# *Financial Mathematics*

Course Organizer: Dr. Chak Hei (Hugo) Lo (C.Lo@ed.ac.uk)

version 1.0

10/01/2021

# General notes

These lecture notes are built from Dr. Gonçalo Dos Reis, who taught this course for the past few years. They are updated by Dr. Chak Hei(Hugo) Lo starting from year 2019-20.

### Support

Writing these lectures notes has been a very nice exercise in restructuring this course. I am specially grateful to Dr. Jose Gregorio Rodriguez Villarreal for all his help in starting the LaTeX files and building the structure for Chapter 3 and 4.

I would also like to thank several other people that contributed in a non-trivial way to the improvement of these notes, either by giving me feedback or taking matters hands-on and refreshing/improving some of the material.

- Mark Alexander (S1 FiM 2015-16 student)

### Sources

These lectures notes were written during the school year 2015/16. The material from which they are sourced is described below. Some of it has been heavily edited, some other remains close to the original.

- Ch 1 & 2: Based on the lectures slides that existed for the course and that were written by Sotirios Sabanis and Chaman Kumar; these slides are partly sourced from [Hul06]. The first section of Chapter 2 was written by Tibor Antal.
- Ch 3: this chapter is based on the first chapters of [Bjo09]. This book is available in electronic format in the Library.
- Ch 4: this chapter mixes content from [Bjo09] and the lecture notes written by David Siska entitled "Brief introduction to Stochastic analysis" and lectures slides that existed for the course and that were written by Sotirios Sabanis and Chaman Kumar.

### References

The main book reference we will use is the book [Bjo09], an eletronic version of it is available in the *Noreen and Kenneth Murray Library* in King's Buildings.

There are many other complementing texts to the course and, in reference to this course's materials, we list them below.

- [Hul06]. ▷ Relates to all chapters of this course.
- [BR98]. ▷ Chapter 1
- [Gar13]. ▷ Chapter 1, 4
- [Ros08]. ▷ Chapter 2
- [Shr04a]. ▷ Chapter 3
- [Shr04b]. ▷ Chapter 4

### Tentative plan to running the FiM course 2019/20 S1

Each week there are two lectures. Every other week there is a tutorial.

W1) Overview of course. Start Chapter 1: interest rates and PV & FV of money; Bonds and PV of bonds;
Workshop 01; Assignment 01

W2) Stocks & market assumptions; Forward contract; Arbitrage; introduction to options.

W3) Review of options; Put-Call parity and pricing principle; overview of hedging & speculation strategies. End of Chapter 1.
Workshop 02; Assignment 02

W4) Intro to Chapter 3: The one-period model; strategy, wealth process, arbitrage and no-arbitrage; Risk-neutral measure.
Contingent claims; replication and market completeness; Example.

W5) Pricing methodology in 1-period market; 1) Risk-neutral and 2) risk-free portfolio. Example. Multi-period model and pricing; Algebraic pricing formulas;
Workshop 03; Assignment 03

W6) American options;
(Review of probability, essential contents of Chapter 2.)

W7) Volatility and calibration; Continuous time limit of multi-period model; End of Chapter 3;
Beginning of Chapter 4: Introduction to Brownian motion (?).
Workshop 04; Assignment 04

W8) More on Brownian motion, simple processes and short introduction to stochastic integration. Ito Isometry.

W9) Ito's formula. Examples. The pricing theorem in continuous time.
Workshop 05; Assignment 05;

W10) More on stochastic analysis. End of chapter Chapter 4;

W11) Review of material
Workshop 06;

# Contents

# Chapter 1

# The Market, Value of Money, Securities and Strategies

## 1.1 Basic Securities

The most basic of financial instruments is the **equity** (also known as **stock** or **share**). The owner of a stock is a **stockholder** or a **shareholder**; she gets both the right to participate in the control of the company and to receive **dividends**.

Dividends are lump sum payments, paid out every quarter or every six months, to the holder of the stock. The amount of the dividend varies from year to year depending on the profitability of the company.

**Commodities** are usually raw products such as precious metals, oil, food products etc. The prices of these products are unpredictable but often show seasonal effects. Commodities are usually traded by people who have no need for the raw material themselves. For example, they may just be speculating on the direction of gold without wanting to stockpile it or make jewelery.

### 1.1.1 Interest Rates and the Time Value of Money

One pound (£1) today is worth more than £1 in a year's time! The simple reason for this is that money can be invested and thus "produce" more money. For example, an investor might deposit their money in a bank and receive **interest** payments, or they might decide to help a family member by investing in their business and agreeing to share in any returns. However, investing money involves different levels of risk depending on the type of investment. Investing money in a bank though is considered to be a *risk-free* investment.

**Simple interest** is the interest received directly on the amount that has been initially invested (*capital*).

**Example 1.1.1.** *Suppose £1 is invested in a bank at a discrete interest rate of* $r = 10\%$ *paid once per annum. Then, at the end of one year the investor will have a total of*

$$\pounds 1 \times (1 + r) = \pounds 1 \times 1.10 = \textit{\pounds 1.10}$$

*The question can also be posed in the opposite direction: if in one year one has* $\pounds 1.10$ *in the bank, and the interest rate is* $r = 5\%$*, how much money was deposited originally?*

But what if we want to carry out interest rate calculations over more than one period? In the example above, we might want to know how much the investor would have if they left their deposit in the account for many years, taking account of the interest that has been accrued in each one. Of course, we could perform the same multiplication for each year in turn, but this quickly becomes tedious and in fact there is a much more straightforward way.

