{-r_t(T-t)}E[e^{-\sigma \int_0^{T-t} (N_r - \lambda r) \, dr}] = e^{-r_t(T-t) + \frac{\sigma}{2} \sigma (T-t)} E[e^{-\sigma \int_0^{T-t} N_r \, dr}].
\]  
(4.12.39)

We need to work out the expectation. Using integration by parts,
\[
-\sigma \int_0^T N_t \, dt = -\sigma \left( TN_T - \int_0^T t \, dN_t \right) = -\sigma \left( T \int_0^T dN_t - \int_0^T t \, dN_t \right) = -\sigma \int_0^T (T-t) \, dN_t,
\]
so
\[ e^{-\sigma \int_0^T N_t \, dt} = e^{-\sigma \int_0^T (T-t) \, dN_t}. \]

Using the formula
\[ E\left[ e^{\int_0^T u(t) \, dN_t} \right] = e^{-\lambda \int_0^T (1-e^{u(t)}) \, dt} \]
we have
\[ E\left[ e^{-\sigma \int_0^T N_t \, dt} \right] = E\left[ e^{-\sigma \int_0^T (T-t) \, dN_t} \right] = e^{-\lambda \int_0^T (1-e^{-\sigma (T-t)}) \, dt} \]
\[ = e^{-\lambda T} e^{-\frac{\lambda}{\sigma}(e^{-\sigma T}-1)} \]
\[ = e^{-\lambda \left( T + \frac{1}{\sigma} (e^{-\sigma T}-1) \right)} \]

Replacing \( T \) by \( T - t \) and \( t \) by \( \tau \) yields
\[ E\left[ e^{-\sigma \int_0^{T-t} N_{\tau} \, d\tau} \right] = e^{-\lambda \left( T - t + \frac{1}{\sigma} (e^{-\sigma (T-t)}-1) \right)}. \]

Substituting in (4.12.39) yields the formula for the bond price
\[ P(t, T) = e^{-r_t(T-t) + \frac{1}{2}\sigma \lambda(T-t)^2} e^{-\lambda \left( T - t + \frac{1}{\sigma} (e^{-\sigma (T-t)}-1) \right)} \]
\[ = \exp\left\{ - (\lambda + r_t)(T-t) + \frac{1}{2}\sigma \lambda(T-t)^2 - \frac{\lambda}{\sigma} (e^{-\sigma (T-t)} - 1) \right\}. \]

**Proposition 4.12.1** Assume the spot rate \( r_t \) satisfies \( dr_t = \sigma dM_t \), with \( \sigma > 0 \) constant. Then the price of the zero-coupon bond that pays off \$1 at time \( T \) is given by
\[ P(t, T) = \exp\left\{ - (\lambda + r_t)(T-t) + \frac{1}{2}\sigma \lambda(T-t)^2 - \frac{\lambda}{\sigma} (e^{-\sigma (T-t)} - 1) \right\}. \]

The yield curve is given by
\[ R(t, T) = -\frac{1}{T-t} \ln P(t, T) \]
\[ = r_t + \lambda - \frac{\sigma \lambda}{2} (T-t) + \frac{\lambda}{\sigma} (e^{-\sigma (T-t)} - 1). \]

**Exercise 4.12.2** Find the instantaneous forward rate \( F(t, T) \).

**Exercise 4.12.3** The price of an interest rate derivative with maturity time \( T \) and payoff \( r_T \) has the price at time \( t = 0 \) given by \( P_0 = E_0[e^{-\int_0^T r_s \, ds} r_T] \). Find the price of this derivative.

**Exercise 4.12.4** Assume the spot rate satisfies the equation \( dr_t = \sigma r_t dM_t \), where \( M_t \) is the compensated Poisson process. Find a solution of the form \( r_t = r_0 e^{\phi(t)} (1 + \sigma)^{N_t} \), where \( N_t \) denotes the Poisson process.
Chapter 5

Modeling Stock Prices

The price of a stock can be modeled by a continuous stochastic process which is the sum of a predictable and an unpredictable part. However, this type of model does not take into account market crashes. If those are to be taken into consideration, the stock price needs to contain a third component which models unexpected jumps. We shall discuss these models in the present chapter.

5.1 Constant Drift and Volatility Model

Let \( S_t \) denote the price of a stock at time \( t \). If \( \mathcal{F}_t \) denotes the market information at time \( t \), then \( S_t \) is a continuous process that is \( \mathcal{F}_t \)-adapted. The return on the stock during the time interval \( \Delta t \) measures the percentage increase in the stock price between instances \( t \) and \( t + \Delta t \) and is given by \( \frac{S_{t+\Delta t} - S_t}{S_t} \).

When \( \Delta t \) is infinitesimally small, we obtain the instantaneous return

\[
\frac{dS_t}{S_t} = \lim_{\Delta t \to 0} \frac{S_{t+\Delta t} - S_t}{S_t}.
\]

This is supposed to be the sum of two components:

- the deterministic part, \( \mu dt \), due to the drift;
- the noisy part, \( \sigma dW_t \), due to unexpected market news and fluctuations.

Adding these parts yields

\[
\frac{dS_t}{S_t} = \mu dt + \sigma dW_t,
\]

which leads to the stochastic differential equation

\[
\frac{dS_t}{S_t} = \mu dt + \sigma S_t dW_t, \tag{5.1.1}
\]

63
The parameters $\mu$ and $\sigma$ are positive constants, which represent the drift rate and volatility of the stock. Since equation (5.1.1) is linear in $S_t$, standard results of existence and uniqueness of solutions yield a unique solution, provided an initial condition $S_0$ is given. The solution can be found using the method of variation of parameters. Considering a solution of the form $S_t = e^{\alpha(t)} e^{\sigma W_t}$, an application of Ito’s formula leads to

$$dS_t = \alpha'(t) e^{\alpha(t)} e^{\sigma W_t} dt + e^{\alpha(t)} \left( \sigma e^{\sigma W_t} dW_t + \frac{1}{2} \sigma^2 e^{\sigma W_t} dt \right)$$

Comparing with (5.1.1) and equating the coefficients of $dt$ implies

$$\alpha'(t) + \frac{1}{2} \sigma^2 = \mu,$$

and hence $\alpha(t) = (\mu - \frac{1}{2} \sigma^2) t + \alpha_0$. Concluding, the solution of (5.1.1) becomes

$$S_t = S_0 e^{(\mu - \frac{1}{2} \sigma^2) t + \sigma W_t},$$

(5.1.2)

where $S_0$ denotes the price of the stock at time $t = 0$. It is worth noting that the stock price is $\mathcal{F}_t$-adapted, positive, and it has a log-normal distribution, with the mean and variance given by

$$E[S_t] = S_0 e^{\mu t},$$

(5.1.3)

$$Var[S_t] = S_0^2 e^{2 \mu t} (e^{\sigma^2 t} - 1).$$

(5.1.4)

Two simulations of the stock price $S_t$ can be seen in Fig.5.1.
Exercise 5.1.1 Let $F_u$ be the information set at time $u$. Find $E[S_t|F_u]$ and $\text{Var}[S_t|F_u]$ for $u \leq t$. How do these formulas become in the case $s = t$?

Exercise 5.1.2 Find the stochastic process followed by $\ln S_t$. What are the values of $E[\ln(S_t)]$ and $\text{Var}[\ln(S_t)]$?

Exercise 5.1.3 Find the stochastic differential equations associated with the following processes:

(a) $\frac{1}{S_t}$
(b) $S^n_t$
(c) $(S_t - 1)^2$.

Exercise 5.1.4 (a) Show that $E[S^2_t] = S^2_0 e^{(2\mu + \sigma^2)t}$.
(b) Find a similar formula for $E[S^n_t]$, with $n$ positive integer.

Exercise 5.1.5 (a) Find the expectation $E[S_tW_t]$.
(b) Find the correlation function $\rho_t = \text{Corr}(S_t, W_t)$. What happens for $t$ large?

5.2 Correlation of two stocks

Consider two stock prices driven by the same novelty term

\[
\begin{align*}
\text{d}S_1 &= \mu_1 S_1 \text{d}t + \sigma_1 S_1 \text{d}W_t \\
\text{d}S_2 &= \mu_2 S_2 \text{d}t + \sigma_2 S_2 \text{d}W_t.
\end{align*}
\]

(5.2.5)

(5.2.6)

Since the underlying Brownian motions are perfectly correlated, one may be tempted to think that the stock prices $S_1$ and $S_2$ behave in a similar way. The following result shows that in general the stock prices are positively correlated, and the stock volatilities play the determinant role in this business:

**Proposition 5.2.1** The correlation coefficient between the stock prices $S_1$ and $S_2$ driven by the same Brownian motion is

\[
\text{Corr}(S_1, S_2) = \frac{e^{\sigma_1 \sigma_2 t} - 1}{(e^{\sigma_1^2 t} - 1)^{1/2}(e^{\sigma_2^2 t} - 1)^{1/2}} > 0.
\]

In particular, if $\sigma_1 = \sigma_2$, then $\text{Corr}(S_1, S_2) = 1$.

**Proof:** Since

\[
S_1(t) = S_1(0)e^{\mu_1 t - \frac{1}{2}\sigma_1^2 t}e^{\sigma_1 W_t}, \quad S_2(t) = S_2(0)e^{\mu_2 t - \frac{1}{2}\sigma_2^2 t}e^{\sigma_2 W_t},
\]

the conditional variance is defined by $\text{Var}(X|\mathcal{F}) = E[X^2|\mathcal{F}] - E[X|\mathcal{F}]^2$. 

---

\[
1\text{The conditional variance is defined by } \text{Var}(X|\mathcal{F}) = E[X^2|\mathcal{F}] - E[X|\mathcal{F}]^2.
\]
from Exercise 5.2.4 and formula $E[e^{kW_t}] = e^{kt/2}$ we have

$$
\text{Corr}(S_1, S_2) = \text{Corr}(e^{\sigma_1 W_t}, e^{\sigma_2 W_t}) = \frac{\text{Cov}(e^{\sigma_1 W_t}, e^{\sigma_2 W_t})}{\sqrt{\text{Var}(e^{\sigma_1 W_t})\text{Var}(e^{\sigma_2 W_t})}} \\
= \frac{E[e^{(\sigma_1+\sigma_2)W_t}] - E[e^{\sigma_1 W_t}]E[e^{\sigma_2 W_t}]}{\sqrt{\text{Var}(e^{\sigma_1 W_t})\text{Var}(e^{\sigma_2 W_t})}} \\
= \frac{e^{\frac{1}{2}(\sigma_1+\sigma_2)^2t} - e^{\frac{1}{2}\sigma_1^2t}e^{\frac{1}{2}\sigma_2^2t}}{\sqrt{(e^{\sigma_1^2t} - 1)(e^{\sigma_2^2t} - 1)}} \\
= \frac{e^{\sigma_1\sigma_2t} - 1}{(e^{\sigma_1^2t} - 1)^{1/2}(e^{\sigma_2^2t} - 1)^{1/2}}.
$$

If $\sigma_1 = \sigma_2 = \sigma$ then the previous formula provides

$$
\text{Corr}(S_1, S_2) = \frac{e^{\sigma^2t} - 1}{e^{\sigma^2t} - 1} = 1,
$$

i.e. the stocks are perfectly correlated if they have the same volatility.  

**Corollary 5.2.2** The stock prices $S_1$ and $S_2$ are positively strongly correlated for small values of $t$:

$$
\text{Corr}(S_1, S_2) \to 1 \quad \text{as} \quad t \to 0.
$$

This fact has the following financial interpretation. If some stocks are driven by the same market fluctuations, when one stock increases, then the other one tends to increase too, at least for a small amount of time. In the case when some bad news affects an entire financial market, the risk becomes systemic, and hence if one stock fails, all the others tend to decrease as well, leading to a severe strain on the financial market.

**Corollary 5.2.3** The stock prices correlation gets weak as $t$ gets large:

$$
\text{Corr}(S_1, S_2) \to 0 \quad \text{as} \quad t \to \infty.
$$

**Proof:** It follows from the asymptotic correspondence

$$
\frac{e^{\sigma_1\sigma_2t} - 1}{(e^{\sigma_1^2t} - 1)^{1/2}(e^{\sigma_2^2t} - 1)^{1/2}} \sim \frac{e^{\sigma_1\sigma_2t}}{e^{\frac{1}{2}(\sigma_1^2+\sigma_2^2)t}} = e^{-\frac{(\sigma_1-\sigma_2)^2}{2}t} \to 0, \quad t \to 0.
$$

It follows that in the long run any two stocks tend to become uncorrelated, see Fig.5.2.
Figure 5.2: The correlation function \( f(t) = \frac{e^{\sigma_1 t} - 1}{(e^{\sigma_1^2 t} - 1)^{1/2}(e^{\sigma_2^2 t} - 1)^{1/2}} \) in the case \( \sigma_1 = 0.15, \sigma_2 = 0.40 \).

Exercise 5.2.4 If \( X \) and \( Y \) are random variables and \( \alpha, \beta \in \mathbb{R} \), show that
\[
\text{Corr}(\alpha X, \beta Y) = \begin{cases} 
\text{Corr}(X, Y), & \text{if } \alpha \beta > 0 \\
-\text{Corr}(X, Y), & \text{if } \alpha \beta < 0.
\end{cases}
\]

Exercise 5.2.5 Find the following
(a) \( \text{Cov}(dS_1(t), dS_2(t)) \);
(b) \( \text{Corr}(dS_1(t), dS_2(t)) \).

5.3 When Does a Stock Hit a Given Barrier?

Let \( a > S_0 \) and consider the first time when the stock reaches the barrier \( a \)
\[
T_a = \inf\{t > 0; S_t \geq a\}.
\]
The random variable \( T_a \) is a stopping time, in the sense that the event \( \{T_a \leq t\} \) belongs to \( \mathcal{F}_t \), i.e. having given the market information at time \( t \), we can decide whether the stock has reached or not the level \( a \) yet.

In the following we shall compute the moment generating function for \( T_a \), find its expectation and derive the formula for its probability law.

The moment generating function can be obtained by a manipulation involving martingales and the Optional Stopping Theorem, as in the following.

Let \( c > 0 \) be a constant. Since \( M_t = e^{-\frac{1}{2}c^2t + cW_t} \) is an \( \mathcal{F}_t \)-martingale and \( T_a \) is a stopping time, the Optional Stopping Theorem provides
\[
E[M_{T_a}] = E[M_0] = 1. \tag{5.3.7}
\]
Denote \( m = \mu - \sigma^2/2 \) and assume that \( m > 0 \) for the rest of this section. Solving for \( W_t \) from the stock price expression
\[
S_t = S_0 e^{mt + \sigma W_t}.
\]
we get
\[ W_t = \frac{1}{\sigma} \left( \ln \frac{S_t}{S_0} - mt \right). \]

Substituting it into the expression of the martingale \( M_t \), we obtain
\[ M_t = \left( \frac{S_t}{S_0} \right)^{\frac{\gamma}{\sigma}} e^{-\left( \frac{1}{2} c^2 + \frac{mc}{\sigma} \right)t}. \]

Using that \( S_{Ta} = a \), then taking the expectations yields
\[ E[M_{Ta}] = \left( \frac{a}{S_0} \right)^{\frac{\gamma}{\sigma}} E\left[ e^{-\left( \frac{1}{2} c^2 + \frac{mc}{\sigma} \right) Ta} \right]. \]

Then equation (5.3.7) can be written as
\[ E\left[ e^{-\left( \frac{1}{2} c^2 + \frac{mc}{\sigma} \right) Ta} \right] = \left( \frac{S_0}{a} \right)^{\frac{\gamma}{\sigma}}. \]

Using the substitution
\[ c = -\frac{m}{\sigma} + \sqrt{2s + \frac{m^2}{\sigma^2}}, \quad s > 0 \]
we obtain
\[ E\left[ e^{-sTa} \right] = \left( \frac{S_0}{a} \right)^{-\frac{m}{\sigma} + \frac{1}{2} \sqrt{2s + \frac{m^2}{\sigma^2}}} = \left( \frac{S_0}{a} \right)^{h(s)}, \quad (5.3.8) \]
with
\[ h(s) = \left( \frac{1}{2} - \frac{\mu}{\sigma^2} \right) + \frac{1}{\sigma} \sqrt{2s + \left( \frac{\mu}{\sigma} - \frac{\sigma^2}{2} \right)^2}. \]

**Exercise 5.3.1** Verify that the function \( h(s) \) has the following properties:
(a) \( h(0) = 0 \);
(b) \( h(\mu) = 1 \);
(c) \( h'(0) = (\mu - \sigma^2/2)^{-1} \);
(d) \( h''(0) = -\sigma^2(\mu - \sigma^2/2)^{-3} \).

**Exercise 5.3.2** Show that
\[ E[e^{-\mu Ta}] = \frac{S_0}{a}. \]

**Exercise 5.3.3** Prove that
(a) \( \left. -\frac{d}{ds} E[e^{-sTa}] \right|_{s=0} = E[T_a] \);
(b) \( \left. \frac{d^2}{ds^2} E[e^{-sTa}] \right|_{s=0} = E[T_a^2] \).
Proposition 5.3.4 The mean of the hitting time $T_a$ is given by

$$E[T_a] = \ln \left( \frac{a}{S_0} \right) \frac{1}{\mu - \sigma^2/2}.$$ 

Proof: Differentiating in (5.3.8) we have

$$\frac{d}{ds} E[e^{-sT_a}] = \left( \frac{S_0}{a} \right)^{h(s)} \ln \frac{S_0}{a} h'(s)$$

and then taking the value at $s = 0$, we get

$$\frac{d}{ds} E[e^{-sT_a}] \bigg|_{s=0} = \left( \frac{S_0}{a} \right)^{h(0)} \ln \frac{S_0}{a} h'(0)$$

$$= - \ln \left( \frac{a}{S_0} \right) \frac{1}{\mu - \sigma^2/2}.$$ 

Exercise 5.3.3 implies the desired result. $lacksquare$

Exercise 5.3.5 Use Exercise 5.3.3 to find

(a) $E[T_a^2]$;

(b) $Var(T_a)$.

The probability density of $T_a$ Denote by $p_a(\tau)$ the probability density of $T_a$. Then

$$E[e^{-sT_a}] = \int_0^\infty e^{-a\tau} p_a(\tau) d\tau$$

is the Laplace transform of $p_a(\tau)$. Hence (5.3.8) becomes

$$\mathcal{L}(p_a(\tau))(s) = \left( \frac{S_0}{a} \right)^{h(s)}.$$ 

The probability density can be retrieved as an inverse Laplace transform

$$p_a(t) = \left( \mathcal{L}^{-1} \left( \frac{S_0}{a} \right)^{h(s)} \right)(t).$$ 

Using that

$$\left( \frac{S_0}{a} \right)^{h(s)} = e^{h(s) \ln \frac{S_0}{a}} = \left( \frac{S_0}{a} \right)^{\frac{1}{2} \frac{\mu}{\sigma^2}} e^{-\frac{1}{2} \ln \frac{S_0}{a} \sqrt{2\sigma^2 + (\frac{\mu}{\sigma^2})^2}},$$

using the formula

$$\mathcal{L}^{-1} \left( e^{-c\sqrt{2s + \alpha^2}} \right)(t) = \frac{ce^{-cnt}}{\sqrt{2\pi t^{3/2}}} e^{-\frac{1}{2}(c-nt)^2}$$
with
\[ c = \frac{1}{\sigma} \ln \frac{a}{S_0}, \quad n = \frac{\mu}{\sigma} - \frac{\sigma}{2}, \]
an algebraic computation provides
\[ \left( e^{-\frac{1}{2} \mu \ln \frac{a}{S_0} - \frac{\sigma^2}{4}} \right)^t = \frac{\ln \frac{a}{S_0}}{\sigma \sqrt{2\pi t}} e^{-\frac{1}{2} \left[ \frac{1}{2} \ln \frac{a}{S_0} - \frac{\sigma^2}{2} \right] t}, \]
and hence the probability density of \( T_a \) is given by
\[ p_a(t) = \frac{\ln \frac{a}{S_0}}{\sigma \sqrt{2\pi t}} e^{-\frac{1}{2} \left[ \frac{1}{2} \ln \frac{a}{S_0} - \frac{\sigma^2}{2} \right] t}, \quad t > 0. \]

**Exercise 5.3.6** Consider the doubling time of a stock
\[ T_2 = \inf \{ t > 0; S_t = 2S_0 \}. \]
(a) Find \( E[T_2] \) and \( \text{Var}(T_2) \). Do these values depend on \( S_0 \)?
(b) The expected return of a stock is \( \mu = 0.15 \) and its volatility \( \sigma = 0.20 \). Find the expected time when the stock doubles its value.

**Exercise 5.3.7** Let \( \overline{S}_t = \max_{u \leq t} S_u \) and \( \underline{S}_t = \min_{u \leq t} S_u \) be the running maximum and minimum of the stock.
(a) Show that the events \( \{ \overline{S}_t \geq a \} \) and \( \{ T_a \leq t \} \) are the same.
(b) Use (a) to obtain the distribution function of \( \overline{S}_t \);
(c) Use a similar argument to find the distribution function of \( \underline{S}_t \).

**Exercise 5.3.8** (a) What is the probability that the stock \( S_t \) reaches level \( a, \ a > S_0, \) before time \( T \)?
(b) Describe the probability that the stock \( S_t \) reaches level \( a, \ a > S_0, \) before time \( T_2 \) and after time \( T_1 \).

### 5.4 Probability to Hit a Barrier

We shall deal next with the probability of a stock reaching a certain barrier. This is related to a result describing the probability of a Brownian motion \( W_t \) that hits the line \( y = \alpha + \gamma t \) in the time interval \([0, T]\) (see also Karatzas and Shreve [11], p.265). First, we note the relation with the probability of the running maximum of a Brownian motion with drift:
\[ P(W_t \geq \alpha + \gamma t, \text{ for some } t \in [0, T]) = P(\max_{0 \leq t \leq T}(W_t - \gamma t) \geq \alpha). \]
The following result will be used shortly.
Lemma 5.4.1 Let $\alpha > 0$. Then
\[
P\left(\max_{0 \leq t \leq T} (W_t - \gamma t) \geq \alpha\right) = 1 - N\left(\gamma \sqrt{T} + \frac{\alpha}{\sqrt{T}}\right) + e^{-2\alpha \gamma} N\left(\gamma \sqrt{T} - \frac{\alpha}{\sqrt{T}}\right),
\]
where $N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-z^2/2} dz$.

Proof: We shall do the proof for $\gamma < 0$. The case $\gamma > 0$ is similar. Consider the Brownian motion with drift, $X_t = W_t - \gamma t$, and let $\tau$ denote the hitting time
\[
\tau = \inf\{t > 0; X_t \geq \alpha\}.
\]
Since the event $\{\max_{0 \leq t \leq T} (W_t - \gamma t) \geq \alpha\}$ occurs if and only if $\tau \leq T$, using Exercise 5.4.5 (c) we have
\[
P\left(\max_{0 \leq t \leq T} (W_t - \gamma t) \geq \alpha\right) = P(\tau \leq T) = \int_{0}^{T} \frac{\alpha}{\sqrt{2\pi T^{3/2}}} e^{-\frac{(\alpha + \gamma T)^2}{2T}} d\tau.
\]
Hence, it suffices to show that
\[
\int_{0}^{T} \frac{\alpha}{\sqrt{2\pi T^{3/2}}} e^{-\frac{(\alpha + \gamma T)^2}{2T}} d\tau = 1 - N\left(\gamma \sqrt{T} + \frac{\alpha}{\sqrt{T}}\right) + e^{-2\alpha \gamma} N\left(\gamma \sqrt{T} - \frac{\alpha}{\sqrt{T}}\right).
\]
(5.4.9)

Even if a direct computation of the integral of the right side exists, for the sake of simplicity we shall employ a shorter method, which serves as a verification of formula (5.4.9). Consider the functions
\[
f(T) = 1 - N\left(\gamma \sqrt{T} + \frac{\alpha}{\sqrt{T}}\right) + e^{-2\alpha \gamma} N\left(\gamma \sqrt{T} - \frac{\alpha}{\sqrt{T}}\right)
\]
\[
g(T) = \int_{0}^{T} \frac{\alpha}{\sqrt{2\pi T^{3/2}}} e^{-\frac{(\alpha + \gamma T)^2}{2T}} d\tau.
\]
If we show that $f(0) = g(0)$ and $f'(T) = g'(T)$ for any $T \geq 0$, then it follows that $f(T) = g(T)$. 

Figure 5.3: The running maximum on the stock.
Using the properties of $N(\cdot)$ we have
\[
\begin{align*}
f(0) &= 1 - \lim_{T \to 0^+} N\left(\gamma \sqrt{T} + \frac{\alpha}{\sqrt{T}}\right) + e^{-2\alpha \gamma} \lim_{T \to 0^+} N\left(\gamma \sqrt{T} - \frac{\alpha}{\sqrt{T}}\right) \\
&= 1 - 1 + 0 = 0, \\
g(0) &= 0.
\end{align*}
\]

The Fundamental Theorem of Calculus yields
\[
g'(T) = \frac{\alpha}{\sqrt{2\pi T^{3/2}}} e^{-\frac{(\alpha + \gamma T)^2}{2T}}. \tag{5.4.10}
\]

Using $N'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, applying the chain rule we have
\[
\begin{align*}
f'(T) &= -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\gamma \sqrt{T} + \alpha/\sqrt{T})^2}\left(\frac{\gamma}{2\sqrt{T}} - \frac{\alpha}{2T^{3/2}}\right) \\
&\quad + e^{-2\alpha \gamma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\gamma \sqrt{T} - \alpha/\sqrt{T})^2}\left(\frac{\gamma}{2\sqrt{T}} + \frac{\alpha}{2T^{3/2}}\right) \\
&= -\frac{1}{\sqrt{2\pi}} e^{-\frac{(\alpha + \gamma T)^2}{2T}}\left(\frac{\gamma}{2\sqrt{T}} - \frac{\alpha}{2T^{3/2}}\right) \\
&\quad + \frac{1}{\sqrt{2\pi}} e^{-\frac{(\alpha + \gamma T)^2}{2T}}\left(\frac{\gamma}{2\sqrt{T}} + \frac{\alpha}{2T^{3/2}}\right) \\
&= \frac{\alpha}{\sqrt{2\pi T^{3/2}}} e^{-\frac{(\alpha + \gamma T)^2}{2T}}. \tag{5.4.11}
\end{align*}
\]

Comparing equations (5.4.10) and (5.4.11) yields $f'(T) = g'(T)$. Hence $f(T) = g(T)$ for any $T \geq 0$.

\section*{Corollary 5.4.2} We have
\[
P\left(W_t \geq \alpha + \gamma t, \text{ for some } t \geq 0\right) = \begin{cases} 
1, & \text{if } \gamma \leq 0 \\
e^{-2\alpha \gamma}, & \text{if } \gamma > 0.
\end{cases}
\]

\textbf{Proof:} Take the limit $T \to \infty$ in Lemma 5.4.1 and use that $N(-\infty) = 0$ and $N(+\infty) = 1$. 

Consider the notations
\[
\begin{align*}
d_6 &= \frac{\ln S_0}{b} + \left(\mu - \frac{a^2}{2}\right)T, \tag{5.4.12} \\
d_8 &= \frac{\ln S_0}{b} + \left(\mu - \frac{a^2}{2}\right)T. \tag{5.4.13}
\end{align*}
\]

The next result shows the connection between the running maximum on the stock, $d_6$ and $d_8$. 
Proposition 5.4.3 Let $S_T = \max_{0 \leq t \leq T} S_t$, and let $b > 0$, see Fig. 5.3. Then

$$P(\bar{S}_T \geq b) = N(d_6) + \left( \frac{b}{S_0} \right)^{\frac{2\mu}{\sigma^2} - 1} N(-d_8).$$  \hspace{1cm} (5.4.14)$$

Proof: Let $m = \mu - \frac{\sigma^2}{2}$. Substituting $\gamma = -\frac{m}{\sigma}$ and $\alpha = \frac{1}{\sigma} \ln \frac{b}{S_0}$ in Lemma 5.4.1 we have

$$P(\bar{S}_T \geq b) = P(\max_{t \leq T} (mt + \sigma W_t) \geq b) = P(\max_{t \leq T} \left( \frac{m}{\sigma} t + W_t \right) \geq \frac{1}{\sigma} \ln \frac{b}{S_0})$$

$$= P(W_t - \gamma t \geq \alpha)$$

$$= 1 - N\left( \gamma \sqrt{T} + \frac{\alpha}{\sqrt{T}} \right) + e^{-2\alpha \gamma} N\left( \gamma \sqrt{T} - \frac{\alpha}{\sqrt{T}} \right).$$  \hspace{1cm} (5.4.15)$$

Since

$$\gamma \sqrt{T} + \frac{\alpha}{\sqrt{T}} = \frac{\gamma T + \alpha}{\sqrt{T}} = -mT + \ln \frac{b}{S_0} = -\ln \frac{S_0}{b} + (\mu - \frac{\sigma^2}{2})T = -d_6;$$

$$\gamma \sqrt{T} - \frac{\alpha}{\sqrt{T}} = \frac{\gamma T - \alpha}{\sqrt{T}} = -mT - \ln \frac{b}{S_0} = -\ln \frac{S_0}{b} + (\mu - \frac{\sigma^2}{2})T = -d_8;$$

$$e^{-2\alpha \gamma} = e^{\frac{2m}{\sigma^2} \ln \frac{b}{S_0}} = \left( \frac{b}{S_0} \right)^{\frac{2m}{\sigma^2}} = \left( \frac{b}{S_0} \right)^{\frac{2\mu}{\sigma^2} - 1},$$

substituting in (5.4.15) yields

$$P(\bar{S}_T \geq b) = 1 - N(-d_6) + \left( \frac{b}{S_0} \right)^{\frac{2\mu}{\sigma^2} - 1} N(-d_8)$$

$$= N(d_6) + \left( \frac{b}{S_0} \right)^{\frac{2\mu}{\sigma^2} - 1} N(-d_8).$$

\[\blacksquare\]

Corollary 5.4.4 Let $b > 0$. The probability that the stock $S_t$ will ever reach the barrier $b$ is

$$P(\sup_{t \geq 0} S_t \geq b) = \begin{cases} \left( \frac{b}{S_0} \right)^{\frac{2\mu}{\sigma^2} - 1}, & \text{if } \mu < \frac{\sigma^2}{2} \\ 1, & \text{if } \mu \geq \frac{\sigma^2}{2}. \end{cases}$$  \hspace{1cm} (5.4.16)$$
Proof: (i) Let $\mu > \frac{\sigma^2}{2}$. Since

$$P(\sup_{t \geq 0} S_t \geq b) = \lim_{T \to \infty} P(\mathcal{S}_T \geq b) = P(\mathcal{S}_\infty \geq b),$$

and $\lim_{T \to \infty} d_6 = +\infty$, $\lim_{T \to \infty} d_8 = +\infty$, using Proposition 5.4.3 we have

$$P(\mathcal{S}_\infty \geq b) = N(+\infty) + \left( \frac{b}{S_0} \right)^{\frac{2}{\sigma^2}} N(-\infty) = 1.$$

(ii) Let $\mu < \frac{\sigma^2}{2}$. Since $\lim_{T \to \infty} d_6 = -\infty$, $\lim_{T \to \infty} d_8 = -\infty$, Proposition 5.4.3 implies

$$P(\mathcal{S}_\infty \geq b) = N(-\infty) + \left( \frac{b}{S_0} \right)^{\frac{2}{\sigma^2}} N(+\infty) = \left( \frac{b}{S_0} \right)^{\frac{2}{\sigma^2}}.$$

The previous formulas will be needed when pricing lookback options.

**Exercise 5.4.5** Let $X_t = W_t + \mu t$ be a Brownian motion with drift, $\mu > 0$, and consider the stopping time

$$\tau = \inf\{t > 0; X_t \geq \alpha\},$$

with $\alpha > 0$.

(a) Apply the Optional Stopping Time to the martingale $M_t = e^{cW_t - \frac{1}{2}c^2t}$ to show that

$$E[e^{-(c\mu + \frac{1}{2}c^2)\tau}] = e^{-c\alpha}.$$

(b) Substitute $s = c\mu + \frac{1}{2}c^2$ to prove $E[e^{-s\tau}] = e^{\frac{1}{2s}(\mu - \sqrt{2s + \mu^2})\alpha}$, $s > 0$.

(c) Using the inverse Laplace transform, show that the probability density of $\tau$ is given by

$$p(t) = \frac{\alpha}{\sqrt{2\pi t^{3/2}}} e^{-\frac{1}{2t}(\alpha - \mu)^2}, \quad t > 0.$$

**Exercise 5.4.6** Let $X_t = W_t + \mu t$, with $\mu < 0$. Using a method similar with the one described in Exercise 5.4.5, prove that:

(a) $E[e^{-s\tau}] = e^{\frac{1}{2s}(\mu + \sqrt{2s + \mu^2})\alpha}$, $s > 0$.

(b) The probability density of $\tau$ is given by

$$p(t) = \frac{\alpha}{\sqrt{2\pi t^{3/2}}} e^{-\frac{1}{2t}(\alpha + \mu)^2}, \quad t > 0.$$
5.5 Multiple Barriers

This section studies the case of the double barrier and deals with the probability that the stock price reaches a certain barrier before another barrier, see Fig. 5.4 a,b. We still work under under the assumption that the stock prices are lognormally distributed.

Theorem 5.5.1 Let $S_u$ and $S_d$ be fixed, such that $S_d < S_0 < S_u$.

(i) The probability that the stock price $S_t$ hits the upper value $S_u$ before the lower value $S_d$ is

$$p = \frac{d^\gamma - 1}{d^\gamma - u^\gamma};$$

(ii) The probability that the stock price $S_t$ hits the lower value $S_d$ before the upper value $S_u$ is

$$q = 1 - p = \frac{1 - u^\gamma}{d^\gamma - u^\gamma},$$

where $S_u/S_0 = u$, $S_d/S_0 = d$, and $\gamma = 1 - 2\mu/\sigma^2$.

Proof: Let $X_t = mt + W_t$. Exercise 5.5.5 provides

$$P(X_t \text{ goes up to } \alpha \text{ before down to } -\beta) = \frac{e^{2m\beta} - 1}{e^{2m\beta} - e^{-2m\alpha}}.$$  \hspace{1cm} (5.5.17)

Choosing the following values for the parameters

$$m = \frac{\mu}{\sigma} - \frac{\sigma}{2}, \quad \alpha = \frac{\ln u}{\sigma}, \quad \beta = -\frac{\ln d}{\sigma},$$
we have the sequence of identities
\[
P(X_t \text{ goes up to } \alpha \text{ before down to } -\beta) = P(\sigma X_t \text{ goes up to } \sigma \alpha \text{ before down to } -\sigma \beta) = P(S_0 e^{\sigma X_t} \text{ goes up to } S_0 e^{\sigma \alpha} \text{ before down to } S_0 e^{-\sigma \beta}) = P(S_t \text{ goes up to } S_u \text{ before down to } S_d).
\]

Using (5.5.17) yields
\[
P(S_t \text{ goes up to } S_u \text{ before down to } S_d) = \frac{e^{2m\beta} - 1}{e^{2m\beta} - e^{-2m\alpha}}.
\] (5.5.18)

Since a computation shows that
\[
e^{2m\beta} = e^{-\left(\frac{2\mu}{\sigma^2} - 1\right) \ln d} = d^{1-\frac{2\mu}{\sigma^2}} = d^\gamma
\]

\[
e^{-2m\alpha} = e^{\left(\frac{2\mu}{\sigma^2} + 1\right) \ln u} = u^{1-\frac{2\mu}{\sigma^2}} = u^\gamma,
\]

formula (5.5.18) becomes
\[
P(S_t \text{ goes up to } S_u \text{ before down to } S_d) = \frac{d^\gamma - 1}{d^\gamma - u^\gamma},
\]

which ends the proof.

