MATH3075/3975

FINANCIAL MATHEMATICS

Christian-Oliver Ewald and Marek Rutkowski

School of Mathematics and Statistics

University of Sydney

Semester 2, 2014

- Students enrolled in MATH3075 are expected to understand and learn all the material, except for the material marked as (MATH3975).
- Students enrolled in MATH3975 are expected to know all the material.

Related courses:

- MATH2070/2970: Optimisation and Financial Mathematics.
- STAT3011/3911: Stochastic Processes and Time Series.
- MSH7: Applied Probability and SDEs.
- AMH4: Advanced Option Pricing.
- AMH3: Interest Rate Models.

Suggested readings:

- J. C. Hull: *Options, Futures and Other Derivatives*. 3rd ed. Prentice-Hall, 1997.
- S. R. Pliska: *Introduction to Mathematical Finance: Discrete Time Models*. Blackwell Publishing, 1997.
- S. E. Shreve: *Stochastic Calculus for Finance. Volume 1: The Binomial Asset Pricing Model*. Springer, 2004.
- J. van der Hoek and R. J. Elliott: *Binomial Models in Finance.* Springer, 2006.
- R. U. Seydel: *Tools for Computational Finance.* 3rd ed. Springer, 2006.

Contents

Chapter 1

Introduction

The goal of this chapter is to give a brief introduction to financial markets.

1.1 Financial Markets

A **financial asset** (or a **financial security**) is a negotiable financial instrument representing financial value. Securities are broadly categorised into: **debt securities** (such as: government bonds, corporate bonds (debentures), municipal bonds), **equities** (e.g., common stocks) and **financial derivatives** (such as: forwards, futures, options, and swaps). We present below a tentative classification of existing financial markets and typical securities traded on them:

1. Equity market - common stocks (ordinary shares) and preferred stocks (preference shares).
2. Equity derivatives market - equity options and forwards.
3. Fixed income market - coupon-bearing bonds, zero-coupon bonds, sovereign debt, corporate bonds, bond options.
4. Futures market - forwards and futures contracts, index futures, futures options.
5. Interest rate market - caps, floors, swaps and swaptions, forward rate agreements.
6. Foreign exchange market - foreign currencies, options and derivatives on foreign currencies, cross-currency and hybrid derivatives.
7. Exotic options market - barrier options, lookback options, compound options and a large variety of tailor made exotic options.
8. Credit market - corporate bonds, credit default swaps (CDSs), collateralised debt obligations (CDOs), bespoke tranches of CDOs.
9. Commodity market - metals, oil, corn futures, etc.

Stocks (shares), options of European or American style, forwards and futures, annuities and bonds are all typical examples of modern financial **securities** that are actively traded on financial markets. There is also a growing interest in the so-called **structured products**, which are typically characterised by a complex design (and sometimes blamed for the market debacles).

Organised exchanges versus OTC markets.

There are two types of financial markets: **exchanges** and **over-the-counter** (OTC) markets. An exchange provides an organised forum for the buying and selling of securities. Contract terms are standardised and the exchange also acts as a clearing house for all transactions. As a consequence, buyers and sellers do not interact with each other directly, since all trades are done via the **market maker**. Prices of listed securities traded on exchanges are publicly quoted and they are nowadays easily available through electronic media. OTC markets are less strictly organised and may simply involve two institutions such as, for instance, a bank and an investment company. This feature of OTC transactions has important practical implications. In particular, privately negotiated prices of OTC traded securities are either not disclosed to other investors or they are harder to obtain.

Long and short positions.

If you hold a particular asset, you take the so-called **long position** in that asset. If, on the contrary, you owe that asset to someone, you take the so-called **short position**. As an example, we may consider the holder (long position) and the writer (short position) of an option. In some cases, e.g. for interest rate swaps, the long and short positions need to be specified by market convention.

Short-selling of a stock.

One notable feature of modern financial markets is that it is not necessary to actually own an asset in order to sell it. In a strategy called **short-selling**, an investor borrows a number of stocks and sells them. This enables him to use the proceeds to undertake other investments. At a predetermined time, he has to buy the stocks back at the market and return them to the original owner from whom the shares were temporarily borrowed. Short-selling practice is particularly attractive to those speculators, who make a bet that the price of a certain stock will fall. Clearly, if a large number of traders do indeed sell short a particular stock then its market price is very likely to fall. This phenomenon has drawn some criticism in the last couple of years and restrictions on short-selling have been implemented in some countries. Short-selling is also beneficial since it enhances the market liquidity and conveys additional information about the investors' outlook for listed companies.

1.2 Trading in Securities

Financial markets are tailored to make trading in financial securities efficient. But why would anyone want to trade financial securities in the first place? There are various reasons for trading in financial securities:

1. **Profits**: A trader believes that the price of a security will go up, and hence follows the ancient wisdom of buying at the low and selling at the high price. Of course, this naive strategy does not always work. The literal meaning of **investment** is “the sacrifice of certain present value for possibly uncertain future value”. If the future price of the security happens to be lower than expected then, obviously, a trader makes a loss. Purely profit driven traders are sometimes referred to as **speculators**.

2. **Protection**: Let us consider, for instance, a protection against the uncertainty of exchange rate. A company, which depends on imports, could wish to fix in advance the exchange rate that will apply to a future trade (for example, the exchange rate 1 AUD = 1.035 USD set on July 25, 2013 for a trade taking place on November 10, 2013). This particular goal can be achieved by entering a forward (or, sometimes, a futures) contract on the foreign currency.

3. **Hedging**: In essence, to **hedge** means to reduce the risk exposure by holding suitable positions in securities. A short position in one security can often be hedged by entering a long position in another security. As we will see in the next chapter, to hedge a short position in the call option, one enters a short position in the money market account and a long position in the underlying asset. Traders who focus on hedging are sometimes termed **hedgers**.

4. **Diversification**: The idea that **risk**, which affects each particular security in a different manner, can be reduced by holding a **diversified portfolio** of many securities. For example, if one ‘unlucky’ stock in a portfolio loses value, another one may appreciate so that the portfolio’s total value remains more stable. **(Never put all your eggs in one basket!)**

The most fundamental forces that drive security prices in the marketplace are **supply** (willingness to **sell** at a given price) and **demand** (willingness to **buy** at a given price). In over-supply or under-demand, security prices will generally fall (**bear market**). In under-supply or over-demand, security prices will generally rise (**bull market**).

Fluctuations in supply and demand are caused by many factors, including: market information, results of fundamental analysis, the rumour mill and, last but not least, the human psychology. Prices which satisfy supply and demand are said to be in market equilibrium. Hence

Market equilibrium: supply = demand.

1.3 Perfect Markets

One of central problems of modern finance is the determination of the 'fair' price of a traded security in an efficient securities market. To address this issue, one proceeds as follows:

1. first, the market prices of the most liquidly traded securities (termed **primary assets**) are modelled using stochastic processes,
2. subsequently, the 'fair' prices of other securities (the so-called **financial derivatives**) are computed in terms of prices of primary assets.

To examine the problem of valuation of derivative securities in some detail, we need first to make several simplifying assumptions in order to obtain the so-called **perfect market**. The first two assumptions are technical, meaning that they can be relaxed, but the theory and computations would become more complex. Much of modern financial theory hinges on assumption 3 below, which postulates that financial markets should behave in such a way that the proverbial **free lunch** (that is, an investment yielding profits with no risk of a loss) should not be available to investors.

We work throughout under the following standing assumptions:

1. The market is **frictionless**: there are no transaction costs, no taxes or other costs (such as the cost of carry) and no penalties for short-selling of assets. The assumption is essential for obtaining a simple dynamics of the wealth process of a portfolio.
2. Security prices are **infinitely divisible**, that is, it is possible to buy or sell any non-integer quantity of a security. This assumption allows us to make all computations using real numbers, rather than integer values.
3. The market is **arbitrage-free**, meaning that there are no arbitrage opportunities available to investors. By an **arbitrage opportunity** (or simply, an **arbitrage**) we mean here the guarantee of certain future profits for current zero investment, that is, at null initial cost. A viable pricing of derivatives in a market model that is not arbitrage-free is not feasible.

The existence of a contract that generates a positive cash flow today with no liabilities in the future is inconsistent with the arbitrage-free property. An immediate consequence of the no-arbitrage assumption is thus the so-called:

Law of One Price: If two securities have the same pattern of future cash flows then they must have the same price today.

Assumptions 1. to 3. describe a general framework in which we will be able to develop mathematical models for pricing and hedging of **financial derivatives**, in particular, European call and put options written on a stock.

1.4 European Call and Put Options

We start by specifying the rules governing the European call option.

Definition 1.4.1. *A* ***European call option*** *is a financial security, which gives its buyer the right (but not the obligation) to buy an asset at a future time T for a price K, known as the strike price. The underlying asset, the maturity time T and the strike (or exercise) price K are specified in the contract.*

We assume that an **underlying asset** is one share of the stock and the maturity date is $T > 0$. The strike price K is an arbitrary positive number. It is useful to think of options in monetary terms. Observe that, at time T, a rational holder of a European call option should proceed as follows:

- If the stock price S_T at time T is higher than K, he should buy the stock at time T for the price K from the seller of the option (an option is then **exercised**) and immediately sell it on the market for the market price S_T, leading to a positive payoff of $S_T - K$.
- If, however, the stock price at time T is lower than K, then it does not make sense to buy the stock for the price K from the seller, since it can be bought for a lower price on the market. In that case, the holder should waive his right to buy (an option is then **abandoned**) and this leads to a payoff of 0.

These arguments show that a European call option is formally equivalent to the following random payoff C_T at time T

$$C_T := \max(S_T - K, 0) = (S_T - K)^+,$$

where we denote $x^+ = \max(x, 0)$ for any real number x.

In contrast to the call option, the put option gives the right to sell an underlying asset.

Definition 1.4.2. *A* ***European put option*** *is a financial security, which gives its buyer the right (but not the obligation) to sell an asset at a future time T for a price K, known as the strike price. The underlying asset, the maturity time T and the strike (or exercise) price K are specified in the contract.*

One can show that a European put option is formally equivalent to the following random payoff P_T at time T

$$P_T := \max(K - S_T, 0) = (K - S_T)^+.$$

It is also easy to see that $C_T - P_T = S_T - K$. This is left an as exercise.

European call and put options are actively traded on organised exchanges. In this course, we will examine European options in various financial market models. A pertinent theoretical question thus arises:

What should be the 'fair' prices of European call or put options?

1.5 Interest Rates and Zero-Coupon Bonds

All participants of financial markets have access to riskless cash through borrowing and lending from the money market (retail banks). An investor who borrows cash must pay the loan back to the lender with **interest** at some time in the future. Similarly, an investor who lends cash will receive from the borrower the loan's nominal value and the interest on the loan at some time in the future. We assume throughout that the **borrowing rate** is equal to the **lending rate**. The simplest traded fixed income security is the zero-coupon bond.

Definition 1.5.1. *The unit **zero-coupon bond** maturing at T is a financial security returning to its holder one unit of cash at time T. We denote by $B(t,T)$ the price at time $t \in [0,T]$ of this bond; in particular, $B(T,T) = 1$.*

1.5.1 Discretely Compounded Interest

In the discrete-time framework, we consider the set of dates $\{0, 1, 2, \dots\}$. Let a real number $r > -1$ represent the **simple** interest rate over each period $[t, t+1]$ for $t = 0, 1, 2, \dots$. Then one unit of cash invested at time 0 in the money market account yields the following amount at time $t = 0, 1, 2, \dots$

$$B_t = (1 + r)^t.$$

Using the Law of One Price, one can show that the bond price satisfies

$$B(t,T) = \frac{B_t}{B_T} = (1 + r)^{-(T-t)}.$$

More generally, if $r(t) > -1$ is the deterministic simple interest rate over the time period $[t, t+1]$ then we obtain, for every $t = 0, 1, 2, \dots$,

$$B_t = \prod_{u=0}^{t-1} (1 + r(u)), \quad B(t,T) = \prod_{u=t}^{T-1} (1 + r(u))^{-1}.$$

1.5.2 Continuously Compounded Interest

In the continuous-time setup, the **instantaneous** interest rate is modelled either as a real number r or a deterministic function $r(t)$, meaning that the money market account satisfies $dB_t = r(t)B_t\,dt$. Hence one unit of cash invested at time 0 in the money market account yields the following amount at time t

$$B_t = e^{rt}$$

or, more generally,

$$B_t = \exp\left(\int_0^t r(u)\,du\right).$$

Hence the price at time t of the unit zero-coupon bond maturing at T equals, for all $0 \le t \le T$,

$$B(t,T) = \frac{B_t}{B_T} = \exp\left(-\int_t^T r(u)\,du\right).$$

Chapter 2

Single-Period Market Models

Single period market models are elementary (but useful) examples of financial market models. They are characterised by the following standing assumptions that are enforced throughout this chapter:

- Only a single trading period is considered.
- The beginning of the period is usually set as time $t = 0$ and the end of the period by time $t = 1$.
- At time $t = 0$, stock prices, bond prices, and prices of other financial assets or specific financial values are recorded and an agent can choose his investment, often a portfolio of stocks and bond.
- At time $t = 1$, the prices are recorded again and an agent obtains a payoff corresponding to the value of his portfolio at time $t = 1$.

It is clear that single-period models are unrealistic, since in market practice trading is not restricted to a single date, but takes place over many periods. However, they will allow us to illustrate and appreciate several important economic and mathematical principles of Financial Mathematics, without facing technical problems that are mathematically too complex and challenging.

In what follows, we will see that more realistic multi-period models in discrete time can indeed be obtained by the concatenation of many single-period models.

Single period models can thus be seen as convenient building blocks when constructing more sophisticated models. This statement can be reformulated as follows:

Single period market models can be seen as 'atoms' of Financial Mathematics in the discrete time setup.

Throughout this chapter, we assume that we deal with a finite **sample space**

$$\boxed{\Omega = \{\omega_1, \omega_2, \dots, \omega_k\}}$$

where $\omega_1, \omega_2, \dots, \omega_k$ represent all possible **states of the world** at time $t = 1$.

The prices of the financial assets at time $t = 1$ are assumed to depend on the state of the world at time $t = 1$ and thus the asset price may take a different value for each particular ω_i. The actual state of the world at time $t = 1$ is yet unknown at time $t = 0$ as, obviously, we can not foresee the future. We only assume that we are given some information (or beliefs) concerning the probabilities of various states. More precisely, we postulate that Ω is endowed with a **probability measure** $\mathbb{P}$ such that $\mathbb{P}(\omega_i) > 0$ for all $\omega_i \in \Omega$.

The probability measure $\mathbb{P}$ may, for instance, represent the beliefs of an agent. Different agents may have different beliefs and thus they may use different underlying probability measures to build and implement a model. This explains why $\mathbb{P}$ is frequently referred to as the **subjective** probability measure. However, it is also called the **statistical** (or the **historical**) probability measure when it is obtained through some statistical procedure based on historical data for asset prices. We will argue that the knowledge of an underlying probability measure $\mathbb{P}$ is irrelevant for solving of some important pricing problems.

2.1 Two-State Single-Period Market Model

The most elementary (non-trivial) market model occurs when we assume that Ω contains only two possible states of the world at the future date $t = 1$ (see Rendleman and Bartter (1979)). We denote these states by $\omega_1 = H$ and $\omega_2 = T$. We may think of the state of the world at time $t = 1$ as being determined by the toss of a (possibly asymmetric) coin, which can result in Head (H) or Tail (T), so that the sample space is given as

$$\boxed{\Omega = \{\omega_1, \omega_2\} = \{H, T\}.}$$

The result of the toss is not known at time $t = 0$ and is therefore considered as a random event. Let us stress that we do not assume that the coin is fair, i.e., that H and T have the same probability of occurrence. We only postulate that both outcomes are possible and thus there exists a number $0 < p < 1$ such that

$$\boxed{\mathbb{P}(\omega_1) = p, \quad \mathbb{P}(\omega_2) = 1 - p.}$$

We will sometimes write $q := 1 - p$.

2.1.1 Primary Assets

We consider a two-state single-period financial model with two primary assets: the stock and the money market account.

The **money market account** pays a deterministic interest rate $r > -1$. We denote by $B_0 = 1$ and $B_1 = 1+r$ the values of the money market account at time 0 and 1. This means that one dollar invested into (or borrowed from) the money market account at time $t = 0$ yields a return (or a liability) of $1 + r$ dollars at time $t = 1$.

By the **stock**, we mean throughout a **non-dividend paying common stock**, which is issued by a company listed on a stock exchange. The price of the stock at time $t = 0$ is assumed to be a known positive number, denoted by S_0. The stock price at time $t = 1$ depends on the state of the world and can therefore take two different values $S_1(\omega_1)$ and $S_1(\omega_2)$, depending on whether the state of the world at time $t = 1$ is represented by ω_1 or ω_2. Formally, S_1 is a random variable, taking the value $S_1(\omega_1)$ with probability p and the value $S_1(\omega_2)$ with probability $1 - p$. We introduce the following notation:

$$\boxed{u := \frac{S_1(\omega_1)}{S_0}, \quad d := \frac{S_1(\omega_2)}{S_0},}$$

where, without loss of generality, we assume that $0 < d < u$. The evolution of the stock price under $\mathbb{P}$ can thus be represented by the following diagram:

$$\begin{tikzcd}
 & S_0 u \\
S_0 \arrow[ur, "p"] \arrow[dr, "1-p"'] & \\
 & S_0 d
\end{tikzcd}$$

2.1.2 Wealth of a Trading Strategy

To complete the specification of a single-period market model, we need to introduce the concept of a **trading strategy** (also called a **portfolio**). An agent is allowed to invest in the money market account and the stock. Her portfolio is represented by a pair (x, φ), which is interpreted as follows:

- $x \in \mathbb{R}$ is the total **initial endowment** in dollars at time $t = 0$,
- $\varphi \in \mathbb{R}$ is the number of shares purchased (or sold short) at time $t = 0$,
- an agent invests the surplus of cash $x - \varphi S_0$ in the money market account if $x - \varphi S_0 > 0$ (or borrows cash from the money market account if $x - \varphi S_0 < 0$).

Let us emphasise that we postulate that φ can take any real value, i.e., $\varphi \in \mathbb{R}$. This assumption covers, for example, the short-selling of stock when $\varphi < 0$, as well as taking an arbitrarily high loan (i.e., an unrestricted borrowing of cash).

The **initial wealth** (or **initial value**) of a trading strategy (x, φ) at time $t = 0$ is clearly x, the initial endowment. Within the period, i.e., between time $t = 0$ and time $t = 1$, an agent does not modify her portfolio. The **terminal wealth** (or **terminal value**) $V(x, \varphi)$ of a trading strategy at time $t = 1$ is given by the total amount of cash collected when positions in shares and the money market account are closed. In particular, the terminal wealth depends on the stock price at time $t = 1$ and is therefore random. In the present setup, it can take exactly two values:

$$V_1(x, \varphi)(\omega_1) = (x - \varphi S_0)(1 + r) + \varphi S_1(\omega_1),$$
$$V_1(x, \varphi)(\omega_2) = (x - \varphi S_0)(1 + r) + \varphi S_1(\omega_2).$$

Definition 2.1.1. *The* ***wealth process*** *(or the* ***value process****) of a trading strategy (x, φ) in the two-state single-period market model is given by the pair $(V_0(x, \varphi), V_1(x, \varphi))$, where $V_0(x, \varphi) = x$ and $V_1(x, \varphi)$ is the random variable on Ω given by*

$$\boxed{V_1(x, \varphi) = (x - \varphi S_0)(1 + r) + \varphi S_1.}$$

2.1.3 Arbitrage Opportunities and Arbitrage-Free Model

An essential feature of an efficient market is that if a trading strategy can turn nothing into something, then it must also run the risk of loss. In other words, we postulate the absence of **free lunches** in the economy.

Definition 2.1.2. *An* ***arbitrage*** *(or a* ***free lunch****) is a trading strategy that begins with no money, has zero probability of losing money, and has a positive probability of making money.*

This definition is not mathematically precise, but it conveys the basic idea of arbitrage. A more formal definition is the following:

Definition 2.1.3. *A trading strategy (x, φ) in the two-state single-period market model $\mathcal{M} = (B, S)$ is said to be an* ***arbitrage opportunity*** *(or* ***arbitrage****) if*

1. *$x = V_0(x, \varphi) = 0$ (that is, a trading strategy needs no initial investment),*
2. *$V_1(x, \varphi) \geq 0$ (that is, there is no risk of loss),*
3. *$\mathbb{E}_{\mathbb{P}}(V_1(x, \varphi)) = pV_1(x, \varphi)(\omega_1) + (1 - p)V_1(x, \varphi)(\omega_2) > 0$.*

Under condition 2., condition 3. is equivalent to $\mathbb{P}(V_1(x, \varphi) > 0) > 0$, meaning that strictly positive profits are possible.

Real markets do sometimes exhibit arbitrage, but this is necessarily temporary and limited in scale; as soon as someone discovers it, the forces of supply and demand take actions that remove it. A model that admits arbitrage cannot be used for an analysis of either pricing or portfolio optimisation problems and thus it is not **viable** (since a positive wealth can be generated out of nothing). The following concept is crucial:

A financial market model is said to be* arbitrage-free *whenever no arbitrage opportunity in the model exists.

To rule out arbitrage opportunities from the two-state model, we must assume that $d < 1 + r < u$. Otherwise, we would have arbitrages in our model, as we will show now.

Proposition 2.1.1. *The two-state single-period market model* $\mathcal{M} = (B, S)$ *is arbitrage-free if and only if* $d < 1 + r < u$.

Proof. If $d \geq 1 + r$ then the following strategy is an arbitrage:

- begin with zero initial endowment and at time $t = 0$ borrow the amount S_0 from the money market in order to buy one share of the stock.

Even in the worst case of a tail in the coin toss (i.e., when $S_1 = S_0 d$) the stock at time $t = 1$ will be worth $S_0 d \geq S_0(1 + r)$, which is enough to pay off the money market debt. Also, the stock has a positive probability of being worth strictly more then the debt's value, since $u > d \geq 1 + r$, and thus $S_0 u > S_0(1 + r)$.

If $u \leq 1 + r$ then the following strategy is an arbitrage:

- begin with zero initial endowment and at time $t = 0$ sell short one share of the stock and invest the proceeds S_0 in the money market account.

Even in the worst case of a head in the coin toss (i.e., when $S_1 = S_0 u$) the cost S_1 of repurchasing the stock at time $t = 1$ will be less than or equal to the value $S_0(1 + r)$ of the money market investment at time $t = 1$. Since $d < u \leq 1 + r$, there is also a positive probability that the cost of buying back the stock will be strictly less than the value of the money market investment. We have therefore established the following implication:

No arbitrage $\Rightarrow$ $d < 1 + r < u$.

The converse is also true, namely,

$d < 1 + r < u$ $\Rightarrow$ ***No arbitrage.***

The proof of the latter implication is left as an exception (it will also follow from the discussion in Section 2.2). ■

The stock price fluctuations observed in practice are obviously much more complicated than the mouvements predicted by the two-state single-period model of the stock price. We examine this model in detail for the following reasons:

- Within this model, the concept of arbitrage pricing and its relationship to the risk-neutral pricing can be easily appreciated.
- A concatenation of several single-period market models yields a reasonably realistic model, which is commonly used by practitioners. It provides a sufficiently precise and computationally tractable approximation of a continuous-time market model.

2.1.4 Contingent Claims

Let us first recall the definition of the call option. It will be considered as a standard example of a derivative product, although in several examples we will also examine the **digital call option** (also known as the **binary call option**), rather than the standard (that is, **plain vanilla**) European call option.

Definition 2.1.4. *A **European call option** is a financial security, which gives its buyer the right, but not the obligation, to buy one share of stock at a future time T for a price K from the option writer. The maturity date T and the strike price K are specified in the contract.*

We already know that a European call option is formally equivalent to a **contingent claim** X represented by the following payoff at time $T = 1$

$$X = C_1 := \max(S_1 - K, 0) = (S_1 - K)^+$$

where we denote $x^+ = \max(x, 0)$ for any real number x. More explicitly,

$$\begin{aligned} C_1(\omega_1) &= (S_1(\omega_1) - K)^+ = (uS_0 - K)^+, \\ C_1(\omega_2) &= (S_1(\omega_2) - K)^+ = (dS_0 - K)^+. \end{aligned}$$

Recall also that a European put option is equivalent to a contingent claim X, which is represented by the following payoff at time $T = 1$

$$X = P_1 := \max(K - S_1, 0) = (K - S_1)^+.$$

The random payoff C_T (or P_T) tells us what the call (or put) option is worth at its maturity date $T = 1$. The pricing problem thus reduces to the following question:

What is the 'fair' value of the call (or put) option at time $t = 0$?

In the remaining part of this section, we will provide an answer to this question based on the idea of replication. Since we will deal with a general contingent claim X, the valuation method will in fact cover virtually any financial contract one might imagine in the present setup.

To solve the valuation problem for European options, we will consider a more general **contingent claim**, which is of the type $X = h(S_1)$ where $h : \mathbb{R}_+ \to \mathbb{R}$ is an arbitrary function. Since the stock price S_1 is a random variable on $\Omega = \{\omega_1, \omega_2\}$, the quantity $X = h(S_1)$ is also a random variable on Ω taking the following two values

$$X(\omega_1) = h(S_1(\omega_1)) = h(S_0 u),$$
$$X(\omega_2) = h(S_1(\omega_2)) = h(S_0 d),$$

where the function h is called the **payoff profile** of the claim $X = h(S_1)$. It is clear that a European call option is obtained by choosing the payoff profile h given by: $h(x) = \max(x-K, 0) = (x-K)^+$. There are many other possible choices for a payoff profile h leading to different classes of traded (exotic) options.

2.1.5 Replication Principle

To price a contingent claim, we employ the following **replication principle**:

- Assume that it is possible to find a trading strategy **replicating** a contingent claim, meaning that the trading strategy guarantees exactly the same payoff as a contingent claim at its maturity date.
- Then the initial wealth of this trading strategy must coincide with the price of a contingent claim at time $t = 0$.
- The replication principle can be seen as a consequence of the Law of One Price, which in turn is known to follow from the no-arbitrage assumption.
- We thus claim that, under the postulate that there is no arbitrage in the original and extended markets, the only possible price for the contingent claim is the initial value of the replicating strategy (provided, of course, that such a strategy exists).

Let us formalise these arguments within the present framework.

Definition 2.1.5. *A* ***replicating strategy*** *(or a* ***hedge****) for a contingent claim $X = h(S_1)$ in the two-state single period market model is a trading strategy (x, φ) which satisfies $V_1(x, \varphi) = h(S_1)$ or, more explicitly,*

$$(x - \varphi S_0)(1 + r) + \varphi S_1(\omega_1) = h(S_1(\omega_1)), \tag{2.1}$$
$$(x - \varphi S_0)(1 + r) + \varphi S_1(\omega_2) = h(S_1(\omega_2)). \tag{2.2}$$

From the above considerations, we obtain:

Proposition 2.1.2. *Let $X = h(S_1)$ be a contingent claim in the two-state single-period market model $\mathcal{M} = (B, S)$ and let (x, φ) be a replicating strategy for X. Then x is the only price for the claim at time $t = 0$, which does not allow arbitrage in the extended market in which X is a traded asset.*

We write $x = \pi_0(X)$ and we say that $\pi_0(X)$ is the **arbitrage price** of X at time 0. This definition will be later extended to more general market models.

In the next proposition, we address the following question:

How to find a replicating strategy for a given claim?

Proposition 2.1.3. *Let $X = h(S_1)$ be an arbitrary contingent claim. The pair (x, φ) given by*

$$\varphi = \frac{h(S_1(\omega_1)) - h(S_1(\omega_2))}{S_1(\omega_1) - S_1(\omega_2)} = \frac{h(uS_0) - h(dS_0)}{S_0(u-d)} \tag{2.3}$$

and

$$\pi_0(X) = \frac{1}{1+r}\big(\widetilde{p}\,h(S_1(\omega_1)) + \widetilde{q}\,h(S_1(\omega_2))\big) \tag{2.4}$$

is the unique solution to the hedging and pricing problem for X.

Proof. We note that equations (2.1) and (2.2) represent a system of two linear equations with two unknowns x and φ. We will show that it has a unique solution for any choice of the function h. By subtracting (2.2) from (2.1), we compute the hedge ratio φ, as given by equality (2.3). This equality is often called the **delta hedging formula** since the hedge ratio φ is frequently denoted as δ. One could now substitute this value for φ in equation (2.1) (or equation (2.2)) and solve for the initial value x. We will proceed in a different way, however. First, we rewrite equations (2.1) and (2.2) as follows:

$$x + \varphi\left(\frac{1}{1+r}S_1(\omega_1) - S_0\right) = \frac{1}{1+r}\,h(S_1(\omega_1)), \tag{2.5}$$

$$x + \varphi\left(\frac{1}{1+r}S_1(\omega_2) - S_0\right) = \frac{1}{1+r}\,h(S_1(\omega_2)). \tag{2.6}$$

Let us denote

$$\widetilde{p} := \frac{1+r-d}{u-d}, \quad \widetilde{q} := 1 - \widetilde{p} = \frac{u-(1+r)}{u-d}. \tag{2.7}$$

Since we assumed that $d < 1 + r < u$, we obtain $0 < \widetilde{p} < 1$ and $0 < \widetilde{q} < 1$. The following relationship is worth noting

$$\begin{aligned}
\frac{1}{1+r}\big(\widetilde{p}S_1(\omega_1) + \widetilde{q}S_1(\omega_2)\big) &= \frac{1}{1+r}\left(\frac{1+r-d}{u-d}\,S_0u + \frac{u-(1+r)}{u-d}\,S_0d\right) \\
&= S_0\frac{(1+r-d)u + (u-(1+r))d}{(u-d)(1+r)} = S_0.
\end{aligned}$$

Upon multiplying equation (2.5) with $\widetilde{p}$ and equation (2.6) with $\widetilde{q}$ and adding them, we obtain

$$x + \varphi\left(\frac{1}{1+r}\big(\widetilde{p}S_1(\omega_1) + \widetilde{q}S_1(\omega_2)\big) - S_0\right) = \frac{1}{1+r}\big(\widetilde{p}\,h(S_1(\omega_1)) + \widetilde{q}\,h(S_1(\omega_2))\big).$$

By the choice of $\widetilde{p}$ and $\widetilde{q}$, the equality above reduces to (2.4) ■

From Proposition 2.1.3, we conclude that we can always find a replicating strategy for a contingent claim in the two-state single-period market model. Models that enjoy this property are called **complete**. We will see that some models are **incomplete** and thus the technique of pricing by the replication principle fails to work for some contingent claims.

2.1.6 Risk-Neutral Valuation Formula

We define the probability measure $\widetilde{\mathbb{P}}$ on $\Omega = \{\omega_1, \omega_2\}$ by setting $\widetilde{\mathbb{P}}(\omega_1) = \widetilde{p}$ and $\widetilde{\mathbb{P}}(\omega_2) = 1 - \widetilde{p}$. Then equation (2.4) yields the following result.

Proposition 2.1.4. *The arbitrage price $\pi_0(X)$ admits the following probabilistic representation*

$$\boxed{\pi_0(X) = \mathbb{E}_{\widetilde{\mathbb{P}}}\Big(\tfrac{X}{1+r}\Big) = \mathbb{E}_{\widetilde{\mathbb{P}}}\Big(\tfrac{h(S_T)}{1+r}\Big).} \tag{2.8}$$

We refer to equation (2.8) as the **risk-neutral valuation formula**. In particular, as we already observed in the proof of Proposition 2.1.3, the equality $S_0 = \mathbb{E}_{\widetilde{\mathbb{P}}}\big(\frac{1}{1+r}\, S_T\big)$ holds.

- The probability measure $\widetilde{\mathbb{P}}$ is called a **risk-neutral probability measure**, since the price of a contingent claim X only depends on the expectation of the payoff under this probability measure, and not on its riskiness.
- As we will see, the risk-neutral probability measure will enable us to compute viable prices for contingent claims also in incomplete markets, where pricing by replication is not always feasible.

Remark 2.1.1. It is clear that the price $\pi_0(X)$ of a contingent claim X does not depend on subjective probabilities p and $1-p$. In particular, the arbitrage price of X usually does not coincide with the expected value of the discounted payoff of a claim under the subjective probability measure $\mathbb{P}$, that is, its **actuarial value**. Indeed, in general, we have that

$$\pi_0(X) = \mathbb{E}_{\widetilde{\mathbb{P}}}\left(\frac{h(S_T)}{1+r}\right) \neq \mathbb{E}_{\mathbb{P}}\left(\frac{h(S_T)}{1+r}\right) = \frac{1}{1+r}\big(ph(S_1(\omega_1)) + qh(S_1(\omega_2))\big).$$

Note that the inequality above becomes equality only if either:
(i) the equality $p = \widetilde{p}$ holds (and thus also $q = \widetilde{q}$) or
(ii) $h(S_1(\omega_1)) = h(S_1(\omega_2))$ so that the claim X is non-random.

Using (2.8) and the equality $C_T - P_T = S_T - K$, we obtain the **put-call parity** relationship

$$\boxed{\textbf{\textit{Price of a call}} \; - \; \textbf{\textit{Price of a put}} \; = C_0 - P_0 = S_0 - \tfrac{1}{1+r}\, K.} \tag{2.9}$$

Remark 2.1.2. Let us stress that the put-call parity is not specific to a single-period model and it can be extended to any arbitrage-free multi-period model. It suffices to rewrite (2.9) as follows: $C_0 - P_0 = S_0 - KB(0,T)$. More generally, in any multi-period model we have

$$C_t - P_t = S_t - KB(t,T)$$

for $t = 0, 1, \ldots, T$ where $B(t,T)$ is the price at time t of the unit zero-coupon bond maturing at T.

Example 2.1.1. Assume the parameters in the two-state market model are given by: $r = \frac{1}{3}$, $S_0 = 1$, $u = 2$, $d = .5$ and $p = .75$. Let us first find the price of the **European call option** with strike price $K = 1$ and maturity $T = 1$. We start by computing the risk-neutral probability $\widetilde{p}$

$$\widetilde{p} = \frac{1 + r - d}{u - d} = \frac{1 + \frac{1}{3} - \frac{1}{2}}{2 - \frac{1}{2}} = \frac{5}{9}.$$

Next, we compute the arbitrage price $C_0 = \pi_0(C_1)$ of the call option, which is formally represented by the contingent claim $C_1 = (S_1 - K)^+$

$$C_0 = \mathbb{E}_{\widetilde{\mathbb{P}}}\left(\frac{C_1}{1+r}\right) = \mathbb{E}_{\widetilde{\mathbb{P}}}\left(\frac{(S_1 - K)^+}{1+r}\right) = \frac{1}{1+\frac{1}{3}}\left(\frac{5}{9}\cdot(2-1) + \frac{4}{9}\cdot 0\right) = \frac{15}{36}.$$

This example makes it obvious once again that the value of the subjective probability p is completely irrelevant for the computation of the option's price. Only the risk-neutral probability $\widetilde{p}$ matters. The value of $\widetilde{p}$ depends in turn on the choice of model parameters u, d and r. □

Example 2.1.2. Using the same parameters as in the previous example, we will now compute the price of the **European put option** (i.e., an option to sell a stock) with strike price $K = 1$ and maturity $T = 1$. A simple argument shows that a put option is represented by the following payoff P_1 at time $t = 1$

$$P_1 := \max(K - S_1, 0) = (K - S_1)^+.$$

Hence the price $P_0 = \pi_0(P_1)$ of the European put at time $t = 0$ is given by

$$P_0 = \mathbb{E}_{\widetilde{\mathbb{P}}}\left(\frac{P_1}{1+r}\right) = \mathbb{E}_{\widetilde{\mathbb{P}}}\left(\frac{(K - S_1)^+}{1+r}\right) = \frac{1}{1+\frac{1}{3}}\left(\frac{5}{9}\cdot 0 + \frac{4}{9}\cdot\left(1 - \frac{1}{2}\right)\right) = \frac{1}{6}.$$

Note that the prices at time $t = 0$ of European call and put options with the same strike price K and maturity T satisfy

$$C_0 - P_0 = \frac{15}{36} - \frac{1}{6} = \frac{1}{4} = 1 - \frac{1}{1+\frac{1}{3}} = S_0 - \frac{1}{1+r}K.$$

This result is a special case of the **put-call parity** relationship. □

2.2 General Single-Period Market Models

In a general single-period model $\mathcal{M} = (B, S^1, \dots, S^n)$:

- The money market account is modeled in the same way as in Section 2.1, that is, by setting $B_0 = 1$ and $B_1 = 1 + r$.
- The price of the ith stock at time $t = 0$ (resp., at time $t = 1$) is denoted by S_0^j (resp., S_1^j). The stock prices at time $t = 0$ are known, but the prices the stocks will have at time $t = 1$ are not known at time $t = 0$, and thus they are considered to be random variables.

We assume that the observed state of the world at time $t = 1$ can be one of the k states $\omega_1, \dots, \omega_k$, which are collected in a set Ω, the so-called state space, so that

$$\boxed{\Omega = \{\omega_1, \dots, \omega_k\}.} \tag{2.10}$$

We assume that a subjective probability measure $\mathbb{P}$ on Ω is given. It tells us about the likelihood $\mathbb{P}(\omega_i)$ of the world being in the the ith state at time $t = 1$, as seen from time $t = 0$. For each $j = 1, \dots, n$, the value of the stock price S_1^j at time $t = 1$ can thus be considered as a random variable on the state space Ω, that is,

$$\boxed{S_1^j : \Omega \to \mathbb{R}.}$$

Then the real number $S_1^j(\omega)$ represents the price of the jth stock at time $t = 1$ if the world happens to be in the state $\omega \in \Omega$ at time $t = 1$. We assume that each state at time $t = 1$ is possible, that is, $\mathbb{P}(\omega) > 0$ for all $\omega \in \Omega$.