**Compound interest** is the answer. It is interest received on the amount that has been initially invested and on interest as well.

**Example 1.1.2.** *Suppose £1 is invested in a bank at a discrete interest rate of $r = 10\%$ paid once per annum. Then, at the end of the second year the investor will have*

$$1 \times (1+r) \times (1+r) = (1+r)^2, \qquad \text{that is £}1.21.$$

It is easy to see that after $n$ years, the investor will have $(1+r)^n$ pounds (**discretely compounded**).

**Definition 1.1.3** (Nominal and effective interest rate)**.** *The interest rate associated with the basic time unit (usually a year) is called the* ***effective interest rate.***[1]

*An interest rate is called* ***nominal*** *if the frequency of compounding (e.g. a month, a week) is not equal to the basic time unit. The nominal interest rate is the periodic interest rate multiplied by the number of periods per basic time unit (which is read from the effective interest rate).*

Let an investor receive $m$ interest payments during a year (say every period of $\widehat{m}$-months, where $m \times \widehat{m} = 12$ months $= 1$ year) at a nominal interest rate of $r$, then the **effective interest rate** for each sub-period of $\widehat{m}$-months is $r/m$. And at the end of $m$-periods (at the end of the year) the capital will have grown to

$$\left(1 + \frac{r}{m}\right)^m. \tag{1.1.1}$$

Note that $r/m$ is **not** the **effective interest rate** over 1 year! See the next example.

**Example 1.1.4** (Nominal and effective rate)**.** *For example, a nominal interest rate of $6\%$ compounded monthly is equivalent to an effective interest rate of $6.17\%$ over one year.*

*$6\%$ compounded monthly is credited as $6\%/12 = 0.005$ every month. After one year, the initial capital is increased by a factor of $(1+0.005)^{12} \approx 1.0617$. In other words, $6\%$ compounded monthly is the same as $6.17\%$ yearly, since*

$$(1 + \frac{0.06}{12})^{12} = (1+0.005)^{12} = (1+0.0617).$$

The effective interest rate is always calculated as if compounded annually. The calculation is carried out in the following way, where $r$ is the effective rate, $i$ the nominal rate (as a decimal, e.g. 12% = 0.12), and $m$ the number of compounding periods per year (for example, $m = 12$ for monthly compounding, with $\widehat{m} = 1$):

$$r = \left(1 + \frac{i}{m}\right)^m - 1.$$

**Note:** *Nominal interest rates* cannot be directly compared unless their compounding periods are the same. *Effective interest rates* correct for this by "converting" nominal rates into annual ones that can be used in a straightforward comparison.

[1]The compounding scheme for an effective interest rate is always discrete compounding.

### Continuous Compounding

If the interest payments happen at increasingly frequent intervals (i.e. $m \to \infty$), then expression (1.1.1) becomes $e^r$. We call this **continuous compounding**.

[Warning Check workshop (Week 01) exercises ]

*Proof.* Using the binomial theorem

$$
\begin{aligned}
\left(1+\frac{r}{m}\right)^m &= \sum_{i=0}^{m} \binom{m}{i} \left(\frac{r}{m}\right)^i \\
&= 1 + \frac{m!}{(m-1)!}\frac{r}{m} + \frac{m!}{2!(m-2)!}\left(\frac{r}{m}\right)^2 + \frac{m!}{3!(m-3)!}\left(\frac{r}{m}\right)^3 + \ldots + \left(\frac{r}{m}\right)^m \\
&= 1 + r + \frac{m-1}{m}\frac{r^2}{2!} + \frac{(m-1)(m-2)}{m^2}\frac{r^3}{3!} + \ldots + \left(\frac{r}{m}\right)^m.
\end{aligned}
$$

Taking limits at both sides results in

$$
\lim_{m\to\infty}\left(1+\frac{r}{m}\right)^m = 1 + r + \frac{r^2}{2!} + \frac{r^3}{3!} + \cdots = e^r.
$$

$\square$

Hence, the value of £$b$ compounded over $T$-years at some interest rate $r$ is given by £$b \times e^{rT}$.

An advantage of continuous compounding is that it makes life simpler in terms of calculations.

*Continuously compounded interest rates* are used in these lecture notes except where stated otherwise.

**Example 1.1.5.** *£500 over 10 year compounding at $r = 5\%$:*

*annual comp.:* £$500 \times (1+0.05)^{10}$; $\leqslant^{\textit{check!}}$ *continuous comp.:* £$500 \times e^{0.05\times 10}$;

### Present Value

Let us assume now that an amount $P(0)$ is deposited in the bank. After $T$ years the investor will have

$$P(T) = P(0)e^{rT},$$

where $r$ is the *interest rate*. Conversely, if it is known that the investor will receive the amount $P(T)$ at time $T$ in the future, its value at an earlier time $t$ is given by the following expression

$$P(t) = P(T)e^{-r(T-t)}, \qquad t \in [0, T]$$

Thus, cash flows in the future can be related to their **present value** (PV) by multiplying by $e^{-r(T-t)}$ where $T - t$ is the length of time between current time $t$ and the maturity time $T$.

**Example 1.1.6** (PV of a future cashflow)**.** *Suppose that the interest rate is 5%. Then the present value of* £1,000,000 *to be received in two years is*

$$\pounds 1{,}000{,}000 \times e^{-0.05\times 2} = \pounds 904{,}837.$$

*Suppose that one receives* $\pounds 250,000$ *at the end of the first year and* $750,000$ *at the end of the second. The PV of this cash flow is then*

$$PV = \pounds 250{,}000 \times e^{-0.05\times 1} + \pounds 750{,}000 \times e^{-0.05\times 2} = \pounds 951{,}229.$$

*(Try the exercise of computing the Future Value (FV) of the bond.)*

This operation is also called **discounting**, as in "we discount a future cash flow to its present value".