**Corollary 5.5.2** Let \( S_u > S_0 > 0 \) be fixed. Then
\[
P(S_t \text{ hits } S_u) = \left(\frac{S_0}{S_u}\right)^{1-\frac{2\mu}{\sigma^2}} \text{ for some } t > 0.
\]

**Proof:** Taking \( d = 0 \) implies \( S_d = 0 \). Since \( S_t \) never reaches zero,
\[
P(S_t \text{ hits } S_u) = P(S_t \text{ goes up to } S_u \text{ before down to } S_d = 0) = \left. \frac{d^\gamma - 1}{d^\gamma - u^\gamma} \right|_{d = 0} = \left. \frac{1}{u^\gamma} \right| \left(\frac{S_0}{S_u}\right)^{1-\frac{2\mu}{\sigma^2}}.
\]

**Exercise 5.5.3** A stock has \( S_0 = \$10, \sigma = 0.15, \mu = 0.20 \). What is the probability that the stock goes up to \$15 before it goes down to \$5?

**Exercise 5.5.4** Let \( 0 < S_0 < S_u \). What is the probability that \( S_t \) hits \( S_u \) for some time \( t > 0 \)?
**Exercise 5.5.5** Let $X_t = mt + W_t$ and consider $\alpha, \beta > 0$. Consider the stopping time

$$T = \inf\{t > 0; X_t \geq \alpha \text{ or } X_t \leq -\beta\},$$

when $X_t$ exits first time the interval $(-\beta, \alpha)$.

(a) Use the Optional Stopping Theorem to show

$$E[e^{cX_T - (c\mu + \frac{1}{2}c^2)T}] = 1, \quad \forall c > 0;$$

(b) Choose a convenient value of the constant $c$ to obtain

$$E[e^{-2mX_T}] = 1;$$

(c) Denote $p_\alpha = P(X_T = \alpha)$, $p_\beta = P(X_T = -\beta)$. Show that

$$p_\alpha = \frac{e^{2m\beta} - 1}{e^{2m\beta} - e^{-2m\alpha}}; \quad p_\beta = \frac{1 - e^{2m\alpha}}{e^{2m\beta} - e^{-2m\alpha}}.$$

### 5.6 Parameters Estimation

Assume the stock price $S_t$, which satisfies the equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

with $\mu > 0$ and $\sigma > 0$ constants, is observed $n$ times. Denote by $s_j$ the observed values of $S_{t_j}$, $j = 1, \ldots, n$. The problem is to estimate the model parameters $\mu$ and $\sigma$ from the observations $s_j$ at times $t_j$. This will be done using the method of maximum likelihood. The process

$$X_t = \ln\frac{S_t}{S_0}$$

is a Brownian motion with drift

$$X_t = (\mu - \frac{1}{2}\sigma^2)t + \sigma W_t.$$

The observations of $X_t$ at time $t_j$ are $y_j = \ln\frac{s_j}{S_0}$. Example 2.4.1 provides the following estimation of parameters from the observed values

$$\mu - \frac{1}{2}\sigma^2 = \frac{\sum_{j=1}^{n} y_j}{\sum_{j=1}^{n} t_j};$$

$$\sigma^2 = \frac{1}{n} \sum_{j=1}^{n} \frac{(y_j - \mu t_j)^2}{t_j}.$$
Eliminating $\sigma^2$ from the previous two equations, we obtain the following quadratic equation satisfied by $\mu$

$$\mu^2 \sum_{j=1}^{n} t_j - 2\mu \left( n + \sum_{j=1}^{n} y_j \right) + \sum_{j=1}^{n} \frac{y_j^2}{t_j} + 2n \sum_{j=1}^{n} y_j^2 t_j = 0. \quad (5.6.19)$$

The positive solution (provided it exists) is the desired $\mu$. Substituting back in the expression of $\sigma^2$ yields an estimation for the volatility parameter.

### 5.7 Time-dependent Drift and Volatility

This model considers the drift $\mu = \mu(t)$ and volatility $\sigma = \sigma(t)$ to be deterministic functions of time. In this case the equation (5.1.1) becomes

$$dS_t = \mu(t)S_t dt + \sigma(t)S_t dW_t. \quad (5.7.20)$$

We shall solve the equation using the method of integrating factor. Multiplying by the integrating factor

$$\rho_t = e^{-\int_0^t \sigma(s) \, dW_s + \frac{1}{2} \int_0^t \sigma^2(s) \, ds}$$

the equation (5.7.20) becomes $d(\rho_t S_t) = \rho_t \mu(t) S_t \, dt$. Substituting $Y_t = \rho_t S_t$ yields the deterministic equation $dY_t = \mu(t)Y_t dt$ with the solution

$$Y_t = Y_0 e^{\int_0^t \mu(s) \, ds}.$$

Substituting back $S_t = \rho_t^{-1} Y_t$, we obtain the closed-form solution of equation (5.7.20)

$$S_t = S_0 e^{\int_0^t (\mu(s) - \frac{1}{2} \sigma^2(s)) \, ds + \int_0^t \sigma(s) \, dW_s}.$$

**Proposition 5.7.1** The solution $S_t$ is $\mathcal{F}_t$-adapted and log-normally distributed, with mean and variance given by

$$E[S_t] = S_0 e^{\int_0^t \mu(s) \, ds}$$

$$Var[S_t] = S_0^2 e^{2 \int_0^t \mu(s) \, ds} \left( e^{\int_0^t \sigma^2(s) \, ds} - 1 \right).$$

**Proof:** Let $X_t = \int_0^t (\mu(s) - \frac{1}{2} \sigma^2(s)) \, ds + \int_0^t \sigma(s) \, dW_s$. Since $X_t$ is a sum of a predictable integral function and a Wiener integral, then it is normally distributed, with

$$E[X_t] = \int_0^t (\mu(s) - \frac{1}{2} \sigma^2(s)) \, ds$$

$$Var[X_t] = Var \left[ \int_0^t \sigma(s) \, dW_s \right] = \int_0^t \sigma^2(s) \, ds.$$
Then the mean and variance of the log-normal random variable $S_t = S_0 e^{X_t}$ are given by

\[
E[S_t] = S_0 e^{\int_0^t (\mu - \frac{\sigma^2}{2}) ds + \frac{1}{2} \int_0^t \sigma^2 ds} = S_0 e^{\int_0^t \mu(s) ds}
\]

\[
Var[S_t] = S_0^2 e^{2 \int_0^t (\mu - \frac{1}{2} \sigma^2) ds + \int_0^t \sigma^2 ds} \left( e^{\int_0^t \sigma^2 ds} - 1 \right)
= S_0^2 e^{2 \int_0^t \mu(s) ds} \left( e^{\int_0^t \sigma^2(s) ds} - 1 \right).
\]

If the average drift and average squared volatility are defined as

\[
\overline{\mu} = \frac{1}{t} \int_0^t \mu(s) ds
\]

\[
\overline{\sigma^2} = \frac{1}{t} \int_0^t \sigma^2(s) ds,
\]

the aforementioned formulas can be also written as

\[
E[S_t] = S_0 e^{\overline{\mu} t}
\]

\[
Var[S_t] = S_0^2 e^{2 \overline{\mu} t} \left( e^{\overline{\sigma^2} t} - 1 \right).
\]

It is worth noting that we have obtained formulas similar to (5.1.3)–(5.1.4).

### 5.8 Models for Stock Price Averages

This section described the stochastic differential equations for several types of averages on stocks. These averages are used as underlying assets in the case of Asian options.

**Discretely sampled averages**

Let $S_{t_1}, S_{t_2}, \ldots, S_{t_n}$ be a sample of stock prices at $n$ instances of time $t_1 < t_2 < \cdots < t_n$. The most common types of discrete averages are:

- The **arithmetic average**

\[
A(t_1, t_2, \cdots, t_n) = \frac{1}{n} \sum_{k=1}^n S_{t_k}.
\]

- The **geometric average**

\[
G(t_1, t_2, \cdots, t_n) = \left( \prod_{k=1}^n S_{t_k} \right)^{\frac{1}{n}}.
\]
• The harmonic average

\[ H(t_1, t_2, \cdots, t_n) = \frac{n}{\sum_{k=1}^{n} \frac{1}{S_{tk}}} \]

The well-known inequality of means states

\[ H(t_1, t_2, \cdots, t_n) \leq G(t_1, t_2, \cdots, t_n) \leq A(t_1, t_2, \cdots, t_n), \quad (5.8.21) \]

with identity in the case of constant stock prices.

**The continuously sampled arithmetic average**

Let \( t_n = t \) and assume \( t_{k+1} - t_k = \frac{t}{n} \). Using the definition of the integral as a limit of Riemann sums, we have

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} S_{tk} = \lim_{n \to \infty} \frac{1}{t} \sum_{k=1}^{n} S_{tk} \frac{t}{n} = \frac{1}{t} \int_{0}^{t} S_u \, du. \]

It follows that the *continuously sampled arithmetic average* of stock prices between 0 and \( t \) is given by

\[
A_t = \frac{1}{t} \int_{0}^{t} S_u \, du.
\]

Assuming that \( S_t \) has constant drift rate and volatility, we obtain

\[
A_t = \frac{S_0}{t} \int_{0}^{t} e^{(\mu - \sigma^2/2)u + \sigma W_u} \, du.
\]

This integral can be computed explicitly only in the case \( \sigma = 0 \). In the absence of a closed form expression for \( A_t \), one can hope to describe \( A_t \) by its probability law. For a long time the probability density of \( A_t \) was unknown, fact that made many authors to proceed using various approximations. Fortunately, a trackable formula for the law of \( A_t \) was found by Yor, see [20], in early 90’s. The fact that \( A_t \) is neither normal nor log-normal makes the price of Asian options on arithmetic averages hard to evaluate.

In the following we shall provide an elementary computation of the first two moments of \( A_t \).

Let \( I_t = \int_{0}^{t} S_u \, du \). The Fundamental Theorem of Calculus implies \( dI_t = S_t \, dt \). Then the quotient rule yields

\[
dA_t = d \left( \frac{I_t}{t} \right) = \frac{dI_t \, t - I_t \, dt}{t^2} = \frac{S_t \, dt \, t - I_t \, dt}{t^2} = \frac{1}{t} (S_t - A_t) \, dt,
\]
i.e. the continuous arithmetic average $A_t$ satisfies
\[ dA_t = \frac{1}{t}(S_t - A_t)dt. \]

If $A_t < S_t$, the right side is positive and hence $dA_t > 0$, i.e. the average $A_t$ goes up. Similarly, if $A_t > S_t$, then the average $A_t$ goes down. This shows that the average $A_t$ tends to trace the stock values $S_t$.

An application of l'Hospital's rule implies
\[ A_0 = \lim_{t \to 0} \frac{I_t}{t} = \lim_{t \to 0} S_t = S_0. \]

Using that the expectation commutes with integrals, we have
\[ E[A_t] = \frac{1}{t} \int_0^t E[S_u] \, du = \frac{1}{t} \int_0^t S_0 e^{\mu u} \, du = S_0 \frac{e^{\mu t} - 1}{\mu t}. \]

Hence the average of the arithmetic average is
\[ E[A_t] = \begin{cases} S_0 \frac{e^{\mu t} - 1}{\mu t}, & \text{if } t > 0 \\ S_0, & \text{if } t = 0. \end{cases} \quad (5.8.22) \]

In the following we shall compute the variance $Var[A_t]$. Since
\[ Var[A_t] = \frac{1}{t^2} E[I_t^2] - E[A_t]^2, \quad (5.8.23) \]

it suffices to find $E[I_t^2]$. We need first the following result:

**Lemma 5.8.1** (i) We have
\[ E[I_t S_t] = \frac{S_0^2}{\mu + \sigma^2} [e^{(2\mu + \sigma^2)t} - e^{\mu t}]. \]

(ii) The processes $A_t$ and $S_t$ are not independent.

**Proof:** (i) Using Ito’s formula
\[ d(I_t S_t) = dI_t S_t + I_t dS_t + dI_t dS_t \]
\[ = S_t^2 dt + I_t (\mu S_t dt + \sigma S_t dW_t) + \underbrace{S_t dt dS_t}_0 \]
\[ = (S_t^2 + \mu I_t S_t) dt + \sigma I_t S_t dW_t. \]

Using $I_0 S_0 = 0$, integrating between 0 and $t$ yields
\[ I_t S_t = \int_0^t (S_u^2 + \mu I_u S_u) \, du + \sigma \int_0^t I_u S_u dW_u. \]
Since the expectation of the Ito integral is zero, we have

\[ E[I_tS_t] = \int_0^t (E[S^2_u] + \mu E[I_uS_u]) \, du. \]

Using Exercise 5.1.4 this becomes

\[ E[I_tS_t] = \int_0^t \left( S_0^2 e^{(2\mu + \sigma^2)u} + \mu E[I_uS_u] \right) \, du. \]

If we denote \( g(t) = E[I_tS_t] \), differentiating yields the ODE

\[ g'(t) = S_0^2 e^{(2\mu + \sigma^2)t} + \mu g(t), \]

with the initial condition \( g(0) = 0 \). This can be solved as a linear differential equation in \( g(t) \) by multiplying by the integrating factor \( e^{-\mu t} \). The solution is

\[ g(t) = \frac{S_0^2}{\mu + \sigma^2} \left[ e^{(2\mu + \sigma^2)t} - e^{\mu t} \right]. \]

\[ \text{(ii) Since } A_t \text{ and } I_t \text{ are proportional, it suffices to show that } I_t \text{ and } S_t \text{ are not independent. This follows from part (i) and the fact that} \]

\[ E[I_tS_t] \neq E[I_t]E[S_t] = \frac{S_0^2}{\mu} (e^{\mu t} - 1)e^{\mu t}. \]

Next we shall find \( E[I_t^2] \). Using \( dI_t = S_t \, dt \), then \( (dI_t)^2 = 0 \) and hence Ito’s formula yields

\[ d(I_t^2) = 2I_t \, dI_t + (dI_t)^2 = 2I_t S_t \, dt. \]

Integrating between 0 and \( t \) and using \( I_0 = 0 \) leads to

\[ I_t^2 = 2 \int_0^t I_uS_u \, du. \]

Taking the expectation and using Lemma 5.8.1 we obtain

\[ E[I_t^2] = 2 \int_0^t E[I_uS_u] \, du = \frac{2S_0^2}{\mu + \sigma^2} \left[ \frac{e^{(2\mu + \sigma^2)t} - 1}{2\mu + \sigma^2} - \frac{e^{\mu t} - 1}{\mu} \right]. \]  \[ \text{(5.8.24)} \]

Substituting into (5.8.23) yields

\[ \text{Var}[A_t] = \frac{S_0^2}{t^2} \left\{ \frac{2}{\mu + \sigma^2} \left[ \frac{e^{(2\mu + \sigma^2)t} - 1}{2\mu + \sigma^2} - \frac{e^{\mu t} - 1}{\mu} \right] - \left( \frac{e^{\mu t} - 1}{\mu} \right)^2 \right\}. \]

Concluding the previous calculations, we have the following result:
Proposition 5.8.2 The arithmetic average $A_t$ satisfies the stochastic equation
\[ dA_t = \frac{1}{t}(S_t - A_t)dt, \quad A_0 = S_0. \]
Its mean and variance are given by
\[ E[A_t] = S_0 e^{\mu t} - \frac{1}{\mu t}, \quad t > 0 \]
\[ \text{Var}[A_t] = \frac{S_0^2}{t^2} \left\{ \frac{2}{\mu + \sigma^2} \left[ e^{(2\mu + \sigma^2)t} - 1 - \frac{e^{\mu t} - 1}{\mu} \right] - \left( \frac{e^{\mu t} - 1)^2}{\mu^2} \right) \right\}. \]

Exercise 5.8.3 Find approximative formulas for $E[A_t]$ and $\text{Var}[A_t]$ for $t$ small, up to the order $O(t^2)$.
(Recall that $f(t) = O(t^2)$ if $\lim_{t \to \infty} \frac{f(t)}{t^2} = c < \infty$).

The continuously sampled geometric average
First, we shall find an integral formula for the continuously sampled geometric average. Dividing the interval $[0, t]$ into equal subintervals of length $t_{k+1} - t_k = \frac{t}{n}$, we have
\[ G(t_1, \ldots, t_n) = \left( \prod_{k=1}^{n} S_{t_k} \right)^{1/n} = e^{\frac{1}{n} \sum_{k=1}^{n} \ln S_{t_k}} = e^{\frac{1}{n} \sum_{k=1}^{n} \ln S_{t_k} \frac{1}{n}}. \]
Using the definition of the integral as a limit of Riemann sums
\[ G_t = \lim_{n \to \infty} \left( \prod_{k=1}^{n} S_{t_k} \right)^{1/n} = \lim_{n \to \infty} e^{\frac{1}{n} \sum_{k=1}^{n} \ln S_{t_k} \frac{1}{n}} = e^{\frac{1}{t} \int_0^t \ln S_u du}. \]
Therefore, the continuously sampled geometric average of stock prices between instances 0 and $t$ is given by
\[ G_t = e^{\frac{1}{t} \int_0^t \ln S_u du}. \]
(5.8.25)

Theorem 5.8.4 $G_t$ has a log-normal distribution, with the mean and variance given by
\[ E[G_t] = S_0 e^{(\mu - \sigma^2 \frac{1}{2}) t} \]
\[ \text{Var}[G_t] = S_0^2 e^{(\mu - \sigma^2 \frac{1}{2}) t} \left( e^{\sigma^2 t} - 1 \right). \]
Proof: Using
\[ \ln S_u = \ln \left( S_0 e^{\left( u - \frac{\sigma^2}{2}\right) u + \sigma W_u} \right) = \ln S_0 + (\mu - \frac{\sigma^2}{2})u + \sigma W_u, \]
then taking the logarithm yields
\[ \ln G_t = \frac{1}{t} \int_0^t \left[ \ln S_0 + (\mu - \frac{\sigma^2}{2})u + \sigma W_u \right] du = \ln S_0 + (\mu - \frac{\sigma^2}{2}) \frac{t}{2} + \sigma \int_0^t W_u du. \]
(5.8.26)
Since the integrated Brownian motion \( Z_t = \int_0^t W_u du \) is Gaussian with \( Z_t \sim N(0, t^3/3) \), it follows that \( \ln G_t \) has a normal distribution
\[ \ln G_t \sim N \left( \ln S_0 + (\mu - \frac{\sigma^2}{2}) \frac{t}{2}, \sigma^2 \frac{t}{3} \right). \]
(5.8.27)
This implies that \( G_t \) has a log-normal distribution. Using Exercise ??, we obtain
\[
\begin{align*}
E[G_t] &= e^{E[\ln G_t] + \frac{1}{2} \text{Var}[\ln G_t]} = e^{\ln S_0 + (\mu - \frac{\sigma^2}{2}) \frac{t}{2} + \frac{\sigma^2 t}{4}} = S_0 e^{(\mu - \frac{\sigma^2}{2}) \frac{t}{2}}, \\
\text{Var}[G_t] &= e^{2E[\ln G_t] + \text{Var}[\ln G_t]} e^{\text{Var}[\ln G_t]} - 1 = e^{2 \ln S_0 + (\mu - \frac{\sigma^2}{2}) t} \left( e^{\frac{\sigma^2 t}{3}} - 1 \right) = S_0^2 e^{(\mu - \frac{\sigma^2}{6}) t} \left( e^{\frac{\sigma^2 t}{3}} - 1 \right).
\end{align*}
\]
\]
Corollary 5.8.5 The geometric average \( G_t \) is given by the closed-form formula
\[ G_t = S_0 e^{\left( \mu - \frac{\sigma^2}{2} \right) \frac{t}{2} + \frac{\sigma^2 t}{4} \int_0^t W_u du}. \]
Proof: Take the exponential in the formula (5.8.26).
\]
An important consequence of the fact that \( G_t \) is log-normal is that Asian options on geometric averages have closed-form solutions.
Exercise 5.8.6 (a) Show that \( \ln G_t \) satisfies the stochastic differential equation
\[ d(\ln G_t) = \frac{1}{t} \left( \ln S_t - \ln G_t \right) dt. \]
(b) Show that \( G_t \) satisfies
\[ dG_t = \frac{1}{t} G_t \left( \ln S_t - \ln G_t \right) dt. \]
The continuously sampled harmonic average

Let $S_{tk}$ be values of a stock evaluated at the sampling dates $t_k$, $i = 1, \ldots, n$. Their harmonic average is defined by

$$H(t_1, \ldots, t_n) = \frac{n}{\sum_{k=1}^{n} \frac{1}{S(t_k)}}.$$

Consider $t_k = \frac{kt}{n}$. Then the continuously sampled harmonic average is obtained by taking the limit as $n \to \infty$ in the aforementioned relation

$$\lim_{n \to \infty} \frac{n}{\sum_{k=1}^{n} \frac{1}{S(t_k)}} = \lim_{n \to \infty} \frac{t}{\sum_{k=1}^{n} \frac{1}{S(t_k)n}} = \int_{0}^{t} \frac{1}{S_u} du.$$

Hence, the continuously sampled harmonic average is defined by

$$H_t = \frac{t}{\int_{0}^{t} \frac{1}{S_u} du}.$$

We may also write $H_t = \frac{t}{I_t}$, where $I_t = \int_{0}^{t} \frac{1}{S_u} du$ satisfies

$$dI_t = \frac{1}{S_t} dt, \quad I_0 = 0, \quad d\left(\frac{1}{I_t}\right) = -\frac{1}{S_t I_t^2} dt.$$

From the l'Hospital's rule we get

$$H_0 = \lim_{t \to 0} H_t = \lim_{t \to 0} \frac{t}{I_t} = S_0.$$

Using the product rule we obtain the following:

$$dH_t = t d\left(\frac{1}{I_t}\right) + \frac{1}{I_t} dt + dt d\left(\frac{1}{I_t}\right)$$

$$= \frac{1}{I_t} \left(1 - \frac{t}{S_t I_t}\right) dt + \frac{1}{t} H_t \left(1 - \frac{H_t}{S_t}\right) dt,$$

so

$$dH_t = \frac{1}{t} H_t \left(1 - \frac{H_t}{S_t}\right) dt.$$ \hspace{1cm} (5.8.28)

If at the instance $t$ we have $H_t < S_t$, it follows from the equation that $dH_t > 0$, i.e. the harmonic average increases. Similarly, if $H_t > S_t$, then $dH_t < 0$, i.e $H_t$ decreases. It is worth noting that the converses are also true. The random variable $H_t$ is not normally distributed nor log-normally distributed.
Exercise 5.8.7 Show that $\frac{H_t}{t}$ is a decreasing function of $t$. What is its limit as $t \to \infty$?

Exercise 5.8.8 Show that the continuous analog of inequality (5.8.21) is

$$H_t \leq G_t \leq A_t.$$ 

Exercise 5.8.9 Let $S_{tk}$ be the stock price at time $t_k$. Consider the power $\alpha$ of the arithmetic average of $S_{tk}^\alpha$

$$A^\alpha(t_1, \cdots, t_n) = \left[ \frac{\sum_{k=1}^{n} S_{tk}^\alpha}{n} \right]^\alpha.$$

(a) Show that the aforementioned expression tends to

$$A_t^\alpha = \left[ \frac{1}{t} \int_0^t S_u^\alpha du \right]^\alpha,$$

as $n \to \infty$.

(b) Find the stochastic differential equation satisfied by $A_t^\alpha$.

(c) What does $A_t^\alpha$ become in the particular cases $\alpha = \pm 1$?

Exercise 5.8.10 The stochastic average of stock prices between 0 and $t$ is defined by

$$X_t = \frac{1}{t} \int_0^t S_u dW_u,$$

where $W_u$ is a Brownian motion process.

(a) Find $dX_t$, $E[X_t]$ and $Var[X_t]$.

(b) Show that $\sigma X_t = R_t - \mu A_t$, where $R_t = \frac{S_t - S_0}{t}$ is the “raw average” of the stock price and $A_t = \frac{1}{t} \int_0^t S_u du$ is the continuous arithmetic average.

5.9 Stock Prices with Rare Events

In order to model the stock price when rare events are taken into account, we shall combine the effect of two stochastic processes:

- the Brownian motion process $W_t$, which models regular events given by infinitesimal changes in the price, and which is a continuous process;
- the Poisson process $N_t$, which is discontinuous and models sporadic jumps in the stock price that corresponds to shocks in the market.
Since $E[dN_t] = \lambda dt$, the Poisson process $N_t$ has a positive drift and we need to "compensate" by subtracting $\lambda t$ from $N_t$. The resulting process $M_t = N_t - \lambda t$ is a martingale, called the compensated Poisson process, that models unpredictable jumps of size 1 at a constant rate $\lambda$. It is worth noting that the processes $W_t$ and $M_t$ involved in modeling the stock price are assumed to be independent.

Let $S_{t-} = \lim_{u \uparrow t} S_u$ denote the value of the stock before a possible jump occurs at time $t$. To set up the model, we assume the instantaneous return on the stock, $\frac{dS_t}{S_{t-}}$, to be the sum of the following three components:

- the predictable part $\mu dt$;
- the noisy part due to unexpected news $\sigma dW_t$;
- the rare events part due to unexpected jumps $\rho dM_t$,

where $\mu$, $\sigma$ and $\rho$ are constants, corresponding to the drift rate of the stock, volatility and instantaneous return jump size.²

Adding yields

$$\frac{dS_t}{S_{t-}} = \mu dt + \sigma dW_t + \rho dM_t.$$ 

Hence, the dynamics of a stock price, subject to rare events, are modeled by the following stochastic differential equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t + \rho S_t dM_t.$$ (5.9.29)

It is worth noting that in the case of zero jumps, $\rho = 0$, the previous equation becomes the classical stochastic equation (5.1.1).

Using that $W_t$ and $M_t$ are martingales, we have

$$E[\rho S_t dM_t | \mathcal{F}_t] = \rho S_t E[dM_t | \mathcal{F}_t] = 0,$$

$$E[\sigma S_t dW_t | \mathcal{F}_t] = \sigma S_t E[dW_t | \mathcal{F}_t] = 0.$$

This shows the unpredictability of the last two terms, i.e. given the information set $\mathcal{F}_t$ at time $t$, it is not possible to predict any future increments in the next interval of time $dt$. The term $\sigma S_t dW_t$ captures regular events of insignificant size, while $\rho S_t dM_t$ captures rare events of large size. The "rare events" term, $\rho S_t dM_t$, incorporates jumps proportional to the stock price and is given in terms of the Poisson process $N_t$ as

$$\rho S_{t-} dM_t = \rho S_{t-} d(N_t - \lambda t) = \rho S_{t-} dN_t - \lambda \rho S_{t-} dt.$$

Substituting into equation (5.9.29) yields

$$dS_t = (\mu - \lambda \rho) S_{t-} dt + \sigma S_{t-} dW_t + \rho S_{t-} dN_t.$$ (5.9.30)

²In this model the jump size is constant; there are models where the jump size is a random variable, see Merton [14].
The constant $\lambda$ represents the rate at which the jumps of the Poisson process $N_t$ occur. This is the same as the rate of rare events in the market, and can be determined from historical data.

The following result provides an explicit solution for the stock price when rare events are taken into account.

**Proposition 5.9.1** *The solution of the stochastic equation (5.9.30) is given by*

$$S_t = S_0 e^{(\mu - \lambda \rho - \frac{\sigma^2}{2})t + \sigma W_t (1 + \rho)^N_t},$$

*(5.9.31)*

where

- $\mu$ is the stock price drift rate;
- $\sigma$ is the volatility of the stock;
- $\lambda$ is the rate at which rare events occur;
- $\rho$ is the size of jump in the expected return when a rare event occurs.

**Proof:** We shall construct first the solution and then show that it verifies the equation (5.9.30). If $t_k$ denotes the $k$th jump time, then $N_{tk} = k$. Since there are no jumps before $t_1$, the stock price just before this time is satisfying the stochastic differential equation

$$dS_t = (\mu - \lambda \rho)S_t dt + \sigma S_t dW_t$$

with the solution given by the usual formula

$$S_{t_1} = S_0 e^{(\mu - \lambda \rho - \frac{\sigma^2}{2})t_1 + \sigma W_{t_1}}.$$

Since $\frac{dS_{t_1}}{S_{t_1}} = \frac{S_{t_1} - S_{t_1-}}{S_{t_1-}} = \rho$, then $S_{t_1} = (1 + \rho)S_{t_1-}$. Substituting in the aforementioned formula yields

$$S_{t_1} = S_{t_1-}(1 + \rho) = S_0 e^{(\mu - \lambda \rho - \frac{\sigma^2}{2})(t_1 - t_1) + \sigma W_{t_1} (1 + \rho)}.$$

Since there is no jump between $t_1$ and $t_2$, a similar procedure leads to

$$S_{t_2} = S_{t_2-}(1 + \rho) = S_{t_1} e^{(\mu - \lambda \rho - \frac{\sigma^2}{2})(t_2 - t_1) + \sigma (W_{t_2} - W_{t_1}) (1 + \rho)} = S_0 e^{(\mu - \lambda \rho - \frac{\sigma^2}{2})(t_2 - t_1) + \sigma W_{t_1} (1 + \rho)^2}.$$

Inductively, we arrive at

$$S_{tk} = S_0 e^{(\mu - \frac{\sigma^2}{2})t_k + \sigma W_{tk} (1 + \rho)^k}.$$

This formula holds for any $t \in [t_k, t_{k+1})$; replacing $k$ by $N_t$ yields the desired formula (5.9.31).
In the following we shall show that (5.9.31) is a solution of the stochastic differential equation (5.9.30). If denote

\[ U_t = S_0e^{(\mu - \lambda \rho - \frac{\sigma^2}{2})t + \sigma W_t}, \quad V_t = (1 + \rho)^{N_t}, \]

we have \( S_t = U_t V_t \) and hence

\[ dS_t = V_t dU_t + U_t dV_t + dU_t dV_t. \]  

(5.9.32)

We already know that \( U_t \) verifies

\[ dU_t = (\mu - \lambda \rho)U_t dt + \sigma U_t dW_t \]

since \( U_t \) is a continuous process, i.e., \( U_t = U_{t-} \). Then the first term of (5.9.32) becomes

\[ V_t dU_t = (\mu - \lambda \rho)S_{t-} dt + \sigma S_{t-} dW_t. \]

In order to compute the second term of (5.9.32) we write

\[ dV_t = V_t - V_{t-} = (1 + \rho)^{N_t} - (1 + \rho)^{N_{t-}} = \begin{cases} (1 + \rho)^{1 + N_{t-}} - (1 + \rho)^{N_{t-}}, & \text{if } t = t_k \\ (1 + \rho)^{N_{t-}} - (1 + \rho)^{N_t}, & \text{if } t \neq t_k \end{cases} \]

\[ = \begin{cases} \rho(1 + \rho)^{N_{t-}}, & \text{if } t = t_k \\ 0, & \text{if } t \neq t_k \end{cases} = \rho(1 + \rho)^{N_{t-}} dN_t, \]

so \( U_t dV_t = \rho U_t(1 + \rho)^{N_{t-}} dN_t = \rho S_{t-} dN_t \). Since \( dt dN_t = dN_t dW_t = 0 \), the last term of (5.9.32) becomes \( dU_t dV_t = 0 \). Substituting back into (5.9.32) yields the equation

\[ dS_t = (\mu - \lambda \rho)S_{t-} dt + \sigma S_{t-} dW_t + \rho S_{t-} dN_t. \]

Formula (5.9.31) provides the stock price at time \( t \) if exactly \( N_t \) jumps have occurred and all jumps in the return of the stock are equal to \( \rho \).

It is worth noting that if \( \rho_k \) denotes the jump of the instantaneous return at time \( t_k \), a similar proof leads to the formula

\[ S_t = S_0e^{(\mu - \lambda \rho - \frac{\sigma^2}{2})t + \sigma W_t} \prod_{k=1}^{N_t} (1 + \rho_k), \]
where $\rho = E[\rho_k]$. The random variables $\rho_k$ are assumed independent and identically distributed. They are also independent of $W_t$ and $N_t$. For more details the reader is referred to Merton [14].

For the following exercises $S_t$ is given by (5.9.31).

**Exercise 5.9.2** Find $E[S_t]$ and $Var[S_t]$.

**Exercise 5.9.3** Find $E[\ln S_t]$ and $Var[\ln S_t]$.

**Exercise 5.9.4** Compute the conditional expectation $E[S_t | F_u]$ for $u < t$.

**Remark 5.9.5** Besides stock, the underlying asset of a derivative can be also a stock index, or foreign currency. When use the risk neutral valuation for derivatives on a stock index that pays a continuous dividend yield at a rate $q$, the drift rate $\mu$ is replace by $r - q$.

In the case of foreign currency that pays interest at the foreign interest rate $r_f$, the drift rate $\mu$ is replace by $r - r_f$. 

Chapter 6

Risk-Neutral Valuation

6.1 The Method of Risk-Neutral Valuation

This valuation method is based on the risk-neutral valuation principle, which states that the price of a derivative on an asset $S_t$ is not affected by the risk preference of the market participants; so we may assume they have the same risk aversion. In this case the valuation of the derivative price $f_t$ at time $t$ is done as in the following:

1. Assume the expected return of the asset $S_t$ is the risk-free rate, $\mu = r$.
2. Calculate the expected payoff of the derivative as of time $t$, under condition 1.
3. Discount at the risk-free rate from time $T$ to time $t$.

The first two steps require considering the expectation as of time $t$ in a risk-neutral world. This expectation is denoted by $\hat{E}_t[\cdot]$ and has the meaning of a conditional expectation given by $E[\cdot | F_t, \mu = r]$. The method states that if a derivative has the payoff $f_T$, its price at any time $t$ prior to maturity $T$ is given by

$$f_t = e^{-r(T-t)} \hat{E}_t[f_T].$$

The rate $r$ is considered constant, but the method can be easily adapted for time dependent rates.

In the following we shall present explicit computations for the most common European type\(^1\) derivative prices using the risk-neutral valuation method.

\(^1\)A derivative is of European type if can be exercised only at the expiration time.
6.2 Call Option

A call option is a contract which gives the buyer the right of buying the stock at time $T$ for the price $K$. The time $T$ is called maturity time or expiration date and $K$ is called the strike price. It is worth noting that a call option is a right (not an obligation!) of the buyer, which means that if the price of the stock at maturity, $S_T$, is less than the strike price, $K$, then the buyer may choose not to exercise the option. If the price $S_T$ exceeds the strike price $K$, then the buyer exercises the right to buy the stock, since he pays $K$ dollars for something which worth $S_T$ in the market. Buying at price $K$ and selling at $S_T$ yields a profit of $S_T - K$.

Consider a call option with maturity date $T$, strike price $K$ with the underlying stock price $S_t$ having constant volatility $\sigma > 0$. The payoff at maturity time is $f_T = \max(S_T - K, 0)$, see Fig.6.1 a. The price of the call at any prior time $0 \leq t \leq T$ is given by the expectation in a risk-neutral world

$$c(t) = \hat{E}_t[e^{-r(T-t)} f_T] = e^{-r(T-t)} \hat{E}_t[f_T].$$  \hspace{1cm} (6.2.1)

If we let $x = \ln(S_T/S_t)$, using the log-normality of the stock price in a risk-neutral world

$$S_T = S_te^{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma(W_T - W_t)},$$

and it follows that $x$ has the normal distribution

$$x \sim N\left((r - \frac{1}{2}\sigma^2)(T-t), \sigma^2(T-t)\right).$$

Then the density function of $x$ is

$$p(x) = \frac{1}{\sigma \sqrt{2\pi(T-t)}} e^{-\frac{(x - (r - \frac{1}{2}\sigma^2)(T-t))^2}{2\sigma^2(T-t)}}.$$

We can write the expectation as

$$\hat{E}_t[f_T] = \hat{E}_t[\max(S_T - K, 0)] = \hat{E}_t[\max(S_t e^x - K, 0)]$$

$$= \int_{-\infty}^{\ln(K/S_t)} \max(S_t e^x - K, 0) p(x) \, dx = \int_{\ln(K/S_t)}^{\infty} (S_t e^x - K) p(x) \, dx$$

$$= I_2 - I_1,$$ \hspace{1cm} (6.2.2)

with notations

$$I_1 = \int_{\ln(K/S_t)}^{\infty} K p(x) \, dx, \quad I_2 = \int_{\ln(K/S_t)}^{\infty} S_t e^x p(x) \, dx.$$
With the substitution \( y = \frac{x-(r+\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \), the first integral becomes

\[
I_1 = K \int_{\ln(K/S_t)}^{\infty} p(x) \, dx = K \int_{-d_2}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \, dy
\]

\[
= K \int_{-\infty}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \, dy = KN(d_2),
\]

where

\[
d_2 = \frac{\ln(S_t/K) + (r - \frac{\sigma^2}{2})(T - t)}{\sigma\sqrt{T - t}},
\]

and

\[
N(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-z^2/2} \, dz
\]

denotes the standard normal distribution function.