Let us now formally define the **trading strategies** (or **portfolios**) that are available to all agents.

Definition 2.2.1. *A **trading strategy** in a single-period market model $\mathcal{M}$ is a pair $(x, \varphi) \in \mathbb{R} \times \mathbb{R}^n$ where x represents the **initial endowment** at time $t = 0$ and $\varphi = (\varphi^1, \dots, \varphi^n) \in \mathbb{R}^n$ is an n-dimensional vector, where φ^j specifies the number of shares of the jth stock purchased (or sold) at time $t = 0$.*

Given a trading strategy (x, φ), it is always assumed that the amount

$$\varphi^0 := x - \sum_{j=1}^n \varphi^j S_0^j$$

is invested at time $t = 0$ in the money market account if it is a positive number (or borrowed if it is a negative number). Note that this investment yields the cash amount $\varphi^0 B_1$ at time $t = 1$.

Definition 2.2.2. *The* ***wealth process*** *of a trading strategy (x,φ) in a single-period market model $\mathcal{M}$ is given by the pair $(V_0(x,\varphi), V_1(x,\varphi))$ where*

$$\boxed{V_0(x,\varphi) := \varphi^0 B_0 + \textstyle\sum_{j=1}^n \varphi^j S_0^j = x} \tag{2.11}$$

and $V_1(x,\varphi)$ is the random variable given by

$$\boxed{V_1(x,\varphi) := \varphi^0 B_1 + \textstyle\sum_{j=1}^n \varphi^j S_1^j = \Big(x - \sum_{j=1}^n \varphi^j S_0^j\Big) B_1 + \sum_{j=1}^n \varphi^j S_1^j.} \tag{2.12}$$

Note that equation (2.11) involves random variables, meaning that the equalities hold in any possible state the world might attend at time $t = 1$, i.e. for all $\omega \in \Omega$. Formally, we may say that equalities in (2.11) hold $\mathbb{P}$**-almost surely**, that is, with probability 1.

Remark 2.2.1. We will sometimes consider also the so-called **gains process** $G(x,\varphi)$ of a trading strategy (x,φ), which is given by

$$\boxed{G_0(x,\varphi) := 0, \quad G_1(x,\varphi) := V_1(x,\varphi) - V_0(x,\varphi)} \tag{2.13}$$

or, equivalently,

$$\boxed{G_1(x,\varphi) := \varphi^0 \Delta B_1 + \textstyle\sum_{j=1}^n \varphi^j \Delta S_1^j = \Big(x - \sum_{j=1}^n \varphi^j S_0^j\Big) \Delta B_1 + \sum_{j=1}^n \varphi^j \Delta S_1^j}$$

where we denote

$$\boxed{\Delta B_1 = B_1 - B_0, \quad \Delta S_1^j := S_1^j - S_0^j.} \tag{2.14}$$

As suggested by its name, $G_1(x,\varphi)$ represents the gains (or losses) the agent obtains from his investment (x,φ).

It is often convenient to study stock prices in relation to the money market account. The **discounted stock prices** $\widehat{S}^j$ are defined as follows

$$\boxed{\begin{aligned} \widehat{S}_0^j &:= \frac{S_0^j}{B_0} = S_0^j, \\ \widehat{S}_1^j &:= \frac{S_1^j}{B_1} = \frac{1}{1+r}\, S_1^j. \end{aligned}}$$

Similarly, we define the **discounted wealth process** $\widehat{V}(x,\varphi)$ of a trading strategy (x,φ) by setting, for $t=0,1$,

$$\boxed{\widehat{V}_t(x,\varphi) := \tfrac{V_t(x,\varphi)}{B_t}.} \tag{2.15}$$

It is easy to check that

$$\boxed{\begin{aligned} \widehat{V}_0(x,\varphi) &= x, \\ \widehat{V}_1(x,\varphi) &= \Big(x - \sum_{j=1}^n \varphi^j S_0^j\Big) + \sum_{j=1}^n \varphi^j \widehat{S}_1^j = x + \sum_{j=1}^n \varphi^j \Delta \widehat{S}_1^j, \end{aligned}}$$

where we denote

$$\boxed{\Delta \widehat{S}_1^j := \widehat{S}_1^j - \widehat{S}_0^j.}$$

The **discounted gains process** $\widehat{G}_1(x,\varphi)$ is given by

$$\boxed{\widehat{G}_0(x,\varphi) := 0, \quad \widehat{G}_1(x,\varphi) := \widehat{V}_1(x,\varphi) - \widehat{V}_0(x,\varphi) = \textstyle\sum_{j=1}^n \varphi^j \Delta \widehat{S}_1^j} \tag{2.16}$$

where we also used the property of $\widehat{V}_1(x,\varphi)$. It follows from (2.16) that $\widehat{G}_1(x,\varphi)$ does not depend on x, so that, in particular, the equalities

$$\boxed{\widehat{G}_1(x,\varphi) = \widehat{G}_1(0,\varphi) = \widehat{V}_1(0,\varphi)}$$

hold for any $x \in \mathbb{R}$ and any $\varphi \in \mathbb{R}^n$.

Example 2.2.1. We consider the single-period market model $\mathcal{M} = (B, S^1, S^2)$ and we assume that the state space $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Let the interest rate be equal to $r = \frac{1}{9}$. Stock prices at time $t=0$ are given by $S_0^1 = 5$ and $S_0^2 = 10$, respectively. Random stock prices at time $t=1$ are given by the following table

	ω_1	ω_2	ω_3
S_1^1	$\frac{60}{9}$	$\frac{60}{9}$	$\frac{40}{9}$
S_1^2	$\frac{40}{3}$	$\frac{80}{9}$	$\frac{80}{9}$

Let us consider a trading strategy (x,φ) with $x \in \mathbb{R}$ and $\varphi = (\varphi^1, \varphi^2) \in \mathbb{R}^2$. Then (2.11) gives

$$V_1(x,\varphi) = (x - 5\varphi^1 - 10\varphi^2)\left(1 + \frac{1}{9}\right) + \varphi^1 S_1^1 + \varphi^2 S_1^2.$$

More explicitly, the random variable $V_1(x,\varphi):\Omega\to\mathbb{R}$ is given by

$$
\begin{aligned}
V_1(x,\varphi)(\omega_1) &= (x-5\varphi^1-10\varphi^2)\left(1+\frac{1}{9}\right)+\frac{60}{9}\varphi^1+\frac{40}{3}\varphi^2,\\
V_1(x,\varphi)(\omega_2) &= (x-5\varphi^1-10\varphi^2)\left(1+\frac{1}{9}\right)+\frac{60}{9}\varphi^1+\frac{80}{9}\varphi^2,\\
V_1(x,\varphi)(\omega_3) &= (x-5\varphi^1-10\varphi^2)\left(1+\frac{1}{9}\right)+\frac{40}{9}\varphi^1+\frac{80}{9}\varphi^2.
\end{aligned}
$$

The increments ΔS_1^j for $j=1,2$ are given by the following table

	ω_1	ω_2	ω_3
ΔS_1^1	$\frac{5}{3}$	$\frac{5}{3}$	$-\frac{5}{9}$
ΔS_1^2	$\frac{10}{3}$	$-\frac{10}{9}$	$-\frac{10}{9}$

and the gains process $G(x,\varphi)$ satisfies: $G_0(x,\varphi)=0$ and

$$
\begin{aligned}
G_1(x,\varphi)(\omega_1) &= \frac{1}{9}(x-5\varphi^1-10\varphi^2)+\frac{15}{9}\varphi^1+\frac{30}{9}\varphi^2,\\
G_1(x,\varphi)(\omega_2) &= \frac{1}{9}(x-5\varphi^1-10\varphi^2)+\frac{15}{9}\varphi^1-\frac{10}{9}\varphi^2,\\
G_1(x,\varphi)(\omega_3) &= \frac{1}{9}(x-5\varphi^1-10\varphi^2)-\frac{5}{9}\varphi^1-\frac{10}{9}\varphi^2.
\end{aligned}
$$

Let us now consider the discounted processes. The discounted stock prices at time $t=1$ are:

	ω_1	ω_2	ω_3
$\widehat{S}_1^1$	6	6	4
$\widehat{S}_1^2$	12	8	8

and the discounted wealth process at time $t=1$ equals

$$
\begin{aligned}
\widehat{V}_1(x,\varphi)(\omega_1) &= (x-5\varphi^1-10\varphi^2)+6\varphi^1+2\varphi^2,\\
\widehat{V}_1(x,\varphi)(\omega_2) &= (x-5\varphi^1-10\varphi^2)+6\varphi^1+8\varphi^2,\\
\widehat{V}_1(x,\varphi)(\omega_3) &= (x-5\varphi^1-10\varphi^2)+4\varphi^1+8\varphi^2.
\end{aligned}
$$

The increments of the discounted stock prices $\Delta\widehat{S}_1^j$ are given by

	ω_1	ω_2	ω_3
$\Delta\widehat{S}_1^1$	1	1	-1
$\Delta\widehat{S}_1^2$	2	-2	-2

The discounted gains process $\widehat{G}_1(x,\varphi)=\widehat{G}_1(0,\varphi)=\widehat{V}_1(0,\varphi)$ satisfies: $\widehat{G}_0(x,\varphi)=0$ and

$$
\begin{aligned}
\widehat{G}_1(x,\varphi)(\omega_1) &= \varphi^1+2\varphi^2,\\
\widehat{G}_1(x,\varphi)(\omega_2) &= \varphi^1-2\varphi^2,\\
\widehat{G}_1(x,\varphi)(\omega_3) &= -\varphi^1-2\varphi^2.
\end{aligned}
$$

2.2.1 Fundamental Theorem of Asset Pricing

Let us return to the study of a general single-period model. Given the definition of the wealth process (2.11) in a general single-period market model, the definition of an arbitrage in this model is very similar to Definition 2.1.3:

Definition 2.2.3. *A trading strategy (x, φ), where x denotes the total initial endowment and $\varphi = (\varphi^1, \dots, \varphi^n)$ with φ^j denoting the number of shares of stock S^j, is called an **arbitrage opportunity** (or simply an **arbitrage**) whenever*

1. $x = V_0(x, \varphi) = 0$*,*
2. $V_1(x, \varphi) \geq 0$*,*
3. *there exists $\omega_i \in \Omega$ such that $V_1(x, \varphi)(\omega_i) > 0$.*

Recall that the wealth $V_1(x, \varphi)$ at time $t = 1$ is given by equation (2.11). The following remark is useful:

Remark 2.2.2. Given that a trading strategy (x, φ) satisfies condition 2. in Definition 2.2.3, condition 3. in this definition is equivalent to the following condition:
3.′ $\mathbb{E}_{\mathbb{P}}(V_1(x, \varphi)) = \sum_{i=1}^{k} V_1(x, \varphi)(\omega_i)\mathbb{P}(\omega_i) > 0$.

The definition of an arbitrage can also be formulated in terms of the discounted wealth process or the discounted gains process. This is sometimes very useful, when one has to check whether a model admits an arbitrage or not. The following proposition gives us such a statement. The proof of Proposition 2.2.1 is left as an exercise.

Proposition 2.2.1. *A trading strategy (x, φ) in a single-period market model $\mathcal{M}$ is an arbitrage opportunity if and only if one of the following equivalent conditions hold:*

1. *Conditions 1.–3. in Definition 2.2.3 are satisfied with $\widehat{V}_t(x, \varphi)$ instead of $V_t(x, \varphi)$ for $t = 0, 1$.*
2. $x = V_0(x, \varphi) = 0$ *and conditions 2.–3. in Definition 2.2.3 are satisfied with $\widehat{G}_1(x, \varphi)$ instead of $V_1(x, \varphi)$.*

Furthermore, condition 3. can be replaced by condition 3.′

We will now return to the analysis of risk-neutral probability measures and their use in arbitrage pricing. Recall that this was already indicated in Section 2.1 (see equation (2.8)).

Definition 2.2.4. *A probability measure $\mathbb{Q}$ on Ω is called a **risk-neutral probability measure** for a single-period market model $\mathcal{M} = (B, S^1, \dots, S^n)$ if*

1. $\mathbb{Q}(\omega_i) > 0$ *for every $i = 1, \dots, k$,*
2. $\mathbb{E}_{\mathbb{Q}}(\Delta \widehat{S}_1^j) = 0$ *for every $j = 1, \dots, n$.*

We denote by $\mathbb{M}$ the set of all risk-neutral probability measures for the model $\mathcal{M}$.

The equality in condition 2. of Definition 2.2.4 can be represented as follows:

$$\mathbb{E}_{\mathbb{Q}}\left(S_1^j\right) = (1+r)S_0^j.$$

We thus see that when Ω consists of two elements only and there is a single traded stock in the model, we obtain exactly what was called a risk-neutral probability measure in Section 2.1.

We say that a model is **arbitrage-free** is no arbitrage opportunities exist. Risk-neutral probability measures are closely related to the question whether a model is arbitrage-free. The following result, which clarifies this connection, is one of the main pillars of Financial Mathematics. It was first established by Harrison and Pliska (1981) and later extended to continuous-time models.

Theorem 2.2.1. ***Fundamental Theorem of Asset Pricing.*** *A single-period market model $\mathcal{M} = (B, S^1, \dots, S^n)$ is arbitrage-free if and only if there exists a risk-neutral probability measure for the model.*

In other words, the FTAP states that the following equivalence is valid:

A single-period model $\mathcal{M}$ is arbitrage-free $\Leftrightarrow \quad \mathbb{M} \neq \emptyset.$

2.2.2 Proof of the FTAP (MATH3975)

Proof of the implication $\Leftarrow$ in Theorem 2.2.1.

Proof. ($\Leftarrow$) We assume that $\mathbb{M} \neq \emptyset$, so that there exists a risk-neutral probability measure, denoted by $\mathbb{Q}$. Let $(0, \varphi)$ be an arbitrary trading strategy with $x = 0$ and let $\widehat{V}_1(0,\varphi)$ be the associated discounted wealth at time $t = 1$. Then

$$\mathbb{E}_{\mathbb{Q}}\big(\widehat{V}_1(0,\varphi)\big) = \mathbb{E}_{\mathbb{Q}}\Big(\sum_{j=1}^n \varphi^j \Delta \widehat{S}_1^j\Big) = \sum_{j=1}^n \varphi^j \underbrace{\mathbb{E}_{\mathbb{Q}}\big(\Delta \widehat{S}_1^j\big)}_{=0} = 0.$$

If we assume that $\widehat{V}_1(0,\varphi) \geq 0$, then the last equation clearly implies that the equality $\widehat{V}_1(0,\varphi)(\omega) = 0$ must hold for all $\omega \in \Omega$. Hence, by part 1. in Proposition 2.2.1, no trading strategy satisfying all conditions of an arbitrage opportunity may exist. ■

The proof of the second implication in Theorem 2.2.1 is essentially geometric and thus some preparation is needed. To start with, it will be very handy to interpret random variables on Ω as vectors in the k-dimensional Euclidean space $\mathbb{R}^k$. This can be achieved through the following formal identification

$$\boxed{X = (X(\omega_1), X(\omega_2), \dots, X(\omega_k)) \in \mathbb{R}^k.}$$

This one-to-one correspondence means that every random variable X on Ω can be interpreted as a vector $X = (x_1, \dots, x_k)$ in $\mathbb{R}^k$ and, conversely, any vector $X \in \mathbb{R}^k$ uniquely specifies a random variable X on Ω. We can therefore identify the set of random variables on Ω with the vector space $\mathbb{R}^k$.

Similarly, any probability measure $\mathbb{Q}$ on Ω can be interpreted as a vector in $\mathbb{R}^k$. The latter identification is simply given by

$$\boxed{\mathbb{Q} = (\mathbb{Q}(\omega_1), \mathbb{Q}(\omega_2), \dots, \mathbb{Q}(\omega_k)) \in \mathbb{R}^k.}$$

It is clear that there is a one-to-one correspondence between the set of all probability measures on Ω and the set $\mathcal{P} \subset \mathbb{R}^k$ of vectors $\mathbb{Q} = (q_1, \dots, q_k)$ with the following two properties:

1. $q_i \geq 0$ for every $i = 1, \dots, k$,
2. $\sum_{i=1}^k q_i = 1$.

Let X be the vector representing the random variable X and $\mathbb{Q}$ be the vector representing the probability measure $\mathbb{Q}$. Then the **expected value** of a random variable X with respect to a probability measure $\mathbb{Q}$ on Ω can be identified with the **inner product** $\langle \cdot, \cdot \rangle$ in the Euclidean space $\mathbb{R}^k$, specifically,

$$\boxed{\mathbb{E}_{\mathbb{Q}}(X) = \textstyle\sum_{i=1}^k X(\omega_i)\mathbb{Q}(\omega_i) = \sum_{i=1}^k x_i q_i = \langle X, \mathbb{Q} \rangle.}$$

Let us define the following subset of $\mathbb{R}^k$

$$\boxed{\mathbb{W} = \left\{ X \in \mathbb{R}^k \,|\, X = \widehat{G}_1(x, \varphi) \text{ for some } (x, \varphi) \in \mathbb{R}^{n+1} \right\}.}$$

Recall that $\widehat{G}_1(x, \varphi) = \widehat{G}_1(0, \varphi) = \widehat{V}_1(0, \varphi)$ for any $x \in \mathbb{R}$ and $\varphi \in \mathbb{R}^n$. Hence $\mathbb{W}$ is the set of all possible discounted values at time $t = 1$ of trading strategies that start with an initial endowment $x = 0$, that is,

$$\boxed{\mathbb{W} = \left\{ X \in \mathbb{R}^k \,|\, X = \widehat{V}_1(0, \varphi) \text{ for some } \varphi \in \mathbb{R}^n \right\}.} \tag{2.17}$$

From equality (2.16), we deduce that $\mathbb{W}$ is a vector subspace of $\mathbb{R}^k$ generated by the vectors $\Delta \widehat{S}_1^1, \dots, \Delta \widehat{S}_1^n$. Of course, the dimension of the subspace $\mathbb{W}$ cannot be greater than k. We observe that in any arbitrage-free model the dimension of $\mathbb{W}$ is less or equal to $k - 1$ (this remark is an immediate consequence of equivalence (2.19)).

Next, we define the following set

$$\boxed{\mathbb{A} = \left\{ X \in \mathbb{R}^k \,|\, X \neq 0, \, x_i \geq 0, \, i = 1, \dots, k \right\}.} \tag{2.18}$$

The set $\mathbb{A}$ is simply the closed nonnegative **orthant** in $\mathbb{R}^k$ (the first **quadrant** when $k = 2$, the first **octant** when $k = 3$, etc.), but with the origin excluded. Indeed, the conditions in (2.18) imply that at least one component of any vector X from $\mathbb{A}$ is a strictly positive number and all other components are non-negative.

Remark 2.2.3. In view of Proposition 2.2.1, we obtain the following useful equivalence

$$\boxed{\mathcal{M} = (B, S^1, \dots, S^n) \ \textbf{\textit{is arbitrage-free}} \quad \Leftrightarrow \quad \mathbb{W} \cap \mathbb{A} = \emptyset.} \tag{2.19}$$

To establish (2.19), it suffices to note that any vector belonging to $\mathbb{W} \cap \mathbb{A}$ can be interpreted as the discounted value at time $t = 1$ of an arbitrage opportunity. The FTAP can now be restated as follows: $\mathbb{W} \cap \mathbb{A} = \emptyset \Leftrightarrow \mathbb{M} \neq \emptyset$.

We will also need the **orthogonal complement** $\mathbb{W}^\perp$ of $\mathbb{W}$, which is given by

$$\boxed{\mathbb{W}^\perp = \big\{ Z \in \mathbb{R}^k \,|\, \langle X, Z \rangle = 0 \text{ for all } X \in \mathbb{W} \big\}.} \tag{2.20}$$

Finally, we define the following subset $\mathcal{P}^+$ of the set $\mathcal{P}$ of all probability measures on Ω

$$\boxed{\mathcal{P}^+ = \big\{ \mathbb{Q} \in \mathbb{R}^k \mid \textstyle\sum_{i=1}^k q_i = 1,\ q_i > 0 \big\}.} \tag{2.21}$$

The set $\mathcal{P}^+$ can be identified with the set of all probability measures on Ω that satisfy property 1. from Definition 2.2.4.

Lemma 2.2.1. *A probability measure $\mathbb{Q}$ is a risk-neutral probability measure for a single-period market model $\mathcal{M} = (B, S^1, \dots, S^n)$ if and only if $\mathbb{Q} \in \mathbb{W}^\perp \cap \mathcal{P}^+$. Hence the set $\mathbb{M}$ of all risk-neutral probability measures satisfies $\mathbb{M} = \mathbb{W}^\perp \cap \mathcal{P}^+$.*

Proof. ($\Rightarrow$) Let us first assume that $\mathbb{Q}$ is a risk-neutral probability measure. Then, by property 1. in Definition 2.2.4, it is obvious that $\mathbb{Q}$ belongs to $\mathcal{P}^+$. Furthermore, using property 2. in Definition 2.2.4 and equality (2.16), we obtain for an arbitrary vector $X = \widehat{V}_1(0, \varphi) \in \mathbb{W}$

$$\langle X, \mathbb{Q} \rangle = \mathbb{E}_{\mathbb{Q}}\big(\widehat{V}_1(0, \varphi)\big) = \mathbb{E}_{\mathbb{Q}}\Big(\sum_{j=1}^n \varphi^j \Delta \widehat{S}_1^j \Big) = \sum_{j=1}^n \varphi^j \underbrace{\mathbb{E}_{\mathbb{Q}}\big(\Delta \widehat{S}_1^j\big)}_{=0} = 0.$$

This means that $\mathbb{Q} \in \mathbb{W}^\perp$ and thus we conclude that $\mathbb{Q} \in \mathbb{W}^\perp \cap \mathcal{P}^+$.

($\Leftarrow$) Assume now that $\mathbb{Q}$ is an arbitrary vector in $\mathbb{W}^\perp \cap \mathcal{P}^+$. We first note that $\mathbb{Q}$ defines a probability measure satisfying condition 1. in Definition 2.2.4. It remains to show that it also satisfies condition 2. in Definition 2.2.4. To this end, for a fixed (but arbitrary) $j = 1, \dots, n$, we consider the trading strategy

$(0,\varphi)$ with $\varphi=(0,\dots,0,1,0,\dots,0)=e_j$. The discounted wealth of this strategy clearly satisfies $\widehat{V}_1(0,\varphi)=\Delta\widehat{S}_1^j$. Since $\widehat{V}_1(0,\varphi)\in\mathbb{W}$ and $\mathbb{Q}\in\mathbb{W}^\perp$, we obtain

$$0=\langle\widehat{V}_1(0,\varphi),\mathbb{Q}\rangle=\langle\Delta\widehat{S}_1^j,\mathbb{Q}\rangle=\mathbb{E}_{\mathbb{Q}}\big(\Delta\widehat{S}_1^j\big).$$

Since j is arbitrary, we see that $\mathbb{Q}$ satisfies condition 2. in Definition 2.2.4. ■

Remark 2.2.4. Using Lemma 2.2.1, we obtain a purely geometric reformulation of the FTAP: $\mathbb{W}\cap\mathbb{A}=\emptyset \Leftrightarrow \mathbb{W}^\perp\cap\mathcal{P}^+\neq\emptyset$.

Separating hyperplane theorem.

In the proof of the implication $\Rightarrow$ in Theorem 2.2.1, we will employ an auxiliary result from the convex analysis, known as the **separating hyperplane theorem**. We state below the most convenient for us version of this theorem. Let us first recall the definition of convexity.

Definition 2.2.5. *A subset D of $\mathbb{R}^k$ is said to be **convex** if for all $d_1,d_2\in D$ and every $\alpha\in[0,1]$, we have that $\alpha d_1+(1-\alpha)d_2\in D$.*

It is interesting to observe that the set $\mathbb{M}$ is convex, that is, if $\mathbb{Q}_1$ and $\mathbb{Q}_2$ belong to $\mathbb{M}$ then the probability measure $\alpha\mathbb{Q}_1+(1-\alpha)\mathbb{Q}_2$ belongs to $\mathbb{M}$ for every $\alpha\in[0,1]$. A verification of this property is left as an exercise.

Proposition 2.2.2. *Let $B,C\subset\mathbb{R}^k$ be nonempty, closed, convex sets such that that $B\cap C=\emptyset$. Assume, in addition, that at least one of these sets is compact (i.e., bounded and closed). Then there exist vectors $a,y\in\mathbb{R}^k$ such that*

$$\langle b-a,y\rangle<0 \quad \textit{for all} \quad b\in B$$

and

$$\langle c-a,y\rangle>0 \quad \textit{for all} \quad c\in C.$$

Proof. The proof of this classic result can be found in any textbook on convex analysis, so we will merely sketch the main steps.

Step 1. In the first step, one shows that if D is a (nonempty) closed, convex set such that the origin 0 is not in D then there exists a non-zero vector $v\in D$ such that for every $d\in D$ we have $\langle d,v\rangle\geq\langle v,v\rangle$ (hence $\langle d,v\rangle>0$ for all $d\in D$). To this end, one checks that the vector v that realises the minimum in the problem $\min_{d\in D}\|d\|$ has the desired properties (note that since D is closed and $0\notin D$ we have that $v\neq 0$).

Step 2. In the second step, we define D as the algebraic difference of B and C, that is, $D=C-B$. More explicitly,

$$D=\big\{x\in\mathbb{R}^k\,|\,x=c-b \text{ for some } b\in B,\, c\in C\big\}.$$

It is clear that $0\notin D$. One can also check that D is convex (this always holds if B and C and convex) and closed (for closedness, we need to postulate that at least one of the sets B and C is compact).

From the first step in the proof, there exists a non-zero vector $y \in \mathbb{R}^k$ such that for all $b \in B$ and $c \in C$ we have that $\langle c - b, y \rangle \geq \langle y, y \rangle$. This in turn implies that for all $b \in B$ and $c \in C$

$$\langle c, y \rangle \geq \langle b, y \rangle + \alpha$$

where $\alpha = \langle y, y \rangle$ is a strictly positive number. Hence there exists a vector $a \in \mathbb{R}^k$ such that

$$\inf_{c \in C} \langle c, y \rangle > \langle a, y \rangle > \sup_{b \in B} \langle b, y \rangle.$$

It is now easy to check that the desired inequalities are satisfied for this choice of vectors a and y. ■

Let $a, y \in \mathbb{R}^k$ be as in Proposition 2.2.2. Observe that y is never a zero vector. We define the $(k-1)$-dimensional hyperplane $H \subset \mathbb{R}^k$ by setting

$$H = a + \left\{ x \in \mathbb{R}^k \,|\, \langle x, y \rangle = 0 \right\} = a + \{y\}^{\perp}.$$

Then we say that the hyperplane H **strictly separates** the convex sets B and C. Intuitively, the sets B and C lie on different sides of the hyperplane H and thus they can be seen as geometrically separated by H. Note that the compactness of at least one of the sets is a necessary condition for the **strict** separation of B and C.

Corollary 2.2.1. *Assume that $B \subset \mathbb{R}^k$ is a vector subspace and set C is a compact convex set such that $B \cap C = \emptyset$. Then there exists a vector $y \in \mathbb{R}^k$ such that*

$$\langle b, y \rangle = 0 \quad \textit{for all} \quad b \in B$$

and

$$\langle c, y \rangle > 0 \quad \textit{for all} \quad c \in C.$$

Proof. Note that any vector subspace of $\mathbb{R}^k$ is a closed, convex set. From Proposition 2.2.2, there exist vectors $a, y \in \mathbb{R}^k$ such that the inequality

$$\langle b, y \rangle < \langle a, y \rangle$$

is satisfied for all vectors $b \in B$. Since B is a vector space, the vector λb belongs to B for any $\lambda \in \mathbb{R}$. Hence for any $b \in B$ and $\lambda \in \mathbb{R}$ we have that

$$\langle \lambda b, y \rangle = \lambda \langle b, y \rangle < \langle a, y \rangle.$$

This in turn implies that $\langle b, y \rangle = 0$ for any vector $b \in B$, meaning that $y \in B^{\perp}$.

To establish the second inequality, we observe that from Proposition 2.2.2, we obtain

$$\langle c, y \rangle > \langle a, y \rangle \quad \text{for all} \quad c \in C.$$

Consequently, for any $c \in C$

$$\langle c, y \rangle > \langle a, y \rangle > \langle b, y \rangle = 0.$$

We conclude that $\langle c, y \rangle > 0$ for all $c \in C$. ■

Corollary 2.2.1 will be used in the proof of the implication $\Rightarrow$ in Theorem 2.2.1.

Proof of the implication $\Rightarrow$ in Theorem 2.2.1.

Proof. ($\Rightarrow$) We now assume that the model is arbitrage-free. We know that this is equivalent to the condition $\mathbb{W} \cap \mathbb{A} = \emptyset$. Our goal is to show that the class $\mathbb{M}$ of risk-neutral probabilities is non-empty. In view of Lemma 2.2.1 (see also Remark 2.2.4), it suffices to show that the following implication is valid

$$\mathbb{W} \cap \mathbb{A} = \emptyset \quad \Rightarrow \quad \mathbb{W}^{\perp} \cap \mathcal{P}^{+} \neq \emptyset.$$

We define the following auxiliary set $\mathbb{A}^{+} = \{X \in \mathbb{A} \,|\, \langle X, \mathbb{P} \rangle = 1\}$. Observe that $\mathbb{A}^{+}$ is a closed, bounded (hence compact) and convex subset of $\mathbb{R}^{k}$. Recall also that $\mathbb{P}$ is the subjective probability measure (although any other probability measure from $\mathcal{P}^{+}$ could have been used to define $\mathbb{A}^{+}$). Since $\mathbb{A}^{+} \subset \mathbb{A}$, it is clear that

$$\mathbb{W} \cap \mathbb{A} = \emptyset \;\Rightarrow\; \mathbb{W} \cap \mathbb{A}^{+} = \emptyset.$$

By applying Corollary 2.2.1 to the sets $B = \mathbb{W}$ and $C = \mathbb{A}^{+}$, we see that there exists a vector $Y \in \mathbb{W}^{\perp}$ such that

$$\langle X, Y \rangle > 0 \quad \text{for all} \quad X \in \mathbb{A}^{+}. \tag{2.22}$$

Our goal is to show that Y can be used to define a risk-neutral probability $\mathbb{Q}$ after a suitable normalisation. We need first to show that $y_i > 0$ for every i. For this purpose, for any fixed $i = 1, \dots, k$, we define the auxiliary vector X_i as the vector in $\mathbb{R}^{k}$ whose ith component equals $1/\mathbb{P}(\omega_i)$ and all other components are zero, that is,

$$X_i = \frac{1}{\mathbb{P}(\omega_i)}\,(0, \dots, 0, 1, 0 \dots, 0) = \frac{1}{\mathbb{P}(\omega_i)}\, e_i.$$

Then clearly

$$\mathbb{E}_{\mathbb{P}}(X_i) = \langle X_i, \mathbb{P} \rangle = \frac{1}{\mathbb{P}(\omega_i)}\, \mathbb{P}(\omega_i) = 1$$

and thus $X_i \in \mathbb{A}^{+}$. Let us denote by y_i the ith component of Y. It then follows from (2.22) that

$$0 < \langle X_i, Y \rangle = \frac{1}{\mathbb{P}(\omega_i)}\, y_i,$$

which means that the inequality $y_i > 0$ holds for all $i = 1, \dots, k$. We will now define a vector $\mathbb{Q} = (q_1, \dots, q_k) \in \mathbb{R}^{k}$ through the normalisation of the vector Y. To this end, we define

$$q_i = \frac{y_i}{y_1 + \dots + y_k} = c y_i$$

and we set $\mathbb{Q}(\omega_i) = q_i$ for $i = 1, \dots, k$. In this way, we defined a probability measure $\mathbb{Q}$ such that $\mathbb{Q} \in \mathcal{P}^{+}$. Furthermore, since $\mathbb{Q}$ is merely a scalar multiple of Y (i.e. $\mathbb{Q} = cY$ for some scalar c) and $\mathbb{W}^{\perp}$ is a vector space, we have that $\mathbb{Q} \in \mathbb{W}^{\perp}$ (recall that $Y \in \mathbb{W}^{\perp}$). We conclude that $\mathbb{Q} \in \mathbb{W}^{\perp} \cap \mathcal{P}^{+}$ so that $\mathbb{W}^{\perp} \cap \mathcal{P}^{+} \neq \emptyset$. By virtue of by Lemma 2.2.1, the probability measure $\mathbb{Q}$ is a risk-neutral probability measure on Ω, so that $\mathbb{M} \neq \emptyset$. ∎

Example 2.2.2. We continue the study of the market model $\mathcal{M} = (B, S^1, S^2)$ introduced in Example 2.2.1. Our aim is to illustrate the fact that the existence of a risk-neutral probability is a necessary condition for the no-arbitrage property of a market model (i.e., the 'only if' implication in Theorem 2.2.1).

Recall that the increments of discounted prices in this example were given as displayed in the following table:

	ω_1	ω_2	ω_3
$\Delta\widehat{S}^1_1$	1	1	-1
$\Delta\widehat{S}^2_1$	2	-2	-2

From the definition of the set $\mathbb{W}$ (see (2.17)) and the equality

$$\widehat{V}_1(0,\varphi) = \varphi^1 \Delta\widehat{S}^1_1 + \varphi^2 \Delta\widehat{S}^2_1,$$

it follows that

$$\mathbb{W} = \left\{ \begin{pmatrix} \varphi^1 + 2\varphi^2 \\ \varphi^1 - 2\varphi^2 \\ -\varphi^1 - 2\varphi^2 \end{pmatrix} \,\middle|\, (\varphi^1, \varphi^2) \in \mathbb{R}^2 \right\}.$$

We note that for any vector $X \in \mathbb{W}$ we have $x_1 + x_3 = 0$, where x_i is the ith component of the vector X.

Conversely, if a vector $X \in \mathbb{R}^3$ is such that $x_1 + x_3 = 0$ than we may choose $\varphi^1 = \frac{1}{2}(x_1 + x_2)$ and $\varphi^2 = \frac{1}{4}(x_1 - x_2)$ and obtain

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \varphi^1 + 2\varphi^2 \\ \varphi^1 - 2\varphi^2 \\ -\varphi^1 - 2\varphi^2 \end{pmatrix}.$$

We conclude that $\mathbb{W}$ is the plane in $\mathbb{R}^3$ given by

$$\mathbb{W} = \{X \in \mathbb{R}^3 \,|\, x_1 + x_3 = 0\} = \{X \in \mathbb{R}^3 \,|\, X = (\gamma, x_2, -\gamma)^\top \text{ for some } \gamma \in \mathbb{R}\}.$$

Hence the orthogonal complement of $\mathbb{W}$ is the line given by

$$\mathbb{W}^\perp = \{Y \in \mathbb{R}^3 \,|\, Y = (\lambda, 0, \lambda)^\top \text{ for some } \lambda \in \mathbb{R}\}.$$

It is now easily seen that $\mathbb{W}^\perp \cap \mathcal{P}^+ = \emptyset$, so that there is no risk-neutral probability measure in this model, that is, $\mathbb{M} = \emptyset$. One can also check directly that the sub-models (B, S^1) and (B, S^2) are arbitrage-free, so that the corresponding classes of risk-neutral probabilities $\mathbb{M}_1$ and $\mathbb{M}_2$ are non-empty, but $\mathbb{M} = \mathbb{M}_1 \cap \mathbb{M}_2 = \emptyset$.

In view of Theorem 2.2.1, we already know at this point that there must be an arbitrage opportunity in the model. To find explicitly an arbitrage opportunity, we use (2.19). If we compare $\mathbb{W}$ and $\mathbb{A}$, we see that

$$\mathbb{W} \cap \mathbb{A} = \{X \in \mathbb{R}^3 \,|\, x_1 = x_3 = 0,\, x_2 > 0\}$$

so that the set $\mathbb{W} \cap \mathbb{A}$ is manifestly non-empty. We deduce once again, but this time using equivalence (2.19), that the considered model is not arbitrage-free.

We will now describe all arbitrage opportunities in the model. We start with any positive number $x_2 > 0$. Since

$$\begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix} \in \mathbb{W} \cap \mathbb{A},$$

we know that there must exist a trading strategy $(0, \varphi)$ such that

$$\widehat{V}_1(0, \varphi) = \begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix}.$$

To identify $\varphi = (\varphi^1, \varphi^2) \in \mathbb{R}^2$, we solve the following system of linear equations:

$$\begin{aligned} \widehat{V}_1(0, \varphi)(\omega_1) &= \varphi^1 + 2\varphi^2 = 0, \\ \widehat{V}_1(0, \varphi)(\omega_2) &= \varphi^1 - 2\varphi^2 = x_2, \\ \widehat{V}_1(0, \varphi)(\omega_3) &= -\varphi^1 - 2\varphi^2 = 0, \end{aligned}$$

where the last equation is manifestly redundant. The unique solution reads

$$\varphi^1 = \frac{x_2}{2}, \quad \varphi^2 = -\frac{x_2}{4}.$$

As we already know, these numbers give us the number of shares of each stock we need to buy in order to obtain the arbitrage. It thus remains to compute how much money we have to invest in the money market account. Since the strategy $(0, \varphi)$ starts with zero initial endowment, the amount φ^0 invested in the money market account satisfies

$$\varphi^0 = 0 - \varphi^1 S_0^1 - \varphi^2 S_0^2 = -\frac{x_2}{2}\, 5 - \left(-\frac{x_2}{4}\right) 10 = 0.$$

This means that any arbitrage opportunity in this model is a trading strategy that only invests in risky assets, that is, stocks S^1 and S^2. One can observe that the return on the first stock dominates the return on the second. $\square$

2.2.3 Arbitrage Pricing of Contingent Claims

In Sections 2.2.3 and 2.2.4, we work under the standing assumption that a general single-period model $\mathcal{M} = (B, S^1, \ldots, S^n)$ is arbitrage-free or, equivalently, that the class $\mathbb{M}$ of risk-neutral probability measures is non-empty. We will address the following question:

How to price contingent claims in a multi-period model?