### 1.1.2 Fixed-Income Securities

When a corporation (or a government) wishes to borrow money from the public on a long-term basis, it usually does so by issuing or selling debt securities that are generally called **bonds**. Thus, governments and companies issue bonds as a form of borrowing. The less creditworthy the issuer of the bond, the higher the interest that they will have to offer in order to attract a buyer.

A bond is normally an interest-only loan, meaning that the issuer will pay the interest every period, but none of the principal will be repaid until the end of the loan. The **face value/par value/nominal** of the bond is the lump sum that is repaid at the end of the term. The date on which the face/par value of a bond is paid is called the **maturity** of the bond.

The interest payments made on a bond are called **coupons**. The **annual coupon** $C$ divided by the face value $F$ of a bond is called the **coupon rate** ($:= C/F$, annual coupon divided by face value). Interest rates are assumed to be constant although that happens rarely in reality.

**Example 1.1.7** (The PV of a bond)**.** *Say BT issues a bond with 10 years to* ***maturity*** *and an* ***annual coupon*** *of £80 (i.e. BT pays £80 per year for the next 10 years in coupon interest). Moreover, in 10 years, BT will pay £1000 (****face/par****) to the owner of the bond. The* Bank interest rate *is assumed to be 8% (annual compounding).*

*What would this bond sell for* now? *What is the PV of the bond?*

*The BT bond's cash flows have an annuity component (the coupons) and a lump sum (the face value paid at maturity). To estimate the market value of the bond we need to compute the present value of these two cash flows, the coupons and the face value.*

$$\textit{PV of the face value} = \frac{\pounds 1000}{1.08^{10}} = \pounds 463.19$$
$$\textit{P.V. of coupon payments} = \frac{\pounds 80}{1.08^{10}} + \frac{\pounds 80}{1.08^{9}} + \cdots + \frac{\pounds 80}{1.08} = \pounds 80\frac{1-\frac{1}{1.08^{10}}}{0.08} = \pounds 536.81.$$

*Note that in the last line we used:*

$$a_0\lambda + a_0\lambda^2 + \ldots + a_0\lambda^n = a_0\lambda(1+\lambda+\ldots+\lambda^{n-1}) = a_0\lambda\frac{1-\lambda^n}{1-\lambda} \qquad \textit{(See Workshop sheet)}$$

*with* $\lambda = (1+r)^{-1} = 1/1.08$ *and* $a_0 = 80$.

*Adding the two PVs together, the bond's PV or market value follows:*

$$PV_{Bond} = PV_{Face} + PV_{Coupon\ cash\ Flow} = \pounds 463.19 + 536.81 = \pounds 1000.$$

This bond sells for exactly its face value. This is not a coincidence. This bond (an interest-only loan) pays exactly 8% interest only when it sells for £1000.

Had the market applied another rate then the result would have been different. Consider the following scenario:

**Example 1.1.8** (The PV of a bond II)**.** *The interest rate in the market has risen to 10% (due to high inflation figures). Then, the present value (at* $t$ *= 0 year and* $T$ *= 10 years) of the face value at 10% is*

$$\text{P.V. of the face value} = \frac{\pounds 1000}{1.1^{10}} = \pounds 385.54$$
$$\text{P.V. of coupon payments} = \pounds 80\frac{1 - \frac{1}{1.1^{10}}}{0.1} = \pounds 491.57$$

*Thus, the PV of the bond or the total bond value = £877.01.*

**Proposition 1.1.9.** *The present value of a bond* ***B*** *with a face value* ***F****, a coupon of* ***C*** *paid per period, with* ***T*** *periods to maturity and an interest rate* ***r*** *per period, is given by*

$$PV_B = C\frac{1}{r}\Big(1 - \frac{1}{(1+r)^T}\Big) + \frac{F}{(1+r)^T}. \tag{1.1.2}$$

*Proof.* This proof is a straightforward adaptation of the previous example. Try to do it yourself! □

A bond that pays no coupons at all is called a **zero coupon bond**. Such bond must be offered at a price that is much lower than its stated value. Note that depositing $B_0$ in a bank account is exactly the same as a zero coupon bond with face value $F$ being the future value of the initial deposit in the bank account, i.e. $F = B_0(1+r)$ (In this case for a year. For general $T$ years the calculation is the same. Write this down and check by yourself).

**Exercise 1.1.10.** *Rewrite the previous result under continuous compounding.*

### 1.1.3 An Introduction to "Arbitrage" Using Bonds

So far we have discussed the cashflows that define a bond (coupon payments and face value) and also computed the present value corresponding to those cash flows (see (1.1.2)). We might ask ourselves then, if someone offers us a bond priced at its present value, does this represent a good deal? Might we expect to find the same bond cheaper elsewhere?

In fact, it is precisely the present value of a bond that represents the 'fair' price for which it is bought and sold. But why should there be only one 'correct' value? To understand this, we must look at the concept of **arbitrage**. Arbitrage means, at a philosophical level, the possibility to make money without any risk. It ensures that there will be one and only one fair price for a given bond, assuming some requirements that we will discuss later.