Using the aforementioned substitution the second integral can be computed by completing the square

\[
I_2 = S_t \int_{\ln(K/S_t)}^{\infty} e^y p(x) \, dx = S_t \int_{-d_2}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2 + y\sigma\sqrt{T-t} + (r - \frac{\sigma^2}{2})(T-t)} \, dy
\]

\[
= S_t \int_{-d_2}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y - \sigma\sqrt{T-t})^2} e^{r(T-t)} \, dy \quad \text{(let } z = y - \sigma\sqrt{T-t})
\]

\[
= S_t e^{r(T-t)} \int_{-d_2 - \sigma\sqrt{T-t}}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \, dz = S_t e^{r(T-t)} \int_{-\infty}^{d_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \, dz
\]

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

where

\[
d_1 = d_2 + \sigma\sqrt{T-t} = \frac{\ln(S_t/K) + (r + \frac{\sigma^2}{2})(T - t)}{\sigma\sqrt{T - t}}.
\]

Substituting back into (6.2.2) and then into (6.2.1) yields

\[
c(t) = e^{-r(T-t)}(I_2 - I_1) = e^{-r(T-t)}[S_t e^{r(T-t)} N(d_1) - KN(d_2)]
\]

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

We have obtained the well known formula of Black and Scholes:

**Proposition 6.2.1** The price of a European call option at time \( t \) is given by

\[
c(t) = S_t N(d_1) - K e^{-r(T-t)} N(d_2).
\]
Exercise 6.2.2 Show that
(a) \( \lim_{S_t \to \infty} \frac{c(t)}{S_t} = 1; \)
(b) \( \lim_{S_t \to 0} c(t) = 0; \)
(c) \( c(t) \sim S_t - Ke^{-r(T-t)} \) for \( S_t \) large.

Exercise 6.2.3 Show that \( \frac{dc(t)}{dS_t} = N(d_1). \) This expression is called the delta of a call option.

6.3 Cash-or-nothing Contract

A financial security that pays 1 dollar if the stock price \( S_T \geq K \) and 0 otherwise, is called a bet contract, or cash-or-nothing contract, see Fig.6.1 b. The payoff can be written as

\[
    f_T = \begin{cases} 
        1, & \text{if } S_T \geq K \\
        0, & \text{if } S_T < K.
    \end{cases}
\]

Substituting \( S_t = e^{X_t} \), the payoff becomes

\[
    f_T = \begin{cases} 
        1, & \text{if } X_T \geq \ln K \\
        0, & \text{if } X_T < \ln K,
    \end{cases}
\]

where \( X_T \) has the normal distribution

\[
    X_T \sim N \left( \ln S_t + (\mu - \frac{\sigma^2}{2})(T-t), \sigma^2(T-t) \right).
\]
The expectation in the risk-neutral world as of time \( t \) is

\[
\hat{E}_t[f_T] = E[f_T|\mathcal{F}_t, \mu = r] = \int_{-\infty}^{\infty} f_T(x)p(x) \, dx
\]

\[
= \int_{\ln K}^{+\infty} \frac{1}{\sqrt{2\pi} \sigma \sqrt{T - t}} e^{-\frac{(x - \ln S_t - (r - \frac{\sigma^2}{2})(T - t))^2}{2\sigma^2(T - t)}} \, dx
\]

\[
= \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dy = \int_{-\infty}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dy = N(d_2),
\]

where we used the substitution 

\[
y = \frac{x - \ln S_t - (r - \frac{\sigma^2}{2})(T - t)}{\sigma \sqrt{T - t}}
\]

and the notation

\[
d_2 = \frac{\ln S_t - \ln K + (r - \frac{\sigma^2}{2})(T - t)}{\sigma \sqrt{T - t}}.
\]

The price at time \( t \) of a bet contract is

\[
f_t = e^{-r(T-t)}\hat{E}_t[f_T] = e^{-r(T-t)}N(d_2). \tag{6.3.3}
\]

**Exercise 6.3.1** Let \( 0 < K_1 < K_2 \). Find the price of a financial derivative which pays at maturity $1 if \( K_1 \leq S_T \leq K_2 \) and zero otherwise, see Fig.6.2 a. This is a “box-bet” and its payoff is given by

\[
f_T = \begin{cases} 
1, & \text{if } K_1 \leq S_T \leq K_2 \\
0, & \text{otherwise}.
\end{cases}
\]

**Exercise 6.3.2** An asset-or-nothing contract pays \( S_T \) if \( S_T > K \) at maturity time \( T \), and pays 0 otherwise, see Fig.6.2 b. Show that the price of the contract at time \( t \) is \( f_t = S_t N(d_1) \).

**Exercise 6.3.3** (a) Find the price at time \( t \) of a derivative which pays at maturity

\[
f_T = \begin{cases} 
S^n_T, & \text{if } S_T \geq K \\
0, & \text{otherwise}.
\end{cases}
\]

(b) Show that the value of the contract can be written as \( f_t = g_t N(d_2 + n\sigma \sqrt{T - t}) \), where \( g_t \) is the value at time \( t \) of a power contract at time \( t \) given by (6.5.5).

(c) Recover the result of Exercise 6.3.2 in the case \( n = 1 \).
Figure 6.2: \( a \) The payoff of a box-bet; \( b \) The payoff of an asset-or-nothing contract.

### 6.4 Log-contract

A financial security that pays at maturity \( f_T = \ln S_T \) is called a log-contract. Since the stock is log-normally distributed,

\[
\ln S_T \sim N \left( \ln S_t + (\mu - \frac{\sigma^2}{2})(T - t), \sigma^2(T - t) \right),
\]

the risk-neutral expectation at time \( t \) is

\[
\hat{E}_t[f_T] = E[\ln S_T | F_t, \mu = r] = \ln S_t + (r - \frac{\sigma^2}{2})(T - t),
\]

and hence the price of the log-contract is given by

\[
f_t = e^{-r(T-t)} \hat{E}_t[f_T] = e^{-r(T-t)} \left( \ln S_t + (r - \frac{\sigma^2}{2})(T - t) \right).
\]

(6.4.4)

**Exercise 6.4.1** Find the price at time \( t \) of a square log-contract whose payoff is given by \( f_T = (\ln S_T)^2 \).

### 6.5 Power-contract

The financial derivative which pays at maturity the \( n \)th power of the stock price, \( S^n_T \), is called a power contract. Since \( S_T \) has a log-normal distribution with

\[
S_T = S_t e^{(\mu - \frac{\sigma^2}{2})(T-t) + \sigma(W_T - W_t)},
\]

the \( n \)th power of the stock, \( S^n_T \), is also log-normally distributed, with
\[ g_T = S_T^n = S_t^n e^{n \left( \mu - \frac{\sigma^2}{2} \right) (T-t)} + n \sigma (W_T - W_t). \]

Then the expectation at time \( t \) in the risk-neutral world is

\[
\hat{E}_t[g_T] = E[g_T | F_t, \mu = r] = S_t^n e^{n \left( \mu - \frac{\sigma^2}{2} \right) (T-t)} E[e^{n \sigma (W_T - W_t)} | F_t] = S_t^n e^{n \left( \mu - \frac{\sigma^2}{2} \right) (T-t)} E[e^{n \sigma W_{T-t}}] = S_t^n e^{n \left( \mu - \frac{\sigma^2}{2} \right) (T-t)} e^{\frac{1}{2} n^2 \sigma^2 (T-t)}.
\]

The price of the power-contract is obtained by discounting to time \( t \)

\[
g_t = e^{-r(T-t)} E[g_T | F_t, \mu = r] = S_t^n e^{-r(T-t)} e^{n \left( \mu - \frac{\sigma^2}{2} \right) (T-t)} e^{\frac{1}{2} n^2 \sigma^2 (T-t)} = S_t^n e^{(n-1)(r+\frac{\sigma^2}{2}) (T-t)}.
\]

Hence the value of a power contract at time \( t \) is given by

\[
g_t = S_t^n e^{(n-1)(r+\frac{\sigma^2}{2}) (T-t)}. \tag{6.5.5}
\]

It is worth noting that if \( n = 1 \), i.e. if the payoff is \( g_T = S_T \), then the price of the contract at any time \( t \leq T \) is \( g_t = S_t \), i.e. the stock price itself. This will be used shortly when valuing forward contracts.

In the case \( n = 2 \), i.e. if the contract pays \( S_T^2 \) at maturity, then the price is \( g_t = S_t^2 e^{(r+\sigma^2)(T-t)} \).

**Exercise 6.5.1** Let \( n \geq 1 \) be an integer. Find the price of a power call option whose payoff is given by \( f_T = \max (S_T^n - K, 0) \).

### 6.6 Forward Contract on the Stock

A forward contract pays at maturity the difference between the stock price, \( S_T \), and the delivery price of the asset, \( K \). The price at time \( t \) is

\[
f_t = e^{-r(T-t)} \hat{E}_t[S_T - K] = e^{-r(T-t)} \hat{E}_t[S_T] - e^{-r(T-t)} K = S_t - e^{-r(T-t)} K,
\]

where we used that \( K \) is a constant and that \( \hat{E}_t[S_T] = S_t e^{r(T-t)} \).

**Exercise 6.6.1** Let \( n \in \{2, 3\} \). Find the price of the power forward contract that pays at maturity \( f_T = (S_T - K)^n \).
6.7 The Superposition Principle

If the payoff of a derivative, \( f_T \), can be written as a linear combination of payoffs

\[
f_T = \sum_{i=1}^{n} c_i h_{i,T}
\]

with \( c_i \) constants, then the price at time \( t \) is given by

\[
f_t = \sum_{i=1}^{n} c_i h_{i,t}
\]

where \( h_{i,t} \) is the price at time \( t \) of a derivative that pays at maturity \( h_{i,T} \). We shall successfully use this method in the situation when the payoff \( f_T \) can be decomposed into simpler payoffs, for which we can evaluate directly the price of the associate derivative. In this case the price of the initial derivative, \( f_t \), is obtained as a combination of the prices of the easier to valuate derivatives.

The reason underlying the aforementioned superposition principle is the linearity of the expectation operator \( \hat{E} \)

\[
f_t = e^{-r(T-t)} \hat{E}[f_T] = e^{-r(T-t)} \hat{E}\left[\sum_{i=1}^{n} c_i h_{i,T}\right]
\]

\[
= e^{-r(T-t)} \sum_{i=1}^{n} c_i \hat{E}[h_{i,T}] = \sum_{i=1}^{n} c_i h_{i,t}.
\]

This principle is also connected with the absence of arbitrage opportunities\(^2\) in the market. Consider two portfolios of derivatives with equal values at the maturity time \( T \)

\[
\sum_{i=1}^{n} c_i h_{i,T} = \sum_{j=1}^{m} a_j g_{j,T}.
\]

If we take this common value to be the payoff of a derivative, \( f_T \), then by the aforementioned principle, the portfolios have the same value at any prior time \( t \leq T \)

\[
\sum_{i=1}^{n} c_i h_{i,t} = \sum_{j=1}^{m} a_j g_{j,t}.
\]

The last identity can also result from the absence of arbitrage opportunities in the market. If there is a time \( t \) at which the identity fails, then buying the

---

\(^2\) An arbitrage opportunity deals with the practice of making profits by taking simultaneous long and short positions in the market.
cheaper portfolio and selling the more expensive one will lead to an arbitrage profit.

The superposition principle can be used to price package derivatives such as spreads, straddles, strips, straps and strangles. We shall deal with these type of derivatives in proposed exercises.

6.8 General Contract on the Stock

By a general contract on the stock we mean a derivative that pays at maturity the amount \( g_T = G(S_T) \), where \( G \) is a given analytic function. In the case \( G(S_T) = S_T^n \), we have a power contract. If \( G(S_T) = \ln S_T \), we have a log-contract. If choose \( G(S_T) = S_T - K \), we obtain a forward contract. Since \( G \) is analytic we shall write \( G(x) = \sum_{n \geq 0} c_n x^n \), where \( c_n = G^{(n)}(0)/n! \).

Decomposing into power contracts and using the superposition principle we have

\[
\hat{E}_t [g_T] = \hat{E}_t [G(S_T)] = \hat{E}_t [\sum_{n \geq 0} c_n S_T^n] \\
= \sum_{n \geq 0} c_n \hat{E}_t [S_T^n] \\
= \sum_{n \geq 0} c_n S_t^n e^{n(r - \frac{\sigma^2}{2})(T-t)} \frac{1}{2} \sigma^2 (T-t) \\
= \sum_{n \geq 0} c_n \left[ S_t e^{n(r - \frac{\sigma^2}{2})(T-t)} \right]^n \left[ \frac{1}{2} \sigma^2 (T-t) \right] n^2.
\]

Hence the value at time \( t \) of a general contract is

\[
g_t = e^{r(T-t)} \sum_{n \geq 0} c_n \left[ S_t e^{n(r - \frac{\sigma^2}{2})(T-t)} \right]^n \left[ \frac{1}{2} \sigma^2 (T-t) \right] n^2.
\]

Exercise 6.8.1 Find the value at time \( t \) of an exponential contract that pays at maturity the amount \( e^{S_T} \).

6.9 Call Option

In the following we price a European call using the superposition principle. The payoff of a call option can be decomposed as

\[
c_T = \max(S_T - K, 0) = h_{1,T} - K h_{2,T},
\]
with
\[ h_{1,T} = \begin{cases} S_T, & \text{if } S_T \geq K \\ 0, & \text{if } S_T < K \end{cases}, \quad h_{2,T} = \begin{cases} 1, & \text{if } S_T \geq K \\ 0, & \text{if } S_T < K \end{cases} \]

These are the payoffs of asset-or-nothing and of cash-or-nothing derivatives. From section 6.3 and Exercise 6.3.2 we have \( h_{1,t} = S_t N(d_1), h_{2,t} = e^{-r(T-t)} N(d_2) \).

By superposition we get the price of a call at time \( t \)
\[
c_t = h_{1,t} - Kh_{2,t} = S_t N(d_1) - Ke^{-r(T-t)} N(d_2). \]

Exercise 6.9.1 (Put option) (a) Consider the payoff \( h_{1,T} = \begin{cases} 1, & \text{if } S_T \leq K \\ 0, & \text{if } S_T > K \end{cases} \). Show that
\[ h_{1,t} = e^{-r(T-t)} N(-d_2), \quad t \leq T. \]

(b) Consider the payoff \( h_{2,T} = \begin{cases} S_T, & \text{if } S_T \leq K \\ 0, & \text{if } S_T > K \end{cases} \). Show that
\[ h_{2,t} = S_t N(-d_1), \quad t \leq T. \]

(c) The payoff of a put is \( p_T = \max(K - S_T, 0) \). Verify that
\[ p_T = Kh_{1,T} - h_{2,T} \]
and use the superposition principle to find the price \( p_t \) of a put.

6.10 General Options on the Stock

Let \( G \) be an increasing analytic function with the inverse \( G^{-1} \) and \( K \) be a positive constant. A general call option is a contract with the payoff
\[
f_T = \begin{cases} G(S_T) - K, & \text{if } S_T \geq G^{-1}(K) \\ 0, & \text{otherwise} \end{cases} \quad (6.10.6)\]

We note the payoff function \( f_T \) is continuous. We shall work out the value of the contract at time \( t \), \( f_t \), using the superposition method. Since the payoff can be written as the linear combination
\[ f_T = h_{1,T} - Kh_{2,T}, \]
with
\[ h_{1,T} = \begin{cases} G(S_T), & \text{if } S_T \geq G^{-1}(K) \\ 0, & \text{otherwise} \end{cases}, \quad h_{2,T} = \begin{cases} 1, & \text{if } S_T \geq G^{-1}(K) \\ 0, & \text{otherwise} \end{cases}, \]
then
\[ f_t = h_{1,t} - Kh_{2,t}. \quad (6.10.7) \]
We had already computed the value \( h_{2,t} \). In this case we have \( h_{2,t} = e^{-r(T-t)} N(d_2^G) \), where \( d_2^G \) is obtained by replacing \( K \) with \( G^{-1}(K) \) in the formula of \( d_2 \)

\[
d_2^G = \frac{\ln S_t - \ln (G^{-1}(K)) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}.
\]  

(6.10.8)

We shall compute in the following \( h_{1,t} \). Let \( G(S_T) = \sum_{n \geq 0} c_n S_T^n \) with \( c_n = G^{(n)}(0)/n! \). Then \( h_{1,T} = \sum_{n \geq 0} c_n f_{T}^{(n)} \), where

\[
f_{T}^{(n)} = \begin{cases} 
S_T^n, & \text{if } S_T \geq G^{-1}(K) \\
0, & \text{otherwise}
\end{cases}
\]

By Exercise 6.3.3 the price at time \( t \) for a contract with the payoff \( f_{T}^{(n)} \) is

\[
f_t^{(n)} = S_t^n e^{(n-1)(r+\frac{\sigma^2}{2})(T-t)} N(d_2^G + n\sigma\sqrt{T-t}).
\]

The value at time \( t \) of \( h_{1,T} \) is given by

\[
h_{1,t} = \sum_{n \geq 0} c_n f_{T}^{(n)} = \sum_{n \geq 0} c_n S_T^n e^{(n-1)(r+\frac{\sigma^2}{2})(T-t)} N(d_2^G + n\sigma\sqrt{T-t}).
\]

Substituting in (6.10.7) we obtain the following value at time \( t \) of a general call option with the payoff (6.10.6)

\[
f_t = \sum_{n \geq 0} c_n S_t^n e^{(n-1)(r+\frac{\sigma^2}{2})(T-t)} N(d_2^G + n\sigma\sqrt{T-t}) - Ke^{-r(T-t)} N(d_2^G),
\]

(6.10.9)

with \( c_n = \frac{G^{(n)}}{n!} \) and \( d_2^G \) given by (6.10.8).

It is worth noting that in the case \( G(S_T) = S_T \) we have \( n = 1 \), \( d_2 = d_2^G \), \( d_1 = d_2 + \sigma(T-t) \) formula (6.10.9) becomes the value of a plain vanilla call option

\[
f_t = S_t N(d_1) - Ke^{-r(T-t)}.
\]

### 6.11 Packages

*Packages* are derivatives whose payoffs are linear combinations of payoffs of options, cash and underlying asset. They can be priced using the superposition principle. Some of these packages are used in hedging techniques.
The Bull Spread Let $0 < K_1 < K_2$. A derivative with the payoff

$$f_T = \begin{cases} 
0, & \text{if } S_T \leq K_1 \\
S_T - K_1, & \text{if } K_1 < S_T \leq K_2 \\
K_2 - K_1, & \text{if } K_2 < S_T 
\end{cases}$$

is called a bull spread, see Fig.6.3 a. A market participant enters a bull spread position when the stock price is expected to increase. The payoff $f_T$ can be written as the difference of the payoffs of two calls with strike prices $K_1$ and $K_2$:

$$f_T = c_1(T) - c_2(T),$$

$$c_1(T) = \begin{cases} 
0, & \text{if } S_T \leq K_1 \\
S_T - K_1, & \text{if } K_1 < S_T \leq K_2 
\end{cases}, \quad c_2(T) = \begin{cases} 
0, & \text{if } S_T \leq K_2 \\
S_T - K_2, & \text{if } K_2 < S_T \end{cases}$$

Using the superposition principle, the price of a bull spread at time $t$ is

$$f_t = c_1(t) - c_2(t)$$

$$= S_t N(d_1(K_1)) - K_1 e^{-r(T-t)} N(d_2(K_1)) - \left( S_t N(d_1(K_2)) - K_2 e^{-r(T-t)} N(d_2(K_2)) \right)$$

$$= S_t [N(d_1(K_1)) - N(d_1(K_2))] - e^{-r(T-t)} \left[ N(d_2(K_1)) - N(d_2(K_2)) \right],$$

with

$$d_2(K_i) = \frac{\ln S_t - \ln K_i + (r - \frac{\sigma^2}{2})(T-t)}{\sigma \sqrt{T-t}}, \quad d_1(K_i) = d_2(K_i) + \sigma \sqrt{T-t}, \quad i = 1, 2.$$

The Bear Spread Let $0 < K_1 < K_2$. A derivative with the payoff

$$f_T = \begin{cases} 
K_2 - K_1, & \text{if } S_T \leq K_1 \\
K_2 - S_T, & \text{if } K_1 < S_T \leq K_2 \\
0, & \text{if } K_2 < S_T 
\end{cases}$$

Figure 6.3: a The payoff of a bull spread; b The payoff of a bear spread.
a

\begin{align*}
&f_T = \begin{cases} 
0, & \text{if } S_T \leq K_1 \\
S_T - K_1, & \text{if } K_1 < S_T \leq K_2 \\
K_3 - S_T, & \text{if } K_2 \leq S_T < K_3 \\
0, & \text{if } K_3 \leq S_T.
\end{cases}
\end{align*}

A short position in a butterfly spread leads to profits when a small move in the stock price occurs, see Fig.6.4 a.

Exercise 6.11.2 Find the price of a butterfly spread at time $t$, with $t < T$.

Straddles A derivative with the payoff $f_T = |S_T - K|$ is called a straddle, see see Fig.6.4 b. A long position in a straddle leads to profits when a move in any direction of the stock price occurs.

Exercise 6.11.3 (a) Show that the payoff of a straddle can be written as

\begin{align*}
&f_T = \begin{cases} 
K - S_T, & \text{if } S_T \leq K \\
S_T - K, & \text{if } K < S_T.
\end{cases}
\end{align*}

(b) Find the price of a straddle at time $t$, with $t < T$. 

Figure 6.4: a The payoff of a butterfly spread; b The payoff of a straddle.
Strangles Let $0 < K_1 < K_2$ and $K = (K_2 + K_1)/2$, $K' = (K_2 - K_1)/2$. A derivative with the payoff $f_T = \max(|S_T - K| - K', 0)$ is called a strangle, see Fig. 6.5 a. A long position in a strangle leads to profits when a small move in any direction of the stock price occurs.

Exercise 6.11.4 (a) Show that the payoff of a strangle can be written as

$$f_T = \begin{cases} K_1 - S_T, & \text{if } S_T \leq K_1 \\ 0, & \text{if } K_1 < S_T \leq K_2 \\ S_T - K_2, & \text{if } K_2 \leq S_T. \end{cases}$$

(b) Find the price of a strangle at time $t$, with $t < T$;
(c) Which is cheaper: a straddle or a strangle?

Exercise 6.11.5 Let $0 < K_1 < K_2 < K_3 < K_4$, with $K_4 - K_3 = K_2 - K_1$. Find the value at time $t$ of a derivative with the payoff $f_T$ given in Fig. 6.5 b.

6.12 Asian Forward Contracts

Forward Contracts on the Arithmetic Average Let $A_T = \frac{1}{T} \int_0^T S_u \, du$ denote the continuous arithmetic average of the asset price between 0 and $T$. It sometimes makes sense for two parties to make a contract in which one party pays the other at maturity time $T$ the difference between the average price of the asset, $A_T$, and the fixed delivery price $K$. The payoff of a forward contract on arithmetic average is

$$f_T = A_T - K.$$
For instance, if the asset is natural gas, it makes sense to make a deal on the average price of the asset, since the price is volatile and can become expensive during the winter season.

Since the risk-neutral expectation at time $t = 0$, $\hat{E}_0[A_T]$, is given by $E[A_T]$ where $\mu$ is replaced by $r$, the price of the forward contract at $t = 0$ is

$$f_0 = e^{-rT} \hat{E}_0[f_T] = e^{-rT}(\hat{E}_0[A_T] - K)$$

$$= e^{-rT}\left(S_0 \frac{e^{rT} - 1}{rT} - K\right) = S_0 \frac{1 - e^{-rT}}{rT} - e^{-rT}K$$

$$= \frac{S_0}{rT} - e^{-rT}\left(\frac{S_0}{rT} + K\right).$$

Hence the value of a forward contract on the arithmetic average at time $t = 0$ is given by

$$f_0 = \frac{S_0}{rT} - e^{-rT}\left(\frac{S_0}{rT} + K\right).$$ (6.12.10)

It is worth noting that the price of a forward contract on the arithmetic average is cheaper than the price of a usual forward contract on the asset. To see that, we substitute $x = rT$ in the inequality $e^{-x} > 1 - x$, $x > 0$, to get

$$\frac{1 - e^{-rT}}{rT} < 1.$$

This implies the inequality

$$S_0 \frac{1 - e^{-rT}}{rT} - e^{-rT}K < - e^{-rT}K.$$

Since the left side is the price of an Asian forward contract on the arithmetic average, while the right side is the price of a usual forward contract, we obtain the desired inequality.

Formula (6.12.10) provides the price of the contract at time $t = 0$. What is the price at any time $t < T$? One might be tempted to say that replacing $T$ by $T - t$ and $S_0$ by $S_t$ in the formula of $f_0$ leads to the corresponding formula for $f_t$. However, this does not hold true, as the next result shows:

**Proposition 6.12.1** The value at time $t$ of a contract that pays at maturity $f_T = A_T - K$ is given by

$$f_t = e^{-r(T-t)}\left(\frac{t}{T}A_t - K\right) + \frac{1}{rT}S_t\left(1 - e^{-r(T-t)}\right).$$ (6.12.11)
Proof: We start by computing the risk-neutral expectation \( \hat{E}_t[A_T] \). Splitting the integral into a predictable and an unpredictable part, we have

\[
\hat{E}_t[A_T] = \hat{E}
\left[
\frac{1}{T} \int_0^T S_u du | F_t
\right]
\]

\[
= \hat{E}
\left[
\frac{1}{T} \int_0^t S_u du + \frac{1}{T} \int_t^T S_u du | F_t
\right]
\]

\[
= \frac{1}{T} \int_0^t S_u du + \hat{E}
\left[
\frac{1}{T} \int_t^T S_u du | F_t
\right]
\]

\[
= \frac{1}{T} \int_0^t S_u du + \frac{1}{T} \int_t^T \hat{E}[S_u | F_t] du
\]

\[
= \frac{1}{T} \int_0^t S_u du + \frac{1}{T} \int_t^T S_t e^{r(u-t)} du,
\]

where we replaced \( \mu \) by \( r \) in the formula \( \hat{E}[S_u | F_t] = S_t e^{\mu(u-t)} \). Integrating we obtain

\[
\hat{E}_t[A_T] = \frac{1}{T} \int_0^t S_u du + \frac{1}{T} \int_t^T S_t e^{r(u-t)} du
\]

\[
= \frac{t}{T} A_t + \frac{1}{T} S_t e^{-rt} \frac{e^{rt} - e^{rT}}{r}
\]

\[
= \frac{t}{T} A_t + \frac{1}{rT} S_t \left( e^{r(T-t)} - 1 \right).
\]

(6.12.12)

Using (6.12.12), the risk-neutral valuation provides

\[
f_t = e^{-r(T-t)} \hat{E}_t[A_T - K] = e^{-r(T-t)} \hat{E}_t[A_T] - e^{-r(T-t)} K
\]

\[
= e^{-r(T-t)} \left( \frac{t}{T} A_t - K \right) + \frac{1}{rT} S_t \left( 1 - e^{-r(T-t)} \right).
\]

Exercise 6.12.2 Show that \( \lim_{t \to 0} f_t = f_0 \), where \( f_t \) is given by (6.12.11) and \( f_0 \) by (6.12.10).

Exercise 6.12.3 Find the value at time \( t \) of a contract that pays at maturity date the difference between the asset price and its arithmetic average, \( f_T = S_T - A_T \).

Forward Contracts on the Geometric Average We shall consider in the following Asian forward contracts on the geometric average. This is a derivative that pays at maturity the difference \( f_T = G_T - K \), where \( G_T \) is the
continuous geometric average of the asset price between 0 and T and K is a
fixed delivery price.

We shall work out first the value of the contract at t = 0. Substituting µ = r in the first relation provided by Theorem 5.8.4, the risk-neutral expectation of GT as of time t = 0 is

$$\hat{E}_0[G_T] = S_0e^{\frac{1}{T}(r - \frac{\sigma^2}{2})T}.$$ 

Then

$$f_0 = e^{-rT}\hat{E}_0[G_T - K] = e^{-rT}\hat{E}_0[G_T] - e^{-rT}K$$

$$= e^{-rT}S_0e^{\frac{1}{2}(r - \frac{\sigma^2}{2})T} - e^{-rT}K$$

$$= S_0e^{-\frac{1}{2}(r + \frac{\sigma^2}{2})T} - e^{-rT}K.$$ 

Thus, the price of a forward contract on geometric average at t = 0 is given by

$$f_0 = S_0e^{-\frac{1}{2}(r + \frac{\sigma^2}{2})T} - e^{-rT}K.$$ (6.12.13)

As in the case of forward contracts on arithmetic average, the value at time 0 < t < T cannot be obtain from (6.12.13) by replacing blindly T and S0 by T − t and St, respectively. The correct relation is given by the following result:

**Proposition 6.12.4** The value at time t of a contract which pays at maturity GT − K is

$$f_t = \frac{G_T}{S_t}S_t^{1 - \frac{1}{T}T}e^{-r(T-t)+\frac{\sigma^2}{2}(T-t)^2+\frac{\sigma^2}{6}(T-t)^3} - e^{-r(T-t)}K.$$ (6.12.14)

**Proof:** Since for t < u

$$\ln S_u = \ln S_t + (\mu - \frac{\sigma^2}{2})(u - t) + \sigma(W_u - W_t),$$

we have

$$\int_0^T \ln S_u \ du = \int_0^t \ln S_u \ du + \int_t^T \ln S_u \ du$$

$$= \int_0^t \ln S_u \ du + \int_t^T \left( \ln S_t + (\mu - \frac{\sigma^2}{2})(u - t) + \sigma(W_u - W_t) \right) \ du$$

$$= \int_0^t \ln S_u \ du + (T - t) \ln S_t + (\mu - \frac{\sigma^2}{2}) \left( \frac{T^2}{2} - \frac{t^2}{2} - t(T - t) \right)$$

$$+ \sigma \int_t^T (W_u - W_t) \ du.$$
The geometric average becomes

\[ G_T = e^\frac{1}{T} \int_0^T \ln S_u \, du \]

\[ = e^\frac{1}{T} \int_0^T \ln S_u \, du \left( S_t^1 - \frac{1}{T} (\mu - \sigma^2) (T - t) e^\frac{1}{T} \int_0^T (W_u - W_t) \, du \right) \]

\[ = \Gamma_t^\mu \left( S_t^1 - \frac{1}{T} \left( \mu - \frac{\sigma^2}{2} \right) (T - t) \right) e^\frac{1}{T} \int_0^T (W_u - W_t) \, du \]

(6.12.15)

where we used that

\[ e^\frac{1}{T} \int_0^T \ln S_u \, du = e^{\frac{1}{T} \ln G_t} = G_t^\frac{\mu}{T}. \]

Relation (6.12.15) provides \( G_T \) in terms of \( G_t \) and \( S_t \). Taking the predictable part out and replacing \( \mu \) by \( r \) we have

\[ \hat{E}_t[G_T] = \Gamma_t^\mu \left( S_t^1 - \frac{1}{T} \left( \mu - \frac{\sigma^2}{2} \right) (T - t) \right) e^\frac{1}{T} \int_0^T (W_u - W_t) \, du \]

(6.12.16)

Since the jump \( W_u - W_t \) is independent of the information set \( \mathcal{F}_t \), the condition can be dropped

\[ E\left[ e^\frac{\sigma}{T} \int_0^T (W_u - W_t) \, du \mid \mathcal{F}_t \right] = E\left[ e^\frac{\sigma}{T} \int_0^T (W_u - W_t) \, du \right]. \]

Integrating by parts yields

\[ \int_t^T (W_u - W_t) \, du = \int_t^T W_u \, du - (T - t)W_t \]

\[ = TW_T - tW_t - \int_t^T u \, dW_u - TW_t + tW_t \]

\[ = T(W_T - W_t) - \int_t^T u \, dW_u = \int_t^T T \, dW_u - \int_t^T u \, dW_u \]

\[ = \int_t^T (T - u) \, dW_u, \]

which is a Wiener integral. This is normally distributed with mean 0 and variance

\[ \int_t^T (T - u)^2 \, du = \frac{(T - t)^3}{3}. \]

Then \( \int_t^T (W_u - W_t) \, du \sim N\left(0, \frac{(T - t)^3}{3}\right) \) and hence

\[ E\left[ e^\frac{\sigma}{T} \int_0^T (W_u - W_t) \, du \right] = E\left[ e^\frac{\sigma}{T} \int_0^T (T - u) \, dW_u \right] = e^{\frac{1}{2} \frac{\sigma^2}{T^2} (T - t)^3} = e^{\frac{\sigma^2}{T} \frac{(T - t)^3}{6}}. \]

Substituting into (6.12.16) yields

\[ \hat{E}_t[G_T] = \Gamma_t^\mu \left( S_t^1 - \frac{1}{T} \left( \mu - \frac{\sigma^2}{2} \right) (T - t) \right) e^{\frac{1}{2} \frac{\sigma^2}{T^2} (T - t)^3} e^{\frac{\sigma^2}{T} \frac{(T - t)^3}{6}}. \]
Hence the value of the contract at time $t$ is given by

$$f_t = e^{-r(T-t)} \hat{E}_t[G_T - K] = G_t^T S_t^{1-\frac{1}{2}} e^{-r(T-t) + \left(\frac{1}{2} \sigma^2 (T-t)^2 + \frac{1}{2} \sigma^2 (T-t)^3\right)} - e^{-r(T-t)} K.$$

**Exercise 6.12.5** Show that $\lim_{t \to 0} f_t = f_0$, where $f_t$ is given by (6.12.14) and $f_0$ by (6.12.13).

**Exercise 6.12.6** Which is cheaper: an Asian forward contract on $A_t$ or an Asian forward contract on $G_t$?

**Exercise 6.12.7** Using Corollary 5.8.5 find a formula of $G_T$ in terms of $G_t$, and then compute the risk-neutral world expectation $\hat{E}_t[G_T]$.

### 6.13 Asian Options

There are several types of Asian options depending on how the payoff is related to the average stock price:

- **Average Price options**:
  - Call: $f_T = \max(S_{\text{ave}} - K, 0)$
  - Put: $f_T = \max(K - S_{\text{ave}}, 0)$.

- **Average Strike options**:
  - Call: $f_T = \max(S_T - S_{\text{ave}}, 0)$
  - Put: $f_T = \max(S_{\text{ave}} - S_T, 0)$.

The average asset price $S_{\text{ave}}$ can be either the arithmetic or the geometric average of the asset price between 0 and $T$.

**Geometric average price options** When the asset is the geometric average, $G_T$, we shall obtain closed form formulas for average price options. Since $G_T$ is log-normally distributed, the pricing procedure is similar with the one used for the usual options on stock. We shall do this by using the superposition principle and the following two results. The first one is a cash-or-nothing type contract where the underlying asset is the geometric mean of the stock between 0 and $T$. 
Lemma 6.13.1 The value at time \( t = 0 \) of a derivative, which pays at maturity \( \$1 \) if the geometric average \( G_T \geq K \) and 0 otherwise, is given by

\[
h_0 = e^{-rT} N(\tilde{d}_2),
\]

where

\[
\tilde{d}_2 = \frac{\ln S_0 - \ln K + (\mu - \frac{\sigma^2}{2}) T}{\sigma \sqrt{T/3}}.
\]

Proof: The payoff can be written as

\[
h_T = \begin{cases} 1, & \text{if } G_T \geq K \\ 0, & \text{if } G_T < K \end{cases} = \begin{cases} 1, & \text{if } X_T \geq \ln K \\ 0, & \text{if } X_T < \ln K, \end{cases}
\]

where \( X_T = \ln G_T \) has the normal distribution

\[
X_T \sim N \left[ \ln S_0 + (\mu - \frac{\sigma^2}{2}) \frac{T}{2}, \frac{\sigma^2 T}{3} \right],
\]

see formula (5.8.27). Let \( p(x) \) be the probability density of the random variable \( X_T \)

\[
p(x) = \frac{1}{\sqrt{2\pi}\sigma\sqrt{T/3}} e^{-\left[ x - \ln S_0 - (\mu - \frac{\sigma^2}{2}) \frac{T}{2} \right]^2 / \left( 2\sigma^2 T / 3 \right)}.
\]  

(6.13.18)

The risk neutral expectation of the payoff at time \( t = 0 \) is

\[
\hat{E}_0[h_T] = \int h_T(x) p(x) \, dx = \int_{\ln K}^{\infty} p(x) \, dx
\]

\[
= \frac{1}{\sqrt{2\pi}\sigma\sqrt{T/3}} \int_{\ln K}^{\infty} e^{-\left[ x - \ln S_0 - (\mu - \frac{\sigma^2}{2}) \frac{T}{2} \right]^2 / \left( 2\sigma^2 T / 3 \right)} \, dx,
\]

where \( \mu \) was replaced by \( r \). Substituting

\[
y = \frac{x - \ln S_0 - (r - \frac{\sigma^2}{2}) \frac{T}{2}}{\sigma \sqrt{T/3}},
\]  

(6.13.19)

yields

\[
\hat{E}_0[h_T] = \frac{1}{\sqrt{2\pi}} \int_{-\tilde{d}_2}^{\infty} e^{-y^2/2} \, dy = \frac{1}{\sqrt{2\pi}} \int_{-\tilde{d}_2}^{\infty} e^{-y^2/2} \, dy
\]

\[
= N(\tilde{d}_2),
\]

where

\[
\tilde{d}_2 = \frac{\ln S_0 - \ln K + (r - \frac{\sigma^2}{2}) \frac{T}{2}}{\sigma \sqrt{T/3}}.
\]
Discounting to the free interest rate yields the price at time $t=0$
\[ h_0 = e^{-rT} \hat{E}_0[h_T] = e^{-rT} N(\tilde{d}_2). \]

The following result deals with the price of an average-or-nothing derivative on the geometric average.