In Section 2.1, we studied claims of the type $h(S_1)$ where h is the payoff's profile, which is an arbitrary function of the single stock price S_1 at time $t = 1$. In our general model, we now have more than one stock and the payoff profiles may be complicated functions of underlying assets. Specifically, any contingent claim can now be described as $h(S_1^1, \ldots, S_1^n)$ where $h : \mathbb{R}^n \to \mathbb{R}$ is an arbitrary function.

It is thus natural to introduce the following definition of a contingent claim.

Definition 2.2.6. *A* ***contingent claim*** *in a single-period market model $\mathcal{M}$ is a random variable X on Ω representing a payoff at time $t = 1$.*

Let us state the arbitrage pricing principle for a general contingent claim.

Definition 2.2.7. *We say that a price x for the contingent claim X* ***complies with the principle of no-arbitrage*** *provided that the extended market model $\widetilde{\mathcal{M}} = (B, S^1, \dots, S^n, S^{n+1})$ consisting of the savings account B, the original stocks $S^1, \dots, S^n$, and an additional asset S^{n+1}, with the price process satisfying $S_0^{n+1} = x$ and $S_1^{n+1} = X$, is arbitrage-free.*

The additional asset S^{n+1} introduced in Definition 2.2.7 may not be interpreted as a stock, in general, since it can take negative values if the contingent claim takes negative values. For the general arbitrage pricing theory developed so far, positiveness of stock prices was not essential, however. It was only essential to assume that the price process of the money market account is strictly positive.

Step A. Pricing of an attainable claim.

To price a contingent claim for which a replicating strategy exists, we apply the replication principle.

Proposition 2.2.3. *Let X be a contingent claim in a general single-period market model and let (x, φ) be a replicating strategy for X, i.e. a trading strategy which satisfies $V_1(x, \varphi) = X$. Then the unique price of X which complies with the no arbitrage principle is $x = V_0(x, \varphi)$. It is called the* ***arbitrage price*** *of X and denoted as $\pi_0(X)$.*

Proof. The proof of this proposition hinges on the same arguments as those used in Section 2.1 and thus it is left as an exercise. ■

Definition 2.2.8. *A contingent claim X is called* ***attainable*** *if there exists a trading strategy (x, φ) which replicates X, that is, satisfies $V_1(x, \varphi) = X$.*

For attainable contingent claims the replication principle applies and it is clear how to price them, namely, the arbitrage price $\pi_0(X)$ is necessarily equal to the initial endowment x needed for a replicating strategy. There might be more than one replicating strategy, in general. However, it can be easily deduced from the no arbitrage principle that the initial endowment x for all strategies replicating a given contingent claim is unique.

In equation (2.8), we established a way to use a risk-neutral probability measure to compute the price of an option in the two-state single-period market model. The next result shows that this probabilistic approach works fine in the general single-period market model as well, at least when we restrict our attention to attainable contingent claims, for which the price is defined as the initial endowment of a replicating strategy.

Proposition 2.2.4. *Let X be an attainable contingent claim and $\mathbb{Q}$ be an arbitrary risk-neutral probability measure, that is, $\mathbb{Q} \in \mathbb{M}$. Then the arbitrage price $\pi_0(X)$ of X at time $t = 0$ satisfies*

$$\boxed{\pi_0(X) = \mathbb{E}_{\mathbb{Q}}\left((1+r)^{-1}X\right).} \tag{2.23}$$

Proof. Let (x, φ) be any replicating strategy for an attainable claim X, so that the equality $X = V_1(x, \varphi)$ is valid. Then we also have that

$$(1+r)^{-1}X = \widehat{V}_1(x, \varphi).$$

Using Definition 2.2.4 of a risk-neutral probability measure, we obtain

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}\big((1+r)^{-1}X\big) &= \mathbb{E}_{\mathbb{Q}}\Big(\widehat{V}_1(x, \varphi)\Big) = \mathbb{E}_{\mathbb{Q}}\bigg(x + \sum_{j=1}^{n} \varphi^j \Delta \widehat{S}_1^j\bigg) \\ &= x + \sum_{j=1}^{n} \varphi^j \underbrace{\mathbb{E}_{\mathbb{Q}}\big(\Delta \widehat{S}_1^j\big)}_{=0} = x \end{aligned}$$

and thus formula (2.23) holds. Note that (2.23) is valid for any choice of a risk-neutral probability measure $\mathbb{Q} \in \mathbb{M}$. ■

Step B. Example of an incomplete model.

A crucial difference between the two-state single-period model and a general single-period market model is that in the latter model a replicating strategy might not exist for some contingent claims. This may happen when there are more sources of randomness than there are stocks to invest in. We will now examine an explicit example of an arbitrage-free single-period model in which some contingent claims are not attainable, that is, an **incomplete** model.

Example 2.2.3. We consider the market model consisting of two traded assets, the money market account B and the stock S so that $\mathcal{M} = (B, S)$. We also introduce an auxiliary quantity, which we call the **volatility** and denote by v. The volatility determines whether the stock price can make big jumps or small jumps. In this model, the volatility is assumed to be random or, in other words, stochastic. Such models are called **stochastic volatility models**. To be more specific, we postulate that the state space consists of four states

$$\Omega := \{\omega_1, \omega_2, \omega_3, \omega_4\}$$

and the volatility v is the random variable on Ω given by

$$v(\omega) := \begin{cases} h & \text{if } \omega = \omega_1, \omega_4, \\ l & \text{if } \omega = \omega_2, \omega_3. \end{cases}$$

We assume here $0 < l < h < 1$, so that l stands for a lower level of the volatility whereas h represents its higher level.

The stock price S_1 is then modeled by the following formula

$$S_1(\omega) := \begin{cases} (1+v(\omega))S_0 & \text{if } \omega = \omega_1, \omega_2, \\ (1-v(\omega))S_0 & \text{if } \omega = \omega_3, \omega_4, \end{cases}$$

where, as usual, a positive number S_0 represents the initial stock price. The stock price can therefore move up or down, as in the two-state single-period market model from Section 2.1. In contrast to the two-state single-period model, the amount by which it jumps is itself random and it is determined by the realised level of the volatility. As usual, the money market account is given by: $B_0 = 1$ and $B_1 = 1 + r$.

Let us consider a **digital call** in this model, i.e., an option with the payoff

$$X = \begin{cases} 1 & \text{if } S_1 > K, \\ 0 & \text{otherwise.} \end{cases}$$

Let us assume that the strike price K satisfies

$$(1+l)S_0 < K < (1+h)S_0.$$

Then a nonzero payoff is only possible if the volatility is high and the stock jumps up, that is, when the state of the world at time $t = 1$ is given by $\omega = \omega_1$. Therefore, the contingent claim X can alternatively be represented as follows

$$X(\omega) = \begin{cases} 1 & \text{if } \omega = \omega_1, \\ 0 & \text{if } \omega = \omega_2, \omega_3, \omega_4. \end{cases}$$

Our goal is to check whether there exists a replicating strategy for this contingent claim, i.e., a trading strategy $(x, \varphi) \in \mathbb{R} \times \mathbb{R}$ satisfying

$$V_1(x, \varphi) = X.$$

Using the definition of $V_1(x, \varphi)$ and our vector notation for random variables, the last equation is equivalent to

$$(x - \varphi S_0) \begin{pmatrix} 1+r \\ 1+r \\ 1+r \\ 1+r \end{pmatrix} + \varphi \begin{pmatrix} (1+h)S_0 \\ (1+l)S_0 \\ (1-l)S_0 \\ (1-h)S_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The existence of a solution (x, φ) to this system is equivalent to the existence of a solution (α, β) to the system

$$\alpha \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} h \\ l \\ -l \\ -h \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

It is easy to see that this system of equations has no solution and thus a digital call is not an attainable contingent claim within the framework of the stochastic volatility model. □

The heuristic explanation of this example is that there is a source of randomness in the volatility, which is not hedgeable, since the volatility is not a directly traded asset. To summarise Example 2.2.3, the stochastic volatility model introduced in this example is incomplete, as for some contingent claims a replicating strategy does not exist.

Step C. Non-uniqueness of a risk-neutral value.

Let us now address the issue of non-attainability of a contingent claim from the perspective of the risk-neutral valuation formula.

- From Propositions 2.2.3 and 2.2.4 we deduce that for all risk-neutral probability measures $\mathbb{Q}$ the model may have, we get the same number when we compute the expected value in equation (2.23), provided that a claim X is attainable.
- The following example shows that the situation changes dramatically if we consider a contingent claim that is not attainable, specifically, a risk-neutral value is no longer unique.

Example 2.2.4. We start by computing the class of all risk-neutral probability measures for the stochastic volatility model $\mathcal{M}$ introduced in Example 2.2.3. To simplify computations, we will now assume that $r = 0$, so that the discounted processes coincide with the original processes. We have

$$\Delta \widehat{S}_1(\omega) := \begin{cases} v(\omega) S_0 & \text{if } \omega = \omega_1, \omega_2, \\ -v(\omega) S_0 & \text{if } \omega = \omega_3, \omega_4, \end{cases}$$

or, using the vector notation,

$$\Delta \widehat{S}_1 = S_0 \begin{pmatrix} h \\ l \\ -l \\ -h \end{pmatrix}.$$

Recall that for any trading strategy (x, φ), the discounted gains satisfies

$$\widehat{G}_1(x, \varphi) = \widehat{G}_1(0, \varphi) = \widehat{V}_1(0, \varphi) = \varphi \Delta \widehat{S}_1.$$

Hence the vector space $\mathbb{W}$ is a one-dimensional vector subspace of $\mathbb{R}^4$ spanned by the vector $\Delta \widehat{S}_1$, that is,

$$\mathbb{W} = \text{span} \left\{ \begin{pmatrix} h \\ l \\ -l \\ -l \end{pmatrix} \right\} = \left\{ \lambda \begin{pmatrix} h \\ l \\ -l \\ -h \end{pmatrix}, \lambda \in \mathbb{R} \right\}.$$

The orthogonal complement of $\mathbb{W}$ is thus the three-dimensional subspace of $\mathbb{R}^4$ given by

$$\mathbb{W}^\perp = \left\{ \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} \in \mathbb{R}^4 \,\middle|\, \langle \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}, \begin{pmatrix} h \\ l \\ -l \\ -h \end{pmatrix} \rangle = 0 \right\}.$$

Recall also that a vector $(q_1, q_2, q_3, q_4)^\top$ belongs to $\mathcal{P}^+$ if and only if the equality $q_1 + q_2 + q_3 + q_4 = 1$ holds and $q_i > 0$ for $i = 1, 2, 3, 4$. Since the set of risk-neutral probability measures is given by $\mathbb{M} = \mathbb{W}^\perp \cap \mathcal{P}^+$, we find that

$$\begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix} \in \mathbb{M} \quad \Leftrightarrow \quad q_1 > 0,\ q_2 > 0,\ q_3 > 0,\ q_4 > 0,\ q_1 + q_2 + q_3 + q_4 = 1$$

$$\text{and} \ \ h(q_1 - q_4) + l(q_2 - q_3) = 0.$$

The class of all risk-neutral probability measures in our stochastic volatility model is therefore given by

$$\mathbb{M} = \left\{ \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ 1 - (q_1 + q_2 + q_3) \end{pmatrix} \,\middle|\, \begin{array}{c} q_1 > 0,\ q_2 > 0,\ q_3 > 0,\ q_1 + q_2 + q_3 < 1 \\ l(q_2 - q_3) = h(1 - (2q_1 + q_2 + q_3)) \end{array} \right\}.$$

Clearly, this set is non-empty and thus we conclude that our stochastic volatility model is arbitrage-free.

In addition, it is not difficult to check that for every $0 < q_1 < \frac{1}{2}$ there exists a probability measure $\mathbb{Q} \in \mathbb{M}$ such that $\mathbb{Q}(\omega_1) = q_1$. Indeed, it suffices to take $q_1 \in (0, \frac{1}{2})$ and to set

$$q_4 = q_1, \quad q_2 = q_3 = \frac{1}{2} - q_1.$$

Let us now compute the (discounted) expected value of the digital call $X = (1, 0, 0, 0)^\top$ under a probability measure $\mathbb{Q} = (q_1, q_2, q_3, q_4)^\top \in \mathbb{M}$. We have

$$\mathbb{E}_{\mathbb{Q}}(X) = \langle X, \mathbb{Q} \rangle = q_1 \cdot 1 + q_2 \cdot 0 + q_3 \cdot 0 + q_4 \cdot 0 = q_1.$$

Since q_1 is here any number from the open interval $(0, \frac{1}{2})$, we obtain many different values as discounted expected values under risk-neutral probability measures. In fact, every value from the open interval $(0, \frac{1}{2})$ can be achieved. □

Note that the situation in Example 2.2.4 is completely different than in Propositions 2.2.3 and 2.2.4. The reason is that, as we already have shown in Example 2.2.3, the digital call is not an attainable contingent claim in the stochastic volatility model and thus it is not covered by Propositions 2.2.3 and 2.2.4.

Step D. Arbitrage pricing of an arbitrary claim.

In view of Example 2.2.4, the next result might be surprising, since it says that for any choice of a risk-neutral probability measure $\mathbb{Q}$, formula (2.23), yields a price which complies with the principle of no-arbitrage. Let us stress once again that we obtain different prices, in general, when we use different risk-neutral probability measures. Hence the number $\pi_0(X)$ appearing in the right-hand side of (2.24) represents a plausible price (rather than the unique price) of a non-attainable claim X. We will show in what follows that:

- an attainable claim is characterised by the uniqueness of the arbitrage price, but
- a non-attainable claim always admits a whole spectrum of prices that comply with the principle of no-arbitrage.

Proposition 2.2.5. *Let X be a possibly non-attainable contingent claim and $\mathbb{Q} \in \mathbb{M}$ be any risk-neutral probability measure for an arbitrage-free single-period market model $\mathcal{M}$. Then the real number $\pi_0(X)$ given by*

$$\pi_0(X) := \mathbb{E}_{\mathbb{Q}}\left((1+r)^{-1}X\right) \tag{2.24}$$

defines a price for the contingent claim X at time $t=0$, which complies with the principle of no-arbitrage. Moreover, the probability measure $\mathbb{Q}$ belongs to the class $\widetilde{\mathbb{M}}$ of risk-neutral probability measures for the extended market model $\widetilde{\mathcal{M}}$.

Proof. In view of the FTAP (see Theorem 2.2.1), it is enough to show that there exists a risk-neutral probability measure for the corresponding model $\widetilde{\mathcal{M}}$, which is extended by S^{n+1} as in Definition 2.2.7. By assumption, $\mathbb{Q}$ is a risk-neutral probability measure for the original model $\mathcal{M}$, consisting of consisting of the savings account B and stocks $S^1, \dots, S^n$. In other words, the probability measure $\mathbb{Q}$ is assumed to satisfy conditions 1. and 2. of Definition 2.2.4 for every $j = 1, \dots, n$. For $j = n+1$, the second condition translates into

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}(\Delta \widehat{S}_1^{n+1}) &= \mathbb{E}_{\mathbb{Q}}\left((1+r)^{-1}X - \pi_0(X)\right) = \mathbb{E}_{\mathbb{Q}}\left((1+r)^{-1}X\right) - \pi_0(X) \\
&= \pi_0(X) - \pi_0(X) = 0.
\end{aligned}$$

Hence, by Definition 2.2.4, the probability measure $\mathbb{Q}$ is a risk-neutral probability measure for the extended market model, that is, $\mathbb{Q} \in \widetilde{\mathbb{M}}$. In view of Theorem 2.2.1, the extended market model $\widetilde{\mathcal{M}}$ is arbitrage-free, so that the price $\pi_0(X)$ complies with the principle of no arbitrage. ∎

When we apply Proposition 2.2.5 to the digital call within the setup of Example 2.2.4, we conclude that any price belonging to $(0, \frac{1}{2})$ does not allow arbitrage and can therefore be considered as 'fair' (or 'viable'). The non-uniqueness of prices is a challenging problem, which was not yet completely resolved.

2.2.4 Completeness of a General Single-Period Model

Let us first characterise the models in which the problem of non-uniqueness of prices does not occur.

Definition 2.2.9. *A financial market model is called* ***complete*** *if for any contingent claim X there exists a replicating strategy $(x, \varphi) \in \mathbb{R}^{n+1}$. A model is* ***incomplete*** *when there exists a claim X for which a replicating strategy does not exist.*

By Proposition 2.2.4, the issue of computing prices of contingent claims in a complete market model is completely solved. But how can we tell whether an arbitrage-free model is complete or not?

Algebraic criterion for market completeness.

The following result gives an algebraic criterion for the market completeness. Note that the vectors $A_0, A_1, \dots, A_n \in \mathbb{R}^k$ represent the columns of the matrix A.

Proposition 2.2.6. *Let us assume that a single-period market model $\mathcal{M} = (B, S^1, \dots, S^n)$ defined on the state space $\Omega = \{\omega_1, \dots, \omega_k\}$ is arbitrage-free. Then this model is complete if and only if the $k \times (n+1)$ matrix A given by*

$$A = \begin{pmatrix} 1+r & S_1^1(\omega_1) & \cdot & \cdot & \cdot & S_1^n(\omega_1) \\ 1+r & S_1^1(\omega_2) & \cdot & \cdot & \cdot & S_1^n(\omega_2) \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ 1+r & S_1^1(\omega_k) & \cdot & \cdot & \cdot & S_1^n(\omega_k) \end{pmatrix} = (A_0, A_1, \dots, A_n)$$

has a full row rank, that is, rank $(A) = k$. Equivalently, a single-period market model $\mathcal{M}$ is complete whenever the linear subspace spanned by the vectors $A_0, A_1, \dots, A_n$ coincides with the full space $\mathbb{R}^k$.

Proof. On the one hand, by the linear algebra, the matrix A has a full row rank if and only if for every $X \in \mathbb{R}^k$ the equation $AZ = X$ has a solution $Z \in \mathbb{R}^{n+1}$.

On the other hand, if we set $\varphi^0 = x - \sum_{j=1}^n \varphi^j S_0^j$ then we have

$$\begin{pmatrix} 1+r & S_1^1(\omega_1) & \cdot & \cdot & \cdot & S_1^n(\omega_1) \\ 1+r & S_1^1(\omega_2) & \cdot & \cdot & \cdot & S_1^n(\omega_2) \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ 1+r & S_1^1(\omega_k) & \cdot & \cdot & \cdot & S_1^n(\omega_k) \end{pmatrix} \begin{pmatrix} \varphi^0 \\ \varphi^1 \\ \cdot \\ \cdot \\ \cdot \\ \varphi^n \end{pmatrix} = \begin{pmatrix} V_1(x,\varphi)(\omega_1) \\ V_1(x,\varphi)(\omega_2) \\ \cdot \\ \cdot \\ \cdot \\ V_1(x,\varphi)(\omega_k) \end{pmatrix}.$$

This shows that computing a replicating strategy for a contingent claim X is equivalent to solving the equation $AZ = X$. Hence the statement of the proposition follows immediately. ∎

Example 2.2.5. We have seen already that the stochastic volatility model discussed in Examples 2.2.3 and 2.2.4 is not complete. Another way to establish this property is to use Proposition 2.2.6. The matrix A in this case has the form

$$A = \begin{pmatrix} 1+r & (1+h)S_0 \\ 1+r & (1-h)S_0 \\ 1+r & (1+l)S_0 \\ 1+r & (1-h)S_0 \end{pmatrix}.$$

The rank of this matrix is 2 and thus, of course, it is not equal to $k = 4$. We therefore conclude that the stochastic volatility model is incomplete. □

Probabilistic criterion for attainability of a claim.

Proposition 2.2.6 offers a simple method of determining whether a given market model is complete, without the need to make explicit computations of replicating strategies. Now, the following question arises: in an incomplete market model, how to check whether a given contingent claim is attainable, without trying to compute the replicating strategy? The following results yields an answer to this question.

Proposition 2.2.7. *A contingent claim X is attainable in an arbitrage-free single-period market model $\mathcal{M} = (B, S^1, \ldots, S^n)$ if and only if the expected value*

$$\mathbb{E}_{\mathbb{Q}}\left((1+r)^{-1}X\right)$$

has the same value for all risk-neutral probability measures $\mathbb{Q} \in \mathbb{M}$.

Proof. $(\Rightarrow)$ It is immediate from Proposition 2.2.4 that if a contingent claim X is attainable then the expected value

$$\mathbb{E}_{\mathbb{Q}}\left((1+r)^{-1}X\right)$$

has the same value for all $\mathbb{Q} \in \mathbb{M}$.

$(\Leftarrow)$ **(MATH3975)** We will prove this implication by contrapositive. Let us thus assume that the contingent claim X is not attainable. Our goal is to find two risk-neutral probabilities, say $\mathbb{Q}$ and $\widehat{\mathbb{Q}}$, for which

$$\mathbb{E}_{\mathbb{Q}}\left((1+r)^{-1}X\right) \neq \mathbb{E}_{\widehat{\mathbb{Q}}}\left((1+r)^{-1}X\right). \tag{2.25}$$

Consider the $(k \times (n+1))$-matrix A introduced in Proposition 2.2.6. Since X is not attainable, there is no solution $Z \in \mathbb{R}^{n+1}$ to the system

$$AZ = X.$$

We define the following subsets of $\mathbb{R}^k$:

$$B = \text{image}\,(A) = \left\{AZ \,|\, Z \in \mathbb{R}^{n+1}\right\} \subset \mathbb{R}^k$$

and $C = \{X\}$. Then B is a subspace of $\mathbb{R}^k$ and, obviously, the set C is convex and compact. Moreover, $B \cap C = \emptyset$.

From Corollary 2.2.1, there exists a vector $Y = (y_1, \dots, y_k) \in \mathbb{R}^k$ such that

$$\begin{aligned} \langle b, Y \rangle &= 0 \ \text{ for all } \ b \in B, \\ \langle c, Y \rangle &> 0 \ \text{ for all } \ c \in C. \end{aligned}$$

In view of the definition of B and C, this means that for $j = 0, \dots, n$

$$\langle A_j, Y \rangle = 0 \ \text{ and } \ \langle X, Y \rangle > 0 \tag{2.26}$$

where A_j is the jth column of the matrix A.

Let $\mathbb{Q} \in \mathbb{M}$ be an arbitrary risk-neutral probability measure. We may choose a real number $\lambda > 0$ to be small enough in order to ensure that for every $i = 1, \dots, k$

$$\widehat{\mathbb{Q}}(\omega_i) := \mathbb{Q}(\omega_i) + \lambda(1+r)y_i > 0. \tag{2.27}$$

We will check that $\widehat{\mathbb{Q}}$ is a risk-neutral probability measure. From the definition of A in Proposition 2.2.6 and the first equality in (2.26) with $j = 0$, we obtain

$$\sum_{i=1}^k \lambda(1+r)y_i = \lambda \langle A_0, Y \rangle = 0.$$

It then follows from (2.27) that

$$\sum_{i=1}^k \widehat{\mathbb{Q}}(\omega_i) = \sum_{i=1}^k \mathbb{Q}(\omega_i) + \sum_{i=1}^k \lambda(1+r)y_i = 1$$

and thus $\widehat{\mathbb{Q}}$ is a probability measure on the space Ω. Moreover, in view of (2.27), it is clear that $\widehat{\mathbb{Q}}$ satisfies condition 1. in Definition 2.2.4.

It remains to check that $\widehat{\mathbb{Q}}$ satisfies also condition 2. in Definition 2.2.4. To this end, we examine the behaviour under $\widehat{\mathbb{Q}}$ of the discounted stock prices $\widehat{S}_1^j$. We have that, for every $j = 1, \dots, n$,

$$\begin{aligned} \mathbb{E}_{\widehat{\mathbb{Q}}}\big(\widehat{S}_1^j\big) &= \sum_{i=1}^k \widehat{\mathbb{Q}}(\omega_i)\widehat{S}_1^j(\omega_i) \\ &= \sum_{i=1}^k \mathbb{Q}(\omega_i)\widehat{S}_1^j(\omega_i) + \lambda \sum_{i=1}^k \widehat{S}_1^j(\omega_i)(1+r)y_i \\ &= \mathbb{E}_{\mathbb{Q}}\big(\widehat{S}_1^j\big) + \lambda \underbrace{\langle A_j, Y \rangle}_{=0} \qquad \text{(in view of (2.26))} \\ &= \widehat{S}_0^j \qquad \text{(since } \mathbb{Q} \in \mathbb{M}\text{)} \end{aligned}$$

We conclude that $\mathbb{E}_{\widehat{\mathbb{Q}}}\big(\Delta \widehat{S}_1^j\big) = 0$ and thus $\widehat{\mathbb{Q}} \in \mathbb{M}$, that is, $\widehat{\mathbb{Q}}$ is a risk-neutral probability measure for the market model $\mathcal{M}$.

From (2.27), it is clear that $\mathbb{Q} \neq \widehat{\mathbb{Q}}$. We have thus proven that if $\mathcal{M}$ is arbitrage-free and incomplete then there exists more than one risk-neutral probability measure.

To complete the proof, it remains to show that inequality (2.25) is satisfied. We observe that

$$\begin{aligned}
\mathbb{E}_{\widehat{\mathbb{Q}}}\left(\frac{X}{1+r}\right) &= \sum_{i=1}^{k}\widehat{\mathbb{Q}}(\omega_i)\frac{X(\omega_i)}{1+r} = \sum_{i=1}^{k}\mathbb{Q}(\omega_i)\frac{X(\omega_i)}{1+r} + \lambda\underbrace{\sum_{i=1}^{k} y_i X(\omega_i)}_{>0} \\
&> \sum_{i=1}^{k}\mathbb{Q}(\omega_i)\frac{X(\omega_i)}{1+r} = \mathbb{E}_{\mathbb{Q}}\left(\frac{X}{1+r}\right)
\end{aligned}$$

where we used the second part of (2.26) and the inequality $\lambda > 0$ in order to conclude that the braced expression is a strictly positive real number. ■

Probabilistic criterion for market completeness.

The following important result complements the Fundamental Theorem of Asset Pricing (see Theorem 2.2.1). It states that for an arbitrage-free model:

Completeness $\Leftrightarrow$ ***Uniqueness of a risk-neutral probability measure.***

Theorem 2.2.2. *Assume that a single-period market model* $\mathcal{M} = (B, S^1, \dots, S^n)$ *is arbitrage-free. Then* $\mathcal{M}$ *is complete if and only if the class* $\mathbb{M}$ *consists of a single element, that is, there exists a unique risk-neutral probability measure.*

Proof. Since $\mathcal{M}$ is assumed to be arbitrage-free, it follows from Theorem 2.2.1 that there exists at least one risk-neutral probability measure, that is, $\mathbb{M} \neq \emptyset$.

($\Leftarrow$) Assume first that a risk-neutral probability measure for $\mathcal{M}$ is unique. Then the condition in Proposition 2.2.7 is trivially satisfied for any claim X. Hence any claim X is attainable and thus the market model is complete.

($\Rightarrow$) Assume $\mathcal{M}$ is complete and consider any two risk-neutral probability measures $\mathbb{Q}$ and $\widehat{\mathbb{Q}}$ from $\mathbb{M}$. For a fixed, but arbitrary, $i = 1, \dots, k$, let the contingent claim X^i be given by

$$X^i(\omega) = \begin{cases} 1+r & \text{if } \omega = \omega_i, \\ 0 & \text{otherwise.} \end{cases}$$

Since $\mathcal{M}$ is now assumed to be complete, the contingent claim X^i is attainable. From Proposition 2.2.4, it thus follows that

$$\mathbb{Q}(\omega_i) = \mathbb{E}_{\mathbb{Q}}\left(\frac{X^i}{1+r}\right) = \pi_0(X^i) = \mathbb{E}_{\widehat{\mathbb{Q}}}\left(\frac{X^i}{1+r}\right) = \widehat{\mathbb{Q}}(\omega_i).$$

Since i was arbitrary, we see that the equality $\mathbb{Q} = \widehat{\mathbb{Q}}$ holds. We have thus shown that the class $\mathbb{M}$ consists of a single risk-neutral probability measure. ■

Chapter 3

Multi-Period Market Models

3.1 General Multi-Period Market Models

The two most important new features of a multi-period market model when compared to a single-period market model are:

- Agents can buy and sell assets not only at the beginning of the trading period, but at any time t out of a discrete set of trading times $t \in \{0, 1, 2, \ldots, T\}$ where $t = 0$ is the beginning of the trading period and $t = T$ is the end.
- Agents can gather information over time, since they can observe prices. For example, they can make their investment decisions at time t dependent on the prices of the asset observed at time t. These are unknown at time $t - 1$ and could not be used in order to choose the investment at time $t - 1$ for $t = 1, \ldots, T$.

These two aspects need special attention. We have to model trading as a dynamic process, as opposed to the static trading approach in single-period market models. We need also to take care about the evolution over time of the information available to investors. The second aspect leads to the probabilistic concepts of a σ-field and a filtration.

Definition 3.1.1. *A collection $\mathcal{F}$ of subsets of the state space Ω is called a **σ-field** (or a **σ-algebra**) whenever the following conditions hold:*

1. $\Omega \in \mathcal{F}$,
2. *if* $A \in \mathcal{F}$ *then* $A^c = \Omega \setminus A \in \mathcal{F}$,
3. *if* $A_i \in \mathcal{F}$ *for* $i \in \mathbb{N}$ *then* $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

When the state space is finite, the third condition is manifestly equivalent to the following condition:

3'. if $A, B \in \mathcal{F}$ *then* $A \cup B \in \mathcal{F}$,

which in turn, in view of condition 2., is equivalent to

3". if $A, B \in \mathcal{F}$ *then* $A \cap B \in \mathcal{F}$.

3.1.1 Static Information: Partitions

A σ-field is supposed to model a certain quantity of information. If $\mathcal{F}$ is chosen to model the level of information of an agent, it is understood that she can distinguish between two events A and B which belong to $\mathcal{F}$, but not necessarily between the particular elements of A or the particular elements of B.

- Intuitively, if an agent looks at the state space Ω then her resolution is not high enough to recognise the actual states ω, but only to see the particular subsets belonging to the σ-field $\mathcal{F}$.
- One can think of particular states ω as atoms, whereas the sets contained in the σ-field can be interpreted as molecules built from these atoms. An agent can only see the molecules, but usually she does not observe particular atoms. The larger the σ-field, the higher the resolution is and thus the information conveyed by the σ-field is richer.

Let I be some index set. By a **partition** of Ω we mean any collection $(A_i)_{i\in I}$ of non-empty subsets of Ω such that the sets A_i are pairwise disjoint, that is, $A_i \cap A_j = \emptyset$ whenever $i \neq j$ and $\bigcup_{i\in I} A_i = \Omega$. It is known that any σ-field on a finite state space Ω can represented by some partition of Ω.

Definition 3.1.2. *We say that a partition $(A_i)_{i\in I}$ **generates** a σ-field $\mathcal{F}$ if every set $A \in \mathcal{F}$ can be written as a union of some of the A_is, that is, $A = \bigcup_{j\in J} A_j$ for some subset $J \subset I$.*

- The sets A_i in a partition satisfy a certain minimality condition. If, for example, $A \in \mathcal{F}$ and $A \subset A_j$ for some j then $A = A_j$, since otherwise it could not be written as a union of some of the A_is. It is not hard to show that, given a σ-field $\mathcal{F}$ on a finite state space Ω, a partition generating this σ-field always exists and is in fact unique.

Let us consider an arbitrary mapping $X : \Omega \to \mathbb{R}$. Recall that the **preimage** (i.e., the inverse image) of a set $U \subset \mathbb{R}$ under X, denoted as $X^{-1}(U)$, is defined by $X^{-1}(U) := \{\omega \in \Omega \,|\, X(\omega) \in U\}$. The following notion of $\mathcal{F}$-measurability formalise the statement that the values of X only depend on the information conveyed by a σ-field $\mathcal{F}$.

Definition 3.1.3. *We say that a mapping $X : \Omega \to \mathbb{R}$ is **$\mathcal{F}$-measurable**, if for every closed interval $[a,b] \subset \mathbb{R}$ the preimage under X belongs to $\mathcal{F}$, that is, $X^{-1}([a,b]) \in \mathcal{F}$. Then X is called a **random variable** on $(\Omega, \mathcal{F})$.*

The extreme case where $a = b$, and thus the interval $[a,b]$ reduces to a single point, is also allowed. If we know the partition corresponding to a σ-field $\mathcal{F}$, then we may establish a handy criterion for the $\mathcal{F}$-measurability. In Proposition 3.1.1, it is not assumed that the c_js all have different values.

Proposition 3.1.1. *Let $(A_i)_{i\in I}$ be a partition generating the σ-field $\mathcal{F}$. Then a map $X : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable if and only if X is constant on each of the sets of the partition, i.e. for every $j \in I$ there exists $c_j \in \mathbb{R}$ such that*

$$X(\omega) = c_j \text{ for all } \omega \in A_j.$$

Proof. ($\Rightarrow$) Assume that $(A_i)_{i\in I}$ is a partition of the σ-field $\mathcal{F}$ and that X is $\mathcal{F}$-measurable. Let $j \in I$ be an index and let an element $\omega \in A_j$ be arbitrary. Define $c_j := X(\omega)$. We wish to show that $X(\omega) = c_j$ for all $\omega \in A_j$. Since X is $\mathcal{F}$-measurable, from Definition 3.1.3 we have that $X^{-1}(c_j) \in \mathcal{F}$. Therefore, by properties 2. and 3. in the definition of a σ-field, we have

$$\emptyset \neq X^{-1}(c_j) \cap A_j \in \mathcal{F}.$$

Since, obviously, the inclusion $X^{-1}(c_j) \cap A_j \subset A_j$ holds, from the aforementioned minimality property of the sets contained in the partition we obtain that

$$X^{-1}(c_j) \cap A_j = A_j.$$

But this means that $X(\omega) = c_j$ for all $\omega \in A_j$ and hence X is constant on A_j. By varying j, we obtain such c_js for all $j \in I$.

($\Leftarrow$) Assume now that $X : \Omega \to \mathbb{R}$ is a function, which is constant on all sets A_j belonging to the partition and that the c_j are given as in the statement of the proposition. Let $[a, b]$ be a closed interval in $\mathbb{R}$. We define

$$C := \big\{ c_j \,|\, j \in I \text{ and } c_j \in [a, b] \big\}.$$

Since $\bigcup_{i\in I} A_i = \Omega$ no other elements then the c_j occur as values of X. Therefore,

$$X^{-1}([a, b]) = X^{-1}(C) = \bigcup_{j \,|\, c_j \in C} X^{-1}(c_j) = \bigcup_{j \,|\, c_j \in C} A_j \in \mathcal{F},$$

where the last equality holds by property 3. of a σ-field. ■

Example 3.1.1.

1. If Ω is a finite state space, then the **power set** 2^Ω is a σ-field. This is the largest possible σ-field on Ω, but beware: if Ω is not finite then the power set of Ω is quite awkward to work with since it is then hard to define a probability measure $\mathbb{P}$ on the power set.

2. The **trivial** σ-field corresponding to a state space Ω is the σ-field $\mathcal{F} = \{\emptyset, \Omega\}$. This set clearly satisfies the conditions in Definition 3.1.1. The trivial σ-field is the smallest σ-field on Ω. Every random variable, which is measurable with respect to the trivial σ-field is necessarily constant, that is, deterministic.

3. On a state space $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ consisting of four elements, the following set is a σ-field:

$$\mathcal{F} = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega\}.$$

   This can be easily verified. A partition of this σ-field, is given by the two sets $A_1 = \{\omega_1, \omega_2\}$ and $A_2 = \{\omega_3, \omega_4\}$.

4. If one has a σ-field $\mathcal{F}$ and a collection $A_i \in \mathcal{F}$ for $i \in I$, one can show that there is always a smallest σ-field which contains the sets A_i. This σ-field is denoted by $\sigma\,(A_i \,|\, i \in I)$ and is called the σ-field **generated** by the collection of sets $(A_i)_{i\in I}$.

3.1.2 Dynamic Information: Filtrations

We now focus on modelling of the development of information over time.

Definition 3.1.4. *A family $(\mathcal{F}_t)_{0\le t\le T}$ of σ-fields on $(\Omega,\mathcal{F})$ is called a **filtration** if $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for all $s \le t$. For brevity, we denote $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$.*

- We interpret the σ-field $\mathcal{F}_t$ as the information available to the agent or observer at time t. In particular, $\mathcal{F}_0$ represents the information available at time 0, that is, the initial information.
- We assume that the information accumulated over time can only grow, so that we never forget anything!

To model the dynamic random behaviour, we use the concept of a **stochastic process** (or simply a **process**), that is, a sequence of random variables indexed by the time parameter $t = 0,1,\dots,T$. Recall that we say that $X : \Omega \to \mathbb{R}$ is a random variable on $(\Omega,\mathcal{F})$ whenever X is $\mathcal{F}$-measurable.