Take Example 1.1.7. Recall that $\text{PV}_{\text{Bond}} = \pounds 1{,}000$. Suppose that Alice is offering to buy the bond at this price. Now suppose Bob is selling the exact same bond at a lower price in another market, call this price £800. What would happen? Is there a way to make a profit from this situation?

The answer is *yes*! For example, one could then invest in the following way:

1) sell the bond to Alice at a price of $\text{PV}_{\text{Bond}} = \pounds 1{,}000$;

2) use the proceeds to immediately buy the bond from Bob at £800;

3) repeat, until rich?

In doing so we see that an instant profit of £200 is made each time. Notice that (in this simple model) no risk was taken in order to make the profit. This type of investment strategy is called an *arbitrage strategy*.

So what stops this from being a viable strategy for us to make a profit in the real world? Well, real bond markets tend to have many participants, and each of them is just as motivated by making a profit as the next. If such an arbitrage opportunity did arise, then we would expect it to vanish very quickly as the participants rush to take advantage.

In this example *arbitrage* is realized at $t = 0$, in the next section arbitrage is realized only at the maturity time.

## 1.2 Market Assumptions, Terminology, and Actions

### Basic Setup

Throughout we will assume that we are trading in a market that contains a generic non-risky asset, say a bank account paying some interest rate. We will also assume throughout the rest of this chapter that this account pays continuously compounded interest.

In this market, there are stocks and commodities being exchanged. We denote the price of any given stock or commodity at time $t \in [0, \infty)$ by $S(t)$. For some $T \in (0, +\infty)$, we denote a price's whole 'evolution' or **price process** by $(S(t))_{t\in[0,T]}$ or $\{S(t) : t \in [0, T]\}$ it being a function of randomness and time, i.e. $S : \Omega \times [0, T] \to [0, \infty)$. The price process is a random variable evolving through time (or what is called a *Stochastic Process*) and one writes: $S$ as $S : \Omega \times [0, T] \to [0, \infty)$. The $\Omega$ represents the randomness inherent in the future price of the stock; since at any point in time that is not now or in the past the price of the stock is uncertain (but surely greater or equal to zero).

**Remark 1.2.1** (Stocks paying dividends)**.** *Some stocks pay* ***dividends****. Dividends are usually annual payments where the company distributes some amount of its profits to its stockholders. The dividend is paid per share. Some companies pay no dividends at all.*

*To simplify our computations we will assume that the stocks we work with pay no dividends.*

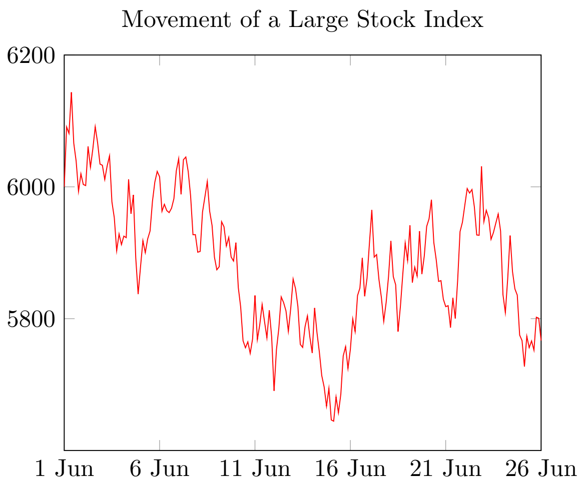



Figure 1.2.1: Example of the price of a stock like financial asset. In this case the Footsie index.

### 1.2.1 Assumptions About the Market

We make a number of assumptions about the market throughout in order to simplify our analysis:

1) There are no **transaction costs**, e.g. no taxes, no bid-ask price, and no brokerage fees are involved in buying or selling. If this assumption holds then we say the market is a *frictionless market* or a *perfectly liquid market*.

2) The same risk-free rate of interest applies for borrowing and lending money (*market efficiency*), and the bank never goes bankrupt.

3) The market participants take advantage of arbitrage opportunities as they occur (effectively causing an *'absence' of arbitrage*, since such opportunities cannot last for long).

4) It is always possible to buy/sell something, in any amount, at its fair price.

### 1.2.2 Market Actions and Terminology

In the previous sections we made use of some key terminology without much explanation. Here is a tentative dictionary.

For some $T \in (0, \infty)$, the price process is denoted by $(S(t))_{t\in[0,T]}$. If the present time is $t = t_0$, then we refer to $S(t_0)$ as the **spot price** (i.e. "the price on the spot where we stand right now"). The time $t_0$ is sometimes referred to as the **spot time**. If one enters a financial contract that expires at time $T$, that time is called the **maturity**, and the time gap between the spot time and the maturity time $(T - t_0)$ is called the **time to maturity**. The contract itself is written in relation to a certain amount of stocks or commodities, and the particular stocks or commodities involved in the contract are referred to as the **underlyings**.

From these basic principles, we can discuss trading stocks and making use of the borrowing and lending facility of our generic bank account. The following keywords are related to trading:

- **Short-selling** involves selling an (unowned or borrowed) asset with the intention of buying it back later. It yields a profit when the price of the asset goes down and a loss when it goes up.
- **Hedging** refers to a strategy which is intended to reduce or minimise risk. For example a futures hedge (see section 1.3) reduces risk by making the outcome more certain.
- **Speculation** is achieved by taking a position in the market where the investor effectively bets (*speculates*) that the price of the underlying asset will go up or down.
- **Arbitrage** involves locking in a profit by simultaneously entering into transactions in two or more markets (usually a trading or portfolio strategy, having zero capital at time $t$, which results in guaranteed profit at time $T$).