**Lemma 6.13.2** The value at time $t=0$ of a derivative, which pays at maturity $G_T$ if $G_T \geq K$ and $0$ otherwise, is given by the formula
\[ g_0 = e^{-\frac{1}{2}(r+\frac{\sigma^2}{2})T} S_t N(\bar{d}_1), \]
where
\[ \bar{d}_1 = \frac{\ln S_0 - \ln K + (r + \frac{\sigma^2}{2})T}{\sigma \sqrt{T/3}}. \]

**Proof:** Since the payoff can be written as
\[ g_T = \begin{cases} G_T, & \text{if } G_T \geq K \\ 0, & \text{if } G_T < K \end{cases} = \begin{cases} e^{X_T}, & \text{if } X_T \geq \ln K \\ 0, & \text{if } X_T < \ln K, \end{cases} \]
with $X_T = \ln G_T$, the risk neutral expectation at the time $t=0$ of the payoff is
\[ \hat{E}_0[g_T] = \int_{-\infty}^{\infty} g_T(x) p(x) \, dx = \int_{\ln K}^{\infty} e^{x} p(x) \, dx, \]
where $p(x)$ is given by (6.13.18), with $\mu$ replaced by $r$. Using the substitution (6.13.19) and completing the square yields
\[ \hat{E}_0[g_T] = \frac{1}{\sqrt{2\pi} \sigma \sqrt{T/3}} \int_{\ln K}^{\infty} e^{x} e^{-\frac{1}{2}[x-\ln S_0 - (r - \frac{\sigma^2}{2})T/2]^{2}/(2\sigma^2 T)} \, dx \]
\[ = \frac{1}{\sqrt{2\pi}} S_0 e^{\frac{1}{2}(r - \frac{\sigma^2}{2})T} \int_{-\bar{d}_2}^{\infty} e^{-\frac{1}{2}(y-\sigma \sqrt{T/3})^2} \, dy \]
If we let
\[ \bar{d}_1 = \bar{d}_2 + \sigma \sqrt{T/3} \]
\[ = \frac{\ln S_0 - \ln K + (r + \frac{\sigma^2}{2})T/2}{\sigma \sqrt{T/3}} + \sigma \sqrt{T/3} \]
\[ = \frac{\ln S_0 - \ln K + (r + \frac{\sigma^2}{6})T}{\sigma \sqrt{T/3}}, \]
the previous integral becomes, after substituting \( z = y - \sigma \sqrt{T/3} \),
\[
\frac{1}{\sqrt{2\pi}} S_0 e^{\frac{1}{2}(r - \frac{\sigma^2}{6})T} \int_{-\tilde{d}_1}^{\infty} e^{-\frac{1}{2}z^2} \, dz = S_0 e^{\frac{1}{2}(r - \frac{\sigma^2}{6})T} N(\tilde{d}_1).
\]
Then the risk neutral expectation of the payoff is
\[
\tilde{E}_0[g_T] = S_0 e^{\frac{1}{2}(r - \frac{\sigma^2}{6})T} N(\tilde{d}_1).
\]
The value of the derivative at time \( t = 0 \) is obtained by discounting at the interest rate \( r \)
\[
g_0 = e^{-rT} \tilde{E}_0[g_T] = e^{-rT} S_0 e^{\frac{1}{2}(r - \frac{\sigma^2}{6})T} N(\tilde{d}_1)
\]
\[
= e^{-\frac{1}{2}(r + \frac{\sigma^2}{6})T} S_0 N(\tilde{d}_1).
\]

**Proposition 6.13.3** The value at time \( t \) of a geometric average price call option is
\[
f_0 = e^{-\frac{1}{2}(r + \frac{\sigma^2}{6})T} S_0 N(\tilde{d}_1) - Ke^{-rT} N(\tilde{d}_2).
\]

**Proof:** Since the payoff \( f_T = \max(G_T - K, 0) \) can be decomposed as
\[
f_T = g_T - Kh_T,
\]
with
\[
g_T = \begin{cases} G_T, & \text{if } G_T \geq K \\ 0, & \text{if } G_T < K \end{cases} \quad h_T = \begin{cases} 1, & \text{if } G_T \geq K \\ 0, & \text{if } G_T < K \end{cases}
\]
applying the superposition principle and Lemmas 6.13.1 and 6.13.2 yields
\[
f_0 = g_0 - Kh_0
\]
\[
= e^{-\frac{1}{2}(r + \frac{\sigma^2}{6})T} S_0 N(\tilde{d}_1) - Ke^{-rT} N(\tilde{d}_2).
\]

**Exercise 6.13.4** Find the value at time \( t = 0 \) of a price put option on a geometric average, i.e. a derivative with the payoff \( f_T = \max(K - G_T, 0) \).

**Arithmetic average price options** There is no simple closed-form solution for a call or for a put on the arithmetic average \( A_t \). However, there is an approximate solution based on computing exactly the first two moments of the distribution of \( A_t \), and applying the risk-neutral valuation assuming that
the distribution is log-normally with the same two moments. This idea was developed by Turnbull and Wakeman [18], and works pretty well for volatilities up to about 20%.

The following result provides the mean and variance of a normal distribution in terms of the first two moments of the associated log-normally distribution.

**Proposition 6.13.5** Let $Y$ be a log-normally distributed random variable, having the first two moments given by

$$ m_1 = E[Y], \quad m_2 = E[Y^2]. $$

Then $\ln Y$ has the normal distribution $\ln Y \sim N(\mu, \sigma^2)$, with the mean and variance given respectively by

$$ \mu = \ln \frac{m_1^2}{\sqrt{m_2}}, \quad \sigma^2 = \ln \frac{m_2}{m_1^2}. \tag{6.13.20} $$

**Proof:** Using Exercise ?? we have

$$ m_1 = E[Y] = e^{\mu + \frac{\sigma^2}{2}}, \quad m_2 = E[Y^2] = e^{2\mu + 2\sigma^2}. $$

Taking a logarithm yields

$$ \mu + \frac{\sigma^2}{2} = \ln m_1, \quad 2\mu + 2\sigma^2 = \ln m_2. $$

Solving for $\mu$ and $\sigma$ yields (6.13.20). \hfill \blacksquare

Assume the arithmetic average $A_t = \frac{I_T}{t}$ has a log-normally distribution. Then $\ln A_t = \ln I_T - \ln t$ is normal, so $\ln I_T$ is normal, and hence $I_T$ is log-normally distributed. Since $I_T = \int_0^T S_u du$, using (5.8.24) yields

$$ m_1 = E[I_T] = S_0 \frac{e^{\mu T} - 1}{\mu}; $$

$$ m_2 = E[I_T^2] = \frac{2S_0^2}{\mu + \sigma^2} \left[ \frac{e^{(2\mu + \sigma^2)T} - 1}{2\mu + \sigma^2} - \frac{e^{\mu T} - 1}{\mu} \right]. $$

Using Proposition 6.13.5 it follows that $\ln A_T$ is normally distributed, with

$$ \ln A_T \sim N\left( \ln \frac{m_1^2}{\sqrt{m_2}} - \ln t, \ln \frac{m_2}{m_1^2} \right). \tag{6.13.21} $$

Relation (6.13.21) represents the normal approximation of $\ln A_T$. We shall price the arithmetic average price call under this condition.

In the next two exercises we shall assume that the distribution of $A_T$ is given by the log-normal distribution (6.13.21).
Exercise 6.13.6 Using a method similar to the one used in Lemma 6.13.1, show that an approximate value at time 0 of a derivative, which pays at maturity $1$ if the arithmetic average $A_T \geq K$ and 0 otherwise, is given by

$$h_0 = e^{-rT}N(d_2),$$

with

$$d_2 = \frac{\ln(m_1^2/\sqrt{m_2^2}) - \ln K - \ln t}{\sqrt{\ln(m_2^2/m_1^2)}},$$

where in the expressions of $m_1$ and $m_2$ we replaced $\mu$ by $r$.

Exercise 6.13.7 Using a method similar to the one used in Lemma 6.13.2, show that the approximate value at time 0 of a derivative, which pays at maturity $A_T$ if $A_T \geq K$ and 0 otherwise, is given by the formula

$$a_0 = S_0 \frac{1 - e^{-rT}}{rt}N(\tilde{d}_1),$$

where (CHECK THIS AGAIN!)

$$\tilde{d}_1 = \ln(m_2^2/m_1^2) + \frac{\ln m_1^2}{\sqrt{\ln(m_2^2/m_1^2)}} + \ln t - \ln K,$$

where in the expressions of $m_1$ and $m_2$ we replaced $\mu$ by $r$.

Proposition 6.13.8 The approximate value at $t = 0$ of an arithmetic average price call is given by

$$f_0 = \frac{S_0(1 - e^{-rT})}{r}N(\tilde{d}_1) - Ke^{-rT}N(\tilde{d}_2),$$

with $\tilde{d}_1$ and $\tilde{d}_2$ given by formulas (6.13.22) and (6.13.23).

(b) How does the formula change if the value is taken at time $t$ instead of time 0?

6.14 Forward Contracts with Rare Events

We shall evaluate the price of a forward contract on a stock which follows a stochastic process with rare events, where the number of events $n = N_T$ until time $T$ is assumed Poisson distributed. As usual, $T$ denotes the maturity of the forward contract.
Let the jump ratios be \( Y_j = S_{t_j}/S_{t_{j-1}} \), where the events occur at times \( 0 < t_1 < t_2 < \cdots < t_n \), where \( n = N_T \). The stock price at maturity is given by Merton’s formula

\[
S_T = S_0 e^{(\mu - \lambda \rho - \frac{1}{2}\sigma^2)T + \sigma W_T} \prod_{j=1}^{N_T} Y_j,
\]

where \( \rho = E[Y_j] \) is the expected jump size, and \( Y_1, \cdots, Y_n \) are considered independent among themselves and also with respect to \( W_t \). Conditioning over \( N_T = n \) yields

\[
E[\prod_{j=1}^{N_T} Y_j] = \sum_{n\geq0} E[\prod_{j=1}^{N_T} Y_j | N_T = n] P(N_T = n)
\]
\[
= \sum_{n\geq0} \prod_{j=1}^{n} E[Y_j] P(N_T = n)
\]
\[
= \sum_{n\geq0} \rho^n \frac{(\lambda T)^n}{n!} e^{-\lambda T} = e^{\lambda \rho T} e^{-\lambda T}.
\]

Since \( W_T \) is independent of \( N_T \) and \( Y_j \) we have

\[
E[S_T] = S_0 e^{(\mu - \lambda \rho - \frac{1}{2}\sigma^2)T} E[e^{\sigma W_T}] E[\prod_{j=1}^{N_T} Y_j]
\]
\[
= S_0 e^{(\mu - \lambda \rho)T} e^{\lambda \rho T} e^{-\lambda T}
\]
\[
= S_0 e^{(\mu - \lambda)T}.
\]

Since the payoff at maturity of the forward contract is \( f_T = S_T - K \), with \( K \) delivery price, the value of the contract at time \( t = 0 \) is obtained by the method of risk neutral valuation

\[
f_0 = e^{r T} E_0 [S_T - K] = e^{-r T} \left( S_0 e^{(r - \lambda)T} - K \right)
\]
\[
= S_0 e^{-\lambda T} - Ke^{-r T}.
\]

Replacing \( T \) with \( T - t \) and \( S_0 \) with \( S_t \) yields the value of the contract time \( t \)

\[
f_t = S_t e^{-\lambda (T-t)} - Ke^{-r(T-t)}, \quad 0 \leq t \leq T.
\]

(6.14.24)

It is worth noting that if the rate of jumps occurrence is \( \lambda = 0 \), we obtain the familiar result

\[
f_t = S_t - e^{-r(T-t)} K.
\]
### 6.15 All-or-Nothing Lookback Options (Needs work!)

Consider a contract that pays the cash amount $K$ at time $T$ if the stock price $S_t$ did ever reach or exceed level $z$ until time $T$, and the amount 0 otherwise. The payoff is

$$V_T = \begin{cases} K, & \text{if } \max_{t \leq T} S_t \geq z \\ 0, & \text{otherwise} \end{cases}$$

where $\max_{t \leq T} S_t$ is the running maximum of the stock.

In order to compute the exact value of the option we need to find the probability density of the running maximum

$$X_t = \max_{s \leq t} S_s = S_0 e^{\max_{s \leq t} (\mu s + \sigma W_s)}$$

where $\mu = r - \frac{\sigma^2}{2}$. Let $Y_t = \max_{s \leq t} (\mu s + \sigma W_s)$.

Let $T_x$ be the first time the process $\mu t + \sigma W_t$ reaches level $x$, with $x > 0$. The probability density of $T_x$ is given by Proposition ??

$$p(\tau) = \frac{x}{\sigma \sqrt{2\pi \tau^3}} e^{-\frac{(x-\mu \tau)^2}{2\sigma^2 \tau}}.$$

The probability function of $Y_t$ is given by

$$P(Y_t \leq x) = 1 - P(Y_t > x) = 1 - P(\max_{0 \leq s \leq t} (\mu s + \sigma W_s) > x)$$

$$= 1 - P(T_x \leq t) = 1 - \int_0^t \frac{x}{\sigma \sqrt{2\pi \tau^3}} e^{-\frac{(x-\mu \tau)^2}{2\sigma^2 \tau}} d\tau$$

$$= \int_t^\infty \frac{t}{\sigma \sqrt{2\pi \tau^3}} e^{-\frac{(x-\mu \tau)^2}{2\sigma^2 \tau}} d\tau.$$

Then the probability function of $X_t$ becomes

$$P(X_t \leq u) = P(e^{Y_t} \leq u/S_0) = P(Y_t \leq \ln(u/S_0))$$

$$= \int_t^\infty \frac{\ln(u/S_0)}{\sigma \sqrt{2\pi \tau^3}} e^{-\frac{(\ln(u/S_0)-\mu \tau)^2}{2\sigma^2 \tau}} d\tau.$$

Let $S_0 < z$. What is the probability that the stock $S_t$ hits the barrier $z$ before time $T$? This can be formalized as the probability $P(\max_{t \leq T} S_t > z)$, which can be computed using the previous probability function:
\[ P(S_T > z) = 1 - P(\max_{t\leq T} S_t \leq z) = 1 - P(X_T \leq z) \]
\[ = 1 - \int_{T}^{\infty} \frac{\ln(z/S_0)}{\sigma \sqrt{2\pi T^3}} e^{-\frac{(\ln(z/S_0) - \mu \tau)^2}{2\sigma^2 \tau}} d\tau \]
\[ = \int_{T}^{0} \frac{\ln(z/S_0)}{\sigma \sqrt{2\pi T^3}} e^{-\frac{(\ln(z/S_0) - \mu \tau)^2}{2\sigma^2 \tau}} d\tau, \]

where \( \mu = r - \sigma^2/2 \). The exact value of the all-or-nothing lookback option at time \( t = 0 \) is

\[ V_0 = e^{-rT}E[V_T] \]
\[ = E[V_T | \max_{t\leq T} S_t \geq z] P(\max_{t\leq T} S_t \geq z) e^{-rT} \]
\[ + E[V_T | \max_{t\leq T} S_t < z] P(\max_{t\leq T} S_t < z) e^{-rT} \]
\[ = e^{-rT} K P(\max_{t\leq T} S_t \geq z) \]
\[ = e^{-rT} K \int_{T}^{0} \frac{\ln(z/S_0)}{\sigma \sqrt{2\pi T^3}} e^{-\frac{(\ln(z/S_0) - \mu \tau)^2}{2\sigma^2 \tau}} d\tau. \] (6.15.25)

Since the previous integral does not have an elementary expression, we shall work out some lower and upper bounds.

**Proposition 6.15.1 (Typos in the proof)** We have

\[ e^{-rT} K \left( 1 - \sqrt{\frac{2}{\pi \sigma^4 T \ln \frac{z}{S_0}}} \right) < V_0 < \frac{K S_0}{z}. \]

**Proof:** First, we shall find an upper bound for the option price using Doob’s inequality. The stock price at time \( t \) is given by

\[ S_t = S_0 e^{(r - \frac{\sigma^2}{2})t + \sigma W_t}. \]

Under normal market conditions we have \( r > \frac{\sigma^2}{2} \), and by Corollary ??, part (b) it follows that \( S_t \) is a submartingale (with respect to the filtration induced by \( W_t \)). Applying Doob’s submartingale inequality, Proposition ??, we have

\[ P(\max_{t\leq T} S_t \geq z) \leq \frac{1}{z} E[S_T] = \frac{S_0}{z} e^{rT}. \]

Then the expected value of \( V_T \) can be estimated as

\[ E[V_T] = K \cdot P(\max_{t\leq T} S_t \geq z) + 0 \cdot P(\max_{t\leq T} S_t < z) \leq \frac{K S_0}{z} e^{rT}. \]
Discounting, we obtain an upper bound for the value of the option at time $t = 0$

$$V_0 = e^{-rT}E[V_T] \leq \frac{K S_0}{z}.$$ 

In order to obtain a lower bound, we need to write the integral away from the singularity at $\tau = 0$, and use that $e^{-x} > 1 - x$ for $x > 0$:

$$V_0 = e^{-rT} K - e^{-rT} K \int_T^\infty \frac{\ln(z/S_0)}{2\pi\sigma^3\tau^3} d\tau$$

$$< e^{-rT} K - e^{-rT} K \int_T^\infty \frac{\ln(z/S_0)}{2 \sqrt{2\pi\sigma^3}} d\tau$$

$$= e^{-rT} K \left( 1 - \sqrt{\frac{2}{\pi\sigma^3}} \ln \frac{z}{S_0} \right).$$

\[\blacksquare\]

**Exercise 6.15.2** Find the value of an asset-or-nothing look-back option whose payoff is

$$V_T = \begin{cases} \overline{S}_T, & \text{if } S_T \geq K \\ 0, & \text{otherwise.} \end{cases}$$

$K$ denotes the strike price.

**Exercise 6.15.3** Find the value of a price call look-back option with the payoff given by

$$V_T = \begin{cases} \overline{S}_T - K, & \text{if } S_T \geq K \\ 0, & \text{otherwise.} \end{cases}$$

**Exercise 6.15.4** Find the value of a price put look-back option whose payoff is given by

$$V_T = \begin{cases} K - \underline{S}_T, & \text{if } S_T < K \\ 0, & \text{otherwise.} \end{cases}$$

**Exercise 6.15.5** Evaluate the following strike look-back options

(a) $V_T = S_T - \underline{S}_T$ (call)

(b) $V_T = \overline{S}_T - S_T$ (put)

It is worth noting that $(S_T - \underline{S}_T)^+ = S_T - \underline{S}_T$; this explains why the payoff is not given as a piece-wise function.

Hint: see p. 739 of Mc Donald for closed form formula.

**Exercise 6.15.6** Find a put-call parity for look-back options.
Exercise 6.15.7 Starting from the formula (see K&S p. 265)

\[ P(\sup_{t \leq T}(W_t - \gamma t) \geq \beta) = 1 - N\left(\gamma \sqrt{T} + \frac{\beta}{\sqrt{T}}\right) + e^{-2\beta \gamma} N\left(\gamma \sqrt{T} - \frac{\beta}{\sqrt{T}}\right), \]

with \( \beta > 0, \gamma \in \mathbb{R} \) and \( N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz \), calculate the price at \( t = 0 \) of a contract that pays \( K \) at time \( T \) if the stock price \( S_t \) did ever reach or exceed level \( z \) before time \( T \).

6.16 Perpetual Look-back Options

Consider a contract that pays \( K \) at the time the stock price \( S_t \) reaches level \( b \) for the first time. What is the price of such contract at any time \( t \geq 0 \)?

This type of contract is called perpetual look-back because the time horizon is infinite, i.e. there is no expiration date for the contract.

The contract pays \( K \) at time \( \tau_b = \inf\{t > 0; S_t \geq b\} \). The value at time \( t = 0 \) is obtained by discounting at the risk-free interest rate \( r \)

\[ V_0 = E[K e^{-r \tau_b}] = K E[e^{-r \tau_b}]. \]  

(6.16.26)

This expectation will be worked out using formula (??). First, we notice that the time \( \tau_b \) at which \( S_t = b \) is the same as the time for which \( S_0 e^{(r-\sigma^2/2)t + \sigma W_t} = b \), or equivalently,

\[ (r - \sigma^2/2)t + \sigma W_t = \ln \frac{b}{S_0}. \]

Applying Proposition ?? with \( \mu = r - \sigma^2/2, s = r, \) and \( x = \ln(b/S_0) \) yields

\[ E[e^{-r \tau_b}] = e^{\frac{1}{2\sigma^2} \left( r - \frac{\sigma^2}{2} - \sqrt{2r\sigma^2 + (r - \frac{\sigma^2}{2})^2} \right) \ln \frac{b}{S_0}} \]

\[ = \left( \frac{b}{S_0} \right)^{\frac{1}{2\sigma^2} \left( r - \frac{\sigma^2}{2} - \sqrt{2r\sigma^2 + (r - \frac{\sigma^2}{2})^2} \right) / \sigma^2} \]

\[ = \left( \frac{b}{S_0} \right)^{\frac{1}{2\sigma^2} - \frac{1}{2} - \sqrt{\frac{2r^2}{\sigma^2} + \frac{1}{\sigma^2}}/\sigma^2}. \]

Substituting in (6.16.26) yields the price of the perpetual look-back option at \( t = 0 \)

\[ V_0 = K \left( \frac{b}{S_0} \right)^{\frac{1}{2\sigma^2} - \frac{1}{2} - \sqrt{\frac{2r^2}{\sigma^2} + \frac{1}{\sigma^2}}/\sigma^2}. \]  

(6.16.27)

Exercise 6.16.1 Find the price of a contract that pays \( K \) when the initial stock price doubles its value.
6.17 Immediate Rebate Options

This contract pays at the time the barrier is hit, $T_a$, the amount $K$. The discounted value at $t = 0$ is

$$f_0 = \begin{cases} e^{-rT_a}, & \text{if } T_a \leq T \\ 0, & \text{otherwise.} \end{cases}$$

The price of the contract at $t = 0$ is

$$V_0 = E[f_0] = E[f_0|T_a < T]P(T_a < T) + \underbrace{E[f_0|T_a \geq T]}_{=0} P(T_a \geq T)$$

$$= KE[e^{-rT_a}]P(T_a < T)$$

$$= K \int_0^\infty e^{-r\tau} p_a(\tau) d\tau \int_0^T p_a(\tau) d\tau,$$

where

$$p_a(\tau) = \frac{\ln \frac{a}{S_0}}{\sqrt{2\pi\sigma^2\tau^3}} e^{-(\ln \frac{a}{S_0} - \alpha \tau)^2/(2\sigma^2)},$$

and $\alpha = r - \frac{\sigma^2}{2}$. Question: Why in McDonald (22.20) the formula is different. Are they equivalent?

6.18 Deferred Rebate Options

The payoff of this contract pays $K$ as long as a certain barrier has been hit. It is called deferred because the payment $K$ is made at the expiration time $T$. If $T_a$ is the first time the stock reaches the barrier $a$, the payoff can be formalized as

$$V_T = \begin{cases} K, & \text{if } T_a < T \\ 0, & \text{if } T_a \geq T. \end{cases}$$

The value of the contract at time $t = 0$ is

$$V_0 = e^{-rT} E[V_T] = e^{-rT} K P(T_a < T)$$

$$= Ke^{-rT} \ln \frac{a}{S_0} \int_0^T \frac{1}{\sigma \sqrt{2\pi \tau^3}} e^{-(\ln \frac{a}{S_0} - \alpha \tau)^2/(2\sigma^2)} d\tau,$$

with $\alpha = r - \frac{\sigma^2}{2}$. 
Chapter 7

Martingale Measures

7.1 Martingale Measures

An $\mathcal{F}_t$-predictable stochastic process $X_t$ on the probability space $(\Omega, \mathcal{F}, P)$ is not always a martingale. However, it might become a martingale with respect to another probability measure $Q$ on $\mathcal{F}$. This is called a martingale measure. The main result of this section is finding a martingale measure with respect to which the discounted stock price is a martingale. This measure plays an important role in the mathematical explanation of the risk-neutral valuation.

7.1.1 Is the stock price $S_t$ a martingale?

Since the stock price $S_t$ is an $\mathcal{F}_t$-predictable and non-explosive process, the only condition which needs to be satisfied to be a $\mathcal{F}_t$-martingale is

$$E[S_t | \mathcal{F}_u] = S_u, \quad \forall u < t.$$  

(7.1.1)

Heuristically speaking, this means that given all information in the market at time $u$, $\mathcal{F}_u$, the expected price of any future stock price is the price of the stock at time $u$, i.e $S_u$. This does not make sense, since in this case the investor would prefer investing the money in a bank at a risk-free interest rate, rather than buying a stock with zero return. Then (7.1.1) does not hold. The next result shows how to fix this problem.

**Proposition 7.1.1** Let $\mu$ be the rate of return of the stock $S_t$. Then

$$E[e^{-\mu t} S_t | \mathcal{F}_u] = e^{-\mu u} S_u, \quad \forall u < t,$$

(7.1.2)

i.e. $e^{-\mu t} S_t$ is an $\mathcal{F}_t$-martingale.

**Proof:** The process $e^{-\mu t} S_t$ is non-explosive since

$$E[|e^{-\mu t} S_t|] = e^{-\mu t} E[S_t] = e^{-\mu t} S_0 e^{\mu t} = S_0 < \infty.$$  

121
Since $S_t$ is $\mathcal{F}_t$-predictable, so is $e^{-\mu t}S_t$. Using formula (5.1.2) and taking out the predictable part yields
\[
E[S_t|\mathcal{F}_u] = E[S_0 e^{(\mu - \frac{1}{2}\sigma^2)u + \sigma W_u} | \mathcal{F}_u] \\
= E[S_0 e^{(\mu - \frac{1}{2}\sigma^2)u + \sigma W_u} e^{(\mu - \frac{1}{2}\sigma^2)(t-u) + \sigma (W_t-W_u)} | \mathcal{F}_u] \\
= E[S_u e^{(\mu - \frac{1}{2}\sigma^2)(t-u) + \sigma (W_t-W_u)} | \mathcal{F}_u] \\
= S_u e^{(\mu - \frac{1}{2}\sigma^2)(t-u)} E[e^{\sigma (W_t-W_u)} | \mathcal{F}_u],
\]  
(7.1.3)

Since the increment $W_t - W_u$ is independent of all values $W_s$, $s \leq u$, then it will also be independent of $\mathcal{F}_u$. By Proposition ??, part 6, the conditional expectation becomes the usual expectation
\[
E[e^{\sigma (W_t-W_u)} | \mathcal{F}_u] = E[e^{\sigma (W_t-W_u)}].
\]

Since $\sigma (W_t - W_u) \sim N(0, \sigma^2(t-u))$, from Exercise ?? (b) we get
\[
E[e^{\sigma (W_t-W_u)}] = e^{\frac{1}{2}\sigma^2(t-u)}.
\]

Substituting back into (7.1.3) yields
\[
E[S_t|\mathcal{F}_u] = S_u e^{(\mu - \frac{1}{2}\sigma^2)(t-u)} e^{\frac{1}{2}\sigma^2(t-u)} = S_u e^{\mu(t-u)},
\]
which is equivalent to
\[
E[e^{-\mu t}S_t | \mathcal{F}_u] = e^{-\mu u} S_u.
\]

The conditional expectation $E[S_t|\mathcal{F}_u]$ can be expressed in terms of the conditional density function as
\[
E[S_t|\mathcal{F}_u] = \int S_t p(S_t|\mathcal{F}_u) \, dS_t,
\]  
(7.1.4)

where $S_t$ is taken as an integration variable.

**Exercise 7.1.2** (a) Find the formula for conditional density function, $p(S_t|\mathcal{F}_u)$, defined by (7.1.4).

(b) Verify the formula
\[
E[S_t|\mathcal{F}_0] = E[S_t]
\]
in two different ways, either by using part (a), or by using the independence of $S_t$ with respect to $\mathcal{F}_0$.

The martingale relation (7.1.2) can be written equivalently as
\[
\int e^{-\mu t}S_t p(S_t|\mathcal{F}_u) \, dS_t = e^{-\mu u} S_u, \quad u < t.
\]

This way, $dP(x) = p(x|\mathcal{F}_u) \, dx$ becomes a martingale measure for $e^{-\mu t} S_t$. 
7.1.2 Risk-neutral World and Martingale Measure

Since the rate of return $\mu$ might not be known from the beginning, and it depends on each particular stock, a meaningful question would be: 

*Under what martingale measure does the discounted stock price, $M_t = e^{-rt}S_t$, become a martingale?*

The constant $r$ denotes, as usual, the risk-free interest rate. Assume such a martingale measure exists. Then we must have

$$\hat{E}_u[e^{-rt}S_t] = \hat{E}[e^{-rt}S_t|\mathcal{F}_u] = e^{-ru}S_u,$$

where $\hat{E}$ denotes the expectation with respect to the requested martingale measure. The previous relation can also be written as

$$e^{-r(t-u)}\hat{E}[S_t|\mathcal{F}_u] = S_u, \quad u < t.$$

This states that the discounted expectation at the risk-free interest rate for the time interval $t - u$ is the price of the stock, $S_u$. Since this does not involve any of the riskiness of the stock, we might think of it as an expectation in the risk-neutral world. The aforementioned formula can be written in the compound mode as

$$\hat{E}[S_t|\mathcal{F}_u] = S_u e^{r(t-u)}, \quad u < t. \quad (7.1.5)$$

This formula can be obtained from the conditional expectation $E[S_t|\mathcal{F}_u] = S_u e^{\mu(t-u)}$ by substituting $\mu = r$ and replacing $E$ by $\hat{E}$, which corresponds to the definition of the expectation in a risk-neutral world. Therefore, the evaluation of derivatives in section 6 is done by using the aforementioned martingale measure under which $e^{-rt}S_t$ is a martingale. In the next section we shall determine this measure explicitly.

**Exercise 7.1.3** Consider the following two games that consist in flipping a fair coin and taking decisions:

A. If the coin lands Heads, you win $2; otherwise you lose $1.
B. If the coin lands Heads, you win $20,000; otherwise you lose $10,000.

(a) Which game involves more risk? Explain your answer.
(b) Which game would you choose to play, and why?
(c) Are you risk-neutral in your decision?

The risk neutral measure is the measure with respect to which investors are not risk-averse. In the case of the previous exercise, it is the measure with respect to which all players are indifferent whether they choose option A or B.
7.1.3 Finding the Risk-Neutral Measure

The solution of the stochastic differential equation of the stock price

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

can be written as

$$S_t = S_0 e^{\mu t} e^{\sigma W_t - \frac{1}{2} \sigma^2 t}.$$

Then $e^{-\mu t} S_t = S_0 e^{\sigma W_t - \frac{1}{2} \sigma^2 t}$ is an exponential process. By Example 7.1.3, particular case 1, this process is an $\mathcal{F}_t$-martingale, where $\mathcal{F}_t = \sigma\{W_u; u \leq t\}$ is the information available in the market until time $t$. Hence $e^{-\mu t} S_t$ is a martingale, which is a result proved also by Proposition 7.1.1. The probability space where this martingale exists is $(\Omega, \mathcal{F}, P)$.

In the following we shall change the rate of return $\mu$ into the risk-free rate $r$ and change the probability measure such that the discounted stock price becomes a martingale. The discounted stock price can be expressed in terms of the Brownian motion with drift (a hat was added in order to distinguish it from the former Brownian motion $W_t$)

$$\hat{W}_t = \frac{\mu - r}{\sigma} t + W_t \quad (7.1.6)$$

as in the following

$$e^{-rt} S_t = e^{-rt} S_0 e^{\mu t} e^{\sigma W_t - \frac{1}{2} \sigma^2 t} = e^{\sigma \hat{W}_t - \frac{1}{2} \sigma^2 t}.$$

If we let $\lambda = \frac{\mu - r}{\sigma}$ in Corollary 7.1.3 of Girsanov’s theorem, it follows that $\hat{W}_t$ is a Brownian motion on the probability space $(\Omega, \mathcal{F}, Q)$, where

$$dQ = e^{-\frac{1}{2} \left( \frac{\mu - r}{\sigma} \right)^2 T - \lambda W_T} dP.$$

As an exponential process, $e^{\sigma \hat{W}_t - \frac{1}{2} \sigma^2 t}$ becomes a martingale on this space. Consequently $e^{-rt} S_t$ is a martingale process w.r.t. the probability measure $Q$. This means

$$E^Q[e^{-rt} S_t | \mathcal{F}_u] = e^{-ru} S_u, \quad u < t.$$

where $E^Q[\cdot | \mathcal{F}_u]$ denotes the conditional expectation in the measure $Q$, and it is given by

$$E^Q[X_t | \mathcal{F}_u] = E^P[X_t e^{-\frac{1}{2} \left( \frac{\mu - r}{\sigma} \right)^2 T - \lambda W_T} | \mathcal{F}_u].$$

The measure $Q$ is called the equivalent martingale measure, or the risk-neutral measure. The expectation taken with respect to this measure is called the expectation in the risk-neutral world. Customarily we shall use the notations

$$\hat{E}[e^{-rt} S_t] = E^Q[e^{-rt} S_t]$$

$$\hat{E}_u[e^{-rt} S_t] = E^Q[e^{-rt} S_t | \mathcal{F}_u]$$
It is worth noting that \( \hat{E}[e^{-rt}S_t] = \hat{E}_0[e^{-rt}S_t] \), since \( e^{-rt}S_t \) is independent of the initial information set \( \mathcal{F}_0 \).

The importance of the process \( \hat{W}_t \) is contained in the following useful result.

**Proposition 7.1.4** The probability measure that makes the discounted stock price, \( e^{-rt}S_t \), a martingale changes the rate of return \( \mu \) into the risk-free interest rate \( r \), i.e

\[
dS_t = rS_t dt + \sigma S_t d\hat{W}_t.
\]

**Proof:** The proof is a straightforward verification using (7.1.6)

\[
dS_t = \mu S_t dt + \sigma S_t d\hat{W}_t = rS_t dt + (\mu - r)S_t dt + \sigma S_t d\hat{W}_t
\]

\[
= rS_t dt + \sigma S_t \left( \frac{\mu - r}{\sigma} dt + d\hat{W}_t \right)
\]

\[
= rS_t dt + \sigma S_t d\hat{W}_t
\]

We note that the solution of the previous stochastic equation is

\[
S_t = S_0 e^{rt} e^{\sigma \hat{W}_t - \frac{1}{2} \sigma^2 t}.
\]

\[
\square
\]

**Exercise 7.1.5** Assume \( \mu \neq r \) and let \( u < t \).

(a) Find \( E^P[e^{-rt}S_t|\mathcal{F}_u] \) and show that \( e^{-rt}S_t \) is not a martingale w.r.t. the probability measure \( P \).