Definition 3.1.5. *A family $X = (X_t)_{0\le t\le T}$ of random variables on $(\Omega,\mathcal{F},\mathbb{P})$ is called a **stochastic process**. A stochastic process X is said to be $\mathbb{F}$-**adapted** if for every $t = 0,1,\dots,T$ the random variable X_t is $\mathcal{F}_t$-measurable.*

If a filtration $\mathbb{F}$ is already known from the context, we also sometimes just write **adapted** instead of $\mathbb{F}$-adapted.

We assume from now on that the sample space Ω is equipped with a fixed σ-field $\mathcal{F}$. Including a probability measure $\mathbb{P}$, which is defined for all sets from $\mathcal{F}$, we denote this dataset as a triple $(\Omega,\mathcal{F},\mathbb{P})$. Given a stochastic process X, we can define the unique filtration, denoted by $\mathbb{F}^X$, such that:
1. X is adapted to the filtration $\mathbb{F}^X$,
2. $\mathbb{F}^X$ represents the information flow induced by observations of X.

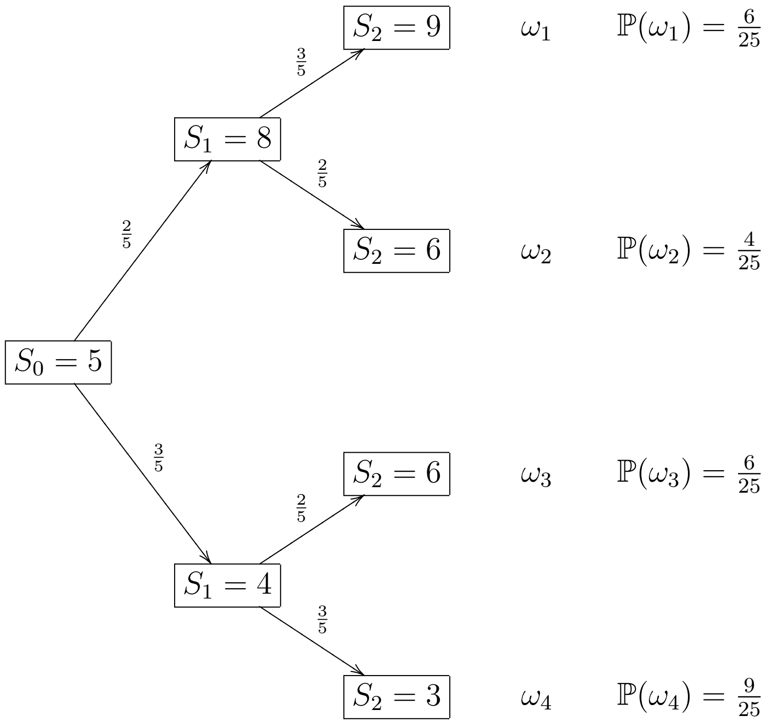
Definition 3.1.6. *Let $X = (X_t)_{0\le t\le T}$ be a stochastic process on $(\Omega,\mathcal{F},\mathbb{P})$. Define for any $t = 0,1,\dots,T$*

$$\boxed{\mathcal{F}^X_t := \sigma\left(X_u^{-1}([a,b]) \,|\, 0 \le u \le t,\ a \le b\right)}$$

*so that $\mathcal{F}^X_t$ is the smallest σ-field containing all the sets $X_u^{-1}([a,b])$ where $0 \le u \le t$ and $a \le b$. Clearly $\mathcal{F}^X_s \subset \mathcal{F}^X_t$ for $s \le t$ and thus $\mathbb{F}^X := (\mathcal{F}^X_t)_{0\le t\le T}$ is a filtration. It follows immediately that a process $(X_t)_{0\le t\le T}$ is $\mathbb{F}^X$-adapted. The filtration $\mathbb{F}^X$ is said to be **generated** by the process X.*

Example 3.1.2. Let S_t be the level of the stock price at time $0 \le t \le T$. Since for $t \ge 1$ the price S_t is not yet known at time 0, on the basis of the initial information at that time, the price S_t is modelled by means of a random variable $S_t : \Omega \to \mathbb{R}_+$. At time t, however, we observe the price of S_t and we thus assume that S_t is $\mathcal{F}_t$-measurable. We thus model the stock price evolution as an $\mathbb{F}$-adapted process S, where the filtration $\mathbb{F}$ is otherwise specified. In most cases, we will set $\mathbb{F} = \mathbb{F}^S$. $\square$

Example 3.1.3. The following diagram describes the evolution of a single stock over two time steps:



The underlying probability space is given by $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and the corresponding σ-field is assumed to be the power set of Ω, i.e. $\mathcal{F} = 2^\Omega$. The number next to an arrow indicates the probability with which a given move occurs.

- At time $t = 0$, the stock price is known and the only one possible value is $S_0 = 5$. Hence the σ-field $\mathcal{F}_0^S$ is the trivial σ-field.

- At time $t = 1$, the stock can take two possible values. The following is easy to verify:

$$S_1^{-1}([a,b]) = \begin{cases} \Omega & \text{if } a \le 4 \text{ and } b \ge 8, \text{ so that } \{4,8\} \subset [a,b] \\ \{\omega_1, \omega_2\} & \text{if } 4 < a \le 8 \text{ and } b \ge 8, \text{ so that } [a,b] \cap \{4,8\} = \{8\} \\ \{\omega_3, \omega_4\} & \text{if } a \le 4 \text{ and } 4 \le b < 8, \text{ so that } [a,b] \cap \{4,8\} = \{4\} \\ \emptyset & \text{otherwise, that is, when } [a,b] \cap \{4,8\} = \emptyset. \end{cases}$$

  We conclude that $\mathcal{F}_1^S = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega\}$. An agent is only able to tell at time $t = 1$, whether the true state of the world belongs to the set $\{\omega_1, \omega_2\}$ or to the set $\{\omega_3, \omega_4\}$.

- At time $t = 2$, an agent is able to pinpoint the true state of the world. The reason for this is as follows: $\mathcal{F}_2^S$ must contain the sets $A_1 := S_2^{-1}([9,10]) = \{\omega_1\}$, $A_2 := S_2^{-1}([6,9]) = \{\omega_1, \omega_2, \omega_3\}$, $A_3 := S_2^{-1}([5,6]) = \{\omega_2, \omega_3\}$ and $A_4 := S_1^{-1}([6,9]) = \{\omega_1, \omega_2\}$ and, since $\mathcal{F}_2^S$ is a σ-field, it must also contain: $\{\omega_1\} = A_1$, $\{\omega_2\} = A_3 \cap A_4$, $\{\omega_3\} = A_2 \setminus A_4$ and $\{\omega_4\} = A_2^c$. Hence $\mathcal{F}_2^S = 2^\Omega$. $\square$

3.1.3 Conditional Expectations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a finite probability space and let X be an arbitrary $\mathcal{F}$-measurable random variable. Assume that $\mathcal{G}$ is a σ-field which is contained in $\mathcal{F}$. Let $(A_i)_{i\in I}$ be the unique partition associated with $\mathcal{G}$.

Definition 3.1.7. *The* ***conditional expectation*** $\mathbb{E}_{\mathbb{P}}(X|\,\mathcal{G})$ *of X with respect to $\mathcal{G}$ is defined as the random variable which satisfies, for every $\omega \in A_i$,*

$$\boxed{\mathbb{E}_{\mathbb{P}}(X|\,\mathcal{G})(\omega) = \tfrac{1}{\mathbb{P}(A_i)} \textstyle\sum_{\omega\in A_i} X(\omega)\mathbb{P}(\omega) = \sum_{x_l} x_l\, \mathbb{P}(X = x_l\,|A_i)}$$

where the summation is over all possible values of X and

$$\mathbb{P}(X = x_l\,|\,A_i) = \frac{\mathbb{P}(\{X = x_l\} \cap A_i)}{\mathbb{P}(A_i)}$$

stands for the conditional probability of the event $\{\omega \in \Omega \,|\, X(\omega) = x_l\}$ given the event A_i. This means that

$$\mathbb{E}_{\mathbb{P}}(X|\,\mathcal{G}) = \sum_{i\in I} \frac{1}{\mathbb{P}(A_i)}\, \mathbb{E}_{\mathbb{P}}(X\mathbb{1}_{A_i})\, \mathbb{1}_{A_i}. \tag{3.1}$$

- Note that $\mathbb{E}_{\mathbb{P}}(X|\,\mathcal{G})$ is well defined by equation (3.1) and, by Proposition 3.1.1, the conditional expectation $\mathbb{E}_{\mathbb{P}}(X|\,\mathcal{G})$ is in fact a $\mathcal{G}$-measurable random variable, meaning here that $\mathbb{E}_{\mathbb{P}}(X|\,\mathcal{G})$ is constant on each event A_i from the partition $(A_i)_{i\in I}$.
- Intuitively, the conditional expectation $\mathbb{E}_{\mathbb{P}}(X|\,\mathcal{G})$ is the best estimate of X given the information represented by the σ-field $\mathcal{G}$.
- One can prove that the following **averaging property** uniquely characterises the conditional expectation (in addition to $\mathcal{G}$-measurability):

$$\boxed{\textstyle\sum_{\omega\in G} X(\omega)\mathbb{P}(\omega) = \sum_{\omega\in G} \mathbb{E}_{\mathbb{P}}(X|\,\mathcal{G})(\omega)\mathbb{P}(\omega), \quad \forall\, G \in \mathcal{G}.} \tag{3.2}$$

- One can represent (3.2) in terms of (discrete) integrals. Then it becomes

$$\int_G X\, d\mathbb{P} = \int_G \mathbb{E}_{\mathbb{P}}(X|\,\mathcal{G})\, d\mathbb{P}, \quad \forall\, G \in \mathcal{G}.$$

  Using this representation, the definition of conditional expectations can be extended to arbitrary random variables and σ-fields. This extension is beyond the scope of this course, however.

Example 3.1.4. We will check that in Example 3.1.3 we have

$$\mathbb{E}_{\mathbb{P}}(S_2|\,\mathcal{F}_1^S) = \begin{cases} \frac{39}{5} & \text{if } \omega \in \{\omega_1, \omega_2\}, \\ \frac{21}{5} & \text{if } \omega \in \{\omega_3, \omega_4\}. \end{cases}$$

We have shown already that $\mathcal{F}_1^S = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega\}$ and that the sets $A_1 = \{\omega_1, \omega_2\}$ and $A_2 = \{\omega_3, \omega_4\}$ form a partition of $\mathcal{F}_1^S$.

We have

$$
\begin{aligned}
\mathbb{P}(S_2 = 9 \,|\, A_1) &= \frac{\mathbb{P}(\{\omega \in A_1\} \cap \{S_2(\omega) = 9\})}{\mathbb{P}(A_1)} = \frac{\mathbb{P}(\omega_1)}{\mathbb{P}(\{\omega_1, \omega_2\})} = \frac{\frac{6}{25}}{\frac{2}{5}} = \frac{3}{5},\\
\mathbb{P}(S_2 = 6 \,|\, A_1) &= \frac{\mathbb{P}(\{\omega \in A_1\} \cap \{S_2(\omega) = 6\})}{\mathbb{P}(A_1)} = \frac{\mathbb{P}(\omega_2)}{\mathbb{P}(\{\omega_1, \omega_2\})} = \frac{\frac{4}{25}}{\frac{2}{5}} = \frac{2}{5},\\
\mathbb{P}(S_2 = 3 \,|\, A_1) &= \frac{\mathbb{P}(\{\omega \in A_1\} \cap \{S_2(\omega) = 3\})}{\mathbb{P}(A_1)} = \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(\{\omega_1, \omega_2\})} = 0,\\
\mathbb{P}(S_2 = 9 \,|\, A_2) &= \frac{\mathbb{P}(\{\omega \in A_2\} \cap \{S_2(\omega) = 9\})}{\mathbb{P}(A_2)} = \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(\{\omega_3, \omega_4\})} = 0,\\
\mathbb{P}(S_2 = 6 \,|\, A_2) &= \frac{\mathbb{P}(\{\omega \in A_2\} \cap \{S_2(\omega) = 6\})}{\mathbb{P}(A_2)} = \frac{\mathbb{P}(\omega_3)}{\mathbb{P}(\{\omega_3, \omega_4\})} = \frac{\frac{6}{25}}{\frac{3}{5}} = \frac{2}{5},\\
\mathbb{P}(S_2 = 3 \,|\, A_2) &= \frac{\mathbb{P}(\{\omega \in A_2\} \cap \{S_2(\omega) = 3\})}{\mathbb{P}(A_2)} = \frac{\mathbb{P}(\omega_4)}{\mathbb{P}(\{\omega_3, \omega_4\})} = \frac{\frac{9}{25}}{\frac{3}{5}} = \frac{3}{5}.
\end{aligned}
$$

It then follows from Definition 3.1.7 that

$$
\begin{aligned}
\mathbb{E}_{\mathbb{P}}(S_2 \,|\, \mathcal{F}_1^S)(\omega) &= 9 \cdot \frac{3}{5} + 6 \cdot \frac{2}{5} + 3 \cdot 0 = \frac{39}{5}, \quad \forall\, \omega \in A_1,\\
\mathbb{E}_{\mathbb{P}}(S_2 \,|\, \mathcal{F}_1^S)(\omega) &= 9 \cdot 0 + 6 \cdot \frac{2}{5} + 3 \cdot \frac{3}{5} = \frac{21}{5}, \quad \forall\, \omega \in A_2.
\end{aligned}
$$

We also note that

$$
\mathbb{E}_{\mathbb{P}}\big(\mathbb{E}_{\mathbb{P}}(S_2 | \, \mathcal{F}_1^S)\big) = \frac{39}{5} \cdot \frac{2}{5} + \frac{21}{5} \cdot \frac{3}{5} = \frac{141}{25} = 9 \cdot \frac{6}{25} + 6 \cdot \frac{4}{25} + 6 \cdot \frac{6}{25} + 3 \cdot \frac{9}{25} = \mathbb{E}_{\mathbb{P}}(S_2).
$$

Proposition 3.1.2. *Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with sub-σ-fields $\mathcal{G}$, $\mathcal{G}_1$ and $\mathcal{G}_2$ of $\mathcal{F}$. Assume furthermore that $\mathcal{G}_1 \subset \mathcal{G}_2$. Then*

1. ***Tower property:*** *If $X : \Omega \to \mathbb{R}$ is an $\mathcal{F}$-measurable random variable then*

$$
\boxed{\mathbb{E}_{\mathbb{P}}(X | \, \mathcal{G}_1) = \mathbb{E}_{\mathbb{P}}\big(\mathbb{E}_{\mathbb{P}}(X | \, \mathcal{G}_2) \,|\, \mathcal{G}_1\big) = \mathbb{E}_{\mathbb{P}}\big(\mathbb{E}_{\mathbb{P}}(X | \, \mathcal{G}_1) \,|\, \mathcal{G}_2\big).} \tag{3.3}
$$

2. ***Pull out property:*** *If $X : \Omega \to \mathbb{R}$ is a $\mathcal{G}$-measurable random variable and $Y : \Omega \to \mathbb{R}$ is an $\mathcal{F}$-measurable random variable then*

$$
\boxed{\mathbb{E}_{\mathbb{P}}(XY | \, \mathcal{G}) = X\, \mathbb{E}_{\mathbb{P}}(Y | \, \mathcal{G}).} \tag{3.4}
$$

3. ***Trivial conditioning:*** *If $X : \Omega \to \mathbb{R}$ is an $\mathcal{F}$-measurable random variable independent of $\mathcal{G}$ then*

$$
\boxed{\mathbb{E}_{\mathbb{P}}(X | \, \mathcal{G}) = \mathbb{E}_{\mathbb{P}}(X).} \tag{3.5}
$$

In particular, if $\mathcal{G} = \{\emptyset, \Omega\}$ is the trivial σ-field then any random variable X is independent of $\mathcal{G}$ and thus (3.5) is valid. By taking $\mathcal{G}_1 = \{\emptyset, \Omega\}$ in (3.3), we obtain $\mathbb{E}_{\mathbb{P}}(\mathbb{E}_{\mathbb{P}}(X | \, \mathcal{G})) = \mathbb{E}_{\mathbb{P}}(X)$ for any sub-σ-field $\mathcal{G}$.

Abstract Bayes formula. (MATH3975) Assume that two equivalent probability measures, $\mathbb{P}$ and $\mathbb{Q}$ say, are given on a space $(\Omega, \mathcal{F})$. Suppose that the Radon-Nikodým density of $\mathbb{Q}$ with respect to $\mathbb{P}$ equals

$$\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega) = L(\omega), \quad \mathbb{P}\text{-a.s.}$$

meaning that, for every $A \in \mathcal{F}$,

$$\int_A X\, d\mathbb{Q} = \int_A X L\, d\mathbb{P}.$$

If Ω is finite then this equality takes the following form

$$\sum_{\omega \in A} X(\omega)\, \mathbb{Q}(\omega) = \sum_{\omega \in A} X(\omega) L(\omega)\, \mathbb{P}(\omega).$$

Note that the random variable L is strictly positive $\mathbb{P}$-a.s. Moreover, L is $\mathbb{P}$-integrable with $\mathbb{E}_{\mathbb{P}}(L) = 1$. Finally, it is clear that the equality $\mathbb{E}_{\mathbb{Q}}(X) = \mathbb{E}_{\mathbb{P}}(XL)$ holds for any $\mathbb{Q}$-integrable random variable X (it suffices to take $A = \Omega$). Recall the following property of the conditional expectation $\mathbb{E}_{\mathbb{P}}(X|\,\mathcal{G})$

$$\boxed{\textstyle\int_G X\, d\mathbb{P} = \int_G \mathbb{E}_{\mathbb{P}}(X|\,\mathcal{G})\, d\mathbb{P}, \quad \forall\, G \in \mathcal{G}.}$$

We take for granted that this property uniquely determines the conditional expectation $\mathbb{E}_{\mathbb{P}}(X|\,\mathcal{G})$. Let us also mention that when Ω is finite then any random variable is $\mathbb{P}$-integrable with respect to any probability measure $\mathbb{P}$ on Ω.

Lemma 3.1.1. *Let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$ and let X be a random variable integrable with respect to $\mathbb{Q}$. Then the following abstract version of the Bayes formula holds*

$$\mathbb{E}_{\mathbb{Q}}(X \,|\, \mathcal{G}) = \frac{\mathbb{E}_{\mathbb{P}}(XL \,|\, \mathcal{G})}{\mathbb{E}_{\mathbb{P}}(L \,|\, \mathcal{G})}.$$

Proof. It is easy to check that $\mathbb{E}_{\mathbb{P}}(L \,|\, \mathcal{G})$ is strictly positive so that the right-hand side of the asserted formula is well defined. By our assumption, the random variable XL is $\mathbb{P}$-integrable. Therefore, it suffices to show that

$$\mathbb{E}_{\mathbb{P}}(XL \,|\, \mathcal{G}) = \mathbb{E}_{\mathbb{Q}}(X \,|\, \mathcal{G})\mathbb{E}_{\mathbb{P}}(L \,|\, \mathcal{G}).$$

Since the right-hand side of the last formula defines a $\mathbb{G}$-measurable random variable, we need to verify that for any set $G \in \mathcal{G}$, we have

$$\int_G XL\, d\mathbb{P} = \int_G \mathbb{E}_{\mathbb{Q}}(X \,|\, \mathcal{G})\mathbb{E}_{\mathbb{P}}(L \,|\, \mathcal{G})\, d\mathbb{P}.$$

But for every $G \in \mathcal{G}$, we obtain

$$\begin{aligned}\int_G XL\, d\mathbb{P} &= \int_G X\, d\mathbb{Q} = \int_G \mathbb{E}_{\mathbb{Q}}(X \,|\, \mathcal{G})\, d\mathbb{Q} = \int_G \mathbb{E}_{\mathbb{Q}}(X \,|\, \mathcal{G}) L\, d\mathbb{P} \\ &= \int_G \mathbb{E}_{\mathbb{P}}\big(\mathbb{E}_{\mathbb{Q}}(X \,|\, \mathcal{G}) L \,\big|\, \mathcal{G}\big)\, d\mathbb{P} = \int_G \mathbb{E}_{\mathbb{Q}}(X \,|\, \mathcal{G})\, \mathbb{E}_{\mathbb{P}}(L \,|\, \mathcal{G})\, d\mathbb{P}\end{aligned}$$

and was required to demonstrate. ■

3.1.4 Self-Financing Trading Strategies

In order to specify a multi-period market model $\mathcal{M} = (B, S^1, \dots, S^n)$, we need to describe traded assets and the class of all trading strategies that the agents in our model are allowed to use. As in Chapter 2, we assume that the market consists of:

- the money market (i.e. savings) account B with deterministic evolution given by $B_t = (1+r)^t$ where r denotes the interest rate,
- n risky assets (stocks) $S^1, \dots, S^n$ with prices assumed to follow $\mathbb{F}$-adapted stochastic processes on the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$.

Definition 3.1.8. *A* ***trading strategy*** *is an $\mathbb{R}^{n+1}$-valued, $\mathbb{F}$-adapted stochastic process $(\varphi_t) = (\varphi_t^0, \varphi_t^1, \dots, \varphi_t^n)$ where φ_t^j, $j = 1, \dots, n$ denotes the number of shares of the jth stock held (or shorted) at time t and $\varphi_t^0 B_t$ represents the amount of money in the savings account at time t. The* ***wealth process*** *(or* ***value process****) of a trading strategy $\varphi = (\varphi_t)_{0 \le t \le T}$ is the stochastic process $V(\varphi) = (V_t(\varphi))_{0 \le t \le T}$ where*

$$\boxed{V_t(\varphi) = \varphi_t^0 B_t + \sum_{j=1}^n \varphi_t^j S_t^j.}$$

In general, it is not reasonable to allow all trading strategies. A useful class of trading strategies is the class of **self-financing** trading strategies.

Definition 3.1.9. *A trading strategy $\varphi = (\varphi_t)_{0 \le t \le T}$ is called* ***self-financing*** *if for all $t = 0, \dots, T-1$*

$$\boxed{\varphi_t^0 B_{t+1} + \sum_{j=1}^n \varphi_t^j S_{t+1}^j = \varphi_{t+1}^0 B_{t+1} + \sum_{j=1}^n \varphi_{t+1}^j S_{t+1}^j.} \tag{3.6}$$

The financial interpretation of equation (3.6) can be summarised as follows:

- An agent invests money at the beginning of the period, i.e. at time t and he only can adjust his portfolio at time $t+1$. Hence he is not allowed to do any trading during the time interval $(t, t+1)$.
- The left-hand side of equation (3.6) represents the value of the agent's portfolio at time $t+1$ before rebalancement, whereas the right-hand side of equation (3.6) represents the value of his portfolio at time $t+1$ after the portfolio was revised.
- The self-financing condition says that these two values must be equal and this means that money was neither withdrawn nor added.
- If we set $t = T-1$ then both sides of formula (3.6) represent the agent's wealth at time T, that is, $V_T(\varphi)$.

It is often useful to represent the self-financing condition in terms of the gains process. Given a trading strategy φ, the corresponding **gains process** $G(\varphi) = (G_t(\varphi))_{0\le t\le T}$ is given by

$$\boxed{G_t(\varphi) := V_t(\varphi) - V_0(\varphi).} \tag{3.7}$$

Definition 3.1.10. *A* ***multi-period market model*** $\mathcal{M} = (B, S^1, \dots, S^n)$ *is given by the following data:*

1. *A probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *endowed with a filtration* $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$.
2. *The money market account* B *given by* $B_t = (1+r)^t$.
3. *A number of financial assets with prices* $S^1, \dots, S^n$*, which are assumed to be* $\mathbb{F}$*-adapted stochastic processes.*
4. *The class* Φ *of all self-financing trading strategies.*

Assume that we are given a multi-period market model, as described in Definition 3.1.10.

Definition 3.1.11. *The* ***increment process*** ΔS^j *corresponding to the* j*th stock is defined by*

$$\boxed{\Delta S^j_{t+1} := S^j_{t+1} - S^j_t \text{ for } t = 0, \dots, T-1.} \tag{3.8}$$

The ***increment process*** ΔB *of the money market account are given by*

$$\boxed{\Delta B_{t+1} := B_{t+1} - B_t = (1+r)^t r = B_t r}$$

for all $t = 0, \dots, T-1$.

As in the case of a single-period market model, it is will be convenient to consider the discounted price processes as well.

Definition 3.1.12. *The* ***discounted stock prices*** *are given by*

$$\boxed{\widehat{S}^j_t := \tfrac{S^j_t}{B_t},} \tag{3.9}$$

so that the increments of discounted prices are

$$\boxed{\Delta\widehat{S}^j_{t+1} := \widehat{S}^j_{t+1} - \widehat{S}^j_t = \tfrac{S^j_{t+1}}{B_{t+1}} - \tfrac{S^j_t}{B_t}.} \tag{3.10}$$

*The **discounted wealth process** $\widehat{V}(\varphi)$ of a trading strategy $\varphi = (\varphi_t)_{0\leq t\leq T}$ is given by*

$$\boxed{\widehat{V}_t(\varphi) := \frac{V_t(\varphi)}{B_t}.} \tag{3.11}$$

*The **discounted gains process** $\widehat{G}(\varphi)$ of a trading strategy $\varphi = (\varphi_t)_{0\leq t\leq T}$ equals*

$$\boxed{\widehat{G}_t(\varphi) := \widehat{V}_t(\varphi) - \widehat{V}_0(\varphi).} \tag{3.12}$$

The self-financing condition (3.6) can now be reformulated as follows:

Proposition 3.1.3. *Consider a multi-period market model $\mathcal{M}$. An $\mathbb{F}$-adapted trading strategy $\varphi = (\varphi_t)_{0\leq t\leq T}$ is self-financing if and only if any of the two equivalent statements hold:*

1. $G_t(\varphi) = \sum_{u=0}^{t-1} \varphi_u^0 \, \Delta B_{u+1} + \sum_{u=0}^{t-1} \sum_{j=1}^{n} \varphi_u^j \, \Delta S_{u+1}^j$ *for all* $1 \leq t \leq T$,
2. $\widehat{G}_t(\varphi) = \sum_{u=0}^{t-1} \sum_{j=1}^{n} \varphi_u^j \, \Delta \widehat{S}_{u+1}^j$ *for all* $1 \leq t \leq T$.

Proof. The proof is elementary and thus it is left as an exercise. Note that the process $\widehat{G}(\varphi)$ given by condition 2. does not depend on the component φ^0 of $\varphi \in \Phi$. ∎

3.1.5 Risk-Neutral Probability Measures

We will now redefine the general concepts of financial mathematics, such as arbitrage opportunity, arbitrage-free model, replicating strategy, arbitrage price, etc., in the framework of a multi-period market model. This is also good revision of the basic ideas from Section 2.2.

Definition 3.1.13. *A trading strategy $\varphi = (\varphi_t)_{0\leq t\leq T} \in \Phi$ is an **arbitrage opportunity** if*

1. $V_0(\varphi) = 0$,
2. $V_T(\varphi)(\omega) \geq 0$ *for all* $\omega \in \Omega$,
3. $V_T(\varphi)(\omega) > 0$ *for some* $\omega \in \Omega$ *or, equivalently,* $\mathbb{E}_{\mathbb{P}}(V_T(\varphi)) > 0$.

*We say that a multi-period market model $\mathcal{M}$ is **arbitrage-free** if no arbitrage opportunities exist.*

As in Chapter 2, one can use either the discounted wealth process or the discounted gains process in order to express the arbitrage conditions. It is also important to note that the conditions 1.–3. in Definition 3.1.13 hold under $\mathbb{P}$ whenever they are satisfied under some probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$.

Let us now come to the point of risk-neutral probability measures. As in Section 2.2, we will find that risk-neutral probability measures are very closely related to the question of arbitrage-free property and completeness of a model.

In the dynamic setup, the concept of a risk-neutral probability measure has to be extended. Using the notion of the conditional expectation, we can now define a risk-neutral probability measure in the multi-period framework.

Definition 3.1.14. *A probability measure $\mathbb{Q}$ on Ω is called a* ***risk-neutral probability measure*** *for a multi-period market model $\mathcal{M} = (B, S^1, \dots, S^n)$ whenever*

1. $\mathbb{Q}(\omega) > 0$ *for all* $\omega \in \Omega$,
2. $\mathbb{E}_{\mathbb{Q}}(\Delta \widehat{S}^j_{t+1} \,|\, \mathcal{F}_t) = 0$ *for all* $j = 1, \dots, n$ *and* $t = 0, \dots, T-1$.

We denote by $\mathbb{M}$ the class of all risk-neutral probability measures for $\mathcal{M}$.

3.1.6 Martingales (MATH3975)

Observe that condition 2. in Definition 3.1.14 is equivalent to the equality, for all $t = 0, \dots, T-1$,

$$\boxed{\mathbb{E}_{\mathbb{Q}}\big(\widehat{S}^j_{t+1} \,|\, \mathcal{F}_t\big) = \widehat{S}^j_t.}$$

The last equation leads us into the world of **martingales**, that is, stochastic processes representing fair games.

Definition 3.1.15. *An $\mathbb{F}$-adapted process $X = (X_t)_{0 \le t \le T}$ on a finite probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a* ***martingale*** *whenever for all $s < t$*

$$\boxed{\mathbb{E}_{\mathbb{P}}(X_t \,|\, \mathcal{F}_s) = X_s.}$$

To establish the equality in Definition 3.1.5, it suffices to check that for every $t = 0, 1, \dots, T-1$

$$\mathbb{E}_{\mathbb{P}}(X_{t+1} \,|\, \mathcal{F}_t) = X_t.$$

Whether a given process X is a martingale obviously depends on the choice of the filtration $\mathbb{F}$ as well as the probability measure $\mathbb{P}$ under which the conditional expectation is evaluated.

Lemma 3.1.2. *Let $\mathbb{Q} \in \mathbb{M}$ be any risk-neutral probability measure for a multi-period market model $\mathcal{M} = (B, S^1, \dots, S^n)$. Then the discounted stock price $\widehat{S}^j$ is martingale under $\mathbb{Q}$ for every $j = 1, \dots, n$.*

Proof. For $s < t$ and arbitrary i we have that

$$\widehat{S}^j_t = \widehat{S}^j_s + \sum_{u=s+1}^{t} \Delta \widehat{S}^j_u$$

where $\Delta\widehat{S}^j_u = \widehat{S}^j_u - \widehat{S}^j_{u-1}$. Hence, using the additivity as well as properties (3.4) and (3.5) of the conditional expectation under $\mathbb{Q}$, we obtain

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}(\widehat{S}^j_t \,|\, \mathcal{F}_s) &= \mathbb{E}_{\mathbb{Q}}\bigg(\widehat{S}^j_s + \sum_{u=s+1}^{t} \Delta\widehat{S}^j_u \,\bigg|\, \mathcal{F}_s\bigg) \\
&= \mathbb{E}_{\mathbb{Q}}\big(\widehat{S}^j_s \,|\, \mathcal{F}_s\big) + \sum_{u=s+1}^{t} \mathbb{E}_{\mathbb{Q}}\big(\Delta\widehat{S}^j_u \,|\, \mathcal{F}_s\big) \\
&= \widehat{S}^j_s + \sum_{u=s+1}^{t} \mathbb{E}_{\mathbb{Q}}\bigg(\underbrace{\mathbb{E}_{\mathbb{Q}}\big(\Delta\widehat{S}^j_u \,|\, \mathcal{F}_{u-1}\big)}_{=0} \,\bigg|\, \mathcal{F}_s\bigg) \\
&= \widehat{S}^j_s.
\end{aligned}$$

Note that to conclude that the braced expression is equal to 0, we used condition 2. in Definition 3.1.14. ■

The previous lemma explains why risk-neutral probability measures are also referred to as **martingale measures**. As the next result shows, one can in fact go one step further.

Proposition 3.1.4. *Let $\varphi \in \Phi$ be a trading strategy. Then the discounted wealth process $\widehat{V}(\varphi)$ and the discounted gains process $\widehat{G}(\varphi)$ are martingales under any risk-neutral probability measure $\mathbb{Q} \in \mathbb{M}$.*

Proof. From equation (3.12), it follows that for all $t = 0, \dots, T$

$$\widehat{V}_t(\varphi) = \widehat{V}_0(\varphi) + \widehat{G}_t(\varphi).$$

Since $\widehat{V}_0(\varphi)$ (the initial endowment) is a constant, it suffices to show that the process $\widehat{G}(\varphi)$ is a martingale under any $\mathbb{Q} \in \mathbb{M}$. From Proposition 3.1.3, we obtain

$$\widehat{G}_t(\varphi) = \widehat{G}_s(\varphi) + \sum_{j=1}^{n} \sum_{u=s}^{t-1} \varphi^j_u \Delta\widehat{S}^j_{u+1}.$$

Therefore, arguing as in the proof of Lemma 3.1.2, we get

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}(\widehat{G}_t(\varphi)|\mathcal{F}_s) &= \widehat{G}_s(\varphi) + \sum_{j=1}^{n} \sum_{u=s}^{t-1} \mathbb{E}_{\mathbb{Q}}\left(\varphi^j_u \Delta\widehat{S}^j_{u+1} \Big|\, \mathcal{F}_s\right) \\
&= \widehat{G}_s(\varphi) + \sum_{j=1}^{n} \sum_{u=s}^{t-1} \mathbb{E}_{\mathbb{Q}}\left(\mathbb{E}_{\mathbb{Q}}\left(\varphi^j_u \Delta\widehat{S}^j_{u+1} \Big|\, \mathcal{F}_u\right) \Big|\, \mathcal{F}_s\right) \\
&= \widehat{G}_s(\varphi) + \sum_{j=1}^{n} \sum_{u=s}^{t-1} \mathbb{E}_{\mathbb{Q}}\bigg(\varphi^j_u \underbrace{\mathbb{E}_{\mathbb{Q}}\big(\Delta\widehat{S}^j_{u+1}|\mathcal{F}_u\big)}_{=0} \,\bigg|\, \mathcal{F}_s\bigg) \\
&= \widehat{G}_s(\varphi),
\end{aligned}$$

where we used the fact that φ^j_u is $\mathcal{F}_u$-measurable, as well as property (3.5) of the conditional expectation. ■

3.1.7 Fundamental Theorem of Asset Pricing

The Fundamental Theorem of Asset Pricing is still valid in the framework of a multi-period market model. The set of allowed trading strategies Φ in a multi-period market model is assumed to be the **full** set of all self-financing and $\mathbb{F}$-adapted trading strategies. In that case, the relationship between the existence of a risk-neutral probability measure and no arbitrage is "if and only if". We will only prove here the following implication:

Existence of a risk-neutral probability measure $\Rightarrow$ ***No arbitrage.***

As in the single-period case, the proof of the inverse implication relies on the **separating hyperplane theorem**, but is more difficult. Hence the proof of part (ii) in Theorem 3.1.1 is omitted.

Theorem 3.1.1. ***Fundamental Theorem of Asset Pricing.*** *Given a multi-period market model $\mathcal{M} = (B, S^1, \dots, S^n)$, the following statements hold:*
(i) if the class $\mathbb{M}$ of risk-neutral probability measures is non-empty then there are no arbitrage opportunities in Φ and thus the model $\mathcal{M}$ is arbitrage-free,
(ii) if there are no arbitrage opportunities in the class Φ of all self-financing trading strategies then there exists a risk-neutral probability measure, so that the class $\mathbb{M}$ is non-empty.

Proof. We will only prove part (i). Let us thus postulate that $\mathbb{Q}$ is a risk-neutral probability measure for a general multi-period market model. Our goal it to show that the model is arbitrage-free. To this end, we argue by contradiction. Let us thus assume that there exists an arbitrage opportunity $\varphi \in \Phi$. Such a strategy would satisfy the following conditions (see Proposition 2.2.1):
1. the initial endowment $\widehat{V}_0(\varphi) = 0$,
2. the discounted gains process $\widehat{G}_T(\varphi) \geq 0$,
3. there exists at least one $\omega \in \Omega$ such that $\widehat{G}_T(\varphi)(\omega) > 0$.

On the one hand, from conditions 2. and 3. above, we deduce easily that

$$\mathbb{E}_{\mathbb{Q}}\big(\widehat{G}_T(\varphi)\big) > 0.$$

On the other hand, using the definition of the discounted gains process, property 2. of a risk-neutral probability measure (see Definition 3.1.14), and properties (3.4) and (3.5) of the conditional expectation, we obtain

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\big(\widehat{G}_T(\varphi)\big) &= \mathbb{E}_{\mathbb{Q}}\bigg(\sum_{j=1}^{n}\sum_{u=0}^{T-1}\varphi_u^j \Delta \widehat{S}_{u+1}^j\bigg) = \sum_{j=1}^{n}\sum_{u=0}^{T-1}\mathbb{E}_{\mathbb{Q}}\big(\varphi_u^j \Delta \widehat{S}_{u+1}^j\big) \\ &= \sum_{j=1}^{n}\sum_{u=0}^{T-1}\mathbb{E}_{\mathbb{Q}}\big(\mathbb{E}_{\mathbb{Q}}(\varphi_u^j \Delta \widehat{S}_{u+1}^j \,|\, \mathcal{F}_u)\big) = \sum_{j=1}^{n}\sum_{u=0}^{T-1}\mathbb{E}_{\mathbb{Q}}\big(\varphi_u^j\, \underbrace{\mathbb{E}_{\mathbb{Q}}(\Delta \widehat{S}_{u+1}^j \,|\, \mathcal{F}_u)}_{=0}\big) = 0.\end{aligned}$$

This clearly contradicts the inequality obtained in the first step. Hence there cannot be an arbitrage in the market model $\mathcal{M} = (B, S^1, \dots, S^n)$. ∎

Remark 3.1.1. (MATH3975) We already know from Proposition 3.1.4 that $\widehat{G}(\varphi)$ is a martingale under $\mathbb{Q}$. Hence we can conclude from (3.5) that

$$\mathbb{E}_{\mathbb{Q}}(\widehat{G}_T(\varphi)) = \mathbb{E}_{\mathbb{Q}}(\widehat{G}_T(\varphi) \,|\, \mathcal{F}_0) = \widehat{G}_0(\varphi) = 0.$$

This observation would, of course, shorten the argument in the proof of Theorem 3.1.1 significantly.