## 1.3 Derivatives - Forwards and Futures

**Derivatives** are financial contracts whose pay-off or value depends (*derives*) from the value(s) of some underlying asset(s). The underlying assets can be a vast array of items including equities/stocks, stock indices, currencies, bonds, commodities, precious metals (gold, silver), energy-related (oil, gas), soft commodities (wheat, sugar, cocoa), financial derivatives (options, futures) and non-tradable variables (volatility, temperature).

This rest of this chapter will introduce some basic types of derivative contract, building of course in complexity. We will cover *forward contracts* (along with their cousins, *futures contracts*), as well as *option contracts* (or *options*). There are many other types of contracts, less standard and hence called *exotic options*.

### 1.3.1 Forward Contracts

A **forward contract** is an agreement between two parties to buy or sell an asset $S(T)$, at a specified future time $T$, and at a specified **delivery price** $K$ (the value $K$ is agreed upon at the time one sets the contract). The specified date is known as the **maturity** or **delivery date**, and participants are *obligated* to buy or sell the asset at maturity.

Note that the price at time zero of a forward contract is zero (for $K$ the *fair delivery price*, we'll see more about this is a page); it costs nothing to *enter* the contract. Nonetheless, the obligation must be met at maturity time $T$.

When a forward contract is traded, one of the parties of the contract assumes a **long position** and agrees to buy the *underlying asset* $S(T)$ at time $T$ for the price $K$. The other party assumes a **short position** and agrees to sell $S(T)$ at time $T$ for the price $K$.

This *long* and *short* terminology is to be found throughout this course and the financial world. In a simple sense, we can think of a long position being one that is *owned*, and of a short position as one that is *owed* (the short position delivers the risky asset while the long position delivers the money).

**Example 1.3.1.** *Forward contracts are commonly used* ***to hedge foreign currency risk****. Suppose it is 10 October 2003 and the treasurer of a U.S. corporation knows that the corporation will receive 1 million pounds sterling in three months (on January 10) and wants to hedge against exchange rate moves.*

*The treasurer could then contact a bank, find out that the exchange rate for a three-month forward contract on sterling is 1.60 (i.e. £1 = \$ 1.60) and agree to sell one million pounds. The corporation then has a short forward contract on sterling. It has agreed that on January 10 it will sell 1 million pounds sterling to the bank for \$ 1.6 million. Both sides (U.S. corporation and Bank) have made a binding commitment.*

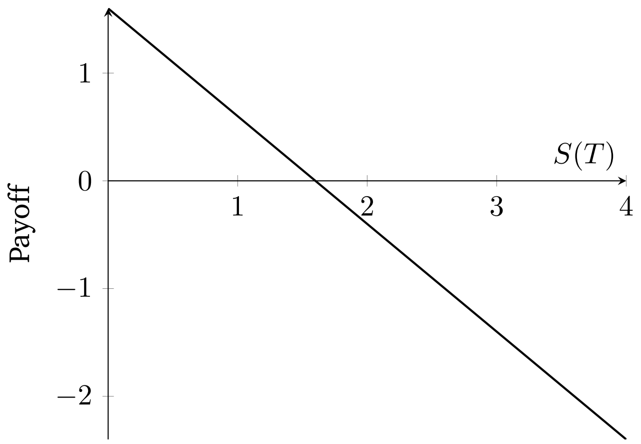



Figure 1.3.1: The payoff from a short position (the position that must sell sterling for dollars) in a forward contract on one unit of an asset (e.g. £1).

### Profit and Loss as a Function of the Underlying Asset Price

Fix a time horizon $T > 0$ and suppose that $(S(t))_{t\in[0,T]}$ represents the price of a traded asset $S$ over the time interval $[0, T]$. Suppose an investor has a long position on a forward contract $F$, which obliges her to buy the security at maturity time $T$ for price $K$. We can describe the investor's profit at time $T$ in terms of $S(T)$ and the delivery price $K$.

Recap: the *long* position pays the delivery price $K$ and the *short* position delivers the (risky) underlying $S(T)$. Computing the payoffs, one has:

▷ Pay-off per unit of asset from *long* position in a forward contract : $S(T) - K$

▷ Pay-off per unit of asset from *short* position in a forward contract : $K - S(T)$

So what is a 'fair' delivery price $K$ for this forward contract?

The price of the underlying asset at time $T$ can (and usually does) differ from the *delivery price* $K$, and so we denote it by £$S(T)$. How much profit the holder of a forward contract makes cannot be known until the value £$S(T)$ is known and that happens only at time $T$.

The question is then: *if there is no arbitrage in the market,* is there any relationship between $K$, $S(t)$, $t$ and $T$?

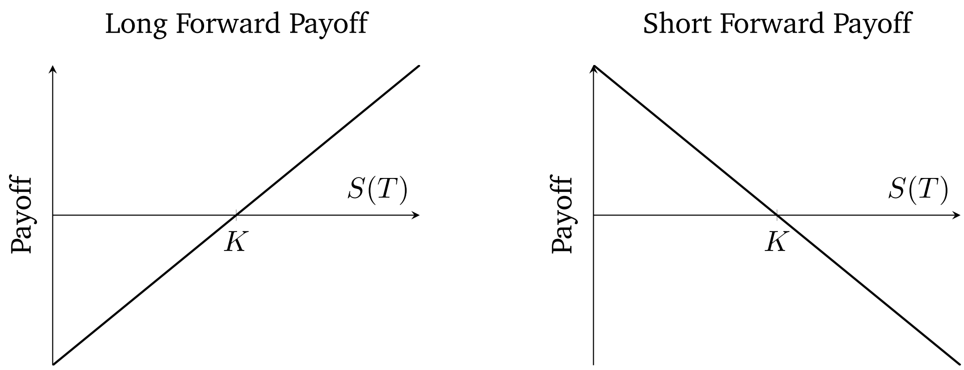



Figure 1.3.2: Payoffs from forward contracts. On the left the Long position and on the right the short position.