(b) Find \( E^Q[e^{-\mu t}S_t|\mathcal{F}_u] \) and show that \( e^{-\mu t}S_t \) is not a martingale w.r.t. the probability measure \( Q \).

### 7.2 Risk-neutral World Density Functions

The purpose of this section is to establish formulas for the densities of Brownian motions \( W_t \) and \( \hat{W}_t \) with respect to both probability measures \( P \) and \( Q \), and discuss their relationship. This will clear some confusions that appear in practical applications when we need to choose the right probability density.

The densities of \( W_t \) and \( \hat{W}_t \) with respect to \( P \) and \( Q \) will be denoted respectively by \( p_P \), \( p_Q \), \( \hat{p}_P \), \( \hat{p}_Q \). Since \( W_t \) and \( \hat{W}_t \) are Brownian motions on the spaces \((\Omega, \mathcal{F}, P)\) and \((\Omega, \mathcal{F}, Q)\), respectively, they have the following normal probability densities

\[
p_P(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} = p(x);
\]

\[
\hat{p}_Q(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} = p(x).
\]
The associated distribution functions are

\[ F_{W_t}(x) = P(W_t \leq x) = \int_{\{W_t \leq x\}} dP(\omega) = \int_{-\infty}^{x} p(u) du; \]

\[ F_{\hat{W}_t}(x) = Q(\hat{W}_t \leq x) = \int_{\{\hat{W}_t \leq x\}} dQ(\omega) = \int_{-\infty}^{x} p(u) du. \]

Expressing \( W_t \) in terms of \( \hat{W}_t \) and using that \( \hat{W}_t \) is normally distributed with respect to \( Q \) we get the distribution function of \( W_t \) with respect to \( Q \) as

\[ F_Q^{W_t}(x) = Q(W_t \leq x) = Q(\hat{W}_t - \eta t \leq x) \]

\[ = Q(\hat{W}_t \leq x + \eta t) = \int_{\{\hat{W}_t \leq x + \eta t\}} dQ(\omega) \]

\[ = \int_{-\infty}^{x+\eta t} p(y) dy = \int_{-\infty}^{x+\eta t} \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}} dy. \]

Differentiating yields the density function

\[ p_Q(x) = \frac{d}{dx} F_Q^{W_t}(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t(y+\eta t)^2}}. \]

It is worth noting that \( p_Q(x) \) can be decomposed as

\[ p_Q(x) = e^{-\eta x - \frac{1}{2} \eta^2 t} p(x), \]

which makes the connection with the Girsanov theorem.

The distribution function of \( \hat{W}_t \) with respect to \( P \) can be worked out in a similar way

\[ F_P^{\hat{W}_t}(x) = P(\hat{W}_t \leq x) = P(W_t + \eta t \leq x) \]

\[ = P(W_t \leq x - \eta t) = \int_{\{W_t \leq x - \eta t\}} dP(\omega) \]

\[ = \int_{-\infty}^{x-\eta t} p(y) dy = \int_{-\infty}^{x-\eta t} \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}} dy, \]

so the density function is

\[ p_P(x) = \frac{d}{dx} F_P^{\hat{W}_t}(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}(y-\eta t)^2}. \]
7.3 Self-financing Portfolios

The value of a portfolio that contains \( \theta_j(t) \) units of stock \( S_j(t) \) at time \( t \) is given by

\[
V(t) = \sum_{j=1}^{n} \theta_j(t)S_j(t).
\]

The portfolio is called self-financing if

\[
dV(t) = \sum_{j=1}^{n} \theta_j(t)dS_j(t).
\]

This means that an infinitesimal change in the value of the portfolio is due to infinitesimal changes in stock values. All portfolios will be assumed self-financing, if otherwise stated.

7.4 The Sharpe Ratio

If \( \mu \) is the expected return on the stock \( S_t \), the risk premium is defined as the difference \( \mu - r \), where \( r \) is the risk-free interest rate. The Sharpe ratio, \( \eta \), is the quotient between the risk premium and stock price volatility

\[
\eta = \frac{\mu - r}{\sigma}.
\]

The following result shows that the Sharpe ratio is an important invariant for the family of stocks driven by the same uncertain source. It is also known under the name of the market price of risk for assets.

**Proposition 7.4.1** Let \( S_1 \) and \( S_2 \) be two stocks satisfying equations (5.2.5) – (5.2.6). Then their Sharpe ratio are equal

\[
\frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2}. \tag{7.4.7}
\]

**Proof:** Eliminating the term \( dW_t \) from equations (5.2.5) – (5.2.6) yields

\[
\frac{\sigma_2}{S_1}dS_1 - \frac{\sigma_1}{S_2}dS_2 = (\mu_1 \sigma_2 - \mu_2 \sigma_1)dt. \tag{7.4.8}
\]

Consider the portfolio \( P(t) = \theta_1(t)S_1(t) - \theta_2(t)S_2(t) \), with \( \theta_1(t) = \frac{\sigma_2(t)}{S_1(t)} \) and \( \theta_2(t) = \frac{\sigma_1(t)}{S_2(t)} \). Using the properties of self-financing portfolios, we have

\[
dP(t) = \theta_1(t)dS_1(t) - \theta_2(t)dS_2(t) = \frac{\sigma_2}{S_1}dS_1 - \frac{\sigma_1}{S_2}dS_2.
\]
Substituting in (7.4.8) yields $dP = (\mu_1 \sigma_2 - \mu_2 \sigma_1)dt$, i.e. $P$ is a risk-less portfolio. Since the portfolio earns interest at the risk-free interest rate, we have $dP = rPdt$. Then equating the coefficients of $dt$ yields

$$\mu_1 \sigma_2 - \mu_2 \sigma_1 = rP(t).$$

Using the definition of $P(t)$, the previous relation becomes

$$\mu_1 \sigma_2 - \mu_2 \sigma_1 = r\theta_1 S_1 - r\theta_2 S_2,$$

that can be transformed to

$$\mu_1 \sigma_2 - \mu_2 \sigma_1 = r\bar{\sigma}_2 - r\sigma_1,$$

which is equivalent with (7.4.7).

Using Proposition 7.1.4, relations (5.2.5) – (5.2.6) can be written as

$$dS_1 = rS_1 dt + \sigma_1 S_1 d\hat{W}_t,$$
$$dS_2 = rS_2 dt + \sigma_2 S_2 d\hat{W}_t,$$

where the risk-neutral process $d\hat{W}_t$ is the same in both equations

$$d\hat{W}_t = \frac{\mu_1 - r}{\sigma_1} dt + dW_t = \frac{\mu_2 - r}{\sigma_2} dt + dW_t.$$

This shows that the process $\hat{W}_t$ is a Brownian motion with drift, where the drift is the Sharpe ratio.

### 7.5 Risk-neutral Valuation for Derivatives

The risk-neutral process $d\hat{W}_t$ plays an important role in the risk neutral valuation of derivatives. In this section we shall prove that if $f_T$ is the price of a derivative at the maturity time, then $f_t = \hat{E}[e^{-r(T-t)} f_T | \mathcal{F}_t]$ is the price of the derivative at the time $t$, for any $t < T$.

In other words, the discounted price of a derivative in the risk-neutral world is the price of the derivative at the new instance of time. This is based on the fact that $e^{-rt}f_t$ is an $\mathcal{F}_t$-martingale with respect to the risk-neutral measure $Q$ introduced previously.

In particular, the idea of the proof can be applied for the stock $S_t$. Applying the product rule

$$d(e^{-rt}S_t) = d(e^{-rt})S_t + e^{-rt}dS_t + d(e^{-rt})dS_t = 0$$

we get

$$d(e^{-rt}S_t) = -re^{-rt}S_t dt + e^{-rt}(rS_t dt + \sigma S_t d\hat{W}_t)$$

or

$$e^{-rt}(rS_t dt + \sigma S_t d\hat{W}_t).$$
If $u < t$, integrating between $u$ and $t$

$$e^{-rt}S_t = e^{-ru}S_u + \int_u^t \sigma e^{-rs}S_s d\hat{W}_s,$$

and taking the risk-neutral expectation with respect to the information set $\mathcal{F}_u$ yields

$$\hat{E}[e^{-rt}S_t | \mathcal{F}_u] = \hat{E}[e^{-ru}S_u + \int_u^t \sigma e^{-rs}S_s d\hat{W}_s | \mathcal{F}_u]$$

$$= e^{-ru}S_u + \hat{E}\left[ \int_u^t \sigma e^{-rs}S_s d\hat{W}_s \right]$$

$$= e^{-ru}S_u,$$

since $\int_u^t \sigma e^{-rs}S_s d\hat{W}_s$ is independent of $\mathcal{F}_u$. It follows that $e^{-rt}S_t$ is an $\mathcal{F}_t$-martingale in the risk-neutral world. The following fundamental result can be shown using a similar proof as the one encountered previously:

**Theorem 7.5.1** If $f_t = f(t, S_t)$ is the price of a derivative at time $t$, then $e^{-rt}f_t$ is an $\mathcal{F}_t$-martingale in the risk-neutral world, i.e.

$$\hat{E}[e^{-rt}f_t | \mathcal{F}_u] = e^{-ru}f_u, \quad \forall 0 < u < t.$$

**Proof:** Using Ito’s formula and the risk neutral process $dS = rS dt + \sigma S d\hat{W}_t$, the process followed by $f_t$ is

$$df_t = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S} dS + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} (dS)^2$$

$$= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S} (rS dt + \sigma S d\hat{W}_t) + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} dt$$

$$= \left( \frac{\partial f}{\partial t} + rS \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} \right) dt + \sigma S \frac{\partial f}{\partial S} d\hat{W}_t$$

$$= rf_t dt + \sigma S \frac{\partial f}{\partial S} d\hat{W}_t,$$

where in the last identity we used that $f$ satisfies the Black-Scholes equation. Applying the product rule we obtain

$$d(e^{-rt}f_t) = d(e^{-rt})f_t + e^{-rt}df_t + d(e^{-rt})df_t$$

$$= -re^{-rt}f_t dt + e^{-rt} \left( rf_t dt + \sigma S \frac{\partial f}{\partial S} d\hat{W}_t \right)$$

$$= e^{-rt} \sigma S \frac{\partial f}{\partial S} d\hat{W}_t.$$
Integrating between $u$ and $t$ we get

$$e^{-rt} f_t = e^{-ru} f_u + \int_u^t e^{-rs} \sigma S \frac{\partial f_s}{\partial S} d\hat{W}_s,$$

which assures that $e^{-rt} f_t$ is a martingale, since $\hat{W}_s$ is a Brownian motion process. Using that $e^{-ru} f_u$ is $\mathcal{F}_u$-predictable, and $\int_u^t e^{-rs} \sigma S \frac{\partial f_s}{\partial S} d\hat{W}_s$ is independent of the information set $\mathcal{F}_u$, we have

$$\hat{E}[e^{-rt} f_t | \mathcal{F}_u] = e^{-ru} f_u + \hat{E}[ \int_u^t e^{-rs} \sigma S \frac{\partial f_s}{\partial S} d\hat{W}_s ] = e^{-ru} f_u.$$

**Exercise 7.5.2** Show the following:
(a) $\hat{E}[e^{\sigma(W_t - W_u)} | \mathcal{F}_u] = e^{(r - \mu + \frac{1}{2} \sigma^2)(t - u)}, \quad u < t$;
(b) $\hat{E}\left[\frac{S_t}{S_u}|\mathcal{F}_u\right] = e^{(\mu - \frac{1}{2}\sigma^2)(t - u)} \hat{E}[e^{\sigma(W_t - W_u)} | \mathcal{F}_u], \quad u < t$;
(c) $\hat{E}\left[\frac{S_t}{S_u}|\mathcal{F}_u\right] = e^{r(t - u)}, \quad u < t$.

**Exercise 7.5.3** Find the following risk-neutral world conditional expectations:
(a) $\hat{E}[\int_0^t S_u du | \mathcal{F}_s], \quad s < t$;
(b) $\hat{E}[S_t \int_0^t S_u du | \mathcal{F}_s], \quad s < t$;
(c) $\hat{E}[\int_0^t S_u dW_u | \mathcal{F}_s], \quad s < t$;
(d) $\hat{E}[S_t \int_0^t S_u dW_u | \mathcal{F}_s], \quad s < t$;
(e) $\hat{E}[(\int_0^t S_u du)^2 | \mathcal{F}_s], \quad s < t$.

**Exercise 7.5.4** Use risk-neutral valuation to find the price of a derivative that pays at maturity the following payoffs:
(a) $f_T = TS_T$;
(b) $f_T = \int_0^T S_u du$;
(c) $f_T = \int_0^T S_u dW_u$. 
Chapter 8

Black-Scholes Analysis

8.1 Heat Equation

This section is devoted to a basic discussion on the heat equation. Its importance resides in the remarkable fact that the Black-Scholes equation, which is the main equation of derivatives calculus, can be reduced to this type of equation.

Let $u(\tau, x)$ denote the temperature in an infinite rod at point $x$ and time $\tau$. In the absence of exterior heat sources the heat diffuses according to the following parabolic differential equation

$$\frac{\partial u}{\partial \tau} - \frac{\partial^2 u}{\partial x^2} = 0,$$

(8.1.1)

called the heat equation. If the initial heat distribution is known and is given by $u(0, x) = f(x)$, then we have an initial value problem for the heat equation.

Solving this equation involves a convolution between the initial temperature $f(x)$ and the fundamental solution of the heat equation $G(\tau, x)$, which will be defined shortly.

**Definition 8.1.1** The function

$$G(\tau, x) = \frac{1}{\sqrt{4\pi \tau}} e^{-\frac{x^2}{4\tau}}, \quad \tau > 0,$$

is called the fundamental solution of the heat equation (8.1.1).

We recall the most important properties of the function $G(\tau, x)$.

- $G(\tau, x)$ has the properties of a probability density$^1$, i.e.

$^1$In fact it is a Gaussian probability density.
1. \( G(\tau,x) > 0, \quad \forall x \in \mathbb{R}, \tau > 0; \)

2. \( \int_{\mathbb{R}} G(\tau,x) \, dx = 1, \quad \forall \tau > 0. \)

- it satisfies the heat equation
  \[
  \frac{\partial G}{\partial \tau} - \frac{\partial^2 G}{\partial x^2} = 0, \quad \tau > 0.
  \]

- \( G \) tends to the \textit{Dirac measure} \( \delta(x) \) as \( \tau \) gets closer to the initial time
  \[
  \lim_{\tau \to 0} G(\tau,x) = \delta(x),
  \]

where the Dirac measure can be defined using integration as
  \[
  \int_{\mathbb{R}} \varphi(x) \delta(x) \, dx = \varphi(0),
  \]

for any smooth function with compact support \( \varphi \). Consequently, we also have
  \[
  \int_{\mathbb{R}} \varphi(x) \delta(x-y) \, dx = \varphi(y).
  \]

One can think of \( \delta(x) \) as a measure with infinite value at \( x = 0 \) and zero for the rest of the values, and with the integral equal to 1, see Fig.8.1.

The physical significance of the fundamental solution \( G(\tau,x) \) is that it describes the heat evolution in the infinite rod after an initial heat impulse of infinite size is applied at \( x = 0 \).

**Proposition 8.1.2** The solution of the initial value heat equation

\[
\frac{\partial u}{\partial \tau} - \frac{\partial^2 u}{\partial x^2} = 0 \\
u(0,x) = f(x)
\]

is given by the convolution between the fundamental solution and the initial temperature

\[
u(\tau,x) = \int_{\mathbb{R}} G(\tau,y-x)f(y) \, dy, \quad \tau > 0.
\]

**Proof:** Substituting \( z = y - x \), the solution can be written as

\[
u(\tau,x) = \int_{\mathbb{R}} G(\tau,z)f(x+z) \, dz.
\] (8.1.2)
Differentiating under the integral yields
\[
\frac{\partial u}{\partial \tau} = \int_{\mathbb{R}} \frac{\partial G(\tau, z)}{\partial \tau} f(x + z) \, dz,
\]
\[
\frac{\partial^2 u}{\partial x^2} = \int_{\mathbb{R}} G(\tau, z) \frac{\partial^2 f(x + z)}{\partial x^2} \, dz = \int_{\mathbb{R}} G(\tau, z) \frac{\partial^2 f(x + z)}{\partial z^2} \, dz
\]
where we applied integration by parts twice and the fact that
\[
\lim_{z \to \infty} G(\tau, z) = \lim_{z \to \infty} \frac{\partial G(\tau, z)}{\partial z} = 0.
\]
Since \( G \) satisfies the heat equation,
\[
\frac{\partial u}{\partial \tau} - \frac{\partial^2 u}{\partial x^2} = \int_{\mathbb{R}} \left[ \frac{\partial G(\tau, z)}{\partial \tau} - \frac{\partial^2 G(\tau, z)}{\partial z^2} \right] f(x + z) \, dz = 0.
\]
Since the limit and the integral commute\(^2\), using the properties of the Dirac measure, we have
\[
u(0, x) = \lim_{\tau \searrow 0} u(\tau, x) = \lim_{\tau \searrow 0} \int_{\mathbb{R}} G(\tau, z) f(x + z) \, dz
\]
\[
= \int_{\mathbb{R}} \delta(z) f(x + z) \, dz = f(x).
\]
\(^2\)This is allowed by the dominated convergence theorem.
Hence (8.1.2) satisfies the initial value heat equation.

It is worth noting that the solution \( u(\tau, x) = \int_{\mathbb{R}} G(y-x, \tau) f(y) \, dy \) provides the temperature at any point in the rod for any time \( \tau > 0 \), but it cannot provide the temperature for \( \tau < 0 \), because of the singularity the fundamental solution exhibits at \( \tau = 0 \). We can reformulate this by saying that the heat equation is semi-deterministic, in the sense that given the present, we can know the future but not the past.

The semi-deterministic character of diffusion phenomena can be exemplified with a drop of ink which starts diffusing in a bucket of water at time \( t = 0 \). We can determine the density of the ink at any time \( t > 0 \) at any point \( x \) in the bucket. However, given the density of ink at a time \( t > 0 \), it is not possible to trace back in time the ink distribution density and to find the initial point where the drop started its diffusion.

The semi-deterministic behavior occurs in the study of derivatives too. In the case of the Black-Scholes equation, which is a backwards heat equation\(^3\), given the present value of the derivative, we can find the past values but not the future ones. This is the capital difficulty in foreseeing the prices of stock market instruments from the present prices. This difficulty will be overcome by working the price from the given final condition, which is the payoff at maturity.

### 8.2 What is a Portfolio?

A portfolio is a position in the market that consists in long and short positions in one or more stocks and other securities. The value of a portfolio can be represented algebraically as a linear combination of stock prices and other securities’ values:

\[
P = \sum_{j=1}^{n} a_j S_j + \sum_{k=1}^{m} b_k F_k.
\]

The market participant holds \( a_j \) units of stock \( S_j \) and \( b_k \) units in derivative \( F_k \). The coefficients are positive for long positions and negative for short positions. For instance, a portfolio given by \( P = 2F - 3S \) means that we buy 2 securities and sell 3 units of stock (a position with 2 securities long and 3 stocks short).

The portfolio is self-financing if

\[
dP = \sum_{j=1}^{n} a_j dS_j + \sum_{k=1}^{m} b_k dF_k.
\]

\(^3\)This comes from the fact that at some point \( \tau \) becomes \( -\tau \) due to a substitution.
8.3 Risk-less Portfolios

A portfolio \( P \) is called risk-less if the increments \( dP \) are completely predictable. In this case the increments’ value \( dP \) should equal the interest earned in the time interval \( dt \) on the portfolio \( P \). This can be written as

\[
dP = rPdt,
\]

(8.3.3)

where \( r \) denotes the risk-free interest rate. For the sake of simplicity the rate \( r \) will be assumed constant throughout this section.

Let’s assume now that the portfolio \( P \) depends on only one stock \( S \) and one derivative \( F \), whose underlying asset is \( S \). The portfolio depends also on time \( t \), so

\[
P = P(t, S, F).
\]

We are interested in deriving the stochastic differential equation followed by the portfolio \( P \). We note that at this moment the portfolio is not assumed risk-less. By Ito’s formula we get

\[
dP = \frac{\partial P}{\partial t}dt + \frac{\partial P}{\partial S}dS + \frac{\partial P}{\partial F}dF + \frac{1}{2} \frac{\partial^2 P}{\partial S^2}dS^2 + \frac{1}{2} \frac{\partial^2 P}{\partial F^2}(dF)^2.
\]

(8.3.4)

The stock \( S \) is assumed to follow the geometric Brownian motion

\[
dS = \mu S dt + \sigma S dW_t,
\]

(8.3.5)

where the expected return rate on the stock \( \mu \) and the stock’s volatility \( \sigma \) are constants. Since the derivative \( F \) depends on time and underlying stock, we can write \( F = F(t, S) \). Applying Ito’s formula, yields

\[
dF = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial S}dS + \frac{1}{2} \frac{\partial^2 F}{\partial S^2}(dS)^2 = \left( \frac{\partial F}{\partial t} + \mu S \frac{\partial F}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} \right) dt + \sigma S \frac{\partial F}{\partial S} dW_t,
\]

(8.3.6)

where we have used (8.3.5). Taking the squares in relations (8.3.5) and (8.3.6), and using the stochastic relations \((dW_t)^2 = dt\) and \(dt^2 = dW_t dt = 0\), we get

\[
(dS)^2 = \sigma^2 S^2 dt
\]

\[
(dF)^2 = \sigma^2 S^2 \left( \frac{\partial F}{\partial S} \right)^2 dt.
\]

Substituting back in (8.3.4), and collecting the predictable and unpredictable parts, yields
\[
dP = \left[ \frac{\partial P}{\partial t} + \mu S \left( \frac{\partial P}{\partial S} + \frac{\partial P}{\partial F} \frac{\partial F}{\partial S} \right) + \frac{\partial P}{\partial F} \left( \frac{\partial F}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} \right) \\
+ \frac{1}{2} \sigma^2 S^2 \left( \frac{\partial^2 P}{\partial S^2} + \frac{\partial^2 P}{\partial F^2} \left( \frac{\partial F}{\partial S} \right)^2 \right) \right] dt \\
+ \sigma S \left( \frac{\partial P}{\partial S} + \frac{\partial P}{\partial F} \frac{\partial F}{\partial S} \right) dW_t. \tag{8.3.7}
\]

Looking at the unpredictable component, we have the following result:

**Proposition 8.3.1** The portfolio \( P \) is risk-less if and only if \( \frac{dP}{dS} = 0 \).

**Proof:** A portfolio \( P \) is risk-less if and only if its unpredictable component is identically zero, i.e.

\[
\frac{\partial P}{\partial S} + \frac{\partial P}{\partial F} \frac{\partial F}{\partial S} = 0.
\]

Since the total derivative of \( P \) is given by

\[
\frac{dP}{dS} = \frac{\partial P}{\partial S} + \frac{\partial P}{\partial F} \frac{\partial F}{\partial S},
\]

the previous relation becomes \( \frac{dP}{dS} = 0 \).

**Definition 8.3.2** The amount \( \Delta_P = \frac{dP}{dS} \) is called the delta of the portfolio \( P \).

The previous result can be reformulated by saying that a portfolio is risk-less if and only if its delta vanishes. In practice this can hold only for a short amount of time, so the portfolio needs to be re-balanced periodically. The process of making a portfolio risk-less involves a procedure called delta hedging, through which the portfolio’s delta becomes zero or very close to this value.

Assume \( P \) is a risk-less portfolio, so

\[
\frac{dP}{dS} = \frac{\partial P}{\partial S} + \frac{\partial P}{\partial F} \frac{\partial F}{\partial S} = 0. \tag{8.3.8}
\]

Then equation (8.3.7) simplifies to

\[
dP = \left[ \frac{\partial P}{\partial t} + \frac{\partial P}{\partial F} \left( \frac{\partial F}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} \right) \\
+ \frac{1}{2} \sigma^2 S^2 \left( \frac{\partial^2 P}{\partial S^2} + \frac{\partial^2 P}{\partial F^2} \left( \frac{\partial F}{\partial S} \right)^2 \right) \right] dt. \tag{8.3.9}
\]
Comparing with (8.3.3) yields

\[
\frac{\partial P}{\partial t} + \frac{\partial P}{\partial F} \left( \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} \right) + \frac{1}{2} \sigma^2 S^2 \left( \frac{\partial^2 P}{\partial S^2} + \frac{\partial^2 P}{\partial F^2} \left( \frac{\partial F}{\partial S} \right)^2 \right) = rP. \tag{8.3.10}
\]

This is the equation satisfied by a risk-free financial instrument, \( P = P(t, S, F) \), that depends on time \( t \), stock \( S \) and derivative price \( F \).

### 8.4 Black-Scholes Equation

This section deals with a parabolic partial differential equation satisfied by all European-type securities, called the Black-Scholes equation. This was initially used by Black and Scholes to find the value of options. This is a deterministic equation obtained by eliminating the unpredictable component of the derivative by making a risk-less portfolio. The main reason for this being possible is the fact that both the derivative \( F \) and the stock \( S \) are driven by the same source of uncertainty.

The next result holds in a market with the following restrictive conditions:

- the risk-free rate \( r \) and stock volatility \( \sigma \) are constant.
- there are no arbitrage opportunities.
- no transaction costs.

**Proposition 8.4.1** If \( F(t, S) \) is a derivative defined for \( t \in [0, T] \), then

\[
\frac{\partial F}{\partial t} + rS \frac{\partial F}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} = rF. \tag{8.4.11}
\]

**Proof:** The equation (8.3.10) works under the general hypothesis that \( P = P(t, S, F) \) is a risk-free financial instrument that depends on time \( t \), stock \( S \) and derivative \( F \). We shall consider \( P \) to be the following particular portfolio

\[ P = F - \lambda S. \]

This means to take a long position in derivative and a short position in \( \lambda \) units of stock (assuming \( \lambda \) positive). The partial derivatives in this case are

\[
\frac{\partial P}{\partial t} = 0, \quad \frac{\partial P}{\partial F} = 1, \quad \frac{\partial P}{\partial S} = -\lambda, \quad \frac{\partial^2 P}{\partial F^2} = 0, \quad \frac{\partial^2 P}{\partial S^2} = 0.
\]
From the risk-less property (8.3.8) we get $\lambda = \frac{\partial F}{\partial S}$. Substituting into equation (8.3.10) yields

$$\frac{\partial F}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} = rF - rS \frac{\partial F}{\partial S},$$

which is equivalent to the desired equation.

However, the Black-Scholes equation is derived most often in a less rigorous way. It is based on the assumption that the number $\lambda = \frac{\partial F}{\partial S}$, which appears in the formula of the risk-less portfolio $P = F - \lambda S$, is considered constant for the time interval $\Delta t$. If we consider the increments over the time interval $\Delta t$

$$\Delta W_t = W_{t+\Delta t} - W_t$$
$$\Delta S = S_{t+\Delta t} - S_t$$
$$\Delta F = F(t + \Delta t, S_t + \Delta S) - F(t, S),$$

then Ito’s formula yields

$$\Delta F = \left(\frac{\partial F}{\partial t}(t, S) + \mu S \frac{\partial F}{\partial S}(t, S) + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial S^2}(t, S)\right) \Delta t$$
$$+ \sigma S \frac{\partial F}{\partial S}(t, S) \Delta W_t.$$

On the other side, the increments in the stock are given by

$$\Delta S = \mu S \Delta t + \sigma S \Delta W_t.$$

Since both increments $\Delta F$ and $\Delta S$ are driven by the same uncertainly source, $\Delta W_t$, we can eliminate it by multiplying the latter equation by $\frac{\partial F}{\partial S}$ and subtract it from the former

$$\Delta F - \frac{\partial F}{\partial S}(t, S) \Delta S = \left(\frac{\partial F}{\partial t}(t, S) + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial S^2}(t, S)\right) \Delta t.$$

The left side can be regarded as the increment $\Delta P$, of the portfolio

$$P = F - \frac{\partial F}{\partial S} S.$$

This portfolio is risk-less because its increments are totally deterministic, so it must also satisfy $\Delta P = rP \Delta t$. The number $\frac{\partial F}{\partial S}$ is assumed constant for small intervals of time $\Delta t$. Even if this assumption is not rigorous enough, the procedure still leads to the right equation. This is obtained by equating the coefficients of $\Delta t$ in the last two equations

$$\frac{\partial F}{\partial t}(t, S) + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial S^2}(t, S) = r\left(F - \frac{\partial F}{\partial S} S\right),$$

which is equivalent to the Black-Scholes equation.
8.5 Delta Hedging

The proof for the Black-Scholes’ equation is based on the fact that the portfolio $P = F - \frac{\partial F}{\partial S}S$ is risk-less. Since the delta of the derivative $F$ is

$$\Delta_F = \frac{dF}{dS} = \frac{\partial F}{\partial S},$$

then the portfolio $P = F - \Delta_F S$ is risk-less. This leads to the delta-hedging procedure, by which selling $\Delta_F$ units of the underlying stock $S$ yields a risk-less investment.

**Exercise 8.5.1** Find the value of the portfolio $P = F - \Delta_F S$ in the case when $F$ is a call option.

$$P = F - \Delta_F S = c - N(d_1)S = -Ke^{-r(T-t)}.$$

8.6 Tradable securities

A derivative $F(t, S)$ that is a solution of the Black-Scholes equation is called tradeable. Its name comes from the fact that it can be traded (either on an exchange or over-the-counter). The Black-Scholes equation constitutes the equilibrium relation that provides the traded price of the derivative. We shall deal next with a few examples of tradable securities.

**Example 8.6.1** (i) It is easy to show that $F = S$ is a solution of the Black-Scholes equation. Hence the stock is a tradable security.

(ii) If $K$ is a constant, then $F = e^{rt}K$ is a tradable security.

(iii) If $S$ is the stock price, then $F = e^S$ is not a tradable security, since $F$ does not satisfy equation (8.4.11).

**Exercise 8.6.1** Show that $F = \ln S$ is not a tradable security.

**Exercise 8.6.2** Find all constants $\alpha$ such that $S^\alpha$ is tradeable.

Substituting $F = S^\alpha$ into equation (8.4.11) we obtain

$$rS\alpha S^{\alpha-1} + \frac{1}{2}\sigma^2 S^2 \alpha(\alpha - 1)S^{\alpha-2} = rS^\alpha.$$

Dividing by $S^\alpha$ yields $r\alpha + \frac{1}{2}\sigma^2 \alpha(\alpha - 1) = r$. This can be factorized as

$$\frac{1}{2}\sigma^2 (\alpha - 1)(\alpha + \frac{2r}{\sigma^2}) = 0,$$
with two distinct solutions \( \alpha_1 = 1 \) and \( \alpha_2 = -\frac{2r}{\sigma^2} \). Hence there are only two tradable securities that are powers of the stock: the stock itself, \( S \), and \( S^{-2r/\sigma^2} \). In particular, \( S^2 \) is not tradeable, since \( -2r/\sigma^2 \neq 2 \) (the left side is negative). The role of these two cases will be clarified by the next result.

**Proposition 8.6.3** The general form of a traded derivative, which does not depend explicitly on time, is given by

\[
F(S) = C_1 S + C_2 S^{-2r/\sigma^2},
\]

(8.6.12)

with \( C_1, C_2 \) constants.

**Proof:** If the derivative depends solely on the stock, \( F = F(S) \), then the Black-Scholes equation becomes the ordinary differential equation

\[
rS \frac{dF}{dS} + \frac{1}{2} \sigma^2 S \frac{d^2F}{dS^2} = rF.
\]

(8.6.13)

This is an Euler-type equation, which can be solved by using the substitution \( S = e^x \). The derivatives \( \frac{d}{dS} \) and \( \frac{d}{dx} \) are related by the chain rule

\[
\frac{d}{dx} = S \frac{d}{dS}.
\]

Since \( \frac{dS}{dx} = \frac{de^x}{dx} = e^x = S \), it follows that \( \frac{d}{dx} = S \frac{d}{dS} \). Using the product rule,

\[
\frac{d^2}{dx^2} = S \frac{d}{dS} \left( S \frac{d}{dS} \right) = S \frac{d}{dS} + S^2 \frac{d^2}{dS^2},
\]

and hence

\[
S^2 \frac{d^2}{dS^2} = \frac{d^2}{dx^2} - \frac{d}{dx}.
\]

Substituting into (8.6.13) yields

\[
\frac{1}{2} \sigma^2 \frac{d^2G(x)}{dx^2} + \left( r - \frac{1}{2} \sigma^2 \right) \frac{dG(x)}{dx} = rG(x).
\]

where \( G(x) = G(e^x) = F(S) \). The associated indicial equation

\[
\frac{1}{2} \sigma^2 \alpha^2 - \left( r - \frac{1}{2} \sigma^2 \right) \alpha = r
\]

has solutions \( \alpha_1 = 1 \), \( \alpha_2 = -r/\sigma^2 \), so the general solution has the form

\[
G(x) = C_1 e^x + C_2 e^{-r/\sigma^2 x},
\]

which is equivalent with (8.6.12).
Exercise 8.6.4  Show that the price of a forward contract, which is given by
\[ F(t, S) = S - Ke^{-r(T-t)}, \]
satisfies the Black-Scholes equation, i.e. a forward contract is a tradable derivative.

Exercise 8.6.5  Show that the bond \( F(t) = e^{-r(T-t)}K \) is a tradable security.

Exercise 8.6.6  Let \( d_1 \) and \( d_2 \) be given by
\[
\begin{align*}
d_1 &= d_2 + \sigma \sqrt{T-t} \\
d_2 &= \frac{\ln(S_t/K) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma \sqrt{T-t}}.
\end{align*}
\]
Show that the following functions satisfy the Black-Scholes equation:
(a) \( F_1(t, S) = SN(d_1) \)
(b) \( F_2(t, S) = e^{-r(T-t)}N(d_2) \)
(c) \( F_2(t, S) = SN(d_1) - Ke^{-r(T-t)}N(d_2) \).

To which well-known derivatives do these formulas correspond?

8.7  Risk-less investment revised

A risk-less investment, \( P(t, S, F) \), which depends on time \( t \), stock price \( S \) and derivative \( F \), and has \( S \) as underlying asset, satisfies equation (8.3.10). Using the Black-Scholes equation satisfied by the derivative \( F \)
\[
\frac{\partial F}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} = rF - rS \frac{\partial F}{\partial S},
\]
equation (8.3.10) becomes
\[
\frac{\partial P}{\partial t} + \frac{\partial P}{\partial F} \left( rF - rS \frac{\partial F}{\partial S} \right) + \frac{1}{2} \sigma^2 S^2 \left[ \frac{\partial^2 P}{\partial S^2} + \frac{\partial^2 P}{\partial F^2} \left( \frac{\partial F}{\partial S} \right)^2 \right] = rP.
\]
Using the risk-less condition (8.3.8)
\[
\frac{\partial P}{\partial S} = -\frac{\partial P}{\partial F} \frac{\partial F}{\partial S}, \quad (8.7.14)
\]
the previous equation becomes
\[
\frac{\partial P}{\partial t} + rS \frac{\partial P}{\partial S} + rF \frac{\partial P}{\partial F} + \frac{1}{2} \sigma^2 S^2 \left[ \frac{\partial^2 P}{\partial S^2} + \frac{\partial^2 P}{\partial F^2} \left( \frac{\partial F}{\partial S} \right)^2 \right] = rP. \quad (8.7.15)
\]
In the following we shall find an equivalent expression for the last term on the left side. Differentiating in (8.7.14) with respect to $F$ yields

\[
\frac{\partial^2 P}{\partial F \partial S} = -\frac{\partial^2 P}{\partial F^2} \frac{\partial F}{\partial S} - \frac{\partial P}{\partial F} \frac{\partial^2 F}{\partial S},
\]

where we used

\[
\frac{\partial^2 F}{\partial F \partial S} = \frac{\partial}{\partial S} \frac{\partial F}{\partial F} = 1.
\]

Multiplying by $\frac{\partial F}{\partial S}$ implies

\[
\frac{\partial^2 P}{\partial F^2} \left( \frac{\partial F}{\partial S} \right)^2 = -\frac{\partial^2 P}{\partial F \partial S} \frac{\partial F}{\partial S}.
\]

Substituting in the aforementioned equation yields the Black-Scholes equation for portfolios

\[
\frac{\partial P}{\partial t} + rS \frac{\partial P}{\partial S} + rF \frac{\partial P}{\partial F} + \frac{1}{2} \sigma^2 S^2 \left[ \frac{\partial^2 P}{\partial S^2} - \frac{\partial P}{\partial F} \frac{\partial^2 F}{\partial S} \right] = rP. \tag{8.7.16}
\]

We have seen in section 8.4 that $P = F - \frac{\partial F}{\partial S}$ is a risk-less investment, in fact a risk-less portfolio. We shall discuss in the following another risk-less investment.