3.1.8 Pricing of European Contingent Claims

Let us extend the concepts of a European contingent claim and its replication. Note that a contingent claim of European style can only be exercised at its maturity date T (as opposed to contingent claims of American style studied in Section 4.4). A **European contingent claim** in a multi-period market model is an $\mathcal{F}_T$-measurable random variable X on Ω to be interpreted as the payoff at the terminal date T. For brevity, European contingent claims will also be referred to as **contingent claims**.

Definition 3.1.16. *A **replicating strategy** (or **hedging strategy**) for a contingent claim X is a trading strategy $\varphi \in \Phi$ such that $V_T(\varphi) = X$, that is, the terminal wealth of the trading strategy matches the claim's payoff.*

The standard method to price a contingent claim is to employ the replication principle. The price will now depend on time t and thus one has to specify a whole price process $\pi(X)$, rather than just an initial price, as in the single-period market model. Therefore, we need first to extend Definition 2.2.7 to a multi-period setup. Obviously, $\pi_T(X) = X$ for any claim X.

Definition 3.1.17. *We say that an $\mathbb{F}$-adapted stochastic process $(\pi_t(X))_{0 \le t \le T}$ is a price process for the contingent claim X that **complies with the principle of no-arbitrage** if there is no $\mathbb{F}$-adapted and self-financing arbitrage strategy in the extended model $\widetilde{\mathcal{M}} = (B, S^1, \dots, S^n, S^{n+1})$ with an additional asset S^{n+1} given by $S^{n+1}_t = \pi_t(X)$ for $0 \le t \le T-1$ and $S^{n+1}_T = X$.*

The following result generalises Proposition 2.2.3. We deal here with an attainable claim, meaning that we assume a priori that a replicating strategy for X exists. The proof of Proposition 3.1.5 is left as an exercise.

Proposition 3.1.5. *Let X be a contingent claim in an arbitrage-free multi-period market model $\mathcal{M}$ and let $\varphi \in \Phi$ be any replicating strategy for X. Then the only price process of X that complies with the principle of no-arbitrage is the wealth process $V(\varphi)$.*

- The arbitrage price of X at time t is unique and it is also denoted as $\pi_t(X)$. Hence the equality $\pi_t(X) = V_t(\varphi)$ holds for any replicating strategy $\varphi \in \Phi$ for X.
- In particular, the price at time $t = 0$ is the initial endowment of any replicating strategy for X, that is, $\pi_0(X) = V_0(\varphi)$ for any strategy $\varphi \in \Phi$ such that $V_T(\varphi) = X$.

Example 3.1.5. Let us consider the two-period model $\mathcal{M} = (B, S)$ with the stock price S specified as in Example 3.1.3. We assume that the interest rate is equal to zero, so that the equality $B_t = 1$ holds for $t = 0, 1, 2$.

We will examine the digital call paying one unit of cash at maturity date $T = 2$ whenever the stock price at this date is greater or equal than 8, that is,

$$X = \begin{cases} 1 & \text{if } S_2 \geq 8, \\ 0 & \text{otherwise.} \end{cases}$$

In our model, the contingent claim X pays off one unit of cash if and only if the state of the world is ω_1 and zero else.

How can we find a self-financing replicating strategy for X, that is, a self-financing strategy φ satisfying

$$V_2(\varphi)(\omega) = \varphi_2^0(\omega) + \varphi_2^1(\omega)S_2(\omega) = X(\omega)?$$

The computational trick now is to decompose the model into single-period market models. In our case, we need to examine three sub-models: one starting at time $t = 0$ going until time $t = 1$ with price $S_0 = 5$, one starting at time $t = 1$ with price $S_1 = 8$ and one starting at time $t = 1$ with price $S_1 = 4$. To solve our problem, we proceed by the **backward induction** with respect to t.

Step 1. We start by solving the hedging problem at time $t = 1$ assuming that $S_1(\omega) = 8$, that is, on the set $\{\omega_1, \omega_2\}$. This problem is in fact the same as in the two-state single-period market model and thus the delta hedging formula applies. Denoting the still unknown hedging strategy with $\varphi = (\varphi_t)_{0 \leq t \leq 2}$, we find

$$\varphi_1^1(\omega) = \frac{1 - 0}{9 - 6} = \frac{1}{3}$$

for $\omega \in \{\omega_1, \omega_2\}$. As usual, the net amount of money available at time $t = 1$, that is, $V_1(\varphi)(\omega) - \frac{1}{3} \cdot 8$ is invested in the money market account. In order to hedge in this situation, we must have

$$\begin{aligned} \left(V_1(\varphi)(\omega_1) - \frac{1}{3} \cdot 8\right) + \frac{1}{3} \cdot 9 &= 1, \\ \left(V_1(\varphi)(\omega_2) - \frac{1}{3} \cdot 8\right) + \frac{1}{3} \cdot 6 &= 0. \end{aligned}$$

It is easily seen that these equations are satisfied for $V_1(\varphi)(\omega_1) = \frac{2}{3} = V_1(\varphi)(\omega_2)$ and thus $\pi_1(X)(\omega) = \frac{2}{3}$ for $\omega \in \{\omega_1, \omega_2\}$. The equality of the wealth process for the two states ω_1 and ω_2 is by no means a coincidence; it must hold, since $V_1(\varphi)$ has to be $\mathcal{F}_1^S$-measurable (see Proposition 3.1.1). We then have, for $\omega \in \{\omega_1, \omega_2\}$,

$$\varphi_1^0(\omega) = \frac{2}{3} - \frac{1}{3} \cdot 8 = -2$$

so that two units of cash are borrowed.

Let us now examine the hedging problem at time $t = 1$ when the stock price is $S_1(\omega) = 4$, that is, on the set $\{\omega_3, \omega_4\}$. There is no chance that the digital call will payoff anything else than zero. The hedging strategy for this payoff is trivial, invest zero in the money market account and invest zero in the stock. We therefore have, for $\omega \in \{\omega_3, \omega_4\}$,

$$\varphi_1^0(\omega) = \varphi_1^1(\omega) = 0$$

and thus $V_1(\varphi)(\omega_3) = 0 = V_1(\varphi)(\omega_4)$. Hence $\pi_1(X)(\omega) = 0$ for $\omega \in \{\omega_3, \omega_4\}$.

Step 2. In the second step, we consider the hedging problem at time $t = 0$ with price $S_0 = 5$. In order to find a hedge of the digital call, it suffices to replicate the contingent claim $V_1(\varphi)$ in the single-period model, where

$$V_1(\varphi) = \pi_1(X) = \begin{cases} \frac{2}{3} & \text{if } \omega \in \{\omega_1, \omega_2\}, \\ 0 & \text{if } \omega \in \{\omega_3, \omega_4\}. \end{cases}$$

The rationale is the following: if we hedge this contingent claim in the period from time $t = 0$ to time $t = 1$ and next we follow the strategy we computed at time $t = 1$, then we are done. Solving the hedging problem at time $t = 0$ is easy. By applying once again the delta hedging formula, we obtain

$$\varphi_0^1 = \frac{\frac{2}{3} - 0}{8 - 4} = \frac{1}{6}.$$

In order to replicate $V_1(\varphi)$, we must have, for all $\omega \in \{\omega_1, \omega_2, \omega_3, \omega_4\}$,

$$\begin{aligned} \left(V_0(\varphi)(\omega) - \frac{1}{6} \cdot 5\right) + \frac{1}{6} \cdot 8 &= \frac{2}{3}, \\ \left(V_0(\varphi)(\omega) - \frac{1}{6} \cdot 5\right) + \frac{1}{6} \cdot 4 &= 0. \end{aligned}$$

It appears that both equations are satisfied for $V_0(\varphi) = \frac{1}{6}$ and thus this number is the price of the digital call in this model. It is also clear that $\varphi_0^0(\omega) = -\frac{4}{5}$. One can now check that the digital call can be dynamically hedged using the self-financing trading strategy computed above. □

As in a single-period market model, there may be contingent claims for which no replicating strategy in Φ exists.

Definition 3.1.18. *A contingent claim X is called **attainable** in $\mathcal{M}$ if there exists a trading strategy $\varphi \in \Phi$ which replicates X, i.e. satisfies $V_T(\varphi) = X$.*

For an attainable contingent claim, the replication principle applies and it is clear how to price such a claim, namely, by the initial endowment needed for a replicating strategy.

Definition 3.1.19. *A multi-period market model $\mathcal{M}$ is called **complete** if and only if for any contingent claim X there exists a replicating strategy φ. A model which is not complete is called **incomplete**.*

Example 3.1.6. One can show that a multi-period version of the stochastic volatility model introduced in Example 2.2.3 is an incomplete multi-period market model. This is left as an exercise. □

3.1.9 Risk-Neutral Valuation of European Claims

Let us now consider a general European contingent claim X (attainable or not) and assume that Φ consists of all self-financing and $\mathbb{F}$-adapted trading strategies. We assume that the market model $\mathcal{M} = (B, S^1, \dots, S^n)$ is arbitrage-free.

Proposition 3.1.6. *Let X be a contingent claim (possibly non-attainable) and $\mathbb{Q}$ any risk-neutral probability measure for the multi-period market model $\mathcal{M}$. Then the **risk-neutral valuation formula***

$$\boxed{\pi_t(X) = B_t\,\mathbb{E}_{\mathbb{Q}}\Big(\frac{X}{B_T}\,\Big|\,\mathcal{F}_t\Big)} \tag{3.13}$$

defines a price process $\pi(X) = (\pi_t(X))_{0 \le t \le T}$ for the contingent claim X that complies with the principle of no-arbitrage.

Proof. The proof of this result is essentially the same as the proof of Proposition 2.2.4 and thus it is left as an exercise. ■

- If a claim X is attainable then the conditional expectation in the right-hand side of (3.13) does not depend on the choice of a risk-neutral probability measure $\mathbb{Q}$.
- If a claim X is not attainable then, as in the case of single-period market models, one faces the problem of non-uniqueness of a price complying with the principle of no-arbitrage. Indeed, unless a claim X is attainable, using Proposition 3.1.6, we obtain an interval of prices $\pi_0(X)$ for the claim X.

As a criterion for completeness, one has the following result, which should be seen as a multi-period counterpart of Theorem 2.2.2.

Theorem 3.1.2. *Assume that a multi-period market model $\mathcal{M} = (B, S^1, \dots, S^n)$ is arbitrage-free. Then $\mathcal{M}$ is complete if and only if there is only one risk-neutral probability measure.*

Example 3.1.7. Let us return to Examples 3.1.3 and 3.1.5. As before, we assume that the interest rate is equal to zero, i.e., $r = 0$. We will re-examine valuation of the digital call that pays one unit of cash at maturity date $T = 2$ whenever the stock price at this date is greater or equal than 8, that is,

$$X = \begin{cases} 1 & \text{if } S_2 \ge 8, \\ 0 & \text{otherwise.} \end{cases}$$

We already know that the arbitrage price equals $\pi_0(X) = \frac{1}{6}$ and that this price is in fact unique, since the claim X can be replicated. We will now compute the arbitrage price of the digital call using the risk-neutral pricing formula (3.13). For this purpose, we will first compute a risk-neutral probability measure $\mathbb{Q} = (q_1, q_2, q_3, q_4)$. The first two conditions for the q_is are obtained by the fact that $\mathbb{Q}$ is a probability measure and condition 1. in Definition 3.1.14, that is:

$$q_1 + q_2 + q_3 + q_4 = 1, \quad q_i > 0, \quad i = 1, 2, 3, 4.$$

Additional conditions for the q_is can be obtained from property 2. in Definition 3.1.14. To this end, we note that since the interest rate is equal to zero, the non-discounted prices agree with the discounted prices. The first condition is obtained by taking $t = 0$

$$5 = \mathbb{E}_{\mathbb{Q}}(\widehat{S}_1) = q_1 \cdot 8 + q_2 \cdot 8 + q_3 \cdot 4 + q_4 \cdot 4.$$

Upon substituting $q_4 = 1 - q_1 - q_2 - q_3$, we obtain

$$5 = 4 + 4(q_1 + q_2),$$

which is equivalent to

$$q_1 + q_2 = \frac{1}{4}, \tag{3.14}$$

so that necessarily

$$q_3 + q_4 = \frac{3}{4}. \tag{3.15}$$

We now examine the conditional expectation for $t = 1$. Using the definition of the conditional expectation, we see that for $\omega \in \{\omega_1, \omega_2\} = A_1$, we have

$$8 = \mathbb{E}_{\mathbb{Q}}(\widehat{S}_2 \,|\, \mathcal{F}_1^S)(\omega) = 9\,\frac{q_1}{q_1 + q_2} + 6\,\frac{q_2}{q_1 + q_2} = 36q_1 + 24q_2,$$

which in turn yields

$$9q_1 + 6q_2 = 2. \tag{3.16}$$

Furthermore, for $\omega \in \{\omega_3, \omega_4\} = A_2$ we must have

$$4 = \mathbb{E}_{\mathbb{Q}}(\widehat{S}_2 \,|\, \mathcal{F}_1^S)(\omega) = 6\,\frac{q_3}{q_3 + q_4} + 3\,\frac{q_4}{q_3 + q_4} = 8q_3 + 4q_4$$

so that

$$2q_3 + q_4 = 1. \tag{3.17}$$

Now, we have four equations (3.14), (3.15), (3.16) and (3.17) with four unknowns q_1, q_2, q_3 and q_4. The unique solution of this system is

$$\mathbb{Q} = (q_1, q_2, q_3, q_4) = \left(\frac{1}{6}, \frac{1}{12}, \frac{1}{4}, \frac{1}{2}\right). \tag{3.18}$$

We conclude that the market model $\mathcal{M} = (B, S)$ is arbitrage-free and complete.

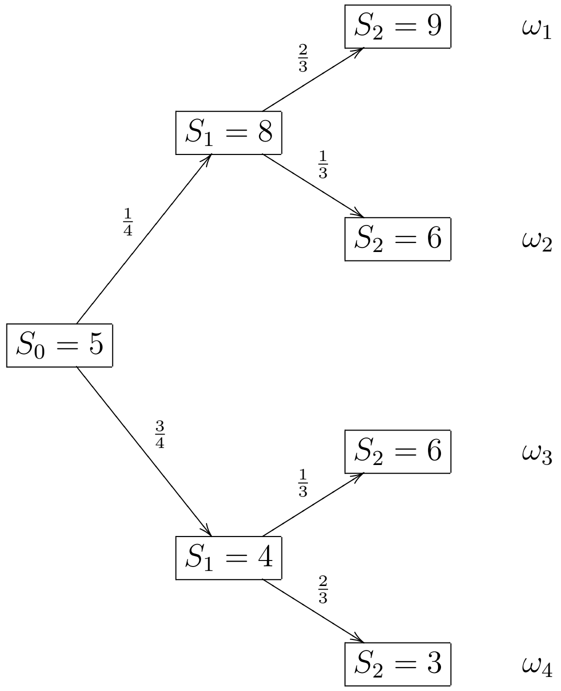
In order to compute the price of the digital call option, which is represented by the claim $X = \mathbb{1}_{\{S_2 \geq 8\}}$, we observe that $X(\omega)$ equals 1 for $\omega = \omega_1$ and it is equal to 0 for all other ωs. Since $r = 0$, the discounting is irrelevant, and thus

$$\pi_0(X) = \mathbb{E}_{\mathbb{Q}}(X) = q_1 \cdot 1 + q_2 \cdot 0 + q_3 \cdot 0 + q_4 \cdot 0 = \frac{1}{6}.$$

One can also compute the price $\pi_1(X)$ using formula (3.13).

Note that the computation of the unique risk-neutral probability $\mathbb{Q}$ can simplified if we focus on the conditional probabilities, that is, the risk-neutral probabilities of upward and downward mouvements of the stock price over one period.

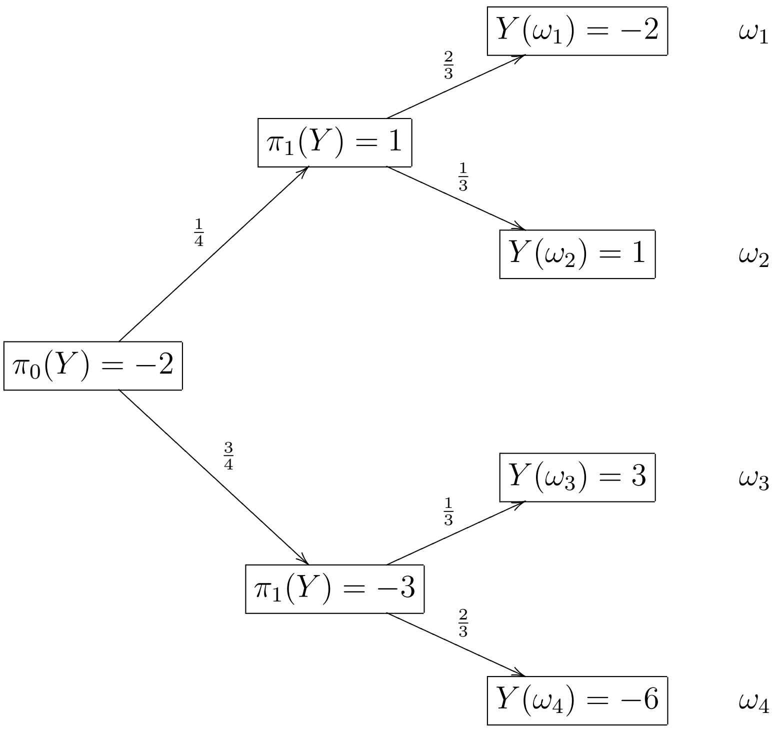
The following diagram describes the evolution of the stock price under $\mathbb{Q}$



where probabilities are the conditional risk-neutral probabilities, for instance,

$$\mathbb{Q}(S_2 = 6 \,|\, S_0 = 5,\ S_1 = 4) = \mathbb{Q}(S_2 = 6 \,|\, S_1 = 4) = \frac{1}{3}.$$

Let us now consider the claim $Y = (Y(\omega_1), Y(\omega_2), Y(\omega_3), Y(\omega_4)) = (-2, 1, 3, -6)$ maturing at $T = 2$. Using formula (3.13), we obtain the price process $\pi(Y)$:



Chapter 4

The Cox-Ross-Rubinstein Model

The most widely used example of a multi-period market model is the **binomial options pricing model**, which we are going to discuss in this section. This model is also commonly known as the **Cox-Ross-Rubinstein model**, since it was proposed in the seminal paper by Cox et al. (1979). For brevity, it is also frequently referred to as the **CRR model**.

As we will see in what follows, the binomial market model is the concatenation of several elementary single-period market models, as discussed in Section 1.1. We assume here that we have one stock S and the money market account B, but a generalization to the case of more than one stock is also possible.

4.1 Multiplicative Random Walk

At each point in time, the stock price is assumed to either go 'up' by a fixed factor u or go 'down' by a fixed factor d. We only assume that $0 < d < u$, but we do not postulate that $d < 1 < u$. The probability of an 'up' mouvement is assumed to be the same $0 < p < 1$ for each period, and is assumed to be independent of all previous stock price mouvements.

Definition 4.1.1. *A stochastic process $X = (X_t)_{1 \leq t \leq T}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called the **Bernoulli process** with parameter $0 < p < 1$ if the random variables $X_1, X_2, \dots, X_T$ are independent and have the following common probability distribution*

$$\boxed{\mathbb{P}(X_t = 1) = 1 - \mathbb{P}(X_t = 0) = p.}$$

*The **Bernoulli counting process** $N = (N_t)_{0 \leq t \leq T}$ is defined by setting $N_0 = 0$ and by postulating that, for every $t = 1, \dots, T$ and $\omega \in \Omega$,*

$$\boxed{N_t(\omega) := X_1(\omega) + X_2(\omega) + \dots + X_t(\omega).}$$

The Bernoulli counting process is a very special case of an **additive random walk**. The stock price process in the CRR model is defined via a deterministic initial value $S_0 > 0$ and for $1 \leq t \leq T$ and all $\omega \in \Omega$

$$\boxed{S_t(\omega) := S_0 u^{N_t(\omega)} d^{\,t-N_t(\omega)}.} \tag{4.1}$$

- The idea behind this construction is that the underlying Bernoulli process X governs the 'up' and 'down' mouvements of the stock. The stock price moves up at time t if $X_t(\omega) = 1$ and moves down if $X_t(\omega) = 0$. The dynamics of the stock price can thus be seen as an example of a **multiplicative random walk**.
- The Bernoulli counting process N counts the up mouvements. Before and including time t, the stock price moves up N_t times and down $t - N_t$ times. Assuming that the stock price can only move up (resp. down) by the factors u (resp. d) we obtain equation (4.1) governing the dynamics of the stock price over time.

Why is this model called the binomial model? The reason is that for each t, the random variable N_t has the **binomial distribution** with parameters p and t. To be more specific, for every $t = 1, \dots, T$ and $k = 0, \dots, t$ we have that

$$\boxed{\mathbb{P}(N_t = k) = \tbinom{t}{k} p^k (1-p)^{t-k}}$$

and thus the probability distribution of the stock price S_t at time t is given by

$$\boxed{\mathbb{P}(S_t = S_0 u^k d^{t-k}) = \tbinom{t}{k} p^k (1-p)^{t-k}} \tag{4.2}$$

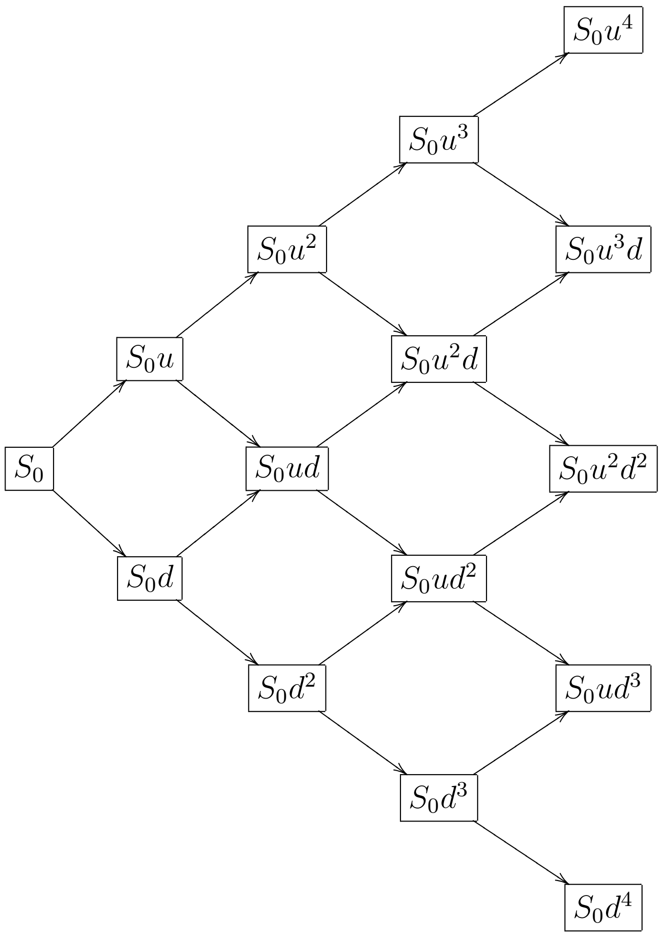
for $k = 0, 1, \dots, t$.

It it easy to show that the filtration $\mathbb{F}^S = (\mathcal{F}_t^S)_{0 \leq t \leq T}$ generated by the stock price $S = (S_t)_{1 \leq t \leq T}$ coincides with the filtration $\mathbb{F}^X = (\mathcal{F}_t^X)_{0 \leq t \leq T}$ generated by the Bernoulli process $X = (X_t)_{1 \leq t \leq T}$, where $\mathcal{F}_0^X$ is set to be the trivial σ-field, by definition. As usual, the money market account B is assumed to be defined via $B_0 = 1$ and

$$\boxed{B_t = (1+r)^t.} \tag{4.3}$$

Definition 4.1.2. *The* ***binomial market model*** *(or the* ***CRR model****) with parameters p, u, d, r and time horizon T on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is the multi-period market model consisting of the stock and the money market account, where the stock price evolution is governed by equation (4.1) and the money market account satisfies (4.3). The underlying filtration $\mathbb{F}$ is assumed to be the filtration $\mathbb{F}^S = (\mathcal{F}_t^S)_{0 \leq t \leq T}$ generated by S and the set of allowed trading strategies Φ is given by all self-financing and $\mathbb{F}^S$-adapted trading strategies.*

Sample paths of the CRR model can be represented through the following lattice, for $T = 4$,



Given a binomial market model, the underlying Bernoulli process X can be recovered from the stock prices. Clearly, $X_t = N_t - N_{t-1}$ and thus

$$\begin{aligned}
\frac{S_t}{S_{t-1}} &= \frac{S_0 u^{N_t} d^{t-N_t}}{S_0 u^{N_{t-1}} d^{t-1-N_{t-1}}} \\
&= u^{N_t - N_{t-1}} d^{1-(N_t - N_{t-1})} \\
&= u^{X_t} d^{1-X_t} \\
&= \begin{cases} u & \text{if } X_t(\omega) = 1, \\ d & \text{if } X_t(\omega) = 0. \end{cases}
\end{aligned}$$

We will now address the crucial question:

Is the CRR market model arbitrage-free?

Proposition 4.1.1. *Assume that $d < 1 + r < u$. Then a probability measure $\widetilde{\mathbb{P}}$ on $(\Omega, \mathcal{F}_T)$ is a risk-neutral probability measure for the CRR model $\mathcal{M} = (B, S)$ with parameters p, u, d, r and time horizon T if and only if:*

1. *$X_1, X_2, X_3, \dots, X_T$ are independent under the probability measure $\widetilde{\mathbb{P}}$,*
2. *$0 < \widetilde{p} := \widetilde{\mathbb{P}}(X_t = 1) < 1$ for all $t = 1, \dots, T$,*
3. *$\widetilde{p}u + (1 - \widetilde{p})d = (1 + r)$,*

where X is the Bernoulli process governing the stock price process S.

The proof of Proposition 4.1.1 is not hard and thus we leave it as an exercise. Note that condition 3. here is the same as in Section 2.1 and it is equivalent to

$$\boxed{\widetilde{p} = \tfrac{1+r-d}{u-d}.} \tag{4.4}$$

We thus see that the binomial model is arbitrage-free whenever $d < 1 + r < u$. This is exactly the same condition as in a single-period model of Section 2.1. As the value for $\widetilde{p}$ in Proposition 4.1.1 is also unique, the risk-neutral probability $\widetilde{\mathbb{P}}$ is unique and thus, from Theorem 2.2.2, the binomial model is complete.

If** $d < 1 + r < u$ **then the CRR market model** $\mathcal{M} = (B, S)$ **is arbitrage-free and complete.

- We assume from now on that $d < 1 + r < u$. Otherwise, the CRR model is not arbitrage-free.
- The sample paths of the CRR process are represented by the **recombining tree** (i.e. the **lattice**). Recall that since N_t can only take $t + 1$ values, so does the stock price S_t for and date $t = 0, 1, \dots, T$. The corresponding number of different sample paths of the stock price between times 0 and t equals 2^t, so that the number of sample paths grows very quickly when the number of periods rises.
- For the sake of computational simplicity, it is sometimes postulated that

$$d = u^{-1}. \tag{4.5}$$

In that case, formula (4.1) simplifies as follows

$$\boxed{S_t = S_0 u^{2N_t - t}.} \tag{4.6}$$

4.2 The CRR Call Pricing Formula

Since the CRR model is complete, the unique arbitrage price for the call option can be computed using the risk-neutral valuation formula (3.13).

Proposition 4.2.1. *The arbitrage price at time $t = 0$ of the European call option $C_T = (S_T - K)^+$ in the binomial market model $\mathcal{M} = (B, S)$ with parameters u, d and r is given by the Cox-Ross-Rubinstein call pricing formula*

$$C_0 = S_0 \sum_{k=\widehat{k}}^{T} \binom{T}{k} \widehat{p}^k (1-\widehat{p})^{T-k} - \frac{K}{(1+r)^T} \sum_{k=\widehat{k}}^{T} \binom{T}{k} \widetilde{p}^k (1-\widetilde{p})^{T-k}$$

where

$$\widetilde{p} = \frac{1+r-d}{u-d}, \quad \widehat{p} = \frac{\widetilde{p}u}{1+r}$$

and $\widehat{k}$ is the smallest integer k such that

$$k \log\left(\frac{u}{d}\right) > \log\left(\frac{K}{S_0 d^T}\right).$$

Proof. The price at time $t = 0$ of the claim $X = C_T = (S_T - K)^+$ can be computed using the risk-neutral valuation under $\widetilde{\mathbb{P}}$

$$C_0 = \mathbb{E}_{\widetilde{\mathbb{P}}}\left(\frac{C_T}{(1+r)^T}\right).$$

In view of Proposition 4.1.1, this can be represented more explicitly as follows

$$C_0 = \frac{1}{(1+r)^T} \sum_{k=0}^{T} \binom{T}{k} \widetilde{p}^k (1-\widetilde{p})^{T-k} \max\left(0, S_0 u^k d^{T-k} - K\right).$$

We note that

$$S_0 u^k d^{T-k} - K > 0 \;\Leftrightarrow\; \left(\frac{u}{d}\right)^k > \frac{K}{S_0 d^T} \;\Leftrightarrow\; k > \frac{\log\left(\frac{K}{S_0 d^T}\right)}{\log\left(\frac{u}{d}\right)}.$$

We define $\widehat{k} = \widehat{k}(S_0, T)$ as the smallest integer k such that this inequality is satisfied. If there are less than $\widehat{k}$ upward mouvements, there is no chance that the option will pay off anything (i.e. it will expire worthless). Therefore, we obtain

$$\begin{aligned}
C_0 &= \frac{1}{(1+r)^T} \sum_{k=0}^{\widehat{k}-1} \binom{T}{k} \widetilde{p}^k (1-\widetilde{p})^{T-k} 0 \\
&\quad + \frac{1}{(1+r)^T} \sum_{k=\widehat{k}}^{T} \binom{T}{k} \widetilde{p}^k (1-\widetilde{p})^{T-k} \left(S_0 u^k d^{T-k} - K\right) \\
&= \frac{S_0}{(1+r)^T} \sum_{k=\widehat{k}}^{T} \binom{T}{k} \widetilde{p}^k (1-\widetilde{p})^{T-k} u^k d^{T-k} - \frac{K}{(1+r)^T} \sum_{k=\widehat{k}}^{T} \binom{T}{k} \widetilde{p}^k (1-\widetilde{p})^{T-k}
\end{aligned}$$

so that

$$\begin{aligned} C_0 &= S_0 \sum_{k=\widehat{k}}^{T} \binom{T}{k} \left(\frac{\widetilde{p}u}{1+r}\right)^k \left(\frac{(1-\widetilde{p})d}{1+r}\right)^{T-k} - \frac{K}{(1+r)^T} \sum_{k=\widehat{k}}^{T} \binom{T}{k} \widetilde{p}^k (1-\widetilde{p})^{T-k} \\ &= S_0 \sum_{k=\widehat{k}}^{T} \binom{T}{k} \widehat{p}^k (1-\widehat{p})^{T-k} - \frac{K}{(1+r)^T} \sum_{k=\widehat{k}}^{T} \binom{T}{k} \widetilde{p}^k (1-\widetilde{p})^{T-k} \end{aligned}$$

where we write $\widehat{p} = \frac{\widetilde{p}u}{1+r}$. ■

Example 4.2.1. (MATH3975) Check that $0 < \widehat{p} = \frac{\widetilde{p}u}{1+r} < 1$ whenever $0 < \widetilde{p} < 1$. Let $\widehat{\mathbb{P}}$ be the probability measure obtained by setting $p = \widehat{p}$ in Definition 4.1.1. Show that the process $\frac{B}{S}$ is an $\mathbb{F}$-martingale under $\widehat{\mathbb{P}}$ and the price of the call satisfies

$$C_0 = S_0 \, \widehat{\mathbb{P}}(D) - KB(0,T) \, \widetilde{\mathbb{P}}(D)$$

where $D = \{\omega \in \Omega : S_T(\omega) > K\}$. Using the abstract Bayes formula of Lemma 3.1.1, establish the second equality in the following formula

$$C_t := B_t \, \mathbb{E}_{\widetilde{\mathbb{P}}}\big(B_T^{-1}(S_T - K)^+ \,|\, \mathcal{F}_t\big) = S_t \, \widehat{\mathbb{P}}(D \,|\, \mathcal{F}_t) - KB(t,T) \, \widetilde{\mathbb{P}}(D \,|\, \mathcal{F}_t).$$

4.2.1 Put-Call Parity

Since $C_T - P_T = S_T - K$, we see that the following *put-call parity* holds at any date $t = 0, 1, \dots, T$

$$C_t - P_t = S_t - K(1+r)^{-(T-t)} = S_t - KB(t,T)$$

where $B(t,T) = (1+r)^{-(T-t)}$ represents the price at time t of the zero-coupon bond paying one unit of cash at time T. Using Proposition 4.2.1 and the put-call parity, one can derive an explicit pricing formula for the European put option with the payoff $P_T = (K - S_T)^+$. This is left as an exercise.

4.2.2 Pricing Formula at Time t

The pricing formula for the European call option, established in Proposition 4.2.1, can be extended to the case of any date $t = 0, 1, \dots, T-1$, specifically,

$$C_t = S_t \sum_{k=\widehat{k}(S_t, T-t)}^{T-t} \binom{T-t}{k} \widehat{p}^k (1-\widehat{p})^{T-t-k} - \frac{K}{(1+r)^{T-t}} \sum_{k=\widehat{k}(S_t, T-t)}^{T-t} \binom{T-t}{k} \widetilde{p}^k (1-\widetilde{p})^{T-t-k}$$

where $\widehat{k}(S_t, T-t)$ is the smallest integer k such that

$$k \log\left(\frac{u}{d}\right) > \log\left(\frac{K}{S_t d^{T-t}}\right).$$

Note that $C_t = C(S_t, T-t)$ meaning that the call option price depends on the time to maturity $T-t$ and the level S_t of the stock price observed at time t, but it is independent of the evolution of the stock price prior to t.

4.2.3 Replicating Strategy

To compute the replicating strategy for the call option, we note that the CRR model can be seen as a concatenation of single-period models. For instance, if we wish to find the replicating portfolio at time $t = 0$, we observe that the replicating portfolio $(\varphi_0^0, \varphi_0^1)$ satisfies

$$\varphi_0^0 + \varphi_0^1 S_0 = V_0(\varphi) = C_0$$

and

$$\begin{aligned} \varphi_0^0(1+r) + \varphi_0^1 S_1^u &= C_1^u, \\ \varphi_0^0(1+r) + \varphi_0^1 S_1^d &= C_1^d, \end{aligned}$$

where in turn $C_1^u = C(uS_0, T-1)$ and $C_1^d = C(dS_0, T-1)$. Consequently,

$$\varphi_0^0 = C_0 - \varphi_0^1 S_0, \quad \varphi_0^1 = \frac{C_1^u - C_1^d}{S_1^u - S_1^d} = \frac{C(uS_0, T-1) - C(dS_0, T-1)}{S_0(u-d)}.$$

4.3 Exotic Options

So far, we considered contingent claims of the type $X = h(S_T)$ where $h : \mathbb{R} \to \mathbb{R}$ is the payoff function. In that case, the value of the claim X at its maturity date T depends on the terminal stock price S_T, but not on the values of stock prices at times strictly before T. Contingent claims of this form are said to be **path-independent**. However, several traded derivatives correspond to the so-called **path-dependent** contingent claims, meaning that they are of the form $X = h(S_0, S_1, \ldots, S_T)$ for some function $h : \mathbb{R}^{T+1} \to \mathbb{R}$. We provide below some examples of path-dependent claims:

1. **Asian option**: An Asian option is simply a European call (or put) option on the average stock price during the option's lifetime. The following payoffs are examples of traded Asian options.

   (a) **Arithmetic average option**:

   $$X = \left(\frac{1}{T+1} \sum_{t=0}^{T} S_t - K \right)^+. \tag{4.7}$$

   (b) **Geometric average option**:

   $$X = \left(\left(\prod_{t=0}^{T} S_t \right)^{\frac{1}{T+1}} - K \right)^+. \tag{4.8}$$

2. **Barrier options**: Barrier options are options that are activated or deactivated if the stock price hits a certain barrier during the option's lifetime. As an example, we consider two examples of the the so-called "knock-out" barrier options.

(a) **Down-and-out call**: This option has the same payoff as a European call with strike K, but only if the stock price always remain over the barrier level of $H < K$. Otherwise, it expires worthless. Hence

$$X = (S_T - K)^+ \mathbb{1}_{\{\min_{0 \le t \le T} S_t > H\}}. \tag{4.9}$$

(b) **Down-and-in call**: This option has the same payoff as a European call with strike K, but only if the stock price once attains a price below the barrier level of $H < K$. Otherwise, it expires worthless. Thus

$$X = (S_T - K)^+ \mathbb{1}_{\{\min_{0 \le t \le T} S_t \le H\}}. \tag{4.10}$$

3. **Lookback options**: A lookback option is an option on a maximum or minimum of a stock price during the option's lifetime.