When the investor enters into the forward contract, although she doesn't pay anything up front, exposes herself to the uncertainty around the value of the asset at maturity (time $T$). By **short selling** the asset, i.e. selling the asset that the investor doesn't own, an amount $S(t)$ is received. This amount can be invested in a bank account (with rate $r$).

Notice though that the investor's net position at time $t$ is zero (see Table 1.3.1).

| **Holding** | **Value at time** $t$ | **Value at maturity** $T$ |
|---|---|---|
| Forward | $+0$ | $S(T) - K$ |
| Stock | $-S(t)$ | $-S(T)$ |
| Bank | $+S(t)$ | $S(t)e^{r(T-t)}$ |
| Total | 0 | $S(t)e^{r(T-t)} - K$ |

Table 1.3.1: Cash flows in a portfolio of assets and a long forward.

At maturity time $T$, the investor hands over the delivery price $K$ and receives the asset $S(T)$. This cancels the short position in the asset *regardless* of the value of $S(T)$. Thus, the *hedged portfolio* at time $T$ (maturity) has the following value:

$$S(t)e^{r(T-t)} - K.$$

It is expected though that the value of the portfolio at time $T$ is actually equal to zero, since it was equal to zero at time $t$ (no *arbitrage*). Therefore, we conclude that

$$K = S(t)e^{r(T-t)}.$$

We emphasize that $K$ is a function of the present time $t$ and the maturity time $T$, i.e. $K(t,T)$.

**Remark 1.3.2** (Bank Vs a zero coupon bond (ZCB))**.** *In the above Table 1.3.1 the bank can be replaced by a bond market where one would buy a zero coupon bond (ZCB) with Present value* $£S(t)$.

### Arbitrage (and forward prices)

Forward contracts are considered to be some of the simplest forms of derivates. Later, we will examine derivatives whose price depends on some assumed set of mathematical models. Forwards are simple precisely because they do not depend on any such model; but only on the initial price of the underlying asset.

**Proposition 1.3.3** (Fair delivery price of a forward contract)**.** *The arbitrage free delivery price (or fair price) of a forward contract is a map* $K : [0, T] \times [0, \infty) \to [0, \infty)$ *with the property*

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

We can examine this proposition in terms of arbitrage opportunities as we have done before.

**Market Rule 1.3.4.** *Agents in the market try to* buy low to sell high*. (In economics language, this means that agents are rational.)*

Thus, if a contract is *overpriced*, we want to sell it while knowing that we can always buy the exact same contract elsewhere for its lower fair price. If a contract is *underpriced*, we want to buy it and sell the exact same contract for its fair price. Once one computes the net position or *portfolio balance* (profits and losses of the portfolio) of these situations, a guaranteed profit emerges.

*Proof of Proposition 1.3.3.* Consider a forward contract written on an underlying company share $S$. Let $S(t)$ be the share price at time $t$. Now suppose a forward contract on this share is over- or under-priced and consider each case in turn:

**Case 1:** An overpriced forward contract: $K > S(t)e^{r(T-t)}$ (a to high delivery price).

Wishing to take advantage of the situation, an investor issues one forward contract and agrees to sell the underlying asset (assumes a short position) for the price $K$ in $T - t$ months. At the same time, she borrows an amount $S(t)$ in cash from the bank (subject to interest at rate $r$) and buys one share. Note that her net position at time $t$ is zero.

Fast forward to time $T$. The investor now holds a share that is worth $S(T)$ on the open market, and also has a debt of $S(t)e^{r(T-t)}$ (to the bank). However, in accordance with the forward contract she sells her share and receives $K$. Now, since $K > S(t)e^{r(T-t)}$, the investor has made a guaranteed profit, without having taken any risk, and with zero starting capital.

| **Holding** | **Value at time** $t$ | **Value at maturity** $T$ |
|---|---|---|
| Forward | $-0$ | $K - S(T) = -(S(T) - K)$ |
| Stock | $+S(t)$ | $S(T)$ |
| Bank | $-S(t)$ | $-S(t)e^{r(T-t)}$ |
| Total | 0 | $K - S(t)\, e^{r(T-t)} > 0$ |

Table 1.3.2: Cash flows in a portfolio of asset and forward.

Naturally, the investor will want to issue as many of these contracts as possible in order to make a fortune without taking any risks. In practice though, the market will react immediately to take advntage of the same price disparity, pushing down the forward price to its 'fair' value. Thus, the arbitrage opportunity could exist only briefly and would disappear quickly before any substantial profits could be made.

**Case 2:** An underpriced forward contract: $K < S(t)e^{r(T-t)}$ (a to low delivery price).

Seeing the opportunity for profit once more, our investor buys one forward contract (assumes a long position), and by selling one share she receives $S(t)$ in cash. She then invests this amount in a bank (subject to interest at rate $r$). Thus, the investor's net position at time $t$ is 0.