**Application 8.7.1** If a risk-less investment $P$ has the variable $S$ and $F$ separable, i.e. it is the sum $P(S, F) = f(F) + g(S)$, with $f$ and $g$ smooth functions, then

\[
P(S, F) = F + c_1 S + c_2 S^{-2r/\sigma^2},
\]

with $c_1$, $c_2$ constants. The derivative $F$ is given by the formula

\[
F(t, S) = -c_1 S - c_2 S^{-2r/\sigma^2} + c_3 e^{rt}, \quad c_3 \in \mathbb{R}.
\]

Since $P$ has separable variables, the mixed derivative term vanishes, and the equation (8.7.16) becomes

\[
Sg'(S) + \frac{\sigma^2}{2r} S^2 g''(S) - g(S) = f(F) - F f'(F).
\]

There is a separation constant $C$ such that

\[
f(F) - F f'(F) = C
\]

\[
Sg'(S) + \frac{\sigma^2}{2r} S^2 g''(S) - g(S) = C.
\]
Dividing the first equation by \( F^2 \) yields the exact equation
\[
\left( \frac{1}{F} f(F) \right)' = -\frac{C}{F^2},
\]
with the solution \( f(F) = c_0 F + C \). To solve the second equation, let \( \kappa = \frac{\sigma^2}{2r} \). Then the substitution \( S = e^x \) leads to the ordinary differential equation with constant coefficients
\[
\kappa h''(x) + (1 - \kappa) h'(x) - h(x) = C,
\]
where \( h(x) = g(e^x) = g(S) \). The associated indicial equation
\[
\kappa \lambda^2 + (1 - \kappa) \lambda - 1 = 0
\]
has the solutions \( \lambda_1 = 1, \lambda_2 = -\frac{1}{\kappa} \). The general solution is the sum between the particular solution \( h_p(x) = -C \) and the solution of the associated homogeneous equation, which is \( h_0(x) = c_1 e^x + c_2 e^{-\frac{1}{\kappa} x} \). Then
\[
h(x) = c_1 e^x + c_2 e^{-\frac{1}{\kappa} x} - C.
\]
Going back to the variable \( S \), we get the general form of \( g(S) \)
\[
g(S) = c_1 S + c_2 S^{-2r/\sigma^2} - C,
\]
with \( c_1, c_2 \) constants. Since the constant \( C \) cancels by addition, we have the following formula for the risk-less investment with separable variables \( F \) and \( S \):
\[
P(S, F) = f(F) + g(S) = c_0 F + c_1 S + c_2 S^{-2r/\sigma^2}.
\]
Dividing by \( c_0 \), we may assume \( c_0 = 1 \). We shall find the derivative \( F(t, S) \) which enters the previous formula. Substituting in (8.7.14) yields
\[
\frac{\partial F}{\partial S} = c_1 - \frac{2r}{\sigma^2} c_2 S^{1-2r/\sigma^2},
\]
which after partial integration in \( S \) gives
\[
F(t, S) = -c_1 S - c_2 S^{-2r/\sigma^2} + \phi(t),
\]
where the integration constant \( \phi(t) \) is a function of \( t \). The sum of the first two terms is the derivative given by formula (8.6.12). The remaining function \( \phi(t) \) has also to satisfy the Black-Scholes equation, and hence it is of the form \( \phi(t) = c_3 e^{rt} \), with \( c_3 \) constant. Then the derivative \( F \) is given by
\[
F(t, S) = -c_1 S - c_2 S^{-2r/\sigma^2} + c_3 e^{rt}.
\]
It is worth noting that substituting in the formula of \( P \) yields \( P = c_3 e^{rt} \), which agrees with the formula of a risk-less investment.
**Example 8.7.1** Find the function $g(S)$ such that the product $P = Fg(S)$ is a risk-less investment, with $F = F(t, S)$ derivative. Find the expression of the derivative $F$ in terms of $S$ and $t$.

**Proof:** Substituting $P = Fg(S)$ into equation (8.7.15) and simplifying by $rF$ yields

$$S \frac{dg(S)}{dS} + \frac{\sigma^2}{2r} S^2 \frac{d^2g(S)}{dS^2} = 0.$$ 

Substituting $S = e^x$, and $h(x) = g(e^x) = g(S)$ yields

$$h''(x) + \left( \frac{2r}{\sigma^2} - 1 \right) h'(x) = 0.$$ 

Integrating leads to the solution

$$h(x) = C_1 + C_2 e^{(1 - \frac{2r}{\sigma^2})x}.$$ 

Going back to variable $S$

$$g(S) = h(\ln S) = C_1 + C_2 e^{(1 - \frac{2r}{\sigma^2})\ln S} = C_1 + C_2 S^{1 - \frac{2r}{\sigma^2}}.$$ 

\[\blacksquare\]

**8.8 Solving the Black-Scholes**

In this section we shall solve the Black-Scholes equation and show that its solution coincides with the one provided by the risk-neutral evaluation in section 6. This way, the Black-Scholes equation provides a variant approach for European-type derivatives by using partial differential equations instead of expectations.

Consider a European-type derivative $F$, with the payoff at maturity $T$ given by $f_T$, which is a function of the stock price at maturity, $S_T$. Then $F(t, S)$ satisfies the following final condition partial differential equation

$$\frac{\partial F}{\partial t} + rS \frac{\partial F}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} = rF$$

$$F(T, S_T) = f_T(S_T).$$

This means the solution is known at the final time $T$ and we need to find its expression at any time $t$ prior to $T$, i.e.

$$f_t = F(t, S_t), \quad 0 \leq t < T.$$
First we shall transform the equation into an equation with constant coefficients. Substituting \( S = e^x \), and using the identities

\[
S \frac{\partial}{\partial S} = \frac{\partial}{\partial x}, \quad S^2 \frac{\partial^2}{\partial S^2} = \frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x}
\]

the equation becomes

\[
\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial x^2} + (r - \frac{1}{2} \sigma^2) \frac{\partial V}{\partial x} = rV,
\]

where \( V(t, x) = F(t, e^x) \). Using the time scaling \( \tau = \frac{1}{2} \sigma^2(T - t) \), the chain rule provides

\[
\frac{\partial}{\partial t} = \frac{\partial \tau}{\partial t} \frac{\partial}{\partial \tau} = -\frac{1}{2} \sigma^2 \frac{\partial}{\partial \tau}.
\]

Denote \( k = \frac{2r}{\sigma^2} \). Substituting in the aforementioned equation yields

\[
\frac{\partial W}{\partial \tau} = \frac{\partial^2 W}{\partial x^2} + (k - 1) \frac{\partial W}{\partial x} - kW, \tag{8.8.17}
\]

where \( W(\tau, x) = V(t, x) \). Next we shall get rid of the last two terms on the right side of the equation by using a crafted substitution.

Consider \( W(\tau, x) = e^\varphi u(\tau, x) \), where \( \varphi = \alpha x + \beta \tau \), with \( \alpha, \beta \) constants that will be determined such that the equation satisfied by \( u(\tau, x) \) has on the right side only the second derivative in \( x \). Since

\[
\frac{\partial W}{\partial x} = e^\varphi \left( \alpha u + \frac{\partial u}{\partial x} \right)
\]

\[
\frac{\partial^2 W}{\partial x^2} = e^\varphi \left( \alpha^2 u + 2\alpha \frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2} \right)
\]

\[
\frac{\partial W}{\partial \tau} = e^\varphi \left( \beta u + \frac{\partial u}{\partial \tau} \right),
\]

substituting in (8.8.17), dividing by \( e^\varphi \) and collecting the derivatives yields

\[
\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2} + (2\alpha + k - 1) \frac{\partial u}{\partial x} + (\alpha^2 + \alpha(k - 1) - k - \beta) u = 0
\]

The constants \( \alpha \) and \( \beta \) are chosen such that the coefficients of \( \frac{\partial u}{\partial x} \) and \( u \) vanish

\[
2\alpha + k - 1 = 0
\]

\[
\alpha^2 + \alpha(k - 1) - k - \beta = 0.
\]
Solving yields

\[ \alpha = -\frac{k - 1}{2} \]
\[ \beta = \alpha^2 + \alpha(k - 1) - k = -\frac{(k + 1)^2}{4}. \]

The function \( u(\tau, x) \) satisfies the heat equation

\[ \frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2} \]

with the initial condition expressible in terms of \( f_T \)

\[ u(0, x) = e^{-\varphi(0, x)} W(0, x) = e^{-\alpha x} V(T, x) = e^{-\alpha x} F(T, e^x) = e^{-\alpha x} f_T(e^x). \]

From the general theory of heat equation, the solution can be expressed as the convolution between the fundamental solution and the initial condition

\[ u(\tau, x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi \tau}} e^{-\frac{(y-x)^2}{4\tau}} u(0, y) \, dy \]

The previous substitutions yield the following relation between \( F \) and \( u \)

\[ F(t, S) = F(t, e^x) = V(t, x) = W(\tau, x) = e^{\varphi(\tau, x)} u(\tau, x), \]

so \( F(T, e^x) = e^{\alpha x} u(0, x) \). This implies

\[ F(t, e^x) = e^{\varphi(\tau, x)} u(\tau, x) = e^{\varphi(\tau, x)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi \tau}} e^{-\frac{(y-x)^2}{4\tau}} u(0, y) \, dy \]
\[ = e^{\varphi(\tau, x)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi \tau}} e^{-\frac{(y-x)^2}{4\tau}} e^{-\alpha y} F(T, e^y) \, dy. \]

With the substitution \( y = x = s \sqrt{2\tau} \) this becomes

\[ F(t, e^x) = e^{\varphi(\tau, x)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2} - \alpha(x + s \sqrt{2\tau})} F(T, e^{x+s\sqrt{2\tau}}) \, ds. \]

Completing the square as

\[ -\frac{s^2}{2} - \alpha(x + s \sqrt{2\tau}) = \frac{1}{2} \left( s - \frac{k - 1}{2} \sqrt{2\tau} \right)^2 + \frac{(k - 1)^2 \tau}{4} + \frac{k - 1}{2} \alpha x, \]

after cancelations, the previous integral becomes

\[ F(t, e^x) = e^{-\frac{(k+1)^2 \tau}{4}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} \left( s - \frac{k - 1}{2} \sqrt{2\tau} \right)^2} e^{\frac{(k-1)^2 \tau}{4}} F(T, e^{x+s\sqrt{2\tau}}) \, ds. \]
Using
\[ e^{-\frac{(k+1)^2}{4} \tau} e^{\frac{(k-1)^2}{4} \tau} = e^{-k \tau} = e^{-r(T-t)}, \]
\[ (k - 1) \tau = (r - \frac{1}{2} \sigma^2) (T - t), \]

after the substitution \( z = x + s \sqrt{2\tau} \) we get
\[
F(t, e^x) = e^{-r(T-t)} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} \frac{(z-x-(k-1)\tau)^2}{2\tau}} F(T, e^z) \frac{1}{\sqrt{2\tau}} dz
\]
\[ = e^{-r(T-t)} \frac{1}{\sqrt{2\pi \sigma^2 (T-t)}} \int_{-\infty}^{\infty} e^{- \frac{(z-x-(r-\frac{1}{2} \sigma^2)(T-t))^2}{2\sigma^2 (T-t)}} F(T, e^z) dz. \]

Since \( e^x = S_t \), considering the probability density
\[
p(z) = \frac{1}{\sqrt{2\pi \sigma^2 (T-t)}} e^{- \frac{(z - \ln S_t - (r-\frac{1}{2} \sigma^2)(T-t))^2}{2\sigma^2 (T-t)}},
\]

the previous expression becomes
\[
F(t, S_t) = e^{-r(T-t)} \int_{-\infty}^{\infty} p(z) f_T(e^z) dz = e^{-r(T-t)} \hat{E}_t[f_T],
\]

with \( f_T(S_T) = F(T, S_T) \) and \( \hat{E}_t \) the risk-neutral expectation operator as of time \( t \), which was introduced and used in section 6.

### 8.9 Black-Scholes and Risk-neutral Valuation

The conclusion of the computation of the last section is of capital importance for derivatives calculus. It shows the equivalence between the Black-Scholes equation and the risk-neutral evaluation. It turns out that instead of computing the risk-neutral expectation of the payoff, as in the case of the risk-neutral evaluation, we may have the choice to solve the Black-Scholes equation directly, and impose the final condition to be the payoff.

In many cases solving a partial differential equation is simpler than evaluating the expectation integral. This is due to the fact that we may look for a solution dictated by the particular form of the payoff \( f_T \). We shall apply that in finding put-call parities for different types of derivatives.

Consequently, all derivatives evaluated by the risk-neutral valuation are solutions of the Black-Scholes equation. The only distinction is their payoff. A few of them are given in the next example.
Example 8.9.1 (a) The price of a European call option is the solution \( F(t, S) \) of the Black-Scholes equation satisfying
\[
f_T(S_T) = \max(S_T - K, 0).
\]

(b) The price of a European put option is the solution \( F(t, S) \) of the Black-Scholes equation with the final condition
\[
f_T(S_T) = \max(K - S_T, 0).
\]

(c) The value of a forward contract is the solution the Black-Scholes equation with the final condition
\[
f_T(S_T) = S_T - K.
\]

It is worth noting that the superposition principle discussed in section 6 can be explained now by the fact that the solution space of the Black-Scholes equation is a linear space. This means that a linear combination of solutions is also a solution.

Another interesting feature of the Black-Scholes equation is its independence of the stock drift rate \( \mu \). Then its solutions must have the same property. This explains why, in the risk-neutral valuation, the value of \( \mu \) does not appear explicitly in the solution.

Asian options satisfy similar Black-Scholes equations, with small differences, as we shall see in the next section.

8.10 Boundary Conditions

We have solved the Black-Scholes equation for a call option, under the assumption that there is a unique solution. The Black-Scholes equation is of first order in the time variable \( t \) and of second order in the stock variable \( S \), so it needs one final condition at \( t = T \) and two boundary conditions for \( S = 0 \) and \( S \to \infty \).

In the case of a call option, the final condition is given by the following payoff:
\[
F(T, S_T) = \max\{S_T - K, 0\}.
\]

When \( S \to 0 \), the option does not get exercised, so the initial boundary condition is
\[
F(t, 0) = 0.
\]

When \( S \to \infty \) the price becomes linear
\[
F(t, S) \sim S - K,
\]
the graph of \( F(\cdot, S) \) having a slant asymptote, see Fig.8.2.
Figure 8.2: NEEDS TO BE REDONE The graph of the option price before maturity in the case $K = 40$, $\sigma = 30\%$, $r = 8\%$, and $T - t = 1$.

### 8.11 Risk-less Portfolios for Rare Events

Consider the derivative $P = P(t, S, F)$, which depends on the time $t$, stock price $S$ and the derivative $F$, whose underlying asset is $S$. We shall find the stochastic differential equation followed by $P$, under the hypothesis that the stock exhibits rare events, i.e.

$$dS = \mu S dt + \sigma S dW_t + \rho S dM_t,$$

where the constant $\mu, \sigma, \rho$ denote the drift rate, volatility and jump in the stock price in the case of a rare event. The processes $W_t$ and $M_t = N_t - \lambda t$ denote the Brownian motion and the compensated Poisson process, respectively. The constant $\lambda > 0$ denotes the rate of occurrence of the rare events in the market.

By Ito’s formula we get

$$dP = \frac{\partial P}{\partial t} dt + \frac{\partial P}{\partial S} dS + \frac{\partial P}{\partial F} dF + \frac{1}{2} \frac{\partial^2 P}{\partial S^2} dS^2 + \frac{1}{2} \frac{\partial^2 P}{\partial F^2} (dF)^2. \quad (8.11.19)$$

We shall use the following stochastic relations:

$$(dW_t)^2 = dt, \quad (dM_t)^2 = dN_t, \quad dt^2 = dt dW_t = dt dM_t = dW_t dM_t = 0,$$

see sections ?? and ??.

Then

$$\begin{align*}
(dS)^2 &= \sigma^2 S^2 dt + \rho^2 S^2 dN_t \\
&= (\sigma^2 + \lambda \rho^2) S^2 dt + \rho^2 S^2 dM_t, \quad (8.11.20)
\end{align*}$$

where we used $dM_t = dN_t - \lambda dt$. It is worth noting that the unpredictable part of $(dS)^2$ depends only on the rare events, and does not depend on the regular daily events.
Exercise 8.11.1 If $S$ satisfies (8.11.18), find the following

(a) $E[(dS)^2]$  
(b) $E[dS]$  
(c) $\text{Var}[dS]$.

Using Ito’s formula, the infinitesimal change in the value of the derivative $F = F(t, S)$ is given by

$$
dF = \frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial S} dS + \frac{1}{2} \frac{\partial^2 F}{\partial S^2} (dS)^2
$$

$$
= \left( \frac{\partial F}{\partial t} + \mu S \frac{\partial F}{\partial S} + \frac{1}{2} (\sigma^2 + \lambda \rho^2) S^2 \frac{\partial^2 F}{\partial S^2} \right) dt
$$

$$
+ \sigma S \frac{\partial F}{\partial S} dW_t
$$

$$
+ \left( \frac{1}{2} \rho^2 S^2 \frac{\partial^2 F}{\partial S^2} + \rho S \frac{\partial F}{\partial S} \right) dM_t,
$$

(8.11.21)

where we have used (8.11.18) and (8.11.20). The increment $dF$ has two independent sources of uncertainty: $dW_t$ and $dM_t$, both with mean equal to 0.

Taking the square, yields

$$
(dF)^2 = \sigma^2 S^2 \left( \frac{\partial F}{\partial S} \right)^2 dt + \left( \frac{1}{2} \rho^2 S^2 \frac{\partial^2 F}{\partial S^2} + \rho S \frac{\partial F}{\partial S} \right) dN_t.
$$

(8.11.22)

Substituting back in (8.11.19), we obtain the unpredictable part of $dP$ as the sum of two components

$$
\sigma S \left( \frac{\partial P}{\partial S} + \frac{\partial P}{\partial F} \frac{\partial F}{\partial S} \right) dW_t
$$

$$
+ \rho S \left( \frac{\partial P}{\partial S} + \frac{\partial P}{\partial F} \frac{\partial F}{\partial S} \right) + \frac{1}{2} \rho^2 S^2 \left( \frac{\partial P}{\partial F} \frac{\partial^2 F}{\partial S^2} + \frac{\partial^2 P}{\partial S^2} \frac{\partial F}{\partial S} + \frac{\partial^2 P}{\partial F^2} \frac{\partial^2 F}{\partial S^2} \right) dM_t.
$$

The risk-less condition for the portfolio $P$ is obtained when the coefficients of $dW_t$ and $dM_t$ vanish

$$
\frac{\partial P}{\partial S} + \frac{\partial P}{\partial F} \frac{\partial F}{\partial S} = 0 \quad \text{(8.11.23)}
$$

$$
\frac{\partial P}{\partial F} \frac{\partial^2 F}{\partial S^2} + \frac{\partial^2 P}{\partial S^2} \frac{\partial F}{\partial S} + \frac{\partial^2 P}{\partial F^2} \frac{\partial^2 F}{\partial S^2} = 0. \quad \text{(8.11.24)}
$$

These relations can be further simplified. If we differentiate in (8.11.23) with respect to $S$

$$
\frac{\partial^2 P}{\partial S^2} + \frac{\partial P}{\partial F} \frac{\partial^2 F}{\partial S^2} = - \frac{\partial^2 P}{\partial S \partial F} \frac{\partial F}{\partial S}.$$
Substituting into (8.11.24) yields
\[
\frac{\partial^2 P}{\partial F^2} \frac{\partial^2 F}{\partial S^2} = \frac{\partial^2 P}{\partial S \partial F} \frac{\partial F}{\partial S}.
\] (8.11.25)

Differentiating in (8.11.23) with respect to \( F \) we get
\[
\frac{\partial^2 P}{\partial F \partial S} = -\frac{\partial^2 P}{\partial F^2} \frac{\partial F}{\partial S} - \frac{\partial P}{\partial F} \frac{\partial^2 F}{\partial F \partial S},
\] (8.11.26)

since
\[
\frac{\partial^2 F}{\partial F \partial S} = \frac{\partial}{\partial S} \left( \frac{\partial F}{\partial F} \right).
\]

Multiplying (8.11.26) by \( \frac{\partial F}{\partial S} \) yields
\[
\frac{\partial^2 P}{\partial F \partial S} \frac{\partial F}{\partial S} = -\frac{\partial^2 P}{\partial F^2} \left( \frac{\partial F}{\partial S} \right)^2,
\]
and substituting in the right side of (8.11.25) leads to the equation
\[
\frac{\partial^2 P}{\partial F^2} \left[ \frac{\partial^2 F}{\partial S^2} + \left( \frac{\partial F}{\partial S} \right)^2 \right] = 0.
\]

We have arrived at the following result:

**Proposition 8.11.2** Let \( F = F(t, S) \) be a derivative with the underlying asset \( S \). The investment \( P = P(t, S, F) \) is risk-less if and only if
\[
\frac{\partial P}{\partial S} + \frac{\partial P}{\partial F} \frac{\partial F}{\partial S} = 0
\]
\[
\frac{\partial^2 P}{\partial F^2} \left[ \frac{\partial^2 F}{\partial S^2} + \left( \frac{\partial F}{\partial S} \right)^2 \right] = 0.
\]

There are two risk-less conditions because there are two unpredictable components in the increments of \( dP \), one due to regular changes and the other due to rare events. The first condition is equivalent with the vanishing total derivative, \( \frac{dP}{dS} = 0 \), and corresponds to offsetting the regular risk.

The second condition vanishes either if \( \frac{\partial^2 P}{\partial F^2} = 0 \) or if \( \frac{\partial^2 F}{\partial S^2} + \left( \frac{\partial F}{\partial S} \right)^2 = 0 \). In the first case \( P \) is linear in \( F \). For instance, if \( P = F - f(S) \), from the first condition yields \( f'(S) = \frac{\partial F}{\partial S} \). In the second case, denote \( U(t, S) = \frac{\partial F}{\partial S} \). Then we need to solve the partial differential equation
\[
\frac{\partial U}{\partial t} + U^2 = 0.
\]

**Future research directions:**
1. Solve the above equation.

2. Find the predictable part of $dP$.

3. Get an analog of the Black-Scholes in this case.

4. Evaluate a call option in this case.

5. Is the risk-neutral valuation still working and why?
Chapter 9

Black-Scholes for Asian Derivatives

In this chapter we shall develop the Black-Scholes equation in the case of Asian derivatives and we shall discuss the particular cases of options and forward contracts on weighted averages. In the case of the latter contracts we obtain closed form solutions, while for the former ones we apply the reduction variable method to decrease the number of variables and discuss the solution.

9.1 Weighted averages

In many practical problems the asset price needs to be considered with a certain weight. For instance, when computing car insurance, more weight is assumed for recent accidents than for accidents that occurred 10 years ago.

In the following we shall define the weight function and provide several examples.

Let \( \rho : [0, T] \to \mathbb{R} \) be a weight function, i.e. a function satisfying

1. \( \rho > 0; \)
2. \( \int_0^T \rho(t) \, dt = 1. \)

The stock weighted average with respect to the weight \( \rho \) is defined as

\[
S_{\text{ave}} = \int_0^T \rho(t) S_t \, dt.
\]

Example 9.1.1 (a) The uniform weight is obtained for \( \rho(t) = \frac{1}{T} \). In this case

\[
S_{\text{ave}} = \frac{1}{T} \int_0^T S_t \, dt
\]
is the continuous arithmetic average of the stock on the time interval \([0, T]\).

(b) The linear weight is obtained if \(\rho(t) = \frac{2t}{T^2}\). In this case the weight is the time

\[
S_{\text{ave}} = \frac{2}{T^2} \int_0^T tS_t \, dt.
\]

(c) The exponential weight is obtained for \(\rho(t) = \frac{ke^{kt}}{e^{kT} - 1}\). If \(k > 0\), the weight is increasing, so recent data are weighted more than old data; if \(k < 0\), the weight is decreasing. The exponential weighted average is given by

\[
S_{\text{ave}} = \frac{k}{e^{kT} - 1} \int_0^T e^{kt} S_t \, dt.
\]

Exercise 9.1.2 Consider the polynomial weighted average

\[
S^{(n)}_{\text{ave}} = \frac{n + 1}{T^{n+1}} \int_0^T t^n S_t \, dt.
\]

Find the limit \(\lim_{n \to \infty} S^{(n)}_{\text{ave}}\) in the cases \(0 < T < 1\), \(T = 1\), and \(T > 1\).

In all previous examples \(\rho(t) = \rho(t, T) = \frac{f(t)}{g(T)}\), with \(\int_0^T f(t) \, dt = g(T)\), so \(g'(T) = f(T)\) and \(g(0) = 0\). The average becomes

\[
S_{\text{ave}}(T) = \frac{1}{g(T)} \int_0^T f(u) S_u \, du = \frac{I_T}{g(T)},
\]

with \(I_t = \int_0^t f(u) S_u \, du\) satisfying \(dI_t = f(t) S_t \, dt\). From the product rule we get

\[
d S_{\text{ave}}(t) = \frac{dI_t}{g(t)^2} \left( g(t) S_t - g'(t) \frac{I_t}{g(t)} \right) \, dt
\]

\[
= \frac{f(t)}{g(t)} \left( S_t - \frac{g'(t)}{f(t)} S_{\text{ave}}(t) \right) \, dt
\]

\[
= \frac{f(t)}{g(t)} \left( S_t - S_{\text{ave}}(t) \right) \, dt,
\]

since \(g'(t) = f(t)\). The initial condition is

\[
S_{\text{ave}}(0) = \lim_{t \searrow 0} S_{\text{ave}}(t) = \lim_{t \searrow 0} \frac{I_t}{g(t)} = \lim_{t \searrow 0} \frac{f(t) S_t}{g'(t)} = S_0 \lim_{t \searrow 0} \frac{f(t)}{g'(t)} = S_0,
\]
Proposition 9.1.3 The weighted average $S_{\text{ave}}(t)$ satisfies the stochastic differential equation

$$dX_t = \frac{f(t)}{g(t)}(S_t - X_t)dt$$

$$X_0 = S_0.$$ 

Exercise 9.1.4 Let $x(t) = E[S_{\text{ave}}(t)]$.

(a) Show that $x(t)$ satisfies the ordinary differential equation

$$x'(t) = \frac{f(t)}{g(t)}(S_0e^{\mu t} - x(t))$$

$$x(0) = S_0.$$ 

(b) Find $x(t)$.

Exercise 9.1.5 Let $y(t) = E[S_{\text{ave}}^2(t)]$.

(a) Find the stochastic differential equation satisfied by $S_{\text{ave}}^2(t)$.

(b) Find the ordinary differential equation satisfied by $y(t)$.

(c) Solve the previous equation to get $y(t)$ and compute $\text{Var}[S_{\text{ave}}]$.

9.2 Setting up the Black-Scholes Equation

Consider an Asian derivative whose value at time $t$, $F(t_t, S_t, S_{\text{ave}}(t))$, depends on variables $t$, $S_t$, and $S_{\text{ave}}(t)$. Using the stochastic process of $S_t$

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

and Proposition 9.1.3, an application of Ito’s formula together with the stochastic formulas

$$d t^2 = 0, \quad (dW_t)^2 = 0, \quad (dS_t)^2 = \sigma^2 S_t^2 dt, \quad (dS_{\text{ave}})^2 = 0$$

yields

$$dF = \frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} (dS_t)^2 + \frac{\partial F}{\partial S_{\text{ave}}} dS_{\text{ave}}$$

$$= \left( \frac{\partial F}{\partial t} + \mu S_t \frac{\partial F}{\partial S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 F}{\partial S_t^2} + f(t) \frac{g(t)}{g(t)} (S_t - S_{\text{ave}}) \frac{\partial F}{\partial S_{\text{ave}}} \right) dt$$

$$+ \sigma S_t \frac{\partial F}{\partial S_t} dW_t.$$
Let $\Delta_F = \frac{\partial F}{\partial t}$. Consider the following portfolio at time $t$

$$P(t) = F - \Delta_F S_t,$$

obtained by buying one derivative $F$ and selling $\Delta_F$ units of stock. The change in the portfolio value during the time $dt$ does not depend on $W_t$

$$dP = dF - \Delta_F dS_t = \left( \frac{\partial F}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 F}{\partial S_t^2} + \frac{f(t)}{g(t)} (S_t - S_{ave}) \frac{\partial F}{\partial S_{ave}} \right) dt \quad (9.2.1)$$

so the portfolio $P$ is risk-less. Since no arbitrage opportunities are allowed, investing a value $P$ at time $t$ in a bank at the risk-free rate $r$ for the time interval $dt$ yields

$$dP = rP dt = \left( rF - rS_t \frac{\partial F}{\partial S_t} \right) dt. \quad (9.2.2)$$

Equating (9.2.1) and (9.2.2) yields the following form of the Black-Scholes equation for Asian derivatives on weighted averages

$$\frac{\partial F}{\partial t} + rS_t \frac{\partial F}{\partial S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 F}{\partial S_t^2} + \frac{f(t)}{g(t)} (S_t - S_{ave}) \frac{\partial F}{\partial S_{ave}} = rF.$$  

9.3 Weighted Average Strike Call Option

In this section we shall use the reduction variable method to decrease the number of variables from three to two. Since $S_{ave}(t) = \frac{I_t}{g(t)}$, it is convenient to consider the derivative as a function of $t, S_t$ and $I_t$

$$V(t, S_t, I_t) = F(t, S_t, S_{ave}).$$

A computation similar to the previous one yields the simpler equation

$$\frac{\partial V}{\partial t} + rS_t \frac{\partial V}{\partial S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2} + f(t)S_t \frac{\partial V}{\partial I_t} = rV. \quad (9.3.3)$$

The payoff at maturity of an average strike call option can be written in the following form

$$V_T = \begin{cases} V(T, S_T, I_T) = \max\{S_T - S_{ave}(T), 0\} \\ = \max\{S_T - \frac{I_T}{g(T)}, 0\} = S_T \max\{1 - \frac{1}{g(T)} \frac{I_T}{S_T}, 0\} \\ = S_T L(T, R_T), \end{cases}$$
where 
\[ R_t = \frac{I_t}{S_t}, \quad L(t, R) = \max\{1 - \frac{1}{g(t)} R, 0\}. \]

Since at maturity the variable \( S_T \) is separated from \( T \) and \( R_T \), we shall look for a solution of equation (9.3.3) of the same type for any \( t \leq T \), i.e. \( V(t, S, I) = SG(t, R) \). Since

\[
\frac{\partial V}{\partial t} = S \frac{\partial G}{\partial t}, \quad \frac{\partial V}{\partial I} = S \frac{\partial G}{\partial R} \frac{1}{SR} = \frac{\partial G}{\partial R}; \\
\frac{\partial V}{\partial S} = G + S \frac{\partial G}{\partial R} \frac{\partial R}{\partial S} = G - R \frac{\partial G}{\partial R}; \\
\frac{\partial^2 V}{\partial S^2} = \frac{\partial}{\partial S}(G - R \frac{\partial G}{\partial R}) = \frac{\partial G}{\partial R} \frac{\partial R}{\partial S} - \frac{\partial R}{\partial S} \frac{\partial G}{\partial R} - R \frac{\partial^2 G}{\partial R^2} \frac{\partial R}{\partial S}; \\
= R \frac{\partial^2 G I}{\partial R^2 S}; \\
S^2 \frac{\partial^2 V}{\partial S^2} = RI \frac{\partial^2 G}{\partial R^2}.
\]

substituting in (9.3.3) and using that \( \frac{RI}{S} = R^2 \), after cancelations yields

\[
\frac{\partial G}{\partial t} + \frac{1}{2} \sigma^2 R^2 \frac{\partial^2 G}{\partial R^2} + (f(t) - rR) \frac{\partial G}{\partial R} = 0. \quad (9.3.4)
\]

This is a partial differential equation in only two variables, \( t \) and \( R \). It can be solved explicitly sometimes, depending on the form of the final condition \( G(T, R_T) \) and expression of the function \( f(t) \).

In the case of a weighted average strike call option the final condition is

\[
G(T, R_T) = \max\{1 - \frac{R_T}{g(T)}, 0\}. \quad (9.3.5)
\]

**Example 9.3.1** In the case of the arithmetic average the function \( G(t, R) \) satisfies the partial differential equation

\[
\frac{\partial G}{\partial t} + \frac{1}{2} \sigma^2 R^2 \frac{\partial^2 G}{\partial R^2} + (1 - rR) \frac{\partial G}{\partial R} = 0
\]

with the final condition \( G(T, R_T) = \max\{1 - \frac{R_T}{T}, 0\} \).

**Example 9.3.2** In the case of the exponential average the function \( G(t, R) \) satisfies the equation

\[
\frac{\partial G}{\partial t} + \frac{1}{2} \sigma^2 R^2 \frac{\partial^2 G}{\partial R^2} + (ke^{kt} - rR) \frac{\partial G}{\partial R} = 0 \quad (9.3.6)
\]

with the final condition \( G(T, R_T) = \max\{1 - \frac{R_T}{e^{kT} - 1}, 0\} \).
Neither of the previous two final condition problems can be solved explicitly.

9.4 Boundary Conditions

The partial differential equation (9.3.4) is of first order in $t$ and second order in $R$. We need to specify one condition at $t = T$ (the payoff at maturity), which is given by (9.3.5), and two conditions for $R = 0$ and $R \to \infty$, which specify the behavior of solution $G(t, R)$ at two limiting positions of the variable $R$.

Taking $R \to 0$ in equation (9.3.4) and using Exercise 9.4.1 yields the first boundary condition for $G(t, R)$

$$
\left. \left( \frac{\partial G}{\partial t} + f \frac{\partial G}{\partial R} \right) \right|_{R=0} = 0.
$$

(9.4.7)

The term $\frac{\partial G}{\partial R}|_{R=0}$ represents the slope of $G(t, R)$ with respect to $R$ at $R = 0$, while $\frac{\partial^2 G}{\partial t^2}|_{R=0}$ is the variation of the price $G$ with respect to time $t$ when $R = 0$.

Another boundary condition is obtained by specifying the behavior of $G(t, R)$ for large values of $R$. If $R_t \to \infty$, we must have $S_t \to 0$, because

$$
R_t = \frac{1}{S_t} \int_0^t f(u)S_u \, du
$$

and $\int_0^t f(u)S_u \, du > 0$ for $t > 0$. In this case we are better off not exercising the option (since otherwise we get a negative payoff), so the boundary condition is

$$
\lim_{R \to \infty} G(R, t) = 0.
$$

(9.4.8)

It can be shown in the theory of partial differential equations that equation (9.3.4) together with the final condition (9.3.5), see Fig.9.1(a), and boundary conditions (9.4.7) and (9.4.8) has a unique solution $G(t, R)$, see Fig.9.1(b).

Exercise 9.4.1 Let $f$ be a bounded differentiable function. Show that

(a) $\lim_{x \to 0} xf'(x) = 0$;

(b) $\lim_{x \to 0} x^2 f''(x) = 0$.

There is no close form solution for the weighted average strike call option. Even in the simplest case, when the average is arithmetic, the solution is just approximative, see section 6.13. In real life the price is worked out using the Monte-Carlo simulation. This is based on averaging a large number, $n$, of simulations of the process $R_t$ in the risk-neutral world, i.e. assuming $\mu = r$. 
For each realization, the associated payoff $G_{T,j} = \max\{1 - \frac{R_{T,j}}{g(T)}\}$ is computed, with $j \leq n$. Here $R_{T,j}$ represents the value of $R$ at time $T$ in the $j$th realization. The average

$$\frac{1}{n} \sum_{j=1}^{n} G_{T,j}$$

is a good approximation of the payoff expectation $E[G_T]$. Discounting under the risk-free rate we get the price at time $t$

$$G(t, R) = e^{-r(T-t)} \left( \frac{1}{n} \sum_{j=1}^{n} G_{T,j} \right).$$

It is worth noting that the term on the right is an approximation of the risk neutral conditional expectation $\hat{E}[G_T|\mathcal{F}_t]$.