   (a) **Call option on a maximum**:

   $$X = \left(\max_{0 \le t \le T} S_t - K \right)^+ .$$

   (b) **Call option on a minimum**:

   $$X = \left(\min_{0 \le t \le T} S_t - K \right)^+ .$$

The crucial difference between the options presented here and the standard (that is, *plain vanilla*) European call or put is that their payoffs depend not only on the terminal stock price S_T, but on the whole path taken by the stock price over the time interval $[0, T]$. Pricing these options, however, works in the same way as for the European call. It suffices to compute the discounted expectation under the risk-neutral probability measure described by Proposition 3.3.1.

4.4 American Contingent Claims in the CRR Model

In contrast to a contingent claim of a European style, a contingent claim of an American style can by exercised by its holder at any date before its expiration date T. We denote by $\mathcal{T}$ the class of all **stopping times** defined on the filtered probability space $(\Omega, \mathbb{F}, \mathbb{P})$, where $\mathcal{F}_t = \mathcal{F}_t^S$ for every $t = 0, 1, \dots, T$.

Definition 4.4.1. *A* **stopping time** *with respect to a filtration $\mathbb{F}$ is a map $\tau : \Omega \to \{0, 1, \dots, T\}$ such that for any $t = 0, 1, \dots, T$ the event $\{\omega \in \Omega \,|\, \tau(\omega) = t\}$ belongs to the σ-field $\mathcal{F}_t$.*

Intuitively, this property means that the decision whether to stop a given process at time t (for instance, whether to exercise an option at time t or not) depends on the stock price fluctuations up to time t only.

Definition 4.4.2. *Let $\mathcal{T}_{[t,T]}$ be the subclass of stopping times τ with respect to $\mathbb{F}$ satisfying the inequalities $t \le \tau \le T$.*

4.4.1 American Call Option

Let us first consider the case of the American call option – that is, the option to buy a specified number of shares, which may be exercised at any time before the option expiry date T, or on that date. The exercise policy of the option holder is necessarily based on the information accumulated to date and not on the future prices of the stock. We denote by C_t^a the arbitrage price at time t of an American call option written on one share of the stock S.

Definition 4.4.3. *By an* **arbitrage price** *of the American call we mean a price process C_t^a, $t \leq T$, such that the extended financial market model – that is, a market with trading in riskless bonds, stocks and an American call option – remains arbitrage-free.*

Lemma 4.4.1. *The price of an American call option in the CRR arbitrage-free market model with $r \geq 0$ coincides with the arbitrage price of a European call option with the same expiry date and strike price.*

Proof. It is sufficient to show that the American call option should never be exercised before maturity, since otherwise the issuer of the option would be able to make riskless profit.

The argument hinges on the following simple inequality

$$C_t \geq (S_t - K)^+, \quad \forall\, t \leq T. \tag{4.11}$$

which can be justified in several ways. An intuitive way of deriving (4.11) is based on no-arbitrage arguments. Notice that since the option's price C_t is always non-negative, it is sufficient to consider the case when the current stock price is greater than the exercise price – that is, when $S_t - K > 0$.

Suppose, on the contrary, that $C_t < S_t - K$ for some t, i.e., $S_t - C_t > K$. Then it would be possible, with zero net initial investment, to buy at time t a call option, short a stock, and invest the sum $S_t - C_t$ in the savings account. By holding this portfolio unchanged up to the maturity date T, we would be able to lock in a riskless profit. Indeed, the value of our portfolio at time T would satisfy (recall that $r \geq 0$)

$$C_T - S_T + (1+r)^{T-t}(S_t - C_t) > (S_T - K)^+ - S_T + (1+r)^{T-t}K \geq 0.$$

We conclude that inequality (4.11) is necessary for the absence of arbitrage opportunities.

Taking (4.11) for granted, we may now deduce the property $C_t^a = C_t$ using simple no-arbitrage arguments. Suppose, on the contrary, that the issuer of an American call is able to sell the option at time 0 at the price $C_0^a > C_0$ (it is evident that, at any time, an American option is worth at least as much as a European option with the same contractual features; in particular, $C_0^a \geq C_0$). In order to profit from this transaction, the option writer establishes a dynamic portfolio that replicates the value process of the European call, and invests the remaining funds in the savings account. Suppose that the holder of the option decides to exercise it at instant t before the expiry date T. Then the issuer of the option locks in a riskless profit, since the value of portfolio satisfies

$$C_t - (S_t - K)^+ + (1+r)^t (C_0^a - C_0) > 0, \quad \forall\, t \le T.$$

The above reasoning implies that the European and American call options are equivalent from the point of view of arbitrage pricing theory; that is, both options have the same price and an American call should never be exercised by its holder before expiry. ∎

4.4.2 American Put Option

An American put is an option to sell a specified number of shares, which may be exercised at any time before or at the expiry date T. Let us consider an American put with strike K and expiry date T. We then have the following valuation result.

Proposition 4.4.1. *The arbitrage price P_t^a of an American put option equals*

$$P_t^a = \max\nolimits_{\tau \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\widetilde{\mathbb{P}}}\big((1+r)^{-(\tau-t)}(K - S_\tau)^+ \,|\, \mathcal{F}_t\big), \quad \forall\, t \le T. \tag{4.12}$$

For any $t \le T$, the stopping time τ_t^* which realizes the maximum in (4.12) *is given by the expression*

$$\tau_t^* = \min\{u \ge t \,|\, (K - S_u)^+ \ge P_u^a\}. \tag{4.13}$$

The stopping time τ_t^* will be referred to as the **rational exercise time** of an American put option that is assumed to be still alive at time t. It should be pointed out that τ_t^* does not solve the optimal stopping problem for any individual, but only for those investors who are risk-neutral.

An application of the classic Bellman principle reduces the optimal stopping problem (4.12) to an explicit recursive procedure for the value process. We state the following corollary to Proposition 4.4.1, in which the **dynamic programming recursion** for the value of an American put option is given.

Corollary 4.4.1. *Let the non-negative adapted process U_t, $t \le T$, be defined recursively by setting $U_T = (K - S_T)^+$ and*

$$U_t = \max\Big\{(K - S_t)^+,\, (1+r)^{-1}\mathbb{E}_{\widetilde{\mathbb{P}}}(U_{t+1} \,|\, \mathcal{F}_t)\Big\}, \quad \forall\, t \le T-1. \tag{4.14}$$

Then the arbitrage price P_t^a of the American put option at time t equals U_t and the rational exercise time after time t admits the following representation

$$\tau_t^* = \min\{u \ge t \,|\, (K - S_u)^+ \ge U_u\}. \tag{4.15}$$

Therefore, $\tau_T^* = T$ and for every $t = 0, 1, \dots, T-1$

$$\tau_t^* = t\mathbb{1}_{\{U_t = (K-S_t)^+\}} + \tau_{t+1}^*\mathbb{1}_{\{U_t > (K-S_t)^+\}}.$$

It is also possible to go the other way around – that is, to first show directly that the price P^a_t satisfies the recursive relationship, for $t \leq T-1$,

$$P^a_t = \max\Big\{(K-S_t)^+,\ (1+r)^{-1}\mathbb{E}_{\widetilde{\mathbb{P}}}\big(P^a_{t+1}\,|\,\mathcal{F}_t\big)\Big\} \tag{4.16}$$

subject to the terminal condition $P^a_T = (K-S_T)^+,$ and subsequently derive the equivalent representation (4.12). In the case of the CRR model, formula (4.16) is the simplest way of valuing American options. The main reason for this is that an apparently difficult valuation problem is thus reduced to the simple single-period case.

To summarise:

- In the case of the CRR model, the arbitrage pricing of an American put reduces to the following simple recursive recipe, for $t \leq T-1$,

$$\boxed{P^a_t = \max\Big\{(K-S_t)^+,\ (1+r)^{-1}\big(\widetilde{p}P^{au}_{t+1} + (1-\widetilde{p})P^{ad}_{t+1}\big)\Big\}} \tag{4.17}$$

  with the terminal condition $P^a_T = (K-S_T)^+$.

- The quantities P^{au}_{t+1} and P^{ad}_{t+1} represent the values of the American put in the next step corresponding to the upward and downward mouvements of the stock price starting from a given node on the CRR lattice.

4.4.3 American Contingent Claims

We assume that a contingent claim of an *American style* (or, briefly, an *American claim*) does not produce any payoff unless it is exercised.

Definition 4.4.4. *An* **American contingent claim** $X^a = (X, \mathcal{T}_{[0,T]})$ *expiring at* T *consists of a sequence of payoffs* $(X_t)_{0\leq t\leq T}$ *where the random variable* X_t *is* $\mathcal{F}_t$*-measurable for* $t = 0,1,\dots,T$ *and the set* $\mathcal{T}_{[0,T]}$ *of admissible exercise policies.*

We interpret X_t as the payoff received by the holder of the claim X^a upon exercising it at time t. The set of admissible exercise policies is restricted to the class $\mathcal{T}_{[0,T]}$ of all stopping times of the filtration $\mathbb{F}^S$ with values in $\{0,1,\dots,T\}$. Let $g : \mathbb{R}\times\{0,1,\dots,T\} \to \mathbb{R}$ be an arbitrary function. We say that X^a is a path-independent American contingent claim with the **payoff function** g if the equality $X_t = g(S_t,t)$ holds for every $t = 0,1,\dots,T$. The arbitrage valuation of an American claim in a discrete-time model is based on a simple recursive procedure. In particular, to price a path-independent American claim in the CRR model, it suffices to move backwards in time along the binomial lattice.

Proposition 4.4.2. *For every* $t \leq T$*, the arbitrage price* $\pi(X^a)$ *of an American claim* X^a *in the CRR model equals*

$$\pi_t(X^a) = \max{}_{\tau\in\mathcal{T}_{[t,T]}}\,\mathbb{E}_{\widetilde{\mathbb{P}}}\big((1+r)^{-(\tau-t)}X_\tau\,|\,\mathcal{F}_t\big).$$

The price process $\pi(X^a)$ satisfies the following recurrence relation, for $t \leq T-1$,

$$\pi_t(X^a) = \max\Big\{X_t,\, \mathbb{E}_{\widetilde{\mathbb{P}}}\big((1+r)^{-1}\pi_{t+1}(X^a)\,|\,\mathcal{F}_t\big)\Big\} \tag{4.18}$$

with $\pi_T(X^a) = X_T$ and the rational exercise time τ_t^* equals

$$\tau_t^* = \min\big\{u \geq t \,|\, X_u \geq \mathbb{E}_{\widetilde{\mathbb{P}}}\big((1+r)^{-1}\pi_{u+1}(X^a)\,|\,\mathcal{F}_u\big)\big\}.$$

For a path-independent American claim X^a with the payoff process $X_t = g(S_t,t)$ we obtain, for every $t \leq T-1$,

$$\pi_t(X^a) = \max\Big\{g(S_t,t),\,(1+r)^{-1}\big(\widetilde{p}\,\pi^u_{t+1}(X^a) + (1-\widetilde{p})\,\pi^d_{t+1}(X^a)\big)\Big\} \tag{4.19}$$

where, for a generic stock price S_t at time t, we denote by $\pi^u_{t+1}(X^a)$ and $\pi^d_{t+1}(X^a)$ the values of the price $\pi_{t+1}(X^a)$ at the nodes corresponding to the upward and downward mouvements of the stock price during the period $[t, t+1]$.

Let us consider a path-independent American claim:

- By a slight abuse of notation, we write X_t^a to denote the arbitrage price at time t of a path-independent American claim X^a.
- Then the pricing formula (4.19) becomes (see (4.17))

$$\boxed{X_t^a = \max\Big\{g(S_t,t),\,(1+r)^{-1}\big(\widetilde{p}\,X^{au}_{t+1} + (1-\widetilde{p})\,X^{ad}_{t+1}\big)\Big\}} \tag{4.20}$$

  with the terminal condition $X_T^a = g(S_T, T)$.
- The risk-neutral valuation formula (4.20) is valid for an arbitrary path-independent American claim with a payoff function g in the CRR binomial model.

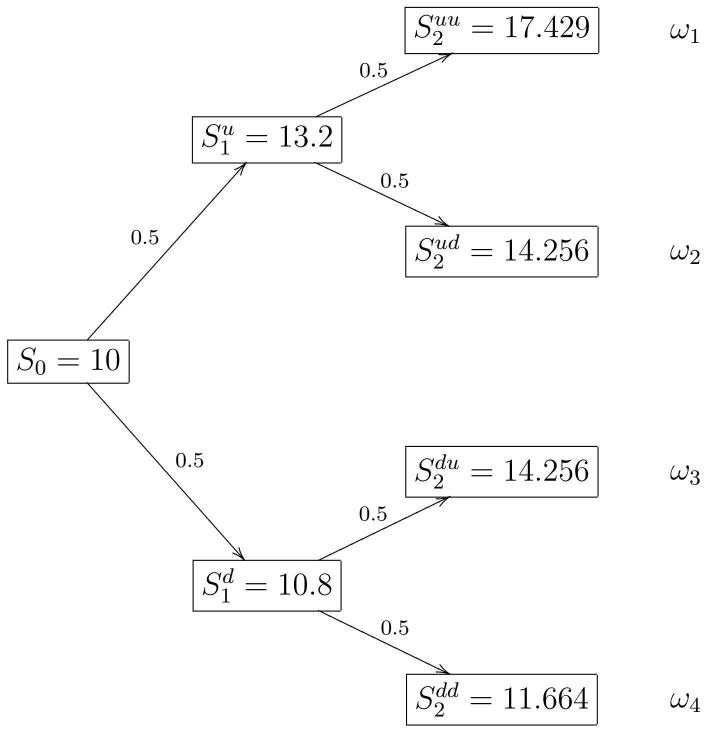
Example 4.4.1. We consider here the CRR binomial model with the horizon date $T=2$, the risk-free rate $r=0.2$, and the following values of the stock price S at times $t=0$ and $t=1$:

$$S_0 = 10, \quad S_1^u = 13.2, \quad S_1^d = 10.8.$$

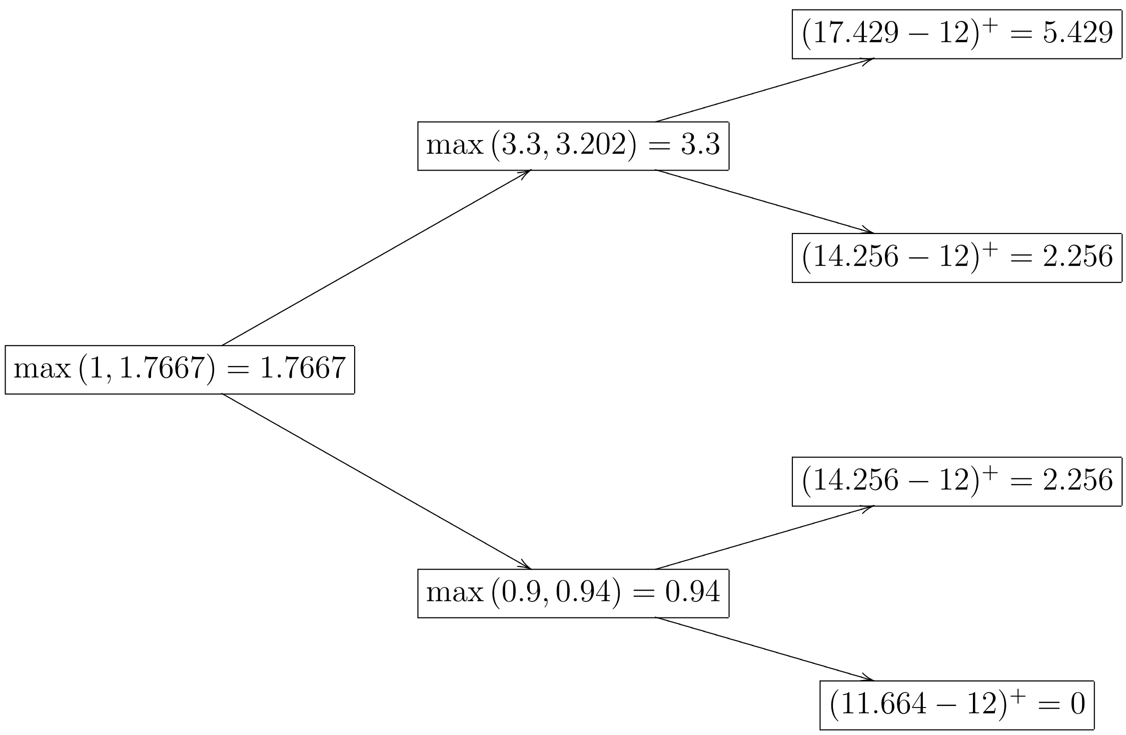
Let X^a be the American call option with maturity date $T=2$ and the following reward process $g(S_t,t) = (S_t - K_t)^+$, where the variable strike K_t satisfies $K_0 = 9$, $K_1 = 9.9$ and $K_2 = 12$. We will first compute the arbitrage price $\pi_t(X^a)$ of this option at times $t = 0,1,2$ and the rational exercise time τ_0^*. Subsequently, we will compute the replicating strategy for X^a up to the rational exercise time τ_0^*. We start by noting that the unique risk-neutral probability measure $\widetilde{\mathbb{P}}$ satisfies

$$\widetilde{p} = \frac{1+r-d}{u-d} = \frac{(1+r)S_0 - S_1^d}{S_1^u - S_1^d} = \frac{12-10.8}{13.2-10.8} = \frac{1.2}{2.4} = \frac{1}{2}.$$

The dynamics of the stock price under $\widetilde{\mathbb{P}}$ are thus given by (note that $S_2^{ud} = S_2^{du}$)



Consequently, the price process $\pi_t(X^a)$ of the American option X^a is given by



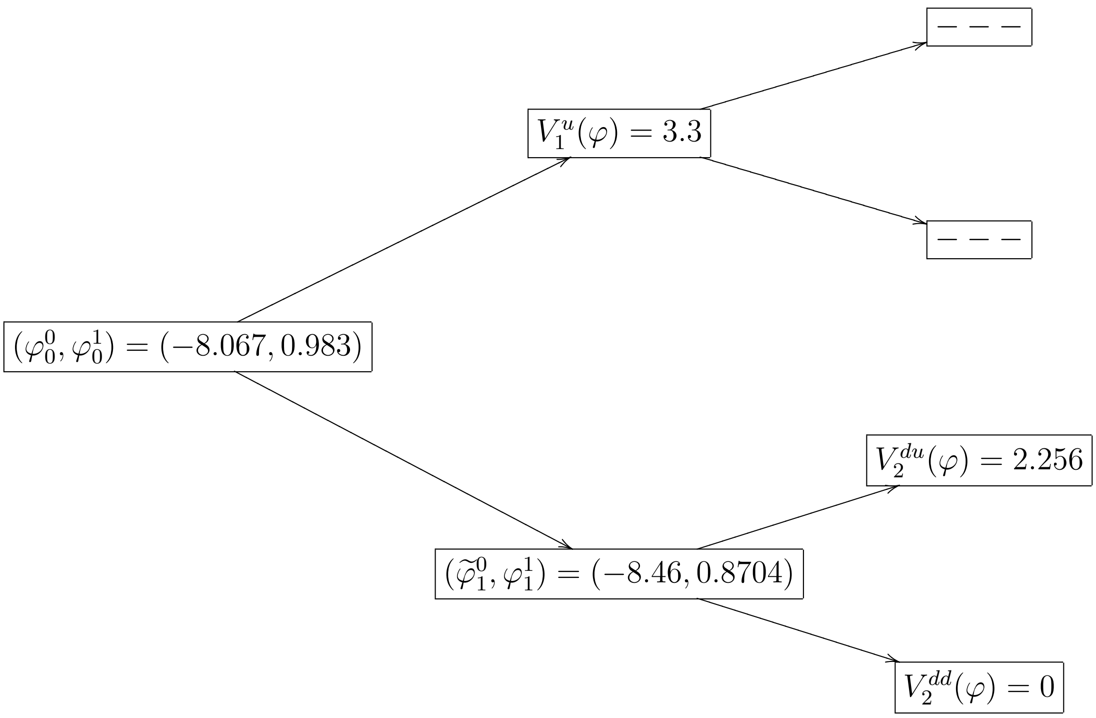
Holder. The rational holder should exercise the American option at time $t = 1$ whenever he observes that the stock price has risen during the first period. Otherwise, he should not exercise the option till time 2. Hence the rational exercise time τ_0^* is a stopping time $\tau_0^* : \Omega \to \{0, 1, 2\}$ given by

$$\tau_0^*(\omega) = 1 \ \text{ for } \ \omega \in \{\omega_1, \omega_2\},$$
$$\tau_0^*(\omega) = 2 \ \text{ for } \ \omega \in \{\omega_3, \omega_4\}.$$

Issuer. We now take the position of the issuer of the option:

- At $t = 0$, we need to solve
$$1.2\,\varphi_0^0 + 13.2\,\varphi_0^1 = 3.3,$$
$$1.2\,\varphi_0^0 + 10.8\,\varphi_0^1 = 0.94.$$
Hence $(\varphi_0^0, \varphi_0^1) = (-8.067, 0.983)$ for all ωs.
- If the stock price has risen during the first period, the option is exercised by its holder. Hence we do not need to compute the strategy at time 1 for $\omega \in \{\omega_1, \omega_2\}$.
- If the stock price has fallen during the first period, we need to solve
$$1.2\,\widetilde{\varphi}_1^0 + 14.256\,\varphi_1^1 = 2.256,$$
$$1.2\,\widetilde{\varphi}_1^0 + 11.664\,\varphi_1^1 = 0.$$
Hence $(\widetilde{\varphi}_1^0, \varphi_1^1) = (-8.46, 0.8704)$ if the stock price has fallen during the first period, that is, for $\omega \in \{\omega_3, \omega_4\}$. Note that $\widetilde{\varphi}_1^0 = -8.46$ is the amount of cash borrowed at time 1, rather than the number of units of the savings account B.

We conclude that the replicating strategy $\varphi = (\varphi^0, \varphi^1)$ is defined at time 0 for all ωs and it is defined at time 1 on the event $\{\omega_3, \omega_4\}$ only:



4.5 Game Contingent Claims (MATH3975)

The concept of a game contingent claim is an extension of an American contingent claim to a situation where both parties to a contract may exercise it prior to its maturity date. The payoff of a game contingent claim depends, in general, not only on the moment when it is exercised, but also on which party takes the decision to exercise. In order to make a clear distinction between the exercise policies of the two parties, in what follows, we will refer to the decision of the seller (also termed the issuer) as the *cancellation policy*, whereas the decision of the buyer (also referred to as the holder) is called the *exercise policy*.

Definition 4.5.1. *A game contingent claim $X^g = (L, H, \mathcal{T}^e, \mathcal{T}^c)$ expiring at time T consists of $\mathbb{F}$-adapted payoff processes L and H, a set $\mathcal{T}^e$ of admissible exercise times, and a set $\mathcal{T}^c$ of admissible cancellation times.*

It is assumed throughout that $L \leq H$, meaning that the $\mathcal{F}_t$-measurable random variables L_t and H_t satisfy the inequality $L_t \leq H_t$ for every $t = 0, 1, \dots, T$. Unless explicitly otherwise stated, the sets of admissible exercise and cancellation times are restricted to the class $\mathcal{T}_{[0,T]}$ of all stopping times of the filtration $\mathbb{F}^S$ with values in $\{0, 1, \dots, T\}$, that is, it is postulated that $\mathcal{T}^e = \mathcal{T}^c = \mathcal{T}_{[0,T]}$.

We interpret L_t as the payoff received by the holder upon exercising a game contingent claim at time t. The random variable H_t represents the payoff received by the holder if a claim is cancelled (i.e., exercised by its issuer) at time t. More formally, if a game contingent claim is exercised by either of the two parties at some date $t \leq T$, that is, on the event $\{\tau = t\} \cup \{\sigma = t\}$, the random payoff equals

$$X_t = \mathbb{1}_{\{\tau = t \leq \sigma\}} L_t + \mathbb{1}_{\{\sigma = t < \tau\}} H_t, \tag{4.21}$$

where τ and σ are the exercise and cancellation times, respectively. Note that both the holder and the issuer may choose freely their exercise and cancellation times. If they decide to exercise their right at the same moment, we adopt the convention that a game contingent claim is exercised, rather than cancelled, so that the payoff is given by the process L. This feature is already reflected in the payoff formula (4.21).

If the class of all admissible cancellation times is assumed to be given as $\mathcal{T}^c = \{T\}$ then a game contingent claim X^g becomes an American claim X^a with the payoff process $X = L$. Therefore, the definition of a game contingent claim covers as a special case the notion of an American contingent claim. If we postulate, in addition, that $\mathcal{T}^e = \{T\}$ then a game contingent claim reduces to a European claim $X = L_T$ maturing at time T.

From the previous section, we know that valuation and hedging of American claims is related to optimal stopping problems. Game contingent claims are associated with the so-called *Dynkin games*, which in turn can be seen as natural extensions of optimal stopping problems.

4.5.1 Dynkin Games

Before analyzing in some detail the game contingent claims, we first present a brief survey of basic results concerning Dynkin games.

We assume that $t = 0, 1, \dots, T$ and we investigate the *Dynkin game* (also known as the *optimal stopping game*) associated with the payoff

$$Z(\sigma, \tau) = \mathbb{1}_{\{\tau \leq \sigma\}} L_\tau + \mathbb{1}_{\{\sigma < \tau\}} H_\sigma,$$

where $L \leq H$ are $\mathbb{F}$-adapted stochastic processes defined on a finite probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with a filtration $\mathbb{F}$.

Definition 4.5.2. For any fixed date $t = 0, 1, \dots, T$, by the **Dynkin game** started at time t and associated with the payoff $Z(\sigma, \tau)$, we mean a stochastic game in which the **min-player**, who controls a stopping time $\sigma \in \mathcal{T}_{[t,T]}$, wishes to minimize the conditional expectation

$$\mathbb{E}_{\mathbb{P}}(Z(\sigma, \tau) \,|\, \mathcal{F}_t), \tag{4.22}$$

while the **max-player**, who controls a stopping time $\tau \in \mathcal{T}_{[t,T]}$, wishes to maximize the conditional expectation (4.22).

Let us fix t. Since the stopping times σ and τ are assumed to belong to the class $\mathcal{T}_{[t,T]}$, formula (4.21) yields

$$\mathbb{E}_{\mathbb{P}}(Z(\sigma, \tau) \,|\, \mathcal{F}_t) = \mathbb{E}_{\mathbb{P}}\Big(\sum_{u=t}^{T} \big(\mathbb{1}_{\{\tau = u \leq \sigma\}} L_u + \mathbb{1}_{\{\sigma = u < \tau\}} H_u \big) \,\Big|\, \mathcal{F}_t \Big).$$

We are interested in finding the *value process* of a Dynkin game and the corresponding *optimal stopping times*. We start by stating the following definition of the upper and lower value processes.

Definition 4.5.3. The $\mathbb{F}$-adapted process $\bar{Y}^u$ given by the formula

$$\bar{Y}^u_t = \min_{\sigma \in \mathcal{T}_{[t,T]}} \max_{\tau \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\mathbb{P}}(Z(\sigma, \tau) \,|\, \mathcal{F}_t)$$

is called the **upper value process**. The **lower value process** $\bar{Y}^l$ is an $\mathbb{F}$-adapted process given by the formula

$$\bar{Y}^l_t = \max_{\tau \in \mathcal{T}_{[t,T]}} \min_{\sigma \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\mathbb{P}}(Z(\sigma, \tau) \,|\, \mathcal{F}_t).$$

Lemma 4.5.1. *Let $L^0(\Omega, \mathcal{F}, \mathbb{P})$ stand for the class of all random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let A and B be two finite sets and let $g : A \times B \to L^0(\Omega, \mathcal{F}, \mathbb{P})$ be an arbitrary map. Then*

$$\min_{a \in A} \max_{b \in B} g(a, b) \geq \max_{b \in B} \min_{a \in A} g(a, b).$$

It follows immediately from Definition 4.5.3 and Lemma 4.5.1 that the upper value process $\bar{Y}^u$ always dominates the lower value process $\bar{Y}^l$, that is, the inequality $\bar{Y}^u_t \geq \bar{Y}^l_t$ holds for every $t = 0, 1, \dots, T$.

Definition 4.5.4. If the equality $\bar{Y}^u = \bar{Y}^l$ is satisfied, we say that the **Stackelberg equilibrium** holds for a Dynkin game. Then the process $\bar{Y} = \bar{Y}^u = \bar{Y}^l$ is called the **value process.**

Definition 4.5.5. We say that the **Nash equilibrium** holds for a Dynkin game if for any t there exist stopping times $\sigma^*_t, \tau^*_t \in \mathcal{T}_{[t,T]}$, such that the inequalities

$$\mathbb{E}_{\mathbb{P}}\big(Z(\sigma^*_t, \tau) \,|\, \mathcal{F}_t\big) \leq \mathbb{E}_{\mathbb{P}}(Z(\sigma^*_t, \tau^*_t) \,|\, \mathcal{F}_t) \leq \mathbb{E}_{\mathbb{P}}\big(Z(\sigma, \tau^*_t) \,|\, \mathcal{F}_t\big) \tag{4.23}$$

are satisfied for arbitrary stopping times $\tau, \sigma \in \mathcal{T}_{[t,T]}$, that is, the pair (σ^*_t, τ^*_t) is a **saddle point** of a Dynkin game.

The next result shows that the existence of a Nash equilibrium for a Dynkin game implies the Stackelberg equilibrium.

Lemma 4.5.2. *Assume that a Nash equilibrium for a Dynkin game exists. Then the Stackelberg equilibrium holds and*

$$\bar{Y}_t = \mathbb{E}_{\mathbb{P}}(Z(\sigma^*_t, \tau^*_t) \,|\, \mathcal{F}_t),$$

so that σ^*_t and τ^*_t are optimal stopping times as of time t.

Proof. From (4.23), we obtain

$$\max_{\tau \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\mathbb{P}}\big(Z(\sigma^*_t, \tau) \,|\, \mathcal{F}_t\big) \leq \mathbb{E}_{\mathbb{P}}(Z(\sigma^*_t, \tau^*_t) \,|\, \mathcal{F}_t) \leq \min_{\sigma \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\mathbb{P}}\big(Z(\sigma, \tau^*_t) \,|\, \mathcal{F}_t\big).$$

Consequently,

$$\begin{aligned}
\bar{Y}^u_t &= \min_{\sigma \in \mathcal{T}_{[t,T]}} \max_{\tau \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\mathbb{P}}(Z(\sigma, \tau) \,|\, \mathcal{F}_t) \leq \mathbb{E}_{\mathbb{P}}(Z(\sigma^*_t, \tau^*_t) \,|\, \mathcal{F}_t) \\
&\leq \max_{\tau \in \mathcal{T}_{[t,T]}} \min_{\sigma \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\mathbb{P}}(Z(\sigma, \tau) \,|\, \mathcal{F}_t) = \bar{Y}^l_t.
\end{aligned}$$

Since the inequality $\bar{Y}^u_t \geq \bar{Y}^l_t$ is known to be always satisfied, we conclude that the value process $\bar{Y}$ is well defined and satisfies $\bar{Y}_t = \mathbb{E}_{\mathbb{P}}(Z(\sigma^*_t, \tau^*_t) \,|\, \mathcal{F}_t)$ for all $t = 0, 1, \dots, T$. ∎

The following definition introduces a plausible candidate for the value process of a Dynkin game.

Definition 4.5.6. The process Y is defined by setting $Y_T = L_T$ and, for any $t = 0, 1, \dots, T-1$,

$$Y_t = \min\Big\{H_t,\, \max\big\{L_t, \mathbb{E}_{\mathbb{P}}(Y_{t+1} \,|\, \mathcal{F}_t)\big\}\Big\}. \tag{4.24}$$

The assumption that $L \le H$ ensures that, for any $t = 0, 1, \dots, T-1$,

$$Y_t = \min\Big\{H_t,\ \max\big\{L_t, \mathbb{E}_{\mathbb{P}}(Y_{t+1}\,|\,\mathcal{F}_t)\big\}\Big\} = \max\Big\{L_t,\ \min\big\{H_t, \mathbb{E}_{\mathbb{P}}(Y_{t+1}\,|\,\mathcal{F}_t)\big\}\Big\}.$$

It is also clear from (4.24) that $L_t \le Y_t \le H_t$ for $t = 0, 1, \dots, T$. In particular, if the equality $L_t = H_t$ holds then necessarily $Y_t = L_t = H_t$.

Theorem 4.5.1. (i) *Let the stopping times σ_t^*, τ_t^* be given by*

$$\sigma_t^* = \min\big\{u \in \{t, t+1, \dots, T\}\,|\,Y_u = H_u\big\}$$

and

$$\tau_t^* = \min\big\{u \in \{t, t+1, \dots, T\}\,|\,Y_u = L_u\big\} \wedge T.$$

Then we have, for arbitrary stopping times $\tau, \sigma \in \mathcal{T}_{[t,T]}$,

$$\mathbb{E}_{\mathbb{P}}\big(Z(\sigma_t^*, \tau)\,|\,\mathcal{F}_t\big) \le Y_t \le \mathbb{E}_{\mathbb{P}}\big(Z(\sigma, \tau_t^*)\,|\,\mathcal{F}_t\big)$$

and thus also

$$\mathbb{E}_{\mathbb{P}}\big(Z(\sigma_t^*, \tau)\,|\,\mathcal{F}_t\big) \le \mathbb{E}_{\mathbb{P}}\big(Z(\sigma_t^*, \tau_t^*)\,|\,\mathcal{F}_t\big) \le \mathbb{E}_{\mathbb{P}}\big(Z(\sigma, \tau_t^*)\,|\,\mathcal{F}_t\big)$$

so that the Nash equilibrium holds.
(ii) *The process Y is the value process of a Dynkin game, that is, for every $t = 0, 1, \dots, T$,*

$$\begin{aligned} Y_t &= \min_{\sigma \in \mathcal{T}_{[t,T]}} \max_{\tau \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\mathbb{P}}\big(Z(\sigma, \tau)\,|\,\mathcal{F}_t\big) = \mathbb{E}_{\mathbb{P}}\big(Z(\sigma_t^*, \tau_t^*)\,|\,\mathcal{F}_t\big) \\ &= \max_{\tau \in \mathcal{T}_{[t,T]}} \min_{\sigma \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\mathbb{P}}\big(Z(\sigma, \tau)\,|\,\mathcal{F}_t\big) = \bar{Y}_t \end{aligned}$$

and thus the stopping times σ_t^* and τ_t^* are optimal as of time t.

4.5.2 Arbitrage Pricing of Game Contingent Claims

We place ourselves within the framework of the CRR model with the unique martingale measure denoted by $\widetilde{\mathbb{P}}$. We define the discounted payoffs $\widehat{L} = B^{-1}L$ and $\widehat{H} = B^{-1}H$ of a game contingent claim X^g and we consider the Dynkin game under $\widetilde{\mathbb{P}}$ with the payoff given by the expression

$$\widehat{Z}(\sigma, \tau) = \mathbb{1}_{\{\tau \le \sigma\}}\, \widehat{L}_\tau + \mathbb{1}_{\{\sigma < \tau\}}\, \widehat{H}_\sigma.$$

In view of the financial interpretation, the max-player and the min-player will now be referred to as the seller (issuer) and the buyer (holder), respectively.

Let us write $U = B\widehat{U}$, where $\widehat{U}$ is the value process of the Dynkin game, that is,

$$\widehat{U}_t = \min_{\sigma \in \mathcal{T}_{[t,T]}} \max_{\tau \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\widetilde{\mathbb{P}}}(\widehat{Z}(\sigma, \tau)\,|\,\mathcal{F}_t) = \max_{\tau \in \mathcal{T}_{[t,T]}} \min_{\sigma \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\widetilde{\mathbb{P}}}(\widehat{Z}(\sigma, \tau)\,|\,\mathcal{F}_t).$$

In view of Theorem 4.5.1, this also means that for $t = 0, 1, \dots, T-1$

$$\widehat{U}_t = \min\Big\{\widehat{H}_t,\ \max\big\{\widehat{L}_t,\ \mathbb{E}_{\widetilde{\mathbb{P}}}\big(\widehat{U}_{t+1}\,|\,\mathcal{F}_t\big)\big\}\Big\}$$

or, equivalently,

$$U_t = \min\Big\{H_t,\, \max\big\{L_t,\, B_t\,\mathbb{E}_{\widetilde{\mathbb{P}}}\big(B_{t+1}^{-1}U_{t+1}\,|\,\mathcal{F}_t\big)\big\}\Big\}$$

with the terminal condition $U_T = L_T$. By convention, we will refer to U (rather than $\widehat{U}$) as the value process of the Dynkin game associated with a game contingent claim X^g. For a fixed $t = 0, 1, \dots, T$, we define the following stopping times

$$\sigma_t^* = \min\big\{u \in \{t, t+1, \dots, T\}\,|\,U_u = H_u\big\}$$

and

$$\tau_t^* = \min\big\{u \in \{t, t+1, \dots, T\}\,|\,U_u = L_u\big\} \wedge T.$$

One can show that they are rational exercise times for the holder and issuer, respectively. Proofs of Propositions 4.5.1 and 4.5.2 are omitted.