Fast forward to time $T$. The investor receives one share and pays $K$ (in accordance with the forward). She also receives interest from her investment with the bank, which gives cash of value $S(t)e^{r(T-t)}$. Now, since $K < S(t)e^{r(T-t)}$, the investor has made a guaranteed profit, again having no capital requirement at time $t$.

| **Holding** | **Value at time $t$** | **Value at maturity $T$** |
|---|---|---|
| Forward | $+0$ | $S(T) - K$ |
| Stock | $-S(t)$ | $-S(T)$ |
| Cash | $+S(t)$ | $S(t)e^{r(T-t)}$ |
| Total | 0 | $S(t)\,e^{r(T-t)} - K > 0$ |

Table 1.3.3: Cash flows in a portfolio of asset and forward.

□

From the formula for the fair no-arbitrage delivery price, we can easily obtain the following result:

**Proposition 1.3.5** (Convergence to the spot price)**.** *The fair no-arbitrage delivery price of a forward contract converges to the spot price of the underlying asset when maturity nears. In other words, let $K(t,T)$ be the arbitrage free delivery price of a Forward contract at time $t$ with maturity $T$ and underlying $S$, i.e. $K(t,T) = S(t)e^{r(T-t)}$, then*

$$\lim_{t \to T} K(t,T) = S(T).$$

*Proof.* From the formula in Proposition 1.3.3, we have $\lim_{t\to T} K(t,T) = S(T)$ . The result then follows.[2]. □

When the delivery period is reached, the forward price equals–or is very close to–the spot price. In case that *the forward price is above the spot price* during the delivery period, a clear arbitrage opportunity will occur:

1. Short a futures contract;
2. Buy the asset;
3. Make delivery.

This is certain to lead to profit equal to the amount by which the forward price exceeds the spot. A similar argument applies for the case where the spot price is below the futures price during the delivery period.

[2] For those paying a bit more attention, for this proof to work we also need for the stock price to be a continuous map in the time variable.

### 1.3.2 Futures Contracts

A **futures contract**, much like a forward contract, is an agreement between two parties to buy or sell an asset at a certain time in the future for a certain price. The main difference is that futures contracts are normally traded on an exchange, which stipulates standard terms for the contract. The vast majority of futures contracts do not lead to delivery. The reason being that most traders choose to **close out** their position prior to the delivery period specified in the contract.

Closing out a position means entering into the opposite type of trade from the original one. For example, a London investor (from the International Financial Futures and Options Exchange - LIFFE) who bought a July corn futures contract on March 5 can close out the position by selling (i.e. shorting) one July corn futures contract on April 20. The Kent investor who sold (i.e. shorted) a July contract on March 5 can close out the position by buying one July contract on April 20. In each case, the investor's total gain or loss is determined by the change in the futures price between March 5 and April 20.

Forward Contracts

- private contract between two parties (OTC)
- not standardized
- exchange restriction
- normally one specified delivery date
- final settlement at maturity, either through physical delivery or cash settlement
- credit risk involved

Futures Contracts

- traded on an exchange
- standardized
- no exchange restriction (daily price limits)
- range of delivery dates
- usually closed out before maturity period
- protection against default risk via a margin account and the role of the clearing house

Forward and futures contracts are often used for *hedging* as well as for *speculation*. In order to avoid foreign currency risk, these derivatives are used to 'lock in' a guaranteed exchange rate (see previous example). Thus, the investor is no longer exposed to exchange rate fluctuations. However, the investor can also speculate about the future growth of the market and thus can benefit by entering into a forward or futures contract. Speculation though can be a very risky business.

## 1.4 Derivatives - Options Contracts

The basic idea behind *options* contracts is to remove the obligatory aspect of the forward contract, hence their name *options*. The contract is essentially the same, but instead of a binding contract one decides if one wants to execute the contract or not.

There are two basic types of option contracts.

▷ A **call option** gives the holder the right, *not the obligation*, to buy the underlying asset $S$ at a certain future date $T$ for a certain price $K$ (called strike).

▷ A **put option** gives the holder the right, *not the obligation*, to sell the underlying asset $S$ at a certain future date $T$ for a certain price $K$ (called strike).

Options on stocks were first traded on 26th April 1973. It was then that the Chicago Board Options Exchange (CBOE) first created standardised, listed options. Initially there were just call options on 16 stocks. Put options were introduced in 1977. Options are now traded on many exchanges (over 50) throughout the world as well as over the counter by banks and other financial institutions. The underlying assets include stocks, stock indices, foreign currencies, debt instruments, commodities and futures contracts.

The date specified in the contract is known as the **expiration date** or *maturity* $T$. The price is known as the **exercise price** or **strike price** $K$ (in the forward contract this was the delivery price; here the quantity $K$ has another meaning). The amount paid initially for purchasing an option is known as the **premium**. Note that the premium of the forward contract is zero - it costs nothing to enter a forward contract. On the other hand, option contracts cost money–the so-called *premium*.

Within options contracts, there are two subtypes of contract:

▷ **European options** can be exercised only on the expiration date itself.

▷ **American options** can be exercised at any time up to the expiration date.

**Exercise 1.4.1.** *Whose premium should be higher? The one for the European or that for the American option?*

European options are easier to analyze and some properties of the American options can be deduced from those of its European counterpart. In this chapter we will not go into details of the American option, we leave that for Chapter 3.

### 1.4.1 Payoffs of Options Contracts

How do we calculate the *premium* of an option contract? Or an option's *fair value*? Or the *no arbitrage price* as we did for the forward contract? For that, we must first understand the payoffs involved in European call and put options.