When simulating the process $R_t$, it is convenient to know its stochastic differential equation. Using

$$dI_t = f(t)S_t dt, \quad d\left(\frac{1}{S_t}\right) = \frac{1}{S_t} \left((\sigma^2 - \mu)dt - \sigma dW_t\right) dt,$$

the product rule yields

$$dR_t = d\left(\frac{I_t}{S_t}\right) = d\left(I_t \frac{1}{S_t}\right)$$

$$= dI_t \frac{1}{S_t} + I_t d\left(\frac{1}{S_t}\right) + dI_t d\left(\frac{1}{S_t}\right)$$

$$= f(t)dt + R_t \left((\sigma^2 - \mu)dt - \sigma dW_t\right).$$
Collecting terms yields the following stochastic differential equation for $R_t$:

$$dR_t = -\sigma R_t dW_t + (f(t) + (\sigma^2 - \mu) R_t) dt$$

(9.4.9)

The initial condition is $R_0 = \frac{I_0}{S_0} = 0$, since $I_0 = 0$.

Can we solve explicitly this equation? Can we find the mean and variance of $R_t$?

We shall start by finding the mean $\mathbb{E}[R_t]$. The equation can be written as

$$dR_t - (\sigma^2 - \mu) R_t dt = f(t) dt - \sigma R_t dW_t.$$

Multiplying by $e^{-(\sigma^2 - \mu)t}$ yields the exact equation

$$d\left(e^{-(\sigma^2 - \mu)t} R_t\right) = e^{-(\sigma^2 - \mu)t} f(t) dt - \sigma e^{-(\sigma^2 - \mu)t} R_t dW_t.$$

Integrating yields

$$e^{-(\sigma^2 - \mu)t} R_t = \int_0^t e^{-(\sigma^2 - \mu)u} f(u) du - \sigma \int_0^t e^{-(\sigma^2 - \mu)u} R_u dW_u.$$

The first integral is deterministic while the second is an Ito integral. Using that the expectations of Ito integrals vanish, we get

$$\mathbb{E}[e^{-(\sigma^2 - \mu)t} R_t] = \int_0^t e^{-(\sigma^2 - \mu)u} f(u) du$$

and hence

$$\mathbb{E}[R_t] = e^{(\sigma^2 - \mu)t} \int_0^t e^{-(\sigma^2 - \mu)u} f(u) du.$$

**Exercise 9.4.2** Find $\mathbb{E}[R_t^2]$ and $\text{Var}[R_t]$.

Equation (9.4.9) is a linear equation of the type discussed in section ??.

Multiplying by the integrating factor

$$\rho_t = e^{\sigma W_t + \frac{1}{2} \sigma^2 t}$$

the equation is transformed into an exact equation

$$d(\rho_t R_t) = (\rho_t f(t) + (\sigma^2 - \mu) \rho_t R_t) dt.$$

Substituting $Y_t = \rho_t R_t$ yields

$$dY_t = (\rho_t f(t) + (\sigma^2 - \mu) Y_t) dt,$$
which can be written as
\[ dY_t - (\sigma^2 - \mu)Y_t dt = \rho_t f(t) dt. \]

Multiplying by \( e^{-(\sigma^2 - \mu)t} \) yields the exact equation
\[ d(e^{-(\sigma^2 - \mu)t} Y_t) = e^{-(\sigma^2 - \mu)t} \rho_t f(t) dt, \]
which can be solved by integration
\[ e^{-(\sigma^2 - \mu)t} Y_t = \int_0^t e^{-(\sigma^2 - \mu)u} \rho_u f(u) du. \]

Going back to the variable \( R_t = Y_t/\rho_t \), we obtain the following closed form expression

\[
R_t = \int_0^t e^{(\mu - \frac{1}{2} \sigma^2)(u-t) + \sigma(W_u - W_t)} f(u) du. \tag{9.4.10}
\]

**Exercise 9.4.3** Find \( E[R_t] \) by taking the expectation in formula (9.4.10).

It is worth noting that we can arrive at formula (9.4.10) directly, without going through solving a stochastic differential equation. We shall show this procedure in the following.

Using the well-known formulas for the stock price
\[
S_u = S_0 e^{(\mu - \frac{1}{2} \sigma^2)u + \sigma W_u}, \quad S_t = S_0 e^{(\mu - \frac{1}{2} \sigma^2)t + \sigma W_t},
\]
and dividing, yields
\[
\frac{S_u}{S_t} = e^{(\mu - \frac{1}{2} \sigma^2)(u-t) + \sigma(W_u - W_t)}.
\]

Then we get
\[
R_t = \frac{I_t}{S_t} = \frac{1}{S_t} \int_0^t S_u f(u) du = \int_0^t \frac{S_u}{S_t} f(u) du = \int_0^t e^{(\mu - \frac{1}{2} \sigma^2)(u-t) + \sigma(W_u - W_t)} f(u) du,
\]
which is formula (9.4.10).

**Exercise 9.4.4** Find an explicit formula for \( R_t \) in terms of the integrated Brownian motion \( Z_t^{(\sigma)} = \int_0^t e^{\sigma W_u} du \), in the case of an exponential weight with \( k = \frac{1}{2} \sigma^2 - \mu \), see Example 9.1.1(c).
Exercise 9.4.5  (a) Find the price of a derivative $G$ which satisfies
\[
\frac{\partial G}{\partial t} + \frac{1}{2}\sigma^2 R^2 \frac{\partial^2 G}{\partial R^2} + (1 - rR) \frac{\partial G}{\partial R} = 0
\]
with the payoff $G(T, R_T) = R_T^2$.
(b) Find the value of an Asian derivative $V_t$ on the arithmetic average, that has the payoff
\[
V_T = V(T, S_T, I_T) = \frac{I_T^2}{S_T},
\]
where $I_T = \int_0^T S_t \, dt$.

Exercise 9.4.6  Use a computer simulation to find the value of an Asian arithmetic average strike option with $r = 4\%$, $\sigma = 50\%$, $S_0 = $40, and $T = 0.5$ years.

9.5 Asian Forward Contracts on Weighted Averages

Since the payoff of this derivative is given by
\[
V_T = S_T - S_{\text{ave}}(T) = S_T \left(1 - \frac{R_T}{g(T)}\right),
\]
the reduction variable method suggests considering a solution of the type $V(t, S_t, I_t) = S_t G(t, R_t)$, where $G(t, T)$ satisfies equation (9.3.4) with the final condition $G(T, R_T) = 1 - \frac{R_T}{g(T)}$. Since this is linear in $R_T$, this implies looking for a solution $G(t, R_t)$ in the following form
\[
G(t, R_t) = a(t)R_t + b(t),
\]
with functions $a(t)$ and $b(t)$ subject to be determined. Substituting into (9.3.4) and collecting $R_t$ yields
\[
(a'(t) - ra(t))R_t + b'(t) + f(t)a(t) = 0.
\]
Since this polynomial in $R_t$ vanishes for all values of $R_t$, then its coefficients are identically zero, so
\[
a'(t) - ra(t) = 0, \quad b'(t) + f(t)a(t) = 0.
\]
When $t = T$ we have
\[
G(T, R_T) = a(T)R_T + b(T) = 1 - \frac{R_T}{g(T)}.
\]
Equating the coefficients of $R_T$ yields the final conditions

$$a(T) = -\frac{1}{g(T)}, \quad b(T) = 1.$$  

The coefficient $a(t)$ satisfies the ordinary differential equation

$$a'(t) = ra(t) \quad a(T) = -\frac{1}{g(T)}$$

which has the solution

$$a(t) = -\frac{1}{g(T)} e^{-r(T-t)}.$$  

The coefficient $b(t)$ satisfies the equation

$$b'(t) = -f(t)a(t) \quad b(T) = 1$$

with the solution

$$b(t) = 1 + \int_t^T f(u)a(u) \, du.$$  

Substituting in (9.5.11) yields

$$G(t,R) = -\frac{1}{g(T)} e^{-r(T-t)} R_t + 1 + \int_t^T f(u)a(u) \, du = 1 - \frac{1}{g(T)} [R_t e^{-r(T-t)} + \int_t^T f(u)e^{-r(T-u)} \, du].$$

Then going back into the variable $I_t = S_t R_t$ yields

$$V(t,S_t,I_t) = S_t G(t,R_t) = S_t - \frac{1}{g(T)} [I_t e^{-r(T-t)} + S_t \int_t^T f(u)e^{-r(T-u)} \, du].$$

Using that $\rho(u) = \frac{f(u)}{g(T)}$ and going back to the initial variable $S_{ave}(t) = I_t/g(t)$ yields

$$F(t,S_t,S_{ave}(t)) = V(t,S_t,I_t) = S_t - \frac{g(t)}{g(T)} S_{ave}(t)e^{-r(T-t)} - S_t \int_t^T \rho(u)e^{-r(T-u)} \, du.$$  

We have arrived at the following result:
Proposition 9.5.1 The value at time $t$ of an Asian forward contract on a weighted average with the weight function $\rho(t)$, i.e. an Asian derivative with the payoff $F_T = S_T - S_{ave}(T)$, is given by

$$F(t, S_t, S_{ave}(t)) = S_t \left(1 - \int_t^T \rho(u)e^{-r(T-u)} du\right) - \frac{g(t)}{g(T)}e^{-r(T-t)}S_{ave}(t).$$

It is worth noting that the previous price can be written as a linear combination of $S_t$ and $S_{ave}(t)$

$$F(t, S_t, S_{ave}(t)) = \alpha(t)S_t + \beta(t)S_{ave}(t),$$

where

$$\alpha(t) = 1 - \int_t^T \rho(u)e^{-r(T-u)} du$$

$$\beta(t) = -\frac{g(t)}{g(T)}e^{-r(T-t)} = -\frac{\int_t^T f(u) du}{\int_t^T f(u) du}e^{-r(T-t)}.$$

In the first formula $\rho(u)e^{-r(T-u)}$ is the discounted weight at time $u$, and $\alpha(t)$ is 1 minus the total discounted weight between $t$ and $T$. One can easily check that $\alpha(T) = 1$ and $\beta(T) = -1$.

Exercise 9.5.2 Find the value at time $t$ of an Asian forward contract on an arithmetic average $A_t = \int_0^t S_u du$.

Exercise 9.5.3 (a) Find the value at time $t$ of an Asian forward contract on an exponential weighted average with the weight given by Example 9.1.1(c).

(b) What happens if $k = -r$? Why?

Exercise 9.5.4 Find the value at time $t$ of an Asian power contract with the payoff $F_T = (\int_0^T S_u du)^n$.
Chapter 10

American Options

American options are options that are allowed to be exercised at any time before maturity. Because of this advantage, they tend to be more expensive than the European counterparts. Exact pricing formulas exist just for perpetuities, while they are missing for finitely lived American options.

10.1 Perpetual American Options

A perpetual American option is an American option that never expires. These contracts can be exercised at any time $t$, $0 \leq t \leq \infty$. Even if finding the optimal exercise time for finite maturity American options is a delicate matter, in the case of perpetual American calls and puts there is always possible to find the optimal exercise time and to derive a close form pricing formula (see Merton, [15]).

10.1.1 Present Value of Barriers

Our goal in this section will be to learn how to compute the present value of a contract that pays a fixed cash amount at a stochastic time defined by the first passage of time of a stock. These formulas will be the main ingredient in pricing perpetual American options over the next couple of sections.

Reaching the barrier from below Let $S_t$ denote the stock price with initial value $S_0$ and consider a positive number $b$ such that $b > S_0$. We recall the first passage of time $\tau_b$ when the stock $S_t$ hits for the first time the barrier $b$, see Fig. 10.1

$$\tau_b = \inf\{t > 0; S_t = b\}.$$

Consider a contract that pays $1 at the time when the stock reaches the barrier $b$ for the first time. Under the constant interest rate assumption, the value of
the contract at time $t = 0$ is obtained by discounting the value of $\$1$ at the rate $r$ for the period $\tau_b$ and taking the expectation in the risk neutral world

$$f_0 = \hat{E}[e^{-r\tau_b}].$$

In the following we shall compute the right side of the previous expression using two different approaches. Using that the stock price in the risk-neutral world is given by the expression $S_t = S_0e^{(r-\frac{\sigma^2}{2})t+\sigma W_t}$, with $r > \frac{\sigma^2}{2}$, then

$$M_t = e^{-rt}S_t = S_0e^{\sigma W_t-\frac{\sigma^2}{2}t}, \quad t \geq 0$$

is a martingale. Applying the Optional Stopping Theorem (Theorem ??) yields $E[M_{\tau_b}] = E[M_0]$, which is equivalent to

$$E[e^{-r\tau_b}S_{\tau_b}] = S_0.$$

Since $S_{\tau_b} = b$, the previous relation implies

$$E[e^{-r\tau_b}] = \frac{S_0}{b},$$

where the expectation is taken in the risk-neutral world. Hence, we arrived at the following result:

**Proposition 10.1.1** The value at time $t = 0$ of $\$1$ received at the time when the stock reaches level $b$ from below is

$$f_0 = \frac{S_0}{b}.$$
Exercise 10.1.2 In the previous proof we had applied the Optional Stopping Theorem (Theorem ??). Show that the hypothesis of the theorem are satisfied.

Exercise 10.1.3 Let $0 < S_0 < b$ and assume $r > \frac{\sigma^2}{2}$.
(a) Show that $P(S_t \text{ reaches } b) = 1$. Compare with Exercise ??.
(b) Prove the identity $P(\omega; \tau_b(\omega) < \infty) = 1$.

The result of Proposition 10.1.1 can be also obtained directly as a consequence of Proposition ??.

Exercise 10.1.4 Let $S_0 < b$. Find the probability density function of the hitting time $\tau_b$.

Exercise 10.1.5 Assume the stock pays continuous dividends at the constant rate $\delta > 0$ and let $b > 0$ such that $S_0 < b$.
(a) Prove that the value at time $t = 0$ of $\$1$ received at the time when the stock reaches level $b$ from below is

$$f_0 = \left(\frac{S_0}{b}\right)^{h_1},$$

where

$$h_1 = \frac{1}{2} \frac{r - \delta}{\sigma^2} + \sqrt{\left(\frac{r - \delta}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2r}{\sigma^2}}.$$

(b) Show that $h_1(0) = 1$ and $h_1(\delta)$ is an increasing function for $\delta > 0$.
(c) Find the limit of the rate of change $\lim_{\delta \to \infty} h_1'(\delta)$.
(d) Work out a formula for the sensitivity of the value $f_0$ with respect to the dividend rate $\delta$ and compute the long run value of this rate.
Reaching the barrier from above Sometimes a stock can reach a barrier \( b \) from above. Let \( S_0 \) be the initial value of the stock \( S_t \) and assume the inequality \( b < S_0 \). Consider again the first passage of time \( \tau_b = \inf\{t > 0; S_t = b\} \), see Fig. 10.2.

In this paragraph we compute the value of a contract that pays $1 at the time when the stock reaches the barrier \( b \) for the first time, which is given by \( f_0 = \mathbb{E}[e^{-r\tau_b}] \). We shall keep the assumption that the interest rate \( r \) is constant and \( r > \frac{\sigma^2}{2} \).

**Proposition 10.1.6** The value at time \( t = 0 \) of $1 received at the time when the stock reaches level \( b \) from above is

\[
f_0 = \left( \frac{S_0}{b} \right)^{-\frac{2r}{\sigma^2}}.
\]

**Proof:** The reader might be tempted to use the Optional Stopping Theorem, but we refrain from applying it in this case (Why?) We should rather use a technique which reduces the problem to Proposition ???. Following this idea, we write

\[
\tau_b = \inf\{t > 0; S_t = b\} = \inf\{t > 0; (r - \frac{\sigma^2}{2})t + \sigma W_t = \ln \frac{b}{S_0}\}
\]

where \( x = \ln \frac{b}{S_0} > 0, \mu = r - \frac{\sigma^2}{2} \). Choosing \( s = r \) in Proposition ?? yields

\[
E[e^{-r\tau_b}] = e^{\frac{1}{\sigma^2} \left( \mu + \sqrt{2\sigma^2 + \mu^2} \right)x} = e^{\frac{1}{\sigma^2} \left( r - \frac{\sigma^2}{2} + \sqrt{2\sigma^2 + (r - \frac{\sigma^2}{2})^2} \right)x}
\]

\[= e^{-\frac{2rx}{\sigma^2}} = e^{-\frac{2r}{\sigma^2} \ln \frac{S_0}{b}} = \left( \frac{S_0}{b} \right)^{-\frac{2r}{\sigma^2}}.\]
In the previous computation we used that
\[
\begin{align*}
  r - \frac{\sigma^2}{2} + \sqrt{2r\sigma^2 + (r - \frac{\sigma^2}{2})^2} &= r - \frac{\sigma^2}{2} + \sqrt{(r + \frac{\sigma^2}{2})^2} \\
  &= r - \frac{\sigma^2}{2} + r + \frac{\sigma^2}{2} = 2r.
\end{align*}
\]

**Exercise 10.1.7** Assume the stock pays continuous dividends at the constant rate \(\delta > 0\) and let \(b > 0\) such that \(b < S_0\).

(a) Use a similar method as in the proof of Proposition 10.1.6 to prove that the value at time \(t = 0\) of $1 received at the time when the stock reaches level \(b\) from above is
\[
f_0 = \left(\frac{S_0}{b}\right)^{h_2},
\]
where
\[
h_2 = \frac{1}{2} - \frac{r - \delta}{\sigma^2} - \sqrt{\left(\frac{r - \delta}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2r}{\sigma^2}}.
\]

(b) What is the value of the contract at any time \(t\), with \(0 \leq t < \tau_b\)?

Perpetual American options have simple exact pricing formulas. This is because of the “time invariance” property of their values. Since the time to expiration for these type of options is the same (i.e infinity), the option exercise problem looks the same at every instance of time. Consequently, their value do not depend on the time to expiration.

**10.1.2 Perpetual American Calls**

A perpetual American call is a call option that never expires, i.e. is a contract that gives the holder the right to buy the stock for the price \(K\) at any instance of time \(0 \leq t \leq +\infty\). The infinity is included to cover the case when the option is never exercised.

When the call is exercised the holder receives \(S_\tau - K\), where \(\tau\) denotes the exercise time. Assume the holder has the strategy to exercise the call whenever the stock \(S_t\) reaches the barrier \(b\), with \(b > K\) subject to be determined later. Then at exercise time \(\tau_b\) the payoff is \(b - K\), where
\[
\tau_b = \inf\{t > 0; S_t = b\}.
\]

We note that it makes sense to choose the barrier such that \(S_0 < b\). The value of this amount at time \(t = 0\) is obtained discounting at the interest rate \(r\) and using Proposition 10.1.1
\[
f(b) = E[(b - K)e^{-r\tau_b}] = (b - K)\frac{S_0}{b} = \left(1 - \frac{K}{b}\right)S_0.
\]
We need to choose the value of the barrier $b > 0$ for which $f(b)$ reaches its maximum. Since $1 - \frac{K}{b}$ is an increasing function of $b$, the optimum value can be evaluated as

$$\max_{b>0} f(b) = \max_{b>0} \left( 1 - \frac{K}{b} \right) S_0 = \lim_{b \to \infty} \left( 1 - \frac{K}{b} \right) S_0 = S_0.$$

This is reached for the optimal barrier $b^* = \infty$, which corresponds to the infinite exercise time $\tau_b = \infty$. Hence, it is never optimal to exercise a perpetual call option on a nondividend paying stock.

The next exercise covers the case of the dividend paying stock. The method is similar with the one described previously.

**Exercise 10.1.8** Consider a stock that pays continuous dividends at rate $\delta > 0$.

(a) Assume a perpetual call is exercised whenever the stock reaches the barrier $b$ from below. Show that the discounted value at time $t = 0$ is

$$f(b) = (b - K) \left( \frac{S_0}{b} \right)^{h_1},$$

where

$$h_1 = \frac{1}{2} - \frac{r - \delta}{\sigma^2} + \sqrt{\left( \frac{r - \delta}{\sigma^2} - \frac{1}{2} \right)^2 + \frac{2r}{\sigma^2}}.$$

(b) Use differentiation to show that the maximum value of $f(b)$ is realized for

$$b^* = K \frac{h_1}{h_1 - 1}.$$

(c) Prove the price of perpetual call

$$f(b^*) = \max_{b>0} f(b) = \frac{K}{h_1 - 1} \left( \frac{h_1 - 1}{h_1} \frac{S_0}{K} \right)^{h_1}.$$

(d) Let $\tau_{b^*}$ be the exercise time of the perpetual call. When do you expect to exercise the call? (Find $E[\tau_{b^*}]$).

**10.1.3 Perpetual American Puts**

A perpetual American put is a put option that never expires, i.e. is a contract that gives the holder the right to sell the stock for the price $K$ at any instance of time $0 \leq t < \infty$.

Assume the put is exercised when $S_t$ reaches the barrier $b$. Then its payoff, $K - S_t = K - b$, has a couple of noteworthy features. First, if we choose $b$ too
large, we loose option value, which eventually vanishes for \( b \geq K \). Second, if
we pick \( b \) too small, the chances that the stock \( S_t \) will hit \( b \) are also small (see
Exercise 10.1.9), fact that diminishes the put value. It follows that the opti-
mum exercise barrier, \( b^* \), is somewhere in between the two previous extreme
values.

**Exercise 10.1.9** Let \( 0 < b < S_0 \) and let \( t > 0 \) fixed.

(a) Show that the following inequality holds in the risk neutral world
\[
P(S_t < b) \leq e^{-\frac{1}{2}\sigma^2 t \left[ \ln(S_t/b) + (r - \sigma^2/2)t \right]^2}.
\]

(b) Use the Squeeze Theorem to show that \( \lim_{b \to 0^+} P(S_t < b) = 0 \).

Since a put is an insurance that gets exercised when the the stock declines,
it makes sense to assume that at the exercise time, \( \tau_b \), the stock reaches the
barrier \( b \) from above, i.e \( 0 < b < S_0 \). Using Proposition 10.1.6 we obtain the
value of the contract that pays \( K - b \) at time \( \tau_b 
\]
\[
f(b) = E[(K - b)e^{-r\tau_b}] = (K - b)E[e^{-r\tau_b}] = (K - b)\left(\frac{S_0}{b}\right)^{-\frac{2r}{\sigma^2}}.
\]
We need to pick the optimal value \( b^* \) for which \( f(b) \) is maximum
\[
f(b^*) = \max_{0 < b < S_0} f(b).
\]
It is useful to notice that the functions \( f(b) \) and \( g(b) = (K - b)b^{\frac{2r}{\sigma^2}} \)
reach the maximum for the same value of \( b \). For the same of simplicity, denote \( \alpha = \frac{2r}{\sigma^2} \).
Then
\[
g(b) = Kb^\alpha - b^{\alpha+1} \\
g'(b) = \alpha Kb^{\alpha-1} - (\alpha + 1)b^\alpha = b^{\alpha-1}[\alpha K - (\alpha + 1)b],
\]
and the equation \( g'(b) = 0 \) has the solution \( b^* = \frac{\alpha}{\alpha + 1}K \). Since \( g'(b) > 0 \) for
\( b < b^* \) and \( g'(b) < 0 \) for \( b > b^* \), it follows that \( b^* \) is a maximum point for the
function \( g(b) \) and hence for the function \( f(b) \). Substituting for the value of \( \alpha \)
the optimal value of the barrier becomes
\[
b^* = \frac{2r/\sigma^2}{2r/\sigma^2 + 1}K = \frac{K}{1 + \frac{\sigma^2}{2r}}.
\](10.1.1)
The condition \( b^* < K \) is obviously satisfied, while the condition \( b^* < S_0 \) is
equivalent with
\[
K < \left(1 + \frac{\sigma^2}{2r}\right)S_0.
\]
The value of the perpetual put is obtained computing the value at $b^*$

$$f(b^*) = \max f(b) = (K - b^*) \left( \frac{S_0}{b^*} \right)^{-\frac{2r}{\sigma^2}} = \left( K - \frac{K}{1 + \frac{\sigma^2}{2r}} \right) \left[ \frac{S_0}{K} \left( 1 + \frac{\sigma^2}{2r} \right) \right]^{-\frac{2r}{\sigma^2}}$$

$$= \frac{K}{1 + \frac{2r}{\sigma^2}} \left[ \frac{S_0}{K} \left( 1 + \frac{\sigma^2}{2r} \right) \right]^{-\frac{2r}{\sigma^2}}.$$

Hence the price of a perpetual put is

$$\frac{K}{1 + \frac{2r}{\sigma^2}} \left[ \frac{S_0}{K} \left( 1 + \frac{\sigma^2}{2r} \right) \right]^{-\frac{2r}{\sigma^2}}.$$

The optimal exercise time of the put, $\tau_{b^*}$, is when the stock hits the optimal barrier

$$\tau_{b^*} = \inf\{t > 0; S_t = b^*\} = \inf\{t > 0; S_t = K/(1 + \frac{\sigma^2}{2r})\}.$$

But what is the expected exercise time of a perpetual American put? To answer this question we need to compute $E[\tau_{b^*}]$. Substituting

$$\tau = \tau_{b^*}, \quad x = \ln \frac{S_0}{b^*}, \quad \mu = r - \frac{\sigma^2}{2} \quad (10.1.2)$$

in Proposition ?? (c) yields

$$E[\tau_{b^*}] = \frac{x}{\mu} e^{-2\mu x} = \frac{\ln \frac{S_0}{b^*}}{r - \frac{\sigma^2}{2}} e^{-\frac{2}{\sigma^2} (r - \frac{\sigma^2}{2}) \ln \frac{S_0}{b^*}}$$

$$= \ln \left[ \left( \frac{S_0}{b^*} \right)^{\frac{1}{2}} \right] e^{-\frac{1}{\sigma^2} \ln \frac{S_0}{b^*}} = \ln \left[ \left( \frac{S_0}{b^*} \right)^{\frac{1}{2}} \right] \left( \frac{S_0}{b^*} \right)^{1-\frac{2r}{\sigma^2}}.$$

Hence the expected time when the holder should exercise the put is given by the exact formula

$$E[\tau_{b^*}] = \ln \left[ \left( \frac{S_0}{b^*} \right)^{\frac{1}{2}} \right] \left( \frac{S_0}{b^*} \right)^{1-\frac{2r}{\sigma^2}},$$

with $b^*$ given by (10.1.1). The probability density function of the optimal exercise time $\tau_{b^*}$ can be found from Proposition ?? (b) using substitutions (10.1.2)

$$p(\tau) = \frac{\ln \frac{S_0}{b^*}}{\sigma \sqrt{2\pi \tau^{3/2}}} e^{-\frac{\left( \ln \frac{S_0}{b^*} + \tau (r - \frac{\sigma^2}{2}) \right)^2}{2\sigma^2 \tau}}, \quad \tau > 0. \quad (10.1.3)$$
Exercise 10.1.10 Consider a stock that pays continuous dividends at rate \( \delta > 0 \).

(a) Assume a perpetual put is exercised whenever the stock reaches the barrier \( b \) from above. Show that the discounted value at time \( t = 0 \) is

\[
g(b) = (K - b) \left( \frac{S_0}{b} \right)^{h_2},
\]

where

\[
h_2 = \frac{1}{2} - \frac{r - \delta}{\sigma^2} - \sqrt{\left( \frac{r - \delta}{\sigma^2} - \frac{1}{2} \right)^2 + \frac{2r}{\sigma^2}}.
\]

(b) Use differentiation to show that the maximum value of \( g(b) \) is realized for

\[
b^* = K \frac{h_2}{h_2 - 1}.
\]

(c) Prove the price of perpetual put

\[
g(b^*) = \max_{b>0} g(b) = K \frac{h_2 - 1}{h_2} \left( \frac{h_2}{h_2 - 1} \right)^{h_2}.
\]

10.2 Perpetual American Log Contract

A perpetual American log contract is a contract that never expires and can be exercised at any time, providing the holder the log of the value of the stock, \( \ln S_t \), at the exercise time \( t \). It is interesting to note that these type of contracts are always optimal to be exercised, and their pricing formula is fairly uncomplicated.

Assume the contract is exercised when the stock \( S_t \) reaches the barrier \( b \), with \( S_0 < b \). If the hitting time of the barrier \( b \) is \( \tau_b \), then its payoff is \( \ln S_{\tau_b} = \ln b \). Discounting at the risk free interest rate, the value of the contract at time \( t = 0 \) is

\[
f(b) = E[e^{-r\tau_b} \ln S_{\tau_b}] = E[e^{-r\tau_b} \ln b] = \frac{\ln b}{b} S_0,
\]

since the barrier is assumed to be reached from below, see Proposition 10.1.1.

The function \( g(b) = \frac{\ln b}{b}, \ b > 0 \), has the derivative \( g'(b) = \frac{1 - \ln b}{b^2} \), so \( b^* = e \) is a global maximum point, see Fig. 10.3. The maximum value is \( g(b^*) = 1/e \). Then the optimal value of the barrier is \( b^* = e \), and the price of the contract at \( t = 0 \) is

\[
f_0 = \max_{b>0} f(b) = S_0 \max_{b>0} g(b) = \frac{S_0}{e}.
\]

In order for the stock to reach the optimum barrier \( b^* \) from below we need to require the condition \( S_0 < e \). Hence we arrived at the following result:
Figure 10.3: The graph of the function $g(b) = \frac{\ln b}{b}$, $b > 0$.

**Proposition 10.2.1** Let $S_0 < e$. Then the optimal exercise price of a perpetual American log contract is \[ \tau = \inf\{t > 0; S_t = e\}, \]
and its value at $t = 0$ is \[ S_0 e. \]

**Remark 10.2.2** If $S_0 > e$, then it is optimal to exercise the perpetual log contract as soon as possible.

**Exercise 10.2.3** Consider a stock that pays continuous dividends at a rate $\delta > 0$, and assume that $S_0 < e^{1/h_1}$, where \[ h_1 = \frac{1}{2} - \frac{r - \delta}{\sigma^2} + \sqrt{\left(\frac{r - \delta}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2r}{\sigma^2}}. \]

(a) Assume a perpetual log contract is exercised whenever the stock reaches the barrier $b$ from below. Show that the discounted value at time $t = 0$ is \[ f(b) = \ln b \left(\frac{S_0}{b}\right)^{h_1}. \]

(b) Use differentiation to show that the maximum value of $f(b)$ is realized for \[ b^* = e^{1/h_1}. \]

(c) Prove the price of perpetual log contract \[ f(b^*) = \max_{b > 0} f(b) = \frac{S_0^{h_1}}{h_1 e}. \]

(d) Show that the higher the dividend rate $\delta$, the lower the optimal exercise time is.
10.3 Perpetual American Power Contract

A perpetual American power contract is a contract that never expires and can be exercised at any time, providing the holder the \( n \)-th power of the value of the stock, \( (S_t)^\alpha \), at the exercise time \( t \), where \( \alpha \neq 0 \). (If \( \alpha = 0 \) the payoff is a constant, which is equal to 1).

1. Case \( \alpha > 0 \). Since we expect the value of the payoff to increase over time, we assume the contract is exercised when the stock \( S_t \) reaches the barrier \( b \), from below. If the hitting time of the barrier \( b \) is \( \tau_b \), then its payoff is \( (S_{\tau})^\alpha \). Discounting at the risk free interest rate, the value of the contract at time \( t = 0 \) is

\[
f(b) = E[e^{-r\tau_b}(S_{\tau})^\alpha] = E[e^{-r\tau_b}b^\alpha] = b^\alpha S_0 \frac{S_0}{b} = b^{\alpha-1} S_0,
\]

where we used Proposition 10.1.1. We shall discuss the following cases:

(i) If \( \alpha > 1 \), then the optimal barrier is \( b^* = \infty \), and hence, it is never optimal to exercise the contract in this case.

(ii) If \( 0 < \alpha < 1 \), the function \( f(b) \) is decreasing, so its maximum is reached for \( b^* = S_0 \), which corresponds to \( \tau_{b^*} = 0 \). Hence it is optimal to exercise the contract as soon as possible.

(iii) In the case \( \alpha = 1 \), the value of \( f(b) \) is constant, and the contract can be exercised at any time.

2. Case \( \alpha < 0 \). The payoff value, \( (S_t)^\alpha \), is expected to decrease, so we assume the exercise occurs when the stock reaches the barrier \( b \) from above. Discounting to initial time \( t = 0 \) and using Proposition 10.1.6 yields

\[
f(b) = E[e^{-r\tau_b}(S_{\tau})^\alpha] = E[e^{-r\tau_b}b^\alpha] = b^\alpha \left( \frac{S_0}{b} \right)^{-\frac{2\alpha}{\sigma^2}} = b^{\alpha-\frac{2\alpha}{\sigma^2}} S_0^{-\frac{2\alpha}{\sigma^2}}.
\]

(i) If \( \alpha < -\frac{2\sigma^2}{\sigma^2} \), then \( f(b) \) is decreasing, so its maximum is reached for \( b^* = S_0 \), which corresponds to \( \tau_{b^*} = 0 \). Hence it is optimal to exercise the contract as soon as possible.

(ii) If \( -\frac{2\sigma^2}{\sigma^2} < \alpha < 0 \), then the maximum of \( f(b) \) occurs for \( b^* = \infty \). Hence it is never optimal to exercise the contract.

Exercise 10.3.1 Consider a stock that pays continuous dividends at a rate \( \delta > 0 \), and assume \( \alpha > h_1 \), with \( h_1 = \frac{1}{2} - \frac{r-\delta}{\sigma^2} + \sqrt{\left( \frac{r-\delta}{\sigma^2} - \frac{1}{2} \right)^2 + \frac{2\delta^2}{\sigma^2}} \). Show that the perpetual power contract with payoff \( S_t^\alpha \) is never optimal to be exercised.

Exercise 10.3.2 (Perpetual American power put) Consider an perpetual American-type contract with the payoff \( (K - S_t)^2 \), where \( K > 0 \) is the strike price. Find the optimal exercise time and the contract value at \( t = 0 \).
10.4 Finitely Lived American Options

Exact pricing formulas are great when they exist and can be easily implemented. Even if we cherish all closed form pricing formulas we can get, there is also a time when exact formulas are not possible and approximations are in order. If we run into a problem whose solution cannot be found explicitly, it would still be very valuable to know something about its approximate quantitative behavior. This will be the case of finitely lived American options.

10.4.1 American Call

The case of Non-dividend Paying Stock

The holder has the right to buy a stock for the price $K$ at any time before or at the expiration $T$. The strike price $K$ and expiration $T$ are specified at the beginning of the contract. The payoff at time $t$ is $(S_t - K)^+ = \max\{S_t - K, 0\}$. The price of the American call at time $t = 0$ is given by

$$ f_0 = \max_{0 \leq \tau \leq T} \hat{E}[e^{-r\tau}(S_\tau - K)^+], $$

where the maximum is taken over all stopping times $\tau$ less than or equal to $T$.

**Theorem 10.4.1** It is not optimal to exercise an American call on a non-dividend paying stock early. It is optimal to exercise the call at maturity, $T$, if at all. Consequently, the price of an American call is equal to the price of the corresponding European call.

**Proof:** The heuristic idea of the proof is based on the observation that the difference $S_t - K$ tends to be larger as time goes on. As a result, there is always hope for a larger payoff and the later we exercise the better. In the following we shall formalize this idea mathematically using the submartingale property of the stock together with the Optional Stopping Theorem.

Let $X_t = e^{-rt}S_t$ and $f(t) = -Ke^{-rt}$. Since $X_t$ is a martingale in the risk-neutral-world, see Proposition 7.1.1, and $f(t)$ is an increasing, integrable function, applying Proposition ?? (c) it follows that

$$ Y_t = X_t + f(t) = e^{-rt}(S_t - K) $$

is an $\mathcal{F}_t$-submartingale, where $\mathcal{F}_t$ is the information set provided by the underlying Brownian motion $W_t$.