Proposition 4.5.1. *The seller's price $\pi_t^s(X^g)$ at time t of a game contingent claim is equal to the seller's price $\pi^s(X^{a,\sigma_t^*})$ of an American claim X^{a,σ_t^*} with the payoff process $X^{\sigma_t^*}$ given by the formula $X_t^{\sigma^*} = Z(\sigma_t^*, t)$ for every $t = 0, 1, \dots, T$. Consequently, $\pi_t^s(X^g) = U_t$ for every $t = 0, 1, \dots, T$, and thus $\pi_t^s(X^g)$ is the solution to the problem*

$$\pi_t^s(X^g) \;=\; \max_{\tau\in\mathcal{T}_{[t,T]}} B_t\,\mathbb{E}_{\widetilde{\mathbb{P}}}(\widehat{Z}(\sigma_t^*, \tau)\,|\,\mathcal{F}_t) = \max_{\tau\in\mathcal{T}_{[t,T]}}\,\min_{\sigma\in\mathcal{T}_{[t,T]}} B_t\,\mathbb{E}_{\widetilde{\mathbb{P}}}(\widehat{Z}(\sigma, \tau)\,|\,\mathcal{F}_t).$$

This also means that

$$\pi_t^s(X^g) = B_t\,\mathbb{E}_{\widetilde{\mathbb{P}}}\big(\widehat{Z}(\sigma_t^*, \tau_t^*)\,|\,\mathcal{F}_t\big).$$

Finally, we examine the buyer's price of a game contingent claim.

Proposition 4.5.2. *The buyer's price $\pi^b(X^g)$ of a game contingent claim X^g is equal to the seller's price $\pi^s(X^g)$, that is, $\pi_t^b(X^g) = U_t$ for every $t = 0, 1, \dots, T$. Hence the arbitrage price $\pi(X^g)$ of a game contingent claim is unique and it is equal to the value process U of the associated Dynkin game. This means that it is given by the recursive formula*

$$\pi_t(X^g) = \min\Big\{H_t,\, \max\big\{L_t,\, B_t\,\mathbb{E}_{\widetilde{\mathbb{P}}}\big(B_{t+1}^{-1}\pi_{t+1}(X^g)\,|\,\mathcal{F}_t\big)\big\}\Big\} \tag{4.25}$$

with $\pi_T(X^g) = L_T$.

Under the assumption that the payoffs $H_t = h(S_t, t)$ and $L_t = \ell(S_t, t)$ are given in terms of the current value of the stock price at time t, it is convenient to denote $X_t^g = \pi_t(X^g)$ and to rewrite the last formula as follows

$$X_t^g = \min\Big\{h(S_t, t),\, \max\big\{\ell(S_t, t),\, (1+r)^{-1}\big(\widetilde{p}X_{t+1}^{gu} + (1-\widetilde{p})X_{t+1}^{gd}\big)\big\}\Big\} \tag{4.26}$$

with the terminal condition $X_T^g = \pi_T(X^g) = \ell(S_T, T)$.

4.6 Implementation of the CRR Model

To implement the CRR model, we proceed as follows:

- We fix maturity T and we assume that the **continuously compounded** interest rate r is such that $B(0,T) = e^{-rT}$. Note that T is now expressed in years. For instance, $T = 2$ months means that $T = 1/6$. In general, for different maturities we may write $B(0,T) = e^{-Y(0,T)T}$ where the function $Y(0,T)$ represents the **yield curve** computed from the bond market data.
- From the stock market data, we take the current stock price S_0 and we estimate the stock price **volatility** σ per one time unit (i.e., one year).
- Note that it was assumed so far that $t = 0, 1, 2, \ldots, T$, meaning that $\Delta t = 1$. In general, the length of each period can be any positive number smaller than 1. We set $n = T/\Delta t$ and we assume that n is an integer.
- Two widely used conventions for obtaining u and d from σ, r and Δt are:
  - **The Cox-Ross-Rubinstein (CRR) parametrisation**:
$$u = e^{\sigma\sqrt{\Delta t}} \quad \text{and} \quad d = \frac{1}{u}.$$
  - **The Jarrow-Rudd (JR) parameterisation**:
$$u = e^{\left(r - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}} \quad \text{and} \quad d = e^{\left(r - \frac{\sigma^2}{2}\right)\Delta t - \sigma\sqrt{\Delta t}}.$$

We first examine the Cox-Ross-Rubinstein parametrisation.

Proposition 4.6.1. *Assume that $B_{k\Delta t} = (1 + r\Delta t)^k$ for every $k = 0, 1, \ldots, n$ and $u = \frac{1}{d} = e^{\sigma\sqrt{\Delta t}}$ in the CRR model. Then the risk-neutral probability measure $\widetilde{\mathbb{P}}$ satisfies*

$$\widetilde{\mathbb{P}}\left(S_{t+\Delta t} = S_t u \,|\, S_t\right) = \frac{1}{2} + \frac{r - \frac{\sigma^2}{2}}{2\sigma}\sqrt{\Delta t} + o(\sqrt{\Delta t})$$

provided that Δt is sufficiently small.

Proof. The risk-neutral probability measure for the CRR model is given by

$$\widetilde{p} = \widetilde{\mathbb{P}}\left(S_{t+\Delta t} = S_t u \,|\, S_t\right) = \frac{1 + r\Delta t - d}{u - d}.$$

Under the CRR parametrisation, we obtain

$$\widetilde{p} = \frac{1 + r\Delta t - d}{u - d} = \frac{1 + r\Delta t - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}.$$

The Taylor expansions up to the second order term are

$$e^{\sigma\sqrt{\Delta t}} = 1 + \sigma\sqrt{\Delta t} + \frac{\sigma^2}{2}\Delta t + o(\Delta t),$$
$$e^{-\sigma\sqrt{\Delta t}} = 1 - \sigma\sqrt{\Delta t} + \frac{\sigma^2}{2}\Delta t + o(\Delta t).$$

By substituting the Taylor expansions into the risk-neutral probability measure, we obtain

$$\begin{aligned}\widetilde{p} &= \frac{1 + r\Delta t - \left(1 - \sigma\sqrt{\Delta t} + \frac{\sigma^2}{2}\Delta t\right) + o(\Delta t)}{\left(1 + \sigma\sqrt{\Delta t} + \frac{\sigma^2}{2}\Delta t\right) - \left(1 - \sigma\sqrt{\Delta t} + \frac{\sigma^2}{2}\Delta t\right) + o(\Delta t)} \\ &= \frac{\sigma\sqrt{\Delta t} + \left(r - \frac{\sigma^2}{2}\right)\Delta t + o(\Delta t)}{2\sigma\sqrt{\Delta t} + o(\Delta t)} \\ &= \frac{1}{2} + \frac{r - \frac{\sigma^2}{2}}{2\sigma}\sqrt{\Delta t} + o(\sqrt{\Delta t})\end{aligned}$$

as was required to show. □

To summarise:

- For Δt sufficiently small, we get
$$\widetilde{p} = \frac{1 + r\Delta t - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}} \approx \frac{1}{2} + \frac{r - \frac{\sigma^2}{2}}{2\sigma}\sqrt{\Delta t}.$$
- Note that $1 + r\Delta t \approx e^{r\Delta t}$ when Δt is sufficiently small.
- Hence we may also represent the risk-neutral probability measure as follows
$$\widetilde{p} \approx \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}.$$

Let us finally note that if we define $\widehat{r}$ such that $(1 + \widehat{r})^n = e^{rT}$ for a fixed T and $n = T/\Delta t$ then $\widehat{r} \approx r\Delta t$ since $\ln(1 + \widehat{r}) = r\Delta t$ and $\ln(1 + \widehat{r}) \approx \widehat{r}$ when $\widehat{r}$ is close to zero.

Example 4.6.1. We consider here the Cox-Ross-Rubinstein parameterisation.

- Let the annualized variance of logarithmic returns be $\sigma^2 = 0.1$.
- The interest rate is set to $r = 0.1$ per annum.
- Suppose that the current stock price is $S_0 = 50$.
- We examine European and American put options with strike price $K = 53$ and maturity $T = 4$ months (that is, $T = \frac{1}{3}$).
- The length of each period is chosen to be $\Delta t = \frac{1}{12}$, that is, one month.
- Hence $n = \frac{T}{\Delta t} = 4$ time steps.
- We adopt the CRR parameterisation to derive the stock price.
- Then $u = 1.0956$ and $d = 1/u = 0.9128$.
- We compute $1 + r\Delta t = 1.00833 \approx e^{r\Delta t}$ and $\widetilde{p} = 0.5228$.

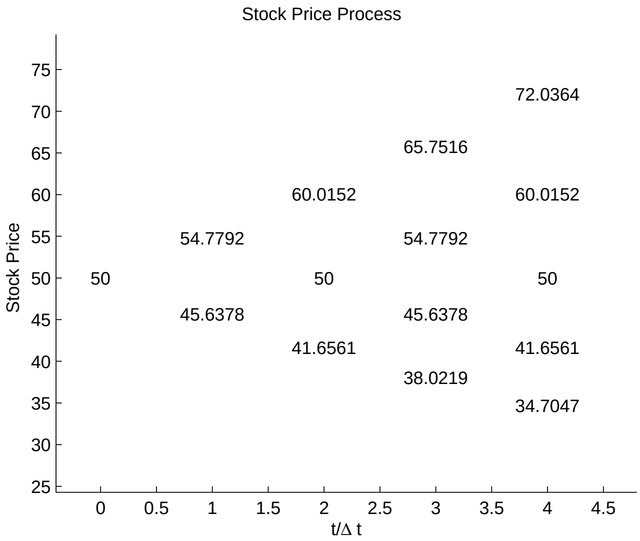


Figure 4.1: Stock Price Process

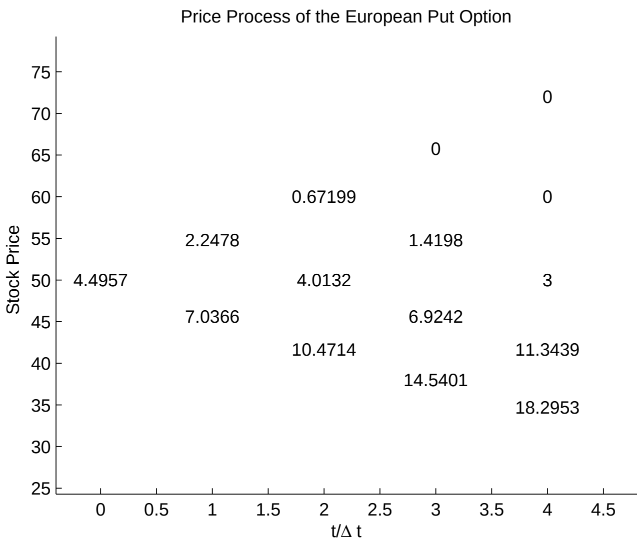


Figure 4.2: European Put Option Price

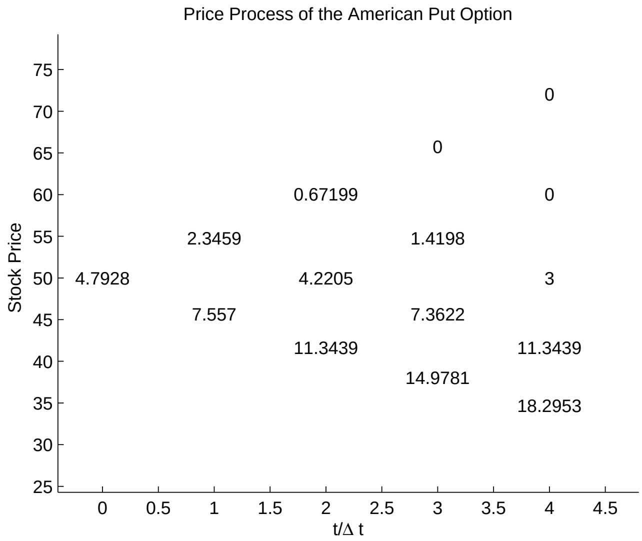


Figure 4.3: American Put Option Price

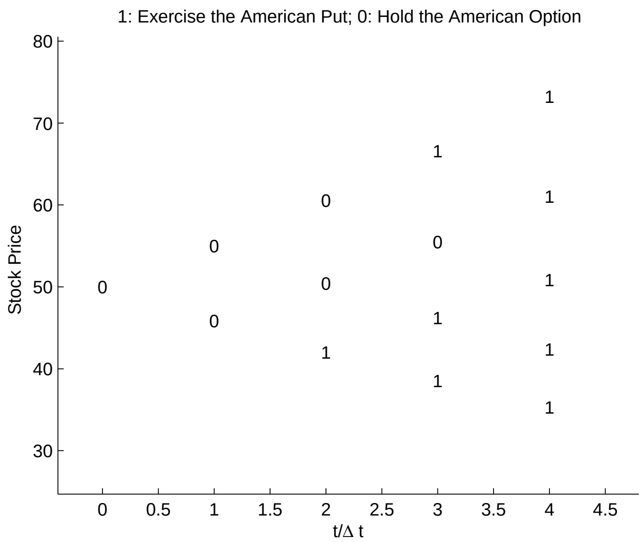


Figure 4.4: Rational Exercise Policy

The next result deals with the Jarrow-Rudd parametrisation.

Proposition 4.6.2. *Let* $B_{k\Delta t} = (1 + r\Delta t)^k$ *for* $k = 0, 1, \ldots, n$. *We assume that*

$$u = e^{\left(r - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}}$$

and

$$d = e^{\left(r - \frac{\sigma^2}{2}\right)\Delta t - \sigma\sqrt{\Delta t}}.$$

Then the risk-neutral probability measure $\widetilde{\mathbb{P}}$ *satisfies*

$$\widetilde{\mathbb{P}}\left(S_{t+\Delta t} = S_t u \,|\, S_t\right) = \frac{1}{2} + o(\Delta t)$$

provided that Δt *is sufficiently small.*

Proof. Under the JR parametrisation, we have

$$\widetilde{p} = \frac{1 + r\Delta t - d}{u - d} = \frac{1 + r\Delta t - e^{\left(r - \frac{\sigma^2}{2}\right)\Delta t - \sigma\sqrt{\Delta t}}}{e^{\left(r - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}} - e^{\left(r - \frac{\sigma^2}{2}\right)\Delta t - \sigma\sqrt{\Delta t}}}.$$

The Taylor expansions up to the second order term are

$$e^{\left(r - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}} = 1 + r\Delta t + \sigma\sqrt{\Delta t} + o(\Delta t)$$
$$e^{\left(r - \frac{\sigma^2}{2}\right)\Delta t - \sigma\sqrt{\Delta t}} = 1 + r\Delta t - \sigma\sqrt{\Delta t} + o(\Delta t)$$

and thus

$$\widetilde{p} = \frac{1}{2} + o(\Delta t)$$

as was required to show. □

Example 4.6.2. We now assume the Jarrow-Rudd parameterisation.

- We consider the same problem as in Example 7.2, but with parameters u and d computed using the JR parameterisation. We obtain $u = 1.1002$ and $d = 0.9166$ (note that $u \neq 1/d$).
- As before, $1 + r\Delta t = 1.00833 \approx e^{r\Delta t}$, but $\widetilde{p} = 0.5$.
- We compute the price processes for the stock, the European put option, the American put option and we find the rational exercise time.
- When we compare with Example 7.2, we see that the results are slightly different than before, although it appears that the rational exercise policy is the same.
- The CRR and JR parameterisations are both set to approach the Black-Scholes model.
- For Δt sufficiently small, the prices computed under the two parametrisations will be very close to one another.

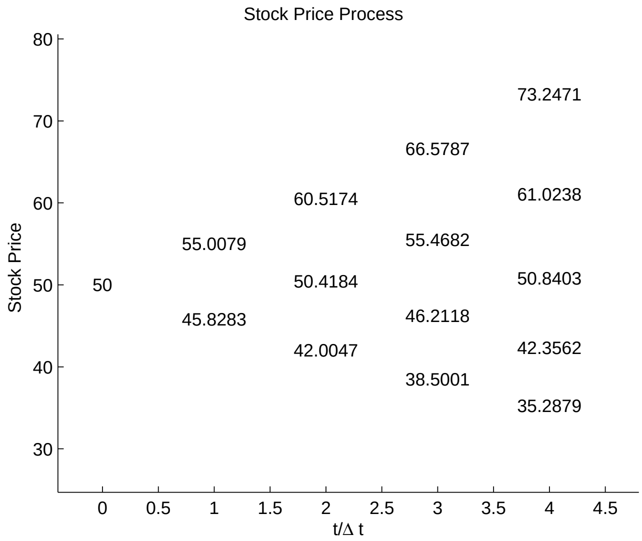


Figure 4.5: Stock Price Process

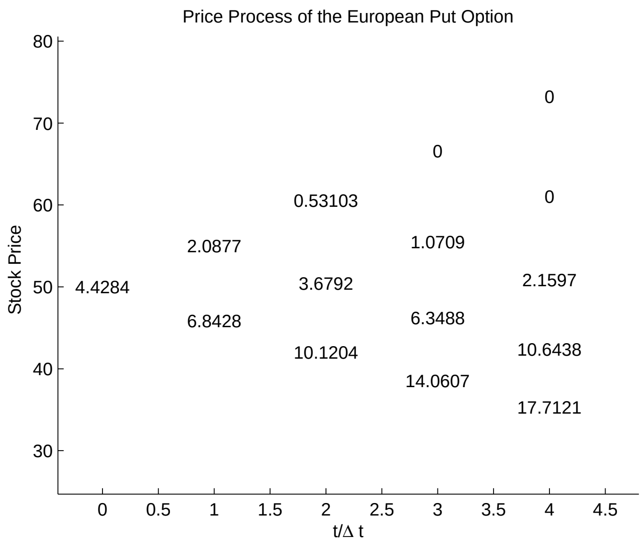


Figure 4.6: European Put Option Price

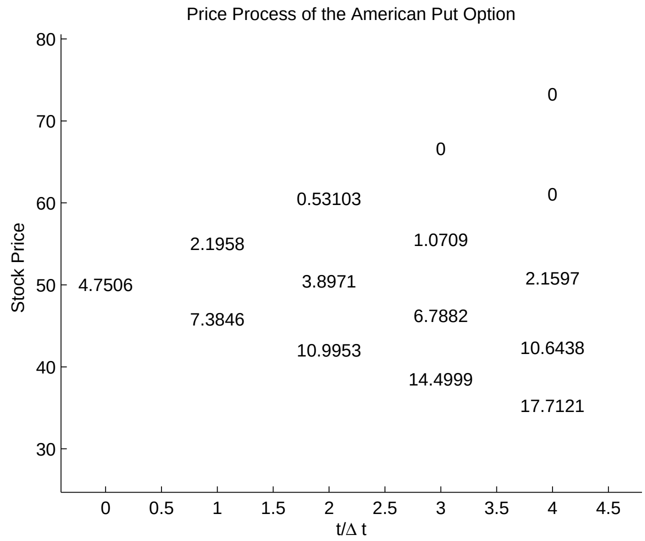


Figure 4.7: American Put Option Price

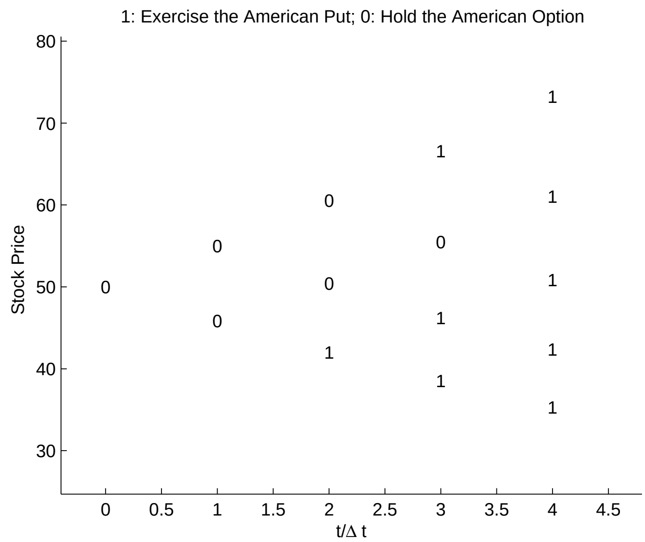


Figure 4.8: Rational Exercise Policy

Chapter 5

The Black-Scholes Model

5.1 The Wiener Process and its Properties

We will now introduce an important example of a continuous-time Markov process, called the **Wiener process** in honour of American mathematician Nobert Wiener (1894–1964). It is also known as the **Brownian motion**, after Scottish botanist Robert Brown (1773–1858). The Wiener process can be seen as a continuous-time counterpart of the Bernoulli counting process, which was introduced in Section 3.3.

Definition 5.1.1. *A stochastic process (W_t) with time parameter $t \in \mathbb{R}_+$ is called the* ***Wiener process*** *(or the* ***Brownian motion****) if: 1. $W_0 = 0$,*
2. the sample paths of the process W, that is, the maps $t \to W_t(\omega)$ are continuous functions,
3. the process W has the Gaussian (i.e. normal) distribution with the expected value $\mathbb{E}_{\mathbb{P}}(W_t) = 0$ for all $t \geq 0$ and the covariance

$$\text{Cov}\,(W_s, W_t) = \min\,(s,t), \quad s,t \geq 0.$$

As opposed to a **Markov chain**, that is, a Markov process taking values in a countable state space, the state space of the Wiener process is the (uncountable) set of all real numbers. This important property can be easily deduced from Definition 5.1.1 since, by assumption, the distribution of the Wiener process at any date t is Gaussian (i.e., normal). We take for granted without proof the following important result, first established by Wiener (1923).

Theorem 5.1.1. *The Wiener process exists, that is, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a process W defined on this space, such that conditions (1)-(3) of Definition 5.1.1 are met.*

Let $N(\mu, \sigma^2)$ denote the Gaussian distribution with the expected value μ and the variance σ^2. From Definition 5.1.1, it follows that for every fixed $t > 0$ we have $W_t \sim N(0,t)$, so that $(\sqrt{t})^{-1}\, W_t \sim N(0,1)$ for every $t > 0$. More explicitly, the random variable W_t has the probability density function $p(x,t)$ given by

$$p(t,x) = \frac{1}{\sqrt{2\pi t}}\, e^{-x^2/2t}.$$

This means that, for any real numbers $a \leq b$,

$$\mathbb{P}(W_t \in [a,b]) = \int_a^b \frac{1}{\sqrt{2\pi t}}\, e^{-x^2/2t}\, dx = \int_{\frac{a}{\sqrt{t}}}^{\frac{b}{\sqrt{t}}} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx = N\left(\frac{b}{\sqrt{t}}\right) - N\left(\frac{a}{\sqrt{t}}\right),$$

where N is the standard Gaussian (i.e. normal) cumulative distribution function, that is, for every $z \in \mathbb{R}$,

$$N(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx = \int_{-\infty}^{z} n(x)\, dx$$

where n is the standard Gaussian probability density function. Since the Wiener process W is Gaussian, the random vector

$$\begin{pmatrix} W_s \\ W_t \end{pmatrix}, \quad s,t > 0,$$

has the two-dimensional Gaussian distribution $N(\mathbf{0}, C)$, where $\mathbf{0}$ denotes the null vector of expected values and C is the covariance matrix

$$C = \begin{pmatrix} s & \min(s,t) \\ \min(s,t) & t \end{pmatrix}.$$

Consequently, the increment $W_t - W_s$ has again Gaussian distribution and manifestly

$$\mathbb{E}_{\mathbb{P}}\left(W_t - W_s\right) = \mathbb{E}_{\mathbb{P}}\left(W_t\right) - \mathbb{E}_{\mathbb{P}}\left(W_s\right) = 0.$$

Moreover, we obtain, for $s < t$,

$$\begin{aligned} \mathbf{Var}\left(W_t - W_s\right) &= \mathbb{E}_{\mathbb{P}}\left(W_t - W_s\right)^2 = \mathbb{E}_{\mathbb{P}}(W_t^2) - 2\mathbb{E}_{\mathbb{P}}\left(W_s W_t\right) + \mathbb{E}_{\mathbb{P}}(W_s^2) \\ &= t - 2\min(s,t) + s = t - 2s + s = t - s. \end{aligned}$$

We conclude that

$$W_t - W_s \sim N(0, t-s), \quad s < t. \tag{5.1}$$

It appears that the Wiener process has a common feature with the Poisson process, specifically, both are processes of independent increments. In fact, they belong to the class of **Lévy processes**, that is, the time-homogeneous processes with independent increments (in other words, the processes with stationary and independent increments).

Proposition 5.1.1. *The Wiener process has stationary and independent increments. Specifically, for every $n = 2, 3, \dots$ and any choice of dates $0 \leq t_1 < t_2 < \dots < t_n$, the random variables*

$$W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2}, \dots, W_{t_n} - W_{t_{n-1}}$$

are independent and they have the same probability distribution as the increments

$$W_{t_2+h} - W_{t_1+h}, W_{t_3+h} - W_{t_2+h}, \dots, W_{t_n+h} - W_{t_{n-1}+h}$$

where h is an arbitrary non-negative number.

Proof. (MATH3975) For simplicity, we only consider the increments $W_{t_2} - W_{t_1}$ and $W_{t_4} - W_{t_3}$. Note first that a random vector $(W_{t_1}, W_{t_2}, W_{t_3}, W_{t_4})$ is Gaussian. Moreover,

$$Y = \begin{pmatrix} W_{t_2} - W_{t_1} \\ W_{t_4} - W_{t_3} \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} W_{t_1} \\ W_{t_2} \\ W_{t_3} \\ W_{t_4} \end{pmatrix},$$

and thus Y is a Gaussian vector as well. We compute

$$\begin{aligned} &\mathrm{Cov}\,(W_{t_2} - W_{t_1}, W_{t_4} - W_{t_3}) = \mathbb{E}_{\mathbb{P}}\,(W_{t_2} - W_{t_1})\,(W_{t_4} - W_{t_3}) \\ &= \mathbb{E}_{\mathbb{P}}\,(W_{t_2}W_{t_4}) - \mathbb{E}_{\mathbb{P}}\,(W_{t_2}W_{t_3}) - \mathbb{E}_{\mathbb{P}}\,(W_{t_1}W_{t_4}) + \mathbb{E}_{\mathbb{P}}\,(W_{t_1}W_{t_3}) \\ &= t_2 - t_2 - t_1 + t_1 = 0. \end{aligned}$$

Hence, by the **normal correlation theorem**, the random variables $W_{t_2} - W_{t_1}$ and $W_{t_4} - W_{t_3}$ are independent. Note also that it follows immediately from (5.1) that the increments of W are stationary. ∎

5.2 Markov Property (MATH3975)

Recall that the state space for the Wiener process is the real line $\mathbb{R}$. In that case, the **Markov property** of W is defined through equality (5.2).

Theorem 5.2.1. *The Wiener process W is a Markov process in the following sense: for every $n \geq 1$, any sequence of times $0 < t_1 < \ldots < t_n < t$ and any collection $x_1, \ldots, x_n$ of real numbers, the following holds:*

$$\mathbb{P}\,(W_t \leq x|\, W_{t_1} = x_1, \ldots, W_{t_n} = x_n) = \mathbb{P}\,(W_t \leq x|\, W_{t_n} = x_n)\,, \tag{5.2}$$

for all $x \in \mathbb{R}$. Moreover,

$$\mathbb{P}\,(W_t \leq y|\, W_s = x) = \int_{-\infty}^{y} p(t-s, z-x)\,dz$$

where

$$p(t-s, z-x) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{(z-x)^2}{2(t-s)}\right)$$

is the transition probability density function of the Wiener process.

Proof. By the independence of increments, we obtain

$$\begin{aligned} &\mathbb{P}\,(W_t \leq x|\, W_{t_1} = x_1, \ldots, W_{t_n} = x_n) = \mathbb{P}\,(W_t - W_{t_n} \leq x - x_n|\, W_{t_1} = x_1, \ldots, W_{t_n} = x_n) \\ &= \mathbb{P}\,(W_t - W_{t_n} \leq x_n)\,. \end{aligned}$$

Similarly,

$$\begin{aligned} &\mathbb{P}\,(W_t \leq x|\, W_{t_n} = x_n) = \mathbb{P}\,(W_t - W_{t_n} \leq x - x_n|\, W_{t_n} = x_n) \\ &= \mathbb{P}\,(W_t - W_{t_n} \leq x - x_n) \end{aligned} \tag{5.3}$$

and the Markov property follows. The last part of the theorem is an immediate consequence of (5.1) and (5.3). ∎

Given the transition density function, we can compute the joint probability distribution of the random vector $(W_{t_1}, \ldots, W_{t_n})$ for $t_1 < t_2 < \ldots < t_n$, using the transition densities only. Let us consider the case $n = 2$. Invoking the independent increments property again, we obtain

$$\begin{aligned}
&\mathbb{P}\left(W_{t_1} \leq x_1, W_{t_2} \leq x_2\right) \\
&= \int_{-\infty}^{x_1} \mathbb{P}\left(W_{t_2} \leq x_2 \middle| W_{t_1} = y\right) p\left(t_1, y\right) dy \\
&= \int_{-\infty}^{x_1} \mathbb{P}\left(W_{t_2} - W_{t_1} \leq x_2 - y\right) p\left(t_1, y\right) dy \\
&= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2 - y} p\left(t_2 - t_1, u\right) p\left(t_1, y\right) du dy \\
&= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} p\left(t_2 - t_1, v - y\right) p\left(t_1, y\right) dv dy.
\end{aligned}$$

Finally,

$$\mathbb{P}\left(W_{t_1} \leq x_1, W_{t_2} \leq x_2\right) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} p\left(t_2 - t_1, v - y\right) p\left(t_1, y\right) dv dy.$$

In the same way, one can compute $\mathbb{P}\left(W_{t_1} \leq x_1, W_{t_2} \leq x_2, W_{t_3} \leq x_3\right)$ and so on.

5.3 Martingale Property (MATH3975)

An important class of stochastic processes are continuous-time **martingales**.

Definition 5.3.1. We say that stochastic process (X_t) is a **continuous-time martingale** with respect to its natural filtration $\mathbb{F}^X$ whenever

$$\mathbb{E}_{\mathbb{P}}\left(X_t \,|\, X_{t_1}, \ldots, X_{t_n}\right) = X_{t_n}$$

for any $n \geq 1$ and $0 \leq t_1 < t_2 < \ldots < t_n \leq t$.

It appears that the Wiener process enjoys the martingale property.

Proposition 5.3.1. *The Wiener process is a continuous-time martingale with respect to its natural filtration $\mathbb{F}^W$, that is,*

$$\mathbb{E}_{\mathbb{P}}\left(W_t \,|\, W_{t_1}, \ldots, W_{t_n}\right) = W_{t_n}$$

for any $n \geq 1$ and $0 \leq t_1 < t_2 < \ldots < t_n \leq t$.

Proof. By the independence of increments, we have

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}\left(W_t \,|\, W_{t_1}, \ldots, W_{t_n}\right) = \mathbb{E}_{\mathbb{P}}\left(W_t - W_{t_n} + W_{t_n} \middle| W_{t_1}, \ldots, W_{t_n}\right) \\
&= \mathbb{E}_{\mathbb{P}}\left(W_t - W_{t_n} \middle| W_{t_1}, \ldots, W_{t_n}\right) + \mathbb{E}_{\mathbb{P}}\left(W_{t_n} \middle| W_{t_1}, \ldots, W_{t_n}\right) \\
&= \mathbb{E}_{\mathbb{P}}\left(W_t - W_{t_n}\right) + W_{t_n} = W_{t_n},
\end{aligned}$$

as required. ■

Another example of a continuous-time martingale associated with the Wiener process is given in Proposition 3.4.3.

5.4 The Black-Scholes Call Pricing Formula

In this section, we apply the Wiener process to model the **stock price** S in the financial market with continuous trading. If you wish to learn more about continuous-time financial models and pertinent results from Stochastic Analysis based on the Itō stochastic integral (see, for instance, Kuo (2006)), you may enrol in the fourth-year course **Advanced Option Pricing**.

Following Black and Scholes (1973), we postulate that the stock price process S can be described under the risk-neutral probability measure $\widetilde{\mathbb{P}}$ by the following **stochastic differential equation**

$$dS_t = rS_t\,dt + \sigma S_t\,dW_t, \quad t \geq 0, \tag{5.4}$$

with a constant initial value $S_0 > 0$. The term $\sigma\,dW_t$ in dynamics (5.4) of the stock price is aimed to give a reasonable description of the uncertainty in the market. In particular, the **volatility** parameter σ measures the size of random fluctuations of the stock price. It turns out that stochastic differential equation (5.4) can be solved explicitly yielding

$$S_t = S_0 \exp\left(\sigma W_t + \Big(r - \frac{1}{2}\sigma^2\Big)t\right). \tag{5.5}$$

This means, in particular, that S_t has the lognormal probability distribution for every $t > 0$. It can also be shown that S is a Markov process. Note, however, that S is not a process of independent increments. We assume that the interest rate r is deterministic, constant and continuously compounded. Hence the **money market account** is given by the formula

$$B_t = B_0 e^{rt}, \quad t \geq 0,$$

where, by the usual convention, we set $B_0 = 1$. It is easy to check that

$$dB_t = rB_t\,dt, \quad t \geq 0.$$

The next result is for the advanced level only.

Proposition 5.4.1. *The discounted stock price, that is, the process $\widehat{S}$ given by the formula*

$$\widehat{S}_t = \frac{S_t}{B_t} = e^{-rt}S_t, \quad t \geq 0, \tag{5.6}$$

is a continuous-time martingale with respect to the filtration $\mathbb{F}^{\widehat{S}}$ under $\widetilde{\mathbb{P}}$, that is, for every $0 \leq s \leq t$,

$$\mathbb{E}_{\widetilde{\mathbb{P}}}(\widehat{S}_t \,|\, \widehat{S}_u,\, u \leq s) = \widehat{S}_s.$$

Proof. It follows immediately from (5.5) and (5.6) that

$$\widehat{S}_t = S_0\, e^{\sigma W_t - \frac{1}{2}\sigma^2 t} = \widehat{S}_s\, e^{\sigma(W_t - W_s) - \frac{1}{2}\sigma^2(t-s)}. \tag{5.7}$$

Hence if we know the value of $\widehat{S}_t$ then we also know the value of W_t and vice versa. This immediately implies that $\mathbb{F}^{\widehat{S}} = \mathbb{F}^W$. Therefore, for any integrable random variable X the following conditional expectations coincide

$$\mathbb{E}_{\widetilde{\mathbb{P}}}(X \,|\, \widehat{S}_u,\, u \leq s) = \mathbb{E}_{\widetilde{\mathbb{P}}}(X \,|\, W_u,\, u \leq s). \tag{5.8}$$

Since the Wiener process W has independent increments, using also the well-known properties of the conditional expectation, we obtain the following chain of equalities

$$
\begin{aligned}
&\mathbb{E}_{\widetilde{\mathbb{P}}}\big(\widehat{S}_t \,\big|\, \widehat{S}_u,\, u \le s\big) \\
&= \mathbb{E}_{\widetilde{\mathbb{P}}}\left(\widehat{S}_s\, e^{\sigma\left(W_t - W_s - \frac{1}{2}\sigma^2(t-s)\right)} \,\Big|\, \widehat{S}_u,\, u \le s\right) && \text{(from (5.7))} \\
&= \widehat{S}_s\, e^{-\frac{1}{2}\sigma^2(t-s)}\, \mathbb{E}_{\widetilde{\mathbb{P}}}\left(e^{\sigma(W_t - W_s)} \,\Big|\, \widehat{S}_u,\, u \le s\right) && \text{(property of conditioning)} \\
&= \widehat{S}_s\, e^{-\frac{1}{2}\sigma^2(t-s)}\, \mathbb{E}_{\widetilde{\mathbb{P}}}\left(e^{\sigma(W_t - W_s)} \,\big|\, W_u,\, u \le s\right) && \text{(from (5.8))} \\
&= \widehat{S}_s\, e^{-\frac{1}{2}\sigma^2(t-s)}\, \mathbb{E}_{\widetilde{\mathbb{P}}}\left(e^{\sigma(W_t - W_s)}\right). && \text{(independence of increments)}
\end{aligned}
$$

Recall also that $W_t - W_s = \sqrt{t-s}\, Z$ where $Z \sim N(0,1)$, and thus

$$\mathbb{E}_{\widetilde{\mathbb{P}}}\big(\widehat{S}_t \,\big|\, \widehat{S}_u,\, u \le s\big) = \widehat{S}_s\, e^{-\frac{1}{2}\sigma^2(t-s)}\, \mathbb{E}_{\widetilde{\mathbb{P}}}\big(e^{\sigma\sqrt{t-s}Z}\big).$$

Let us finally observe that if $Z \sim N(0,1)$ then for any real a

$$\mathbb{E}_{\widetilde{\mathbb{P}}}\big(e^{aZ}\big) = e^{a^2/2}.$$

By setting $a = \sigma\sqrt{t-s}$, we finally obtain

$$\mathbb{E}_{\widetilde{\mathbb{P}}}\big(\widehat{S}_t \,\big|\, \widehat{S}_u,\, u \le s\big) = \widehat{S}_s\, e^{-\frac{1}{2}\sigma^2(t-s)}\, e^{\frac{1}{2}\sigma^2(t-s)} = \widehat{S}_s,$$

which shows that $\widehat{S}$ is indeed a martingale with respect to $\mathbb{F}^{\widehat{S}}$ under $\widetilde{\mathbb{P}}$. ∎

Recall that the **European call option** written on the stock is a traded security, which pays at its maturity T the random amount

$$C_T = (S_T - K)^+,$$

where $x^+ = \max(x, 0)$ and $K > 0$ is a fixed **strike** (or **exercise price**). We take for granted that for $t \le T$ the price $C_t(x)$ of the call option when $S_t = x$ is given by the **risk-neutral pricing formula**

$$C_t(x) = e^{-r(T-t)}\, \mathbb{E}_{\widetilde{\mathbb{P}}}\left((S_T - K)^+ \,\big|\, S_t = x\right). \tag{5.9}$$

In fact, this formula can be supported by the replication principle. However, this argument requires the knowledge of the stochastic integration theory with respect to the Brownian motion, as developed by Itō (1944, 1946). The following call option pricing result was established in the seminal paper by Black and Scholes (1973). Recall that $\ln = \log_e$.