#### Example with a Call Option

Recall that a call option is just like a forward contract but where the long position has the right to not execute the contract!

Consider the situation of an investor who **buys** (goes **long**) one European **call** option on BT stock with a strike price $K = £10$ and with an expiration date in three months. Thus, this particular call gives the holder the right to purchase 100 BT shares (one option contract is usually an agreement to buy or sell blocks of 100 shares) on the expiration date for £10 each.

At maturity, there are two possibilities for $S(T)$:

a) $S(T)$ is *less* than the strike $K$:

   If the market price of the BT stock $S(T)$ on the expiration date $T$ is less than £10, then the investor will clearly choose not to exercise the option. He or she can buy the stock in the market for a better price. Therefore, there is no profit to be made from the option and the investor has actually lost money by paying for the *premium*.

b) $S(T)$ is *more* than the strike $K$:

   If the market price of the BT stock $S(T)$ on the expiration date $T$ is more than £10, for example £12, then the investor will exercise the option and buy 100 shares for the price of £10 each and thus will make a profit (£200 = 100 $\times$ (£12 - £10) - *premium*) by selling them immediately.

**Conclusion:** One can easily compute the **payoff of the long European call** from the above. It is given by $\max\{S(T) - K, 0\}$. For any $K \in \mathbb{R}$, we denote by **call function** the map $f : \mathbb{R} \to \mathbb{R}$ with the property $f(x) = \max\{x - K, 0\} = (x - K)^+$.

Following the above, if it is assumed that the *premium* of the call is $£1$ per share, then the profit diagram of the agent's portfolio (he paid $£1$ for the call and receives the call's payoff at maturity time $T$) is given in (left side) Figure 1.4.2.

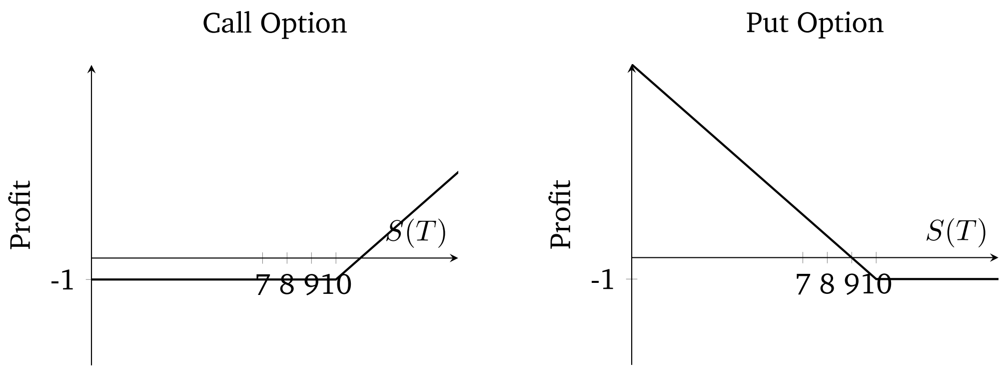



Figure 1.4.1: Payoff from the BT long call and long put option (per share), taking into account the initial price paid.

### Example with a Put Option

Similarly to the above, let us consider now the case where the investor **buys** (**long**) one European **put** option on BT stock with a strike price $K = £10$ and with an expiration date in three months. Thus, the purchaser of the put option has the right to sell 100 BT shares on the expiration date for $£10$ each.

Again, there are two possibilities for $S(T)$ at maturity:

a) $S(T)$ is *less* than the strike $K$:

   If the market price of the BT stock $S(T)$ on the expiration date $T$ is less than $£10$, for example $£8$, then the investor can buy $100$ shares for the price of $£8$ each and sell them, under the terms of the put option, for $£10$ pounds each and thus make a profit (£200 = 100×(£10 - £8) - *premium*).

b) $S(T)$ is *more* than the strike $K$:

   If the market price of the BT stock $S(T)$ on the expiration date $T$ is more than $£10$, then the investor will clearly choose not to exercise the option, and a loss would occur (the put holder has already paid an amount as a premium in order to buy the option).

**Conclusion:** One can easily compute the **payoff of the long European Put** from the above. It is given by $\max\{K - S(T), 0\}$. For any $K \in \mathbb{R}$, we denote by **put function** the map $f : \mathbb{R} \to \mathbb{R}$ with the property $f(x) = \max\{K - x, 0\} = (K - x)^+$.

Following the above, if it is assumed that the *premium* of the put option is $£1$ per share, then the profit diagram of the agent's portfolio (he paid $£1$ for the put and receives the put's payoff at maturity time $T$) is given in (right side) Figure 1.4.2.

### Market Positions

In general, the purchaser of a call option is hoping that the stock price will increase (he only makes profit if $S(T) > K$) whereas the purchaser of a put option is hoping that it will decrease (he only makes profit if $S(T) < K$). Four basic positions are possible by either going long (buying the option) or short (selling the option): 1) long call; 2) long put; 3) short call; and 4) short put option.

Take a call and a put option written on the underlying asset $S$, with maturity $T$, strike $K$ and premium price $P$ then the possible positions are summarized in Table 1.4.1 and the portfolio payoffs in Figures 1.4.2 and 1.4.3.

| Contract | Position | Payoff |
|---|---|---|
| Call | Long | $\max\{S(T) - K, 0\}$ |
| Call | Short | $-\max\{S(T) - K, 0\}$ |
| Put | Long | $\max\{K - S(T), 0\}$ |
| Put | Short | $-\max\{K - S(T), 0\}$ |

Table 1.4.1: Payoffs