Since the hokey-stick function $\phi(x) = x^+ = \max\{x, 0\}$ is convex, then by Proposition ?? (b), the process $Z_t = \phi(Y_t) = e^{-rt}(S_t - K)^+$ is a submartingale.
Applying Doob’s stopping theorem (see Theorem 23) for stopping times $\tau$ and $T$, with $\tau \leq T$, we obtain $\hat{E}[Z_\tau] \leq \hat{E}[Z_T]$. This means
\[
\hat{E}[e^{-r\tau}(S_\tau - K)^+] \leq \hat{E}[e^{-rT}(S_T - K)^+],
\]
i.e. the maximum of the American call price is realized for the optimum exercise time $\tau^* = T$. The maximum value is given by the right side, which denotes the price of an European call option.

With a slight modification in the proof we can treat the problem of American power contract.

**Proposition 10.4.2 (American power contract)** Consider a contract with maturity date $T$, which pays, when exercised, $S^n_t$, where $n > 1$, and $t \leq T$.

Then it is not optimal to exercise this contract early.

**Proof:** Using that $M_t = e^{-rt}S_t$ is a martingale (in the risk neutral world), then $X_t = M^n_t$ is a submartingale. Since $Y_t = e^{-rt}S^n_t = (e^{-rt}S_t)^n e^{(n-1)rt} = X_te^{(n-1)rt}$, then for $s < t$
\[
E[Y_t|\mathcal{F}_s] = E[X_t e^{(n-1)rs}|\mathcal{F}_s] = E[X_t e^{(n-1)rs}|\mathcal{F}_s] \geq e^{(n-1)rs} X_s = Y_s,
\]
so $Y_t$ is a submartingale. Applying Doob’s stopping theorem (see Theorem 23) for stopping times $\tau$ and $T$, with $\tau \leq T$, we obtain $\hat{E}[Y_\tau] \leq \hat{E}[Y_T]$, or equivalently, $\hat{E}[e^{-r\tau}S^n_\tau] \leq \hat{E}[e^{-rT}S^n_T]$, which implies
\[
\max_{\tau \leq T} \hat{E}[e^{-r\tau}S^n_\tau] = \hat{E}[e^{-rT}S^n_T].
\]
Then it is optimal to exercise the contract at maturity $T$.

**Exercise 10.4.3** Consider an American future contract with maturity date $T$ and delivery price $K$, i.e. a contract with payoff at maturity $S_T - K$, which can be exercised at any time $t \leq T$. Show that it is not optimal to exercise this contract early.

**Exercise 10.4.4** Consider an American option contract with maturity date $T$, and time-dependent strike price $K(t)$, i.e. a contract with payoff at maturity $S_T - K(T)$, which can be exercised at any time.

(a) Show that it is not optimal to exercise this contract early in the following two cases:

(i) If $K(t)$ is a decreasing function;

(ii) If $K(t)$ is an increasing function with $K(t) < e^{rt}$.

(b) What happens if $K(t)$ is increasing and $K(t) > e^{rt}$?
The case of Dividend Paying Stock

When the stock pays dividends it is optimal to exercise the American call early. An exact solution of this problem is hard to get explicitly, or might not exist. However, there are some asymptotic solutions that are valid close to expiration (see Wilmott) and analytic approximations given by MacMillan, Barone-Adesi and Whaley.

In the following we shall discuss why it is difficult to find an exact optimal exercise time for an American call on a dividend paying stock. First, consider two contracts:

1. Consider a contract by which one can acquire a stock, $S_t$, at any time before or at time $T$. This can be seen as a contract with expiration $T$, that pays when exercised the stock price, $S_t$. Assume the stock pays dividends at a continuous rate $\delta > 0$. When should the contract be exercised in order to maximize its value?

The value of the contract at time $t = 0$ is $\max_{\tau \leq T} \left\{ e^{-r\tau} S_\tau \right\}$, where the maximum is taken over all stopping times $\tau$ less than or equal to $T$. Since the stock price, which pays dividends at rate $\delta$, is given by

$$S_t = S_0 e^{(r-\delta-\frac{\sigma^2}{2})t + \sigma W_t},$$

then $M_t = e^{-(r-\delta)t} S_t$ is a martingale (in the risk neutral world). Therefore $e^{-rt} S_t = e^{-\delta t} M_t$. Let $X_t = e^{-rt} S_t$. Then for $0 < s < t$

$$\tilde{E}[X_t|F_s] = \tilde{E}[e^{-\delta t} M_t|F_s] < \tilde{E}[e^{-\delta s} M_s|F_s] = e^{-\delta s} \tilde{E}[M_t|F_s] = e^{-\delta s} M_s = e^{-rs} S_s = M_s,$$

so $X_t$ is a supermartingale (i.e. $-X_t$ is a submartingale). Applying the Optional Stopping Theorem for the stopping time $\tau$ we obtain

$$\tilde{E}[X_\tau] \leq \tilde{E}[X_0] = S_0.$$

Hence it is optimal to exercise the contract at the initial time $t = 0$. This makes sense, since in this case we have a longer period of time during which dividends are collected.

2. Consider a contract by which one has to pay the amount of cash $K$ at any time before or at time $T$. Given the time value of money

$$e^{-rt} K > e^{-rT} K, \quad t < T,$$

it is always optimal to defer the payment until time $T$. 

3. Now consider a combination of the previous two contracts. This new contract pays $S_t - K$ and can be exercised at any time $t$, with $t \leq T$. Since it is not clear when is optimal to exercise this contract, we shall consider two limiting cases. Let $\tau^*$ denote its optimal exercise time.

(i) When $K \to 0^+$, then $\tau^* \to 0^+$, because we approach the conditions of case 1, and also assume continuity conditions on the price.

(ii) If $K \to \infty$, then the latter the pay day, the better, i.e. $\tau^* \to T^-$.

The optimal exercise time, $\tau^*$, is somewhere between 0 and $T$, with the tendency of moving towards $T$ as $K$ gets large.

10.4.2 American Put

10.4.3 Mac Millan-Barone-Adesi-Whaley Approximation

TO COME

10.4.4 Black’s Approximation

TO COME

10.4.5 Roll-Geske-Whaley Approximation

TO COME

10.4.6 Other Approximations
Chapter 11
Hints and Solutions

Chapter 3

3.3.4 (a) Since \( r_t \sim N(\mu, s^2) \), with \( \mu = b + (r_0 - b)e^{-at} \) and \( s^2 = \frac{\sigma^2}{2a}(1 - e^{-2at}) \). Then

\[
P(r_t < 0) = \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi s}} e^{-\frac{(x-\mu)^2}{2s^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\mu/s} e^{-v^2/2} dv = N(-\mu/s),
\]

where by computation

\[
-\frac{\mu}{s} = -\frac{1}{\sigma} \sqrt{\frac{2a}{e^{2at} - 1}} (r_0 + b(e^{at} - 1)).
\]

(b) Since

\[
\lim_{t \to \infty} \frac{\mu}{s} = \frac{\sqrt{2a}}{\sigma} \lim_{t \to \infty} \frac{r_0 + b(e^{at} - 1)}{\sqrt{e^{2at} - 1}} = \frac{b\sqrt{2a}}{\sigma},
\]

then

\[
\lim_{t \to \infty} P(r_t < 0) = \lim_{t \to \infty} N(-\mu/s) = N(-\frac{b\sqrt{2a}}{\sigma}).
\]

It is worth noting that the previous probability is less than 0.5.
(c) The rate of change is

\[
\frac{d}{dt}P(r_t < 0) = -\frac{1}{\sqrt{2\pi}} \frac{d}{ds} \left( \frac{\mu}{s} \right) = -\frac{1}{\sqrt{2\pi}} \frac{\sigma^2}{2a} \frac{e^{2at}[b(e^{2at} - 1)(e^{-at} - 1) - r_0]}{(e^{2at} - 1)^{3/2}}.
\]

3.3.5 By Ito’s formula

\[
d(r_t^n) = \left( na(b - r_t)r_t^{n-1} + \frac{1}{2} n(n - 1)\sigma^2 r_t^{n-1} \right) dt + n\sigma r_t^{n-\frac{1}{2}} dW_t.
\]
Integrating between 0 and $t$ and taking the expectation yields

$$\mu_n(t) = r_0^n + \int_0^t [nab\mu_{n-1}(s) - na\mu_n(s) + \frac{n(n-1)}{2}\sigma^2\mu_{n-1}(s)]\,ds$$

Differentiating yields

$$\mu'_n(t) + na\mu_n(t) = (nab + \frac{n(n-1)}{2}\sigma^2)\mu_{n-1}(t).$$

Multiplying by the integrating factor $e^{nat}$ yields the exact equation

$$[e^{nat}\mu_n(t)]' = e^{nat}(nab + \frac{n(n-1)}{2}\sigma^2)\mu_{n-1}(t).$$

Integrating yields the following recursive formula for moments

$$\mu_n(t) = r_0^n e^{-nat} + (nab + \frac{n(n-1)}{2}\sigma^2) \int_0^t e^{-na(t-s)}\mu_{n-1}(s)\,ds.$$  

3.5.1 (a) The spot rate $r_t$ follows the process

$$d(\ln r_t) = \theta(t)\,dt + \sigma\,dW_t.$$  

Integrating yields

$$\ln r_t = \ln r_0 + \int_0^t \theta(u)\,du + \sigma W_t,$$

which is normally distributed. (b) Then $r_t = r_0 e^{\int_0^t \theta(u)\,du + \sigma W_t}$ is log-normally distributed. (c) The mean and variance are

$$\mathbb{E}[r_t] = r_0 e^{\int_0^t \theta(u)\,du} e^{\frac{1}{2}\sigma^2 t}, \quad \text{Var}(r_t) = r_0^2 e^{2\int_0^t \theta(u)\,du} e^{\sigma^2 t}(e^{\sigma^2 t} - 1).$$

3.5.2 Substitute $u_t = \ln r_t$ and obtain the linear equation

$$du_t + a(t)u_t\,dt = \theta(t)\,dt + \sigma(t)\,dW_t.$$  

The equation can be solved multiplying by the integrating factor $e^{\int_0^t a(s)\,ds}$.

Chapter 5

4.7.1 Apply the same method used in the Application 4.7. The price of the bond is: $P(t, T) = e^{-r_T(T-t)}\mathbb{E}[e^{-\sigma\int_0^{T-t} K_s\,ds}]$.

4.9.2 The price of an infinitely lived bond at time $t$ is given by $\lim_{T \to \infty} P(t, T)$.

Using $\lim_{T \to \infty} B(t, T) = \frac{1}{a}$, and

$$\lim_{T \to \infty} A(t, T) = \begin{cases} +\infty, & \text{if } b < \sigma^2/(2a) \\ 1, & \text{if } b > \sigma^2/(2a) \\ \left(\frac{1}{a} + t\right)(b - \frac{\sigma^2}{2a^2}) - \frac{\sigma^2}{4a^2}, & \text{if } b = \sigma^2/(2a), \end{cases}$$
we can get the price of the bond in all three cases.

Chapter 6

5.1.1 Dividing the equations

\[
S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t},
\]

\[
S_u = S_0 e^{(\mu - \frac{\sigma^2}{2})u + \sigma W_u}
\]
yields

\[
S_t = S_u e^{(\mu - \frac{\sigma^2}{2})(t-u) + \sigma (W_t - W_u)}.
\]

Taking the predictable part out and dropping the independent condition we obtain

\[
E[S_t|\mathcal{F}_u] = S_u e^{(\mu - \frac{\sigma^2}{2})(t-u)} E[e^{\sigma (W_t - W_u)}|\mathcal{F}_u]
= S_u e^{(\mu - \frac{\sigma^2}{2})(t-u)} E[e^{\sigma W_t - u}]
= S_u e^{(\mu - \frac{\sigma^2}{2})(t-u)} e^{\frac{1}{2} \sigma^2 (t-u)}
= S_u e^{\mu (t-u)}
\]

Similarly we obtain

\[
E[S_t^2|\mathcal{F}_u] = S_u^2 e^{2(\mu - \frac{\sigma^2}{2})(t-u)} E[e^{2 \sigma (W_t - W_u)}|\mathcal{F}_u]
= S_u^2 e^{2\mu (t-u)} e^{\sigma^2 (t-u)}
\]

Then

\[
Var(S_t|\mathcal{F}_u) = E[S_t^2|\mathcal{F}_u] - E[S_t|\mathcal{F}_u]^2
= S_u^2 e^{2\mu (t-u)} e^{\sigma^2 (t-u)} - S_u^2 e^{2\mu (t-u)}
= S_u^2 e^{2\mu (t-u)} \left( e^{\sigma^2 (t-u)} - 1 \right).
\]

When \( s = t \) we get \( E[S_t|\mathcal{F}_t] = S_t \) and \( Var(S_t|\mathcal{F}_t) = 0. \)

5.1.2 By Ito’s formula

\[
d(\ln S_t) = \frac{1}{S_t} dS_t - \frac{1}{2 S_t^2} (dS_t)^2 = (\mu - \frac{\sigma^2}{2})dt + \sigma dW_t,
\]

so \( \ln S_t = \ln S_0 + (\mu - \frac{\sigma^2}{2})t + \sigma W_t \), and hence \( \ln S_t \) is normally distributed with \( E[\ln S_t] = \ln S_0 + (\mu - \frac{\sigma^2}{2})t \) and \( Var(\ln S_t) = \sigma^2 t. \)
5.1.3 (a) \( d\left(\frac{1}{S_t}\right) = (\sigma^2 - \mu) \frac{1}{S_t^2} dt - \frac{\sigma}{S_t} dW_t; \)
(b) \( d(S_t^n) = n(\mu + \frac{n-1}{2}\sigma^2) dt + n\sigma S_t^b dW_t; \)
(c) \[
\begin{align*}
\quad d((S_t - 1)^2) &= d(S_t^2) - 2dS_t = 2S_tdS_t + (dS_t)^2 - 2dS_t \\
&= \left(2\mu + \sigma^2\right)S_t^2 dt + 2\sigma S_t(S_t - 1)dW_t.
\end{align*}
\]

5.1.4 (b) Since \( S_t = S_0e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t} \), then \( S_t^n = S_0^n e^{(n\mu - \frac{n\sigma^2}{2})t + n\sigma W_t} \) and hence
\[
E[S_t^n] = S_0^n e^{(n\mu - \frac{n\sigma^2}{2})t} E[e^{n\sigma W_t}] = S_0^n e^{(n\mu - \frac{n\sigma^2}{2})t} e^{\frac{1}{2}n^2\sigma^2 t} = S_0^n e^{(n\mu + \frac{n(n-1)}{2}\sigma^2) t}.
\]
(a) Let \( n = 2 \) in (b) and obtain \( E[S_t^2] = S_0e^{(2\mu + \sigma^2) t}. \)

5.1.5 (a) Let \( X_t = S_t W_t \). The product rule yields
\[
\begin{align*}
d(X_t) &= W_t dS_t + S_t dW_t + dS_t dW_t \\
&= W_t (\mu S_t dt + \sigma S_t dW_t) + S_t dW_t + (\mu S_t dt + \sigma S_t dW_t) dW_t \\
&= S_t (\mu W_t + \sigma) dt + (1 + \sigma W_t) S_t dW_t.
\end{align*}
\]
Integrating
\[
X_s = \int_0^s S_t (\mu W_t + \sigma) dt + \int_0^s (1 + \sigma W_t) S_t dW_t.
\]
Let \( f(s) = E[X_s] \). Then
\[
f(s) = \int_0^s \left(\mu f(t) + \sigma S_0 e^{\mu t}\right) dt.
\]
Differentiating yields
\[
f'(s) = \mu f(s) + \sigma S_0 e^{\mu s}, \quad f(0) = 0,
\]
with the solution \( f(s) = \sigma S_0 e^{\mu s} \). Hence \( E[W_t S_t] = \sigma S_0 e^{\mu t} \).
(b) \( \text{Cov}(S_t, W_t) = E[S_t W_t] - E[S_t] E[W_t] = E[S_t W_t] = \sigma S_0 e^{\mu t} \). Then
\[
\text{Corr}(S_t, W_t) = \frac{\text{Cov}(S_t, W_t)}{\sigma S_t \sigma W_t} = \frac{\sigma S_0 e^{\mu t}}{S_0 e^{\mu t} \sqrt{e^{2\sigma^2 t} - 1} \sqrt{t}} = \sigma \frac{t}{\sqrt{e^{2\sigma^2 t} - 1}} \to 0, \quad t \to \infty.
\]
5.5.3 $d = S_d/S_0 = 0.5$, $u = S_u/S_0 = 1.5$, $\gamma = -6.5$, $p = 0.98$.

5.3.7 Let $T_a$ be the time $S_t$ reaches level $a$ for the first time. Then

$$F(a) = P(S_t < a) = P(T_a > t)$$

$$= \ln \frac{a}{S_0} \int_t^{\infty} \frac{1}{\sqrt{2\pi \sigma^2 \tau}} e^{-\frac{(\ln \frac{a}{S_0} - \alpha \tau)^2}{2 \tau \sigma^2}} d\tau,$$

where $\alpha = \mu - \frac{1}{2} \sigma^2$.

5.3.8 (a) $P(T_a < T) = \ln \frac{a}{S_0} T_0^T \frac{1}{\sqrt{2\pi \sigma^2 \tau}} e^{-\frac{(\ln \frac{a}{S_0} - \alpha \tau)^2}{(2\tau \sigma^2)}} d\tau$.

(b) $\int_{T_1}^{T_2} p(\tau) d\tau$.

5.8.3 $E[A_t] = S_0(1 + \frac{\mu}{\sigma^2}) + O(t^2)$, $Var(A_t) = S_0^2 \sigma^2 t + O(t^2)$.

5.8.6 (a) Since $\ln G_t = \frac{1}{t} \int_0^t \ln S_u du$, using the product rule yields

$$d(\ln G_t) = \frac{1}{t} \int_0^t \ln S_u du + \frac{1}{t} d\left( \int_0^t \ln S_u du \right)$$

$$= -\frac{1}{t^2} \left( \int_0^t \ln S_u du \right) dt + \frac{1}{t} \ln S_t dt$$

$$= \frac{1}{t} (\ln S_t - \ln G_t) dt.$$

(b) Let $X_t = \ln G_t$. Then $G_t = e^{X_t}$ and Ito’s formula yields

$$dG_t = de^{X_t} = e^{X_t} dX_t + \frac{1}{2} e^{X_t} (dX_t)^2 = e^{X_t} dX_t$$

$$= G_t d(\ln G_t) = \frac{G_t}{t} (\ln S_t - \ln G_t) dt.$$

5.8.7 Using the product rule we have

$$d\left( \frac{H_t}{t} \right) = d\left( \frac{1}{t} \right) H_t + \frac{1}{t} dH_t$$

$$= -\frac{1}{t^2} H_t dt + \frac{1}{t^2} H_t \left( 1 - \frac{H_t}{S_t} \right) dt$$

$$= -\frac{1}{t^2} H_t^2 dt,$$

so, for $dt > 0$, then $d\left( \frac{H_t}{t} \right) < 0$ and hence $\frac{H_t}{t}$ is decreasing. For the second part, try to apply l’Hospital rule.

5.8.8 By continuity, the inequality (5.8.21) is preserved under the limit.
5.8.9 (a) Use that \( \frac{1}{n} \sum S_{t_k} = \frac{1}{t} \sum S_{t_k} \cdot \frac{t}{n} = \int_0^t S_u^\alpha du \). (b) Let \( I_t = \int_0^t S_u^\alpha du \).

Then
\[
d\left( \frac{1}{t} I_t \right) = d\left( \frac{1}{t} I_t \right) + \frac{1}{t} dI_t = \frac{1}{t} \left( S_t^\alpha - I_t \right) dt.
\]

Let \( X_t = A_t^\alpha \). Then by Ito’s formula we get
\[
dX_t = \alpha \left( \frac{1}{t} I_t \right)^{\alpha-1} d\left( \frac{1}{t} I_t \right) + \frac{1}{2} \alpha(\alpha - 1) \left( \frac{1}{t} I_t \right)^{\alpha-2} \left( d\left( \frac{1}{t} I_t \right) \right)^2
\]
\[
= \alpha \left( \frac{1}{t} I_t \right)^{\alpha-1} d\left( \frac{1}{t} I_t \right) = \alpha \left( \frac{1}{t} I_t \right)^{\alpha-1} \frac{1}{t} \left( S_t^\alpha - I_t \right) dt
\]
\[
= \alpha \left( S_t^\alpha (A_t^\alpha)^{1-\alpha} - A_t^\alpha \right) dt.
\]

(c) If \( \alpha = 1 \) we get the continuous arithmetic mean, \( A_t^1 = \frac{1}{t} \int_0^t S_u du \). If \( \alpha = -1 \) we obtain the continuous harmonic mean, \( A_t^{-1} = \frac{t}{\int_0^t S_u du} \).

5.8.10 (a) By the product rule we have
\[
dX_t = d\left( \frac{1}{t} \right) \int_0^t S_u dW_u + \frac{1}{t} d\left( \int_0^t S_u dW_u \right)
\]
\[
= -\frac{1}{t^2} \left( \int_0^t S_u dW_u \right) dt + \frac{1}{t} S_t dW_t
\]
\[
= -\frac{1}{t} X_t dt + \frac{1}{t} S_t dW_t.
\]

Using the properties of Ito integrals, we have
\[
E[X_t] = \frac{1}{t} E\left[ \int_0^t S_u dW_u \right] = 0
\]
\[
Var(X_t) = E[X_t^2] - E[X_t]^2 = E[X_t^2]
\]
\[
= \frac{1}{t^2} E\left[ \left( \int_0^t S_u dW_u \right) \left( \int_0^t S_u dW_u \right) \right]
\]
\[
= \frac{1}{t^2} \int_0^t E[S_u^2] du = \frac{1}{t^2} \int_0^t S_0^2 e^{(2\mu + \sigma^2)u} du
\]
\[
= \frac{S_0^2 e^{(2\mu + \sigma^2)t} - 1}{2\mu + \sigma^2}.
\]

(b) The stochastic differential equation of the stock price can be written in the form
\[
\sigma S_t dW_t = dS_t - \mu S_t dt.
\]

Integrating between 0 and \( t \) yields
\[
\sigma \int_0^t S_u dW_u = S_t - S_0 - \mu \int_0^t S_u du.
\]
Dividing by $t$ yields the desired relation.

5.9.2 Using independence $\mathbb{E}[S_t] = S_0 e^{\mu t} \mathbb{E}[(1 + \rho)^{N_t}]$. Then use

$$\mathbb{E}[(1 + \rho)^{N_t}] = \sum_{n \geq 0} \mathbb{E}[(1 + \rho)^n | N_t = n] P(N_t = n) = \sum_{n \geq 0} (1 + \rho)^n \frac{(\lambda t)^n}{n!} = e^{(1 + \rho)t}.$$ 

5.9.3 $\mathbb{E}[\ln S_t] = \ln S_0 \left( \mu - \rho - \frac{\sigma^2}{2} + \lambda \ln(\rho + 1) \right) t$.

5.9.4 $\mathbb{E}[S_t | \mathcal{F}_u] = S_u e^{(\mu - \lambda \rho - \frac{\sigma^2}{2})(t-u)} (1 + \rho)^{N_{t-u}}$.

Chapter 7

6.2.2 (a) Substitute $\lim_{S_t \to \infty} N(d_1) = \lim_{S_t \to \infty} N(d_2) = 1$ and $\lim_{S_t \to \infty} \frac{K}{S_t} = 0$ in

$$\frac{c(t)}{S_t} = N(d_1) - \frac{K}{S_t} e^{-r(T-t)} N(d_2).$$

(b) Use $\lim_{S_t \to 0} N(d_1) = \lim_{S_t \to \infty} N(d_2) = 0$. (c) It comes from an analysis similar to the one used at (a).

6.2.3 Differentiating in $c(t) = S_t N(d_1) - K e^{-r(T-t)} N(d_2)$ yields

$$\frac{dc(t)}{dS_t} = N(d_1) \frac{dN(d_1)}{dS_t} - K e^{-r(T-t)} \frac{dN(d_2)}{dS_t}$$

$$= N(d_1) + S_t N'(d_1) \frac{d^2}{dS_t^2} - K e^{-r(T-t)} N'(d_2) \frac{d^2}{dS_t^2}$$

$$= N(d_1) + S_t \frac{1}{\sqrt{2\pi}} \frac{e^{-d_1^2/2}}{\sigma \sqrt{T-t}} \frac{1}{S_t} - K e^{-r(T-t)} \frac{1}{\sqrt{2\pi}} \frac{e^{-d_2^2/2}}{\sigma \sqrt{T-t}} \frac{1}{S_t}$$

$$= N(d_1) + \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma \sqrt{T-t}} \frac{1}{S_t} e^{-d_2^2/2} \left[ S_t e^{(d_2^2-d_1^2)/2} - K e^{-r(T-t)} \right].$$

It suffices to show

$$S_t e^{(d_2^2-d_1^2)/2} - K e^{-r(T-t)} = 0.$$

Since

$$d_2^2 - d_1^2 = (d_2 - d_1)(d_2 + d_1) = -\sigma \sqrt{T-t} \frac{2 \ln(S_t/K) + 2r(T-t)}{\sigma \sqrt{T-t}}$$

$$= -2 \left( \ln S_t - \ln K + r(T-t) \right),$$

then we have

$$e^{(d_2^2-d_1^2)/2} = e^{-\ln S_t + \ln K - r(T-t)} = e^{\ln \frac{K}{S_t} e^{-r(T-t)}} = \frac{1}{S_t} K e^{-r(T-t)}.$$
Therefore
\[ S_t e^{(d_2^2 - d_1^2)/2} - Ke^{-r(T-t)} = S_t \frac{1}{S_t} Ke^{-r(T-t)} - Ke^{-r(T-t)} = 0, \]
which shows that \( \frac{dc(t)}{dS_t} = N(d_1) \).

6.3.1 The payoff is
\[ f_T = \begin{cases} 1, & \text{if } \ln K_1 \leq X_T \leq \ln K_2 \\ 0, & \text{otherwise}. \end{cases} \]
where \( X_T \sim N \left( \ln S_t + (\mu - \frac{\sigma^2}{2})(T-t), \sigma^2(T-t) \right) \).

\[ \hat{E}_t[f_T] = E[f_T | \mathcal{F}_t, \mu = r] = \int_{-\infty}^{\infty} f_T(x)p(x) \, dx \]
\[ = \int_{\ln K_1}^{\ln K_2} \frac{1}{\sqrt{2\pi} \sqrt{T-t}} e^{-\frac{(x-\ln S_t-(r-\frac{\sigma^2}{2})(T-t))^2}{2\sigma^2(T-t)}} \, dx \]
\[ = \int_{d_2(K_1)}^{d_2(K_2)} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \, dy = \int_{d_2(K_1)}^{d_2(K_2)} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \, dy \]
\[ = N(d_2(K_1)) - N(d_2(K_2)), \]
where
\[ d_2(K_1) = \frac{\ln S_t - \ln K_1 + (r - \frac{\sigma^2}{2})(T-t)}{\sigma \sqrt{T-t}} \]
\[ d_2(K_2) = \frac{\ln S_t - \ln K_2 + (r - \frac{\sigma^2}{2})(T-t)}{\sigma \sqrt{T-t}}. \]

Hence the value of a box-contract at time \( t \) is \( f_t = e^{-r(T-t)} \left[ N(d_2(K_1)) - N(d_2(K_2)) \right] \).

6.3.2 The payoff is
\[ f_T = \begin{cases} e^{X_T}, & \text{if } X_T > \ln K \\ 0, & \text{otherwise.} \end{cases} \]
\[ f_t = e^{-r(T-t)} \hat{E}_t[f_T] = e^{-r(T-t)} \int_{\ln K}^{\infty} e^x p(x) \, dx. \]
The computation is similar with the one of the integral \( I_2 \) from Proposition 6.2.1.

6.3.3 (a) The payoff is
\[ f_T = \begin{cases} e^{nX_T}, & \text{if } X_T > \ln K \\ 0, & \text{otherwise.} \end{cases} \]
\[ \hat{E}_t[f_T] = \int_{\ln K}^{\infty} e^{nx} p(x) \, dx = \frac{1}{\sigma \sqrt{2\pi(T-t)}} \int_{\ln K}^{\infty} e^{nx} e^{-\frac{1}{2} \frac{(x-\ln S_t-(r-\frac{\sigma^2}{2})(T-t))^2}{\sigma^2(T-t)}} \, dx \]

\[ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{n\sigma \sqrt{T-t}y} e^{\ln S_t} e^{n(r-\frac{\sigma^2}{2})(T-t)} e^{-\frac{1}{2}y^2} \, dy \]

\[ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sigma} e^{n(r-\frac{\sigma^2}{2})(T-t)} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2+n\sigma \sqrt{T-t}y} \, dy \]

\[ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sigma} e^{(nr+n(n-1)\frac{\sigma^2}{2})(T-t)} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}+ny\sigma \sqrt{T-t}} \, dy \]

\[ = S_t^n e^{(nr+n(n-1)\frac{\sigma^2}{2})(T-t)} \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \, dy \]

\[ = S_t^n \frac{e^{(nr+n(n-1)\frac{\sigma^2}{2})(T-t)} N(d_2+n\sigma \sqrt{T-t})}{\sigma \sqrt{2\pi}}. \]

The value of the contract at time \( t \) is

\[ f_t = e^{-r(T-t)} \hat{E}_t[f_T] = S_t^n e^{(n-1)(r+\frac{\sigma^2}{2})(T-t)} N(d_2+n\sigma \sqrt{T-t}). \]

(b) Using \( g_t = S_t^n e^{(n-1)(r+\frac{\sigma^2}{2})(T-t)} \), we get \( f_t = g_t N(d_2+n\sigma \sqrt{T-t}) \).

(c) In the particular case \( n = 1 \), we have \( N(d_2+\sigma \sqrt{T-t}) = N(d_1) \) and the value takes the form \( f_t = S_t N(d_1) \).

6.4.1 Since \( \ln S_T \sim N \left( \ln S_t + (\mu - \frac{\sigma^2}{2}(T-t)), \sigma^2(T-t) \right) \), we have

\[ \hat{E}_t[f_T] = \mathbb{E} \left[ (\ln S_T)^2 | F_t, \mu = r \right] \]

\[ = Var \left( \ln S_T | F_t, \mu = r \right) + \mathbb{E} \left[ \ln S_T | F_t, \mu = r \right]^2 \]

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

so

\[ f_t = e^{-r(T-t)} \left[ \sigma^2(T-t) + \left[ \ln S_t + (r - \frac{\sigma^2}{2})(T-t) \right]^2 \right]. \]

6.5.1 \( \ln S_T^n \) is normally distributed with

\[ \ln S_T^n = n \ln S_T \sim N \left( n \ln S_t + n(\mu - \frac{\sigma^2}{2}(T-t)), n^2 \sigma^2(T-t) \right). \]

Redo the computation of Proposition 6.2.1 in this case.

6.6.1 Let \( n = 2 \).

\[ f_t = e^{-r(T-t)} \hat{E}_t[(S_T - K)^2] \]

\[ = e^{-r(T-t)} \hat{E}_t[S_T^2] + e^{-r(T-t)} \left( K^2 - 2K \hat{E}_t[S_T] \right) \]

\[ = S_t^2 e^{(r+\sigma^2)(T-t)} + e^{-r(T-t)} K^2 - 2K S_t. \]
Let \( n = 3 \). Then
\[
f_t = e^{-r(T-t)} \tilde{E}_t[(S_T - K)^3] \\
= e^{-r(T-t)} \tilde{E}_t[S_T^3] - 3Ks_T^2 + 3K^2S_T - K^3 \\
= e^{-r(T-t)} \tilde{E}_t[S_T^3] - 3Ke^{-r(T-t)} \tilde{E}_t[S_T^2] + e^{-r(T-t)}(3K^2 \tilde{E}_t[S_T] - K^3) \\
= e^{2(r+3\sigma^2/2)(T-t)} S_t^3 - 3K e^{(r+\sigma^2)(T-t)} S_t^2 + 3K^2 S_t - e^{-r(T-t)} K^3.
\]

6.8.1 Choose \( c_n = 1/n! \) in the general contract formula.

6.9.1 Similar computation with the one done for the call option.

6.11.1 Write the payoff as a difference of two puts, one with strike price \( K_1 \) and the other with strike price \( K_2 \). Then apply the superposition principle.

6.11.2 Write the payoff as \( f_T = c_1 + c_3 - 2c_2 \), where \( c_i \) is a call with strike price \( K_i \). Apply the superposition principle.

6.11.4 (a) By inspection. (b) Write the payoff as the sum between a call and a put both with strike price \( K \), and then apply the superposition principle.

6.11.4 (a) By inspection. (b) Write the payoff as the sum between a call with strike price \( K_2 \) and a put with strike price \( K_1 \), and then apply the superposition principle. (c) The strangle is cheaper.

6.11.5 The payoff can be written as a sum between the payoffs of a \((K_3, K_4)\)-bull spread and a \((K_1, K_2)\)-bear spread. Apply the superposition principle.

6.12.2 By computation.

6.12.3 The risk-neutral valuation yields
\[
f_t = e^{-r(T-t)} \tilde{E}_t[S_T - A_T] = e^{-r(T-t)} \tilde{E}_t[S_T] - e^{-r(T-t)} \tilde{E}_t[A_T] \\
= S_t - e^{-r(T-t)} \frac{t}{T} A_t - \frac{1}{rT} S_t \left( 1 - e^{-r(T-t)} \right) \\
= S_t \left( 1 - \frac{1}{rT} + \frac{1}{rT} e^{-r(T-t)} \right) - e^{-r(T-t)} \frac{t}{T} A_t.
\]

6.12.5 We have
\[
\lim_{t \to 0} f_t = S_0 e^{-rT + (r - \frac{\sigma^2}{2}) T + \frac{\sigma^2 T}{6}} - e^{-rT} K \\
= S_0 e^{-\frac{1}{2}(r + \frac{\sigma^2}{6}) T} - e^{-rT} K.
\]

6.12.6 Since \( G_T < A_T \), then \( \tilde{E}_t[G_T] < \tilde{E}_t[A_T] \) and hence
\[
e^{-r(T-t)} \tilde{E}_t[G_T - K] < e^{-r(T-t)} \tilde{E}_t[A_T - K],
\]
so the forward contract on the geometric average is cheaper.

6.12.7 Dividing

\[ G_t = S_0 e^{(\mu - \frac{\sigma^2}{2}) t} e^{\frac{\sigma}{\sqrt{T}} \int_0^t W_u du} \]

\[ G_T = S_0 e^{(\mu - \frac{\sigma^2}{2}) \frac{T-t}{2} e^{\frac{\sigma}{\sqrt{T}} \int_0^t W_u du}} \]

yields

\[ \frac{G_T}{G_t} = e^{(\mu - \frac{\sigma^2}{2}) \frac{T-t}{2} e^{\frac{\sigma}{\sqrt{T}} \int_0^t W_u du}}. \]

An algebraic manipulation leads to

\[ G_T = G_t e^{(\mu - \frac{\sigma^2}{2}) \frac{T-t}{2} e^{\frac{\sigma}{\sqrt{T}} \int_0^t W_u du}}. \]

6.15.7 Use that

\[ P(\max_{t\leq T} S_t \geq z) = P\left( \sup_{t\leq T} [(\mu - \frac{\sigma^2}{2}) t + \sigma W_t] \geq \ln \left( \frac{z}{S_0} \right) \right) \]

\[ = P\left( \sup_{t\leq T} \left[ \frac{1}{\sigma} (\mu - \frac{\sigma^2}{2}) t + W_t \right] \geq \frac{1}{\sigma} \ln \left( \frac{z}{S_0} \right) \right) \]

Chapter 9

8.5.1 \( P = F - \Delta_F S = c - N(d_1) S = -Ke^{-r(T-t)} \).

Chapter 10

10.1.9

\[ P(S_t < b) = P\left( (r - \frac{\sigma^2}{2}) t + \sigma W_t < \ln \left( \frac{b}{S_0} \right) \right) = P\left( \sigma W_t < \ln \left( \frac{b}{S_0} \right) - (r - \frac{\sigma^2}{2}) t \right) \]

\[ = P(\sigma W_t < -\lambda) = P(\sigma W_t > \lambda) \leq e^{-\frac{\lambda^2}{2\sigma^2}} \]

\[ = e^{-\frac{1}{2\sigma^2} \ln(S_0/b) + (r-\sigma^2/2)t^2}, \]

where we used Proposition ?? with \( \mu = 0 \).
Bibliography


Index

determinism, 3, 154

semi-determinism, 4, 154