Theorem 5.4.1. ***Black-Scholes Call Pricing Formula.*** *The arbitrage price of the call option at time $t \le T$ equals*

$$C_t(S_t) = S_t N\big(d_+(S_t, T-t)\big) - K e^{-r(T-t)} N\big(d_-(S_t, T-t)\big) \tag{5.10}$$

where

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

and N is the standard Gaussian (i.e. normal) cumulative distribution function.

Proof. (MATH3975) Using (5.5), we can represent the stock price S_T as follows

$$S_T = S_t\, e^{\left(r-\frac{1}{2}\sigma^2\right)(T-t)+\sigma(W_T-W_t)}.$$

As in the proof of Proposition 5.4.1, we write $W_T - W_t = \sqrt{T-t}Z$ where Z has the standard Gaussian probability distribution, that is, $Z \sim N(0,1)$.

Using (5.9) and the independence of increments of the Wiener process W, we obtain, for a generic value $x > 0$ of the stock price S_t at time t,

$$\begin{aligned}
C_t(x) &= e^{-r(T-t)}\,\mathbb{E}_{\widetilde{\mathbb{P}}}\left(\left(S_t\, e^{\left(r-\frac{1}{2}\sigma^2\right)(T-t)+\sigma(W_T-W_t)} - K\right)^+ \,\Big|\, S_t = x\right)\\
&= e^{-r(T-t)}\,\mathbb{E}_{\widetilde{\mathbb{P}}}\left(x\, e^{\left(r-\frac{1}{2}\sigma^2\right)(T-t)+\sigma\sqrt{T-t}Z} - K\right)^+\\
&= e^{-r(T-t)}\int_{-\infty}^{\infty}\left(x\, e^{\left(r-\frac{1}{2}\sigma^2\right)(T-t)+\sigma\sqrt{T-t}z} - K\right)^+ n(z)\,dz
\end{aligned}$$

where n is the standard Gaussian probability density function.

It is clear that the function under the integral sign is non-zero if and only if the inequality

$$x\, e^{\left(r-\frac{1}{2}\sigma^2\right)(T-t)+\sigma\sqrt{T-t}z} - K > 0$$

holds. This in turn is equivalent to the following inequality

$$z \geq \frac{\ln\frac{K}{x} - \left(r-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} = -d_-(x, T-t).$$

Hence, denoting $d_+ = d_+(x, T-t)$ and $d_- = d_-(x, T-t)$, we obtain

$$\begin{aligned}
C_t(x) &= e^{-r(T-t)}\int_{-d_-}^{\infty}\left(x\, e^{\left(r-\frac{1}{2}\sigma^2\right)(T-t)+\sigma\sqrt{T-t}z} - K\right) n(z)\,dz\\
&= x\, e^{-\frac{1}{2}\sigma^2(T-t)}\int_{-d_-}^{\infty} e^{\sigma\sqrt{T-t}z} n(z)\,dz - Ke^{-r(T-t)}\int_{-d_-}^{\infty} n(z)\,dz\\
&= x\, e^{-\frac{1}{2}\sigma^2(T-t)}\int_{-d_-}^{\infty} e^{\sigma\sqrt{T-t}z}\frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z^2}\,dz - Ke^{-r(T-t)}\left(1 - N\left(-d_-\right)\right)\\
&= x\int_{-d_-}^{\infty}\frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(z-\sigma\sqrt{T-t}\right)^2}\,dz - Ke^{-r(T-t)}N\left(d_-\right)\\
&= x\int_{-d_- - \sigma\sqrt{T-t}}^{\infty}\frac{1}{\sqrt{2\pi}}\, e^{-u^2/2}\,du - Ke^{-r(T-t)}N\left(d_-\right)\\
&= x\int_{-d_+}^{\infty}\frac{1}{\sqrt{2\pi}}\, e^{-u^2/2}\,du - Ke^{-r(T-t)}N\left(d_-\right)\\
&= xN\left(d_+\right) - Ke^{-r(T-t)}N\left(d_-\right).
\end{aligned}$$

Upon substitution $x = S_t$, this establishes formula (5.10). ■

The arbitrage price of the put option in the Black-Scholes model can now be computed without difficulties, as shown by the following result.

Corollary 5.4.1. *From the put-call parity*

$$C_t - P_t = S_t - Ke^{-r(T-t)}$$

we deduce that the arbitrage price at time $0 \leq t < T$ *of the European put option equals*

$$P_t = K\,e^{-r(T-t)}N\big(-d_-(S_t, T-t)\big) - S_t N\big(-d_+(S_t, T-t)\big).$$

Example 5.4.1. Suppose that the current stock price equals \$31, the stock price volatility is $\sigma = 10\%$ per annum, and the risk-free interest rate is $r = 5\%$ per annum with continuous compounding.

Call option. Let us consider a call option on the stock S, with strike price \$30 and with 3 months to expiry. We may assume, without loss of generality, that $t = 0$ and $T = 0.25$. We obtain (approximately) $d_+(S_0, T) = 0.93$, and thus

$$d_-(S_0, T) = d_+(S_0, T) - \sigma\sqrt{T} = 0.88.$$

The Black-Scholes call option pricing formula now yields (approximately)

$$C_0 = 31N(0.93) - 30e^{-0.05/4}\,N(0.88) == 25.42 - 23.9 = 1.52$$

since $N(0.93) \approx 0.82$ and $N(0.88) \approx 0.81$. Let $C_t = \varphi^0_t B_t + \varphi^1_t S_t$. The hedge ratio for the call option is known to be given by the formula $\varphi^1_t = N(d_+(S_t, T-t))$. Hence the replicating portfolio for the call option at time $t = 0$ is given by

$$\varphi^0_0 = -23.9, \quad \varphi^1_0 = N(d_+(S_0, T)) = 0.82.$$

This means that to hedge the short position in the call option, which was sold at the arbitrage price $C_0 = \$1.52$, the option's writer needs to purchase at time 0 the number $\delta = 0.82$ shares of stock. This transaction requires an additional borrowing of 23.9 units of cash.

Note that the **elasticity** at time 0 of the call option price with respect to the stock price equals

$$\eta^c_0 := \frac{\partial C}{\partial S}\left(\frac{C_0}{S_0}\right)^{-1} = \frac{N\big(d_+(S_0, T)\big)S_0}{C_0} = 16.72.$$

Suppose that the stock price rises immediately from \$31 to \$31.2, yielding a return rate of 0.65% flat. Then the option price will move by approximately 16.5 cents from \$1.52 to \$1.685, giving a return rate of 10.86% flat. The option has nearly 17 times the return rate of the stock; of course, this also means that it will drop 17 times as fast. If an investor's portfolio involves 5 long call options (each on a round lot of 100 shares of stock), the position delta equals $500 \times 0.82 = 410$, so that it is the same as for a portfolio involving 410 shares of the underlying stock.

Put option. Let us now assume that an option is a put. The price of a put option at time 0 equals

$$P_0 = 30\,e^{-0.05/4}\,N(-0.88) - 31N(-0.93) = 5.73 - 5.58 = 0.15.$$

The hedge ratio corresponding to a short position in the put option equals approximately -0.18 (since $N(-0.93) \approx 0.18$). Therefore, to hedge the exposure an investor needs to short 0.18 shares of stock for one put option. The proceeds from the option and share-selling transactions, which amount to \$5.73, should be invested in risk-free bonds.

Notice that the elasticity of the put option is several times larger than the elasticity of the call option. In particular, if the stock price rises immediately from \$31 to \$31.2 then the price of the put option will drop to less than 12 cents.

5.5 The Black-Scholes PDE (MATH3975)

We introduce here an alternative method for valuing European claims within the framework of the Black-Scholes model. Proposition 5.5.1 can be used to value an arbitrary path-independent contingent claim of European style.

Proposition 5.5.1. *Let $g : \mathbb{R} \to \mathbb{R}$ be a function such that the random variable $X = g(S_T)$ is integrable under $\widetilde{\mathbb{P}}$. Then the arbitrage price in the Black-Scholes model $\mathcal{M}$ of the claim X which settles at time T is given by the equality $\pi_t(X) = v(S_t, t)$, where the function $v : \mathbb{R}_+ \times [0, T] \to \mathbb{R}$ solves the Black-Scholes PDE*

$$\frac{\partial v}{\partial t} + \frac{1}{2}\,\sigma^2 s^2 \frac{\partial^2 v}{\partial s^2} + rs\frac{\partial v}{\partial s} - rv = 0, \quad \forall\,(s,t) \in (0,\infty) \times (0,T),$$

with the terminal condition $v(s,T) = g(s)$.

Proof. The derivation of the Black-Scholes PDE hinges on results from Stochastic Analysis (the Feynman-Kac formula and the Itō lemma) and thus it is beyond the scope of this course. ■

It can be checked that arbitrage prices of call and put options satisfy this PDE. If we denote by $c(s,\tau)$ the function $c : \mathbb{R}_+ \times [0,T] \to \mathbb{R}$ such that $C_t = c(S_t, T-t)$ then it is not hard to verify by straightforward calculations that

$$\begin{aligned}
c_s &= N(d_+) = \delta > 0,\\
c_{ss} &= \frac{n(d_+)}{s\sigma\sqrt{\tau}} = \gamma > 0,\\
c_\tau &= \frac{s\sigma}{2\sqrt{\tau}}\,n(d_+) + Kre^{-r\tau}N(d_-) = \theta > 0,\\
c_\sigma &= s\sqrt{\tau}n(d_+) = \lambda > 0,\\
c_r &= \tau Ke^{-r\tau}N(d_-) = \rho > 0,\\
c_K &= -e^{-r\tau}N(d_-) < 0,
\end{aligned}$$

where $d_+ = d_+(s,\tau)$, $d_- = d_-(s,\tau)$ and n stands for the standard Gaussian probability density function.

5.6 Random Walk Approximations

Our final goal is to examine a judicious approximation of the Black-Scholes model by a sequence of CRR models:

- In the first step, we will first examine an approximation of the Wiener process by a sequence of symmetric random walks.
- In the next step, we will use this result in order to show how to approximate the Black-Scholes stock price process by a sequence of the CRR stock price models.
- We will also recognise that the proposed approximation of the stock price leads to the Jarrow-Rudd parametrisation of the CRR model in terms of the short term rate r and the stock price volatility σ.

5.6.1 Approximation of the Wiener Process

We first define the symmetric random walk on integers.

Definition 5.6.1. *A process $Y = (Y_k,\ k = 0, 1, \dots)$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called the **symmetric random walk** starting at zero if $Y_0 = 0$ and $Y_k = \sum_{i=1}^{k} X_i$ where the random variables $X_1, X_2, \dots$ are independent with the following common probability distribution*

$$\mathbb{P}(X_i = 1) = 0.5 = \mathbb{P}(X_i = -1).$$

Next, we consider the scaled random walk Y^h.

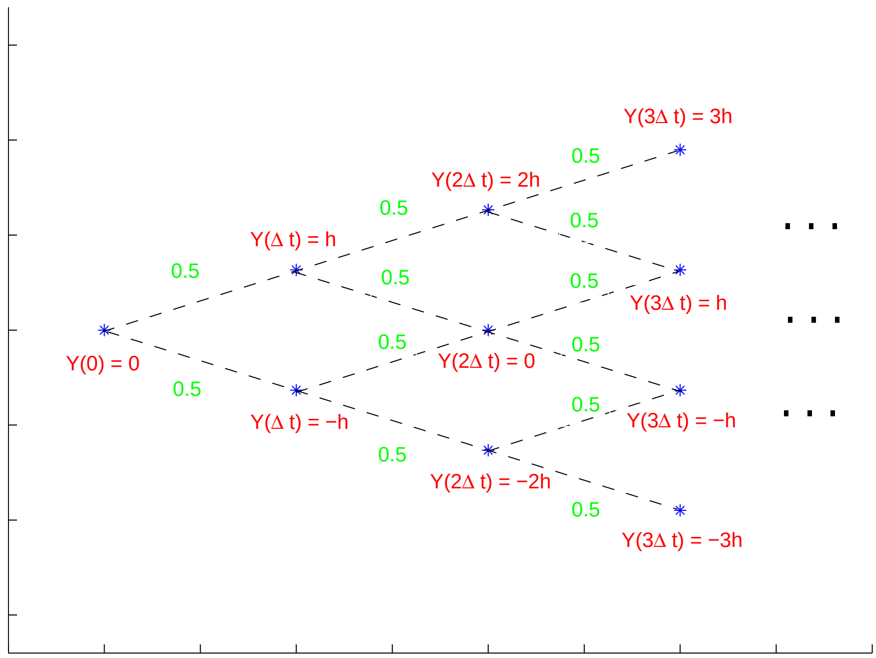


Figure 5.1: Representation of the scaled random walk Y^h

The **scaled random walk** Y^h represented in Figure 5.1 is obtained from the symmetric random walk Y as follows: we fix $h = \sqrt{\Delta t}$ and we set, for every $k = 0, 1, \dots,$

$$Y^h_{k\Delta t} = \sqrt{\Delta t} Y_k = \sum_{i=1}^{k} \sqrt{\Delta t} X_i.$$

Of course, for $h = \sqrt{\Delta t} = 1$ we obtain $Y^h_{k\Delta t} = Y^1_k = Y_k$. The following result is an easy consequence of the classic Central Limit Theorem (CLT) for sequences of independent and identically distributed (i.i.d.) random variables.

Theorem 5.6.1. *Let Y^h_t for $t = 0, \Delta t, \dots,$ be a random walk starting at 0 with increments $\pm h = \pm\sqrt{\Delta t}$. If*

$$\mathbb{P}\left(Y^h_{t+\Delta t} = y + h \,\middle|\, Y^h_t = y\right) = \mathbb{P}\left(Y^h_{t+\Delta t} = y - h \,\middle|\, Y^h_t = y\right) = 0.5$$

then, for any fixed $t \geq 0$, the limit $\lim_{h\to 0} Y^h_t$ exists in the sense of probability distribution. Specifically, $\lim_{h\to 0} Y^h_t \sim W_t$ where W is the Wiener process and the symbol $\sim$ denotes the equality of probability distributions. In other words, $\lim_{h\to 0} Y^h_t \sim N(0,t)$.

Proof. We fix $t > 0$ and we set $k = t/\Delta t$. Hence if $\Delta t \to 0$ then $k \to \infty$. We recall that $h = \sqrt{\Delta t}$ and

$$Y^h_{k\Delta t} = \sum_{i=1}^{k} \sqrt{\Delta t} X_i.$$

Since $\mathbb{E}_{\mathbb{P}}(X_i) = 0$ and $\mathrm{Var}\,(X_i) = \mathbb{E}_{\mathbb{P}}(X_i^2) = 1$, we obtain

$$\mathbb{E}_{\mathbb{P}}(Y^h_{k\Delta t}) = \sqrt{\Delta t} \sum_{i=1}^{k} \mathbb{E}_{\mathbb{P}}(X_i) = 0$$

and

$$\mathrm{Var}\,(Y^h_{k\Delta t}) = \sum_{i=1}^{k} \Delta t \,\mathrm{Var}\,(X_i) = \sum_{i=1}^{k} \Delta t = k\Delta t = t.$$

We conclude using a minor extension of the CLT (see Theorem 6.3.2). □

The sequence of random walks Y^h approximates the Wiener process W when $h = \sqrt{\Delta t} \to 0$ meaning that:

- For any fixed $t \geq 0$, the convergence $\lim_{h\to 0} Y^h_t \sim W_t$ holds, where $\sim$ denotes the equality of probability distributions on $\mathbb{R}$. This follows from Theorem 5.6.1.
- For any fixed n and any dates $0 \leq t_1 < t_2 < \dots < t_n$, we have
  $$\lim_{h\to 0} (Y^h_{t_1}, \dots, Y^h_{t_n}) \sim (W_{t_1}, \dots, W_{t_n})$$
  where $\sim$ denotes the equality of probability distributions on the space $\mathbb{R}^n$.
- The sequence of linear versions of the random walk processes Y^h converge to a continuous time process W in the sense of the weak convergence of stochastic processes on the space of continuous functions. This result is known as Donsker's theorem.

5.6.2 Approximation of the Stock Price

Recall that the JR parameterisation for the CRR binomial model postulates that

$$u = e^{\left(r-\frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}} \text{ and } d = e^{\left(r-\frac{\sigma^2}{2}\right)\Delta t - \sigma\sqrt{\Delta t}},$$

whereas under the CRR convention we set $u = e^{\sigma\sqrt{\Delta t}} = 1/d$. We will show that the JR choice hinges on a particular approximation of the stock price process S. To this end, we recall that for all $0 \leq s < t$

$$\begin{aligned} S_t &= S_0 e^{\left(r-\frac{\sigma^2}{2}\right)t+\sigma W_t} \\ &= S_0 e^{\left(r-\frac{\sigma^2}{2}\right)s+\sigma W_s} e^{\left(r-\frac{\sigma^2}{2}\right)(t-s)+\sigma(W_t-W_s)} \\ &= S_s e^{\left(r-\frac{\sigma^2}{2}\right)(t-s)+\sigma(W_t-W_s)}. \end{aligned} \tag{5.11}$$

Let us set $t - s = \Delta t$ and let us replace the Wiener process W by the random walk Y^h in equation (5). Then

$$W_{t+\Delta t} - W_t \approx Y^h_{t+\Delta t} - Y^h_t = \pm h = \pm\sqrt{\Delta t}.$$

Consequently, we obtain the following approximation

$$S_{t+\Delta t} \approx \begin{cases} S_t e^{\left(r-\frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}} & \text{if the price increases,} \\ S_t e^{\left(r-\frac{\sigma^2}{2}\right)\Delta t - \sigma\sqrt{\Delta t}} & \text{if the price decreases.} \end{cases}$$

More explicitly, for $k = 0, 1, \ldots$

$$S^h_{k\Delta t} = S_0 e^{\left(r-\frac{\sigma^2}{2}\right)k\Delta t + \sigma Y^h_{k\Delta t}}.$$

If we denote $t = k\Delta t$ then

$$S^h_t = S_0 e^{\left(r-\frac{\sigma^2}{2}\right)t+\sigma Y^h_t}.$$

We observe that this approximation of the stock price process leads to the Jarrow-Rudd parameterisation

$$u = e^{\left(r-\frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}} \quad \text{and} \quad d = e^{\left(r-\frac{\sigma^2}{2}\right)\Delta t - \sigma\sqrt{\Delta t}}.$$

We conclude by making the following comments:

- The convergence of the sequence of random walks Y^h to the Wiener process W (Donsker's Theorem) implies that the sequence S^h of CRR stock price models converges to the Black-Scholes stock price model S.
- The convergence of S^h to the stock price process S justifies the claim that the JR parametrisation is more suitable than the CRR method.
- This is especially important when dealing with valuation and hedging of path-dependent and American contingent claims.

Chapter 6

Appendix: Probability Review

We provide here a brief review of basic concepts of probability theory.

6.1 Discrete and Continuous Random Variables

For any random variable X, the **cumulative distribution function** (cdf) of X is defined by the equality

$$F_X(x) = \mathbb{P}(X \leq x), \quad \forall\, x \in \mathbb{R}.$$

Definition 6.1.1. A random variable X is called **discrete** if there exists a countable set of real numbers $x_1, x_2, \ldots$ such that

$$p_i = \mathbb{P}(X = x_i) > 0 \quad \text{for } i = 1, 2, \ldots \text{ and } \quad \sum_{i=1}^{\infty} p_i = 1.$$

Example 6.1.1. Discrete random variables: binomial, geometric, Poisson variables, etc.

The m**th moment** of a discrete random variable X is defined by

$$\mathbb{E}_{\mathbb{P}}(X^m) := \sum_{i=1}^{\infty} x_i^m\, \mathbb{P}(X = x_i) = \sum_{i=1}^{\infty} x_i^m p_i$$

provided that the series converges absolutely, that is, if

$$\mathbb{E}_{\mathbb{P}}\big(|X|^m\big) := \sum_{i=1}^{\infty} |x_i|^m\, \mathbb{P}(X = x_i) = \sum_{i=1}^{\infty} |x_i|^m p_i < \infty.$$

Definition 6.1.2. A random variable X is called **continuous** if there exists a function $f_X : \mathbb{R} \to \mathbb{R}_+$ such that

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\, dy, \quad \forall\, x \in \mathbb{R}.$$

The function f_X is called the **probability density function** (pdf) of X.

We will sometimes write F (f resp.) instead of F_X (f_X resp.) if no confusion may arise. If a random variable X is continuous then

$$\mathbb{P}(X \in B) = \int_B f(x)\,dx$$

for every Borel set B. Recall that intervals (a,b), $(a,b]$, $[a,b]$, etc, are examples of Borel sets. Since

$$F(x) = \int_{-\infty}^{x} f(y)\,dy, \quad \forall\, x \in \mathbb{R},$$

it follows that $f(x) = \frac{d}{dx}F(x)$.

Example 6.1.2. Continuous random variables: uniform, normal (Gaussian), exponential, Gamma, Beta, Cauchy variables, etc.

The mth moment of a continuous random variable X is defined by

$$\mathbb{E}_{\mathbb{P}}(X^m) = \int_{-\infty}^{\infty} x^m\, f_X(x)\,dx,$$

provided that the integral converges absolutely, that is, if

$$\mathbb{E}_{\mathbb{P}}\big(|X|^m\big) = \int_{-\infty}^{\infty} |x|^m f_X(x)\,dx < \infty.$$

Definition 6.1.3. The **expectation** or **mean** of a random variable X is denoted by $\mathbb{E}_{\mathbb{P}}(X)$ and is defined by (if it exists)

$$\mathbb{E}_{\mathbb{P}}(X) = \int_{-\infty}^{\infty} x\,dF_X(x) = \begin{cases} \int_{-\infty}^{\infty} x f_X(x)\,dx, & \text{if } X \text{ is continuous,} \\ \sum_{i=1}^{\infty} x_i\,\mathbb{P}(X = x_i), & \text{if } X \text{ is discrete.} \end{cases} \tag{6.1}$$

Equation (6.1) also defines the expectation of any function of X say $g(X)$. Since $Y = g(X)$ is itself random variable, it follows from equation (6.1) that if $\mathbb{E}_{\mathbb{P}}[g(X)]$ exists then it equals

$$\mathbb{E}_{\mathbb{P}}[g(X)] = \mathbb{E}_{\mathbb{P}}(Y) = \int_{-\infty}^{\infty} y\,dF_Y(y),$$

where F_Y is the cumulative distribution function of the random variable $Y = g(X)$. However, you can easily show that

$$\mathbb{E}_{\mathbb{P}}[g(X)] = \int_{-\infty}^{\infty} g(x)\,dF_X(x),$$

that is,

$$\mathbb{E}_{\mathbb{P}}[g(X)] = \begin{cases} \int_{-\infty}^{\infty} g(x) f_X(x)\,dx, & \text{if } X \text{ is continuous,} \\ \sum_{i=1}^{\infty} g(x_i)\,\mathbb{P}(X = x_i), & \text{if } X \text{ is discrete.} \end{cases} \tag{6.2}$$

Definition 6.1.4. The **variance** of the random variable X is defined by (if it exists)

$$\mathrm{Var}(X) = \mathbb{E}_{\mathbb{P}}[X - \mathbb{E}_{\mathbb{P}}(X)]^2 = \mathbb{E}_{\mathbb{P}}\big(X^2\big) - (\mathbb{E}_{\mathbb{P}}(X))^2. \tag{6.3}$$

6.2 Multivariate Random Variables

Definition 6.2.1. The **joint distribution** $F_{X,Y}$ of two random variables X and Y is defined by

$$F_{X,Y}(x,y) = \mathbb{P}(X \leq x, Y \leq y)$$

for all real numbers x and y.

The **marginal distributions** of X and Y can be obtained in the following way

$$F_X(x) = \lim_{y\to\infty} F_{X,Y}(x,y), \quad F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x,y).$$

Definition 6.2.2. The random variables X and Y are said to be **independent** if

$$F_{X,Y}(x,y) = F_X(x)F_Y(y)$$

for all real x and y.

It is also not hard to show that X and Y are independent if and only if

$$\mathbb{E}_{\mathbb{P}}(g(X)h(Y)) = \mathbb{E}_{\mathbb{P}}(g(X))\,\mathbb{E}_{\mathbb{P}}(h(Y))$$

for all functions g and h for which the expectations exists.

Definition 6.2.3. Two random variables X and Y are said to be **uncorrelated** if their **covariance**, which is defined by

$$\text{Cov}(X,Y) = \mathbb{E}_{\mathbb{P}}[(X - \mathbb{E}_{\mathbb{P}}(X))(Y - \mathbb{E}_{\mathbb{P}}(Y))] = \mathbb{E}_{\mathbb{P}}(XY) - \mathbb{E}_{\mathbb{P}}(X)\,\mathbb{E}_{\mathbb{P}}(Y),$$

is zero.

It is easy to see that independent random variables are uncorrelated, but the converse is not true, as the following example shows.

Example 6.2.1. Let

$$\mathbb{P}(X = \pm 1) = 1/4, \quad \mathbb{P}(X = \pm 2) = 1/4,$$

and $Y = X^2$. Here $\text{Cov}(X,Y) = 0$ by symmetry, but there is a functional dependence of Y on X.

Example 6.2.2. Let U and V have the same probability distribution and let $X = U + V$ and $Y = U - V$. Show that $\text{Cov}(X,Y) = 0$. Are they independent?

We say that two random variables X and Y with finite variances are **correlated** if the **correlation coefficient** if

$$\rho = \rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}$$

exists and is non-zero. It is known that $-1 \leq \rho \leq 1$. Note that X and Y are uncorrelated whenever $\rho(X,Y) = 0$.

Note that the correlation coefficient $\rho(X,Y)$ is not a general measure of dependence between X and Y. Indeed, if X and Y are uncorrelated then they may be either dependent or independent (although if X and Y are independent then they are always uncorrelated). The correlation coefficient can be seen as a measure of the linear dependency of X and Y in the context of **linear regression**.

The random variables X and Y are **jointly continuous** if there exists a function $f_{X,Y}(x,y)$, called the **joint probability density function** such that

$$\mathbb{P}(X \in [a,b], Y \in [c,d]) = \int_a^b \int_c^d f_{X,Y}(x,y)\, dydx$$

for all real numbers $a < b$ and $c < d$.

Definition 6.2.4. The **joint distribution** $F_{X_1,\dots,X_n}$ of a finite collection of random variables $X_1, X_2, \dots, X_n$ is defined by

$$F_{X_1,\dots,X_n}(x_1, x_2, \dots, x_n) = \mathbb{P}(X_1 \le x_1, X_2 \le x_2, \dots, X_n \le x_n)$$

for all real numbers $x_1, x_2, \dots, x_n$.

The **marginal distribution** of X_i can be obtained in the following way

$$F_{X_i}(x_i) = \mathbb{P}(X_i \le x_i) = \lim_{x_j \to \infty,\, j \ne i} F_{X_1,\dots,X_n}(x_1, x_2, \dots, x_n).$$

Definition 6.2.5. We say that the n random variables $X_1, X_2, \dots, X_n$ are **independent** if for all real numbers $x_1, x_2, \dots, x_n$ the following equality holds

$$\mathbb{P}(X_1 \le x_1, X_2 \le x_2, \dots X_n \le x_n) = \mathbb{P}(X_1 \le x_1)\, \mathbb{P}(X_2 \le x_2) \cdots \mathbb{P}(X_n \le x_n)$$

or, equivalently,

$$F_{X_1,\dots,X_n}(x_1, x_2, \dots, x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \cdots F_{X_n}(x_n).$$

6.3 Limit Theorems for Independent Sequences

Some of the most important results in probability are limit theorems. We recall here two important limit theorems: Law of Large Numbers (LLN) and Central Limit Theorem (CLT).

Theorem 6.3.1. [Law of Large Numbers] *If $X_1, X_2, \dots$ are independent and identically distributed random variables with mean μ then, with probability one,*

$$\frac{X_1 + \cdots + X_n}{n} \to \mu \quad \textit{as } n \to \infty.$$

Theorem 6.3.2. [Central Limit Theorem] *If $X_1, X_2, \dots$ are independent and identically distributed random variables with mean μ and variance $\sigma^2 > 0$ then for all real x*

$$\lim_{n \to \infty} \mathbb{P}\left\{ \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le x \right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\, e^{-u^2/2}\, du.$$

6.4 Conditional Distributions and Expectations

Definition 6.4.1. For two random variables X_1 and X_2 and an arbitrary set A such that $\mathbb{P}(X_2 \in A) \neq 0$, we define the **conditional probability**

$$\mathbb{P}(X_1 \in A_1 \,|\, X_2 \in A_2) = \frac{\mathbb{P}(X_1 \in A_1, X_2 \in A_2)}{\mathbb{P}(X_2 \in A_2)}$$

and the **conditional expectation**

$$\mathbb{E}_{\mathbb{P}}(X_1 \,|\, X_2 \in A) = \frac{\mathbb{E}_{\mathbb{P}}(X_1 \mathbb{1}_{X_2 \in A})}{\mathbb{P}(X_2 \in A)}$$

where $\mathbb{1}_{\{X_2 \in A\}} : \Omega \to \{0,1\}$ is the **indicator function** of $\{X_2 \in A\}$, that is,

$$\mathbb{1}_{\{X_2 \in A\}}(\omega) = \begin{cases} 1, & \text{if } X_2(\omega) \in A, \\ 0, & \text{otherwise.} \end{cases}$$

The following result is useful

$$\mathbb{P}(X_1 \in A, X_3 \in C \,|\, X_2 \in B) = \mathbb{P}(X_1 \in A \,|\, X_2 \in B, X_3 \in C)\, \mathbb{P}(X_3 \in C \,|\, X_2 \in B).$$

Discrete case. Assume that X and Y are discrete random variables

$$p_i = \mathbb{P}(X = x_i) > 0, \quad i = 1, 2, \ldots \quad \text{and} \quad \sum_{i=1}^{\infty} p_i = 1,$$

$$\widehat{p}_j = \mathbb{P}(Y = y_j) > 0, \quad j = 1, 2, \ldots \quad \text{and} \quad \sum_{j=1}^{\infty} \widehat{p}_j = 1.$$

Then

$$p_{X|Y}(x_i \,|\, y_j) = \mathbb{P}(X = x_i \,|\, Y = y_j) = \frac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(Y = y_j)}$$

and

$$\mathbb{E}_{\mathbb{P}}(X \,|\, Y = y_j) = \sum_{i=1}^{\infty} x_i \frac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(Y = y_j)} = \sum_{i=1}^{\infty} x_i \,\mathbb{P}(X = x_i \,|\, Y = y_j).$$

It is easy to check that

$$\mathbb{P}(X = x_i) = \sum_{j=1}^{\infty} \mathbb{P}(X = x_i \,|\, Y = y_j)\, \mathbb{P}(Y = y_j)$$

and

$$\mathbb{E}_{\mathbb{P}}(X) = \sum_{j=1}^{\infty} \mathbb{E}_{\mathbb{P}}(X \,|\, Y = y_j)\, \mathbb{P}(Y = y_j).$$

Definition 6.4.2. The **conditional cdf** $F_{X|Y}(\cdot \,|\, y_j)$ of X given Y is defined for all y_j such that $\mathbb{P}(Y = y_j) > 0$ by

$$F_{X|Y}(x \,|\, y_j) = \mathbb{P}(X \leq x \,|\, Y = y_j) = \sum_{x_i \leq x} p_{X|Y}(x_i \,|\, y_j)$$

Note that the conditional expectation $\mathbb{E}_{\mathbb{P}}(X \,|\, Y = y_j)$ is the mean of the conditional distribution $F_{X|Y}(\cdot \,|\, y_j)$.

Continuous case. Assume that X and Y have a joint pdf $f_{X,Y}(x,y)$.

Definition 6.4.3. The **conditional pdf** of Y given X is defined for all x such that $f_X(x) > 0$ and equals

$$f_{Y|X}(y\,|\,x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

and the **conditional cdf** of Y given X equals

$$F_{Y|X}(y\,|\,x) = \mathbb{P}(Y \leq y\,|\,X = x) = \int_{-\infty}^{y} \frac{f_{X,Y}(x,u)}{f_X(x)}\,du.$$

Definition 6.4.4. The **conditional expectation** of Y given X is defined for all x such that $f_X(x) > 0$ by

$$\mathbb{E}_{\mathbb{P}}(Y\,|\,X = x) = \int_{-\infty}^{\infty} y\,dF_{Y|X}(y\,|\,x) = \int_{-\infty}^{\infty} y\frac{f_{X,Y}(x,y)}{f_X(x)}\,dy.$$

An important property of conditional expectation is that

$$\mathbb{E}_{\mathbb{P}}(Y) = \mathbb{E}_{\mathbb{P}}(\mathbb{E}_{\mathbb{P}}(Y\,|\,X)) = \int_{-\infty}^{\infty} \mathbb{E}_{\mathbb{P}}(Y\,|\,X = x) f_X(x)\,dx.$$

Hence the expectation of Y can be determined by first conditioning on X (to give $\mathbb{E}_{\mathbb{P}}(Y|X)$) and then integrating with respect to the pdf of X.

6.5 Exponential Distribution

The following functional equation occurs often in probability theory

$$f(s+t) = f(s)f(t) \quad \text{for all } s,t \geq 0. \tag{6.4}$$

It is known that the only (measurable) solution to equation (6.4) is $f(t) = e^{\alpha t}$ for some constant α. Let us recall the definition of a **memoryless** random variable.

Definition 6.5.1. A non-negative random variable X is **memoryless** if

$$\mathbb{P}(X > s+t\,|\,X > t) = \mathbb{P}(X > s) \quad \text{for all } s,t \geq 0. \tag{6.5}$$

We will now use equation (6.4) to prove that the exponential distribution is the unique memoryless distribution. Suppose that X is memoryless and let $\bar{F}(t) = \mathbb{P}(X > t)$ for $t \geq 0$. From equation (6.5), we have

$$\frac{\mathbb{P}(X > s+t)}{\mathbb{P}(X > t)} = \mathbb{P}(X > s).$$

or $\bar{F}(s+t) = \bar{F}(s)\bar{F}(t)$ which in turn implies $\bar{F}(t) = e^{-\lambda t}$ for some constant $\lambda > 0$. This is the tail of the distribution of an exponential random variable. It is also clear that the exponential distribution is memoryless.

Table of Distributions

	probability/density	domain	mean	variance
Bernoulli	$\mathbb{P}(X=1)=p$ $\mathbb{P}(X=0)=1-p$	$\{0,1\}$	p	$p(1-p)$
Uniform (discrete)	n^{-1}	$\{1,2,\ldots,n\}$	$\frac{1}{2}(n+1)$	$\frac{1}{12}(n^2-1)$
Binomial	$\binom{n}{k}p^k(1-p)^{n-k}$	$\{0,1,\ldots,n\}$	np	$np(1-p)$
Geometric	$p(1-p)^{k-1}$	$k=1,2,\ldots$	p^{-1}	$(1-p)p^{-2}$
Poisson	$\frac{\lambda^k}{k!}\,e^{-\lambda}$	$k=0,1,\ldots$	λ	λ
Uniform (continuous)	$(b-a)^{-1}$	$[a,b]$	$\frac{1}{2}(a+b)$	$\frac{1}{12}(b-a)^2$
Exponential	$\lambda e^{-\lambda x}$	$[0,\infty)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normal (Gaussian) $N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}}\,\exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big)$	$\mathbb{R}$	μ	σ^2
Standard Normal $N(0,1)$	$\frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{x^2}{2}\Big)$	$\mathbb{R}$	0	1
Gamma $\Gamma(n,\lambda)$	$\frac{1}{\Gamma(n)}\,\lambda^n x^{n-1}e^{-\lambda x}$	$[0,\infty)$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$
Beta $\beta(a,b)$	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\,x^{a-1}(1-x)^{b-1}$	$[0,1]$	$\frac{a}{a+b}$	$\frac{ab(a+b)^2}{a+b+1}$
Cauchy	$\frac{1}{\pi(1+x^2)}$	$\mathbb{R}$		

Bibliography

[1] R. E. Bellman: On the theory of dynamic programming. *Proceedings of the National Academy of Sciences* 38 (1952), 716–719.

[2] F. Black and M. Scholes: The pricing of options and corporate liabilities. *Journal of Political Economy* 81 (1973), 637–654.

[3] J. C. Cox, S. A. Ross and M. Rubinstein: Option pricing: a simplified approach. *Journal of Financial Economics* 7 (1979), 229–263.

[4] M.D. Donsker: An invariance principle for certain probability limit theorems. *Memoirs of the American Mathematical Society* 25 (1951), 1–12.

[5] J. M. Harrison and S. R. Pliska: Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications* 11 (1981), 215–260.

[6] R. J. Elliott and P. E. Kopp: *Mathematics of Financial Markets.* Springer, 1999.

[7] K. Itō: Stochastic integral. *Proceedings of the Imperial Academy (Tokyo)* 20 (1944), 519–524.

[8] K. Itō: On stochastic integral equations. *Proceedings of the Japan Academy* 22 (1946), 32–35.

[9] R. A. Jarrow and A. Rudd: Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics* 10 (1982), 347–369.

[10] Y. Kifer: Game options. *Finance and Stochastics* 4 (2000), 443–463.

[11] H. Kuo: *Introduction to Stochastic Integration.* Springer, 2006.

[12] M. Musiela and M. Rutkowski: *Martingale Methods in Financial Modelling.* 2nd edition, Springer, 2005.

[13] R. J. Rendleman and B. J. Bartter: Two-state option pricing. *Journal of Finance* 34 (1979), 1093–1110.

[14] R. T. Rockafellar: *Convex Analysis.* Princeton University Press, 1970.

[15] N. Wiener: Differential space. *Journal of Mathematics and Physics* 58 (1923), 131–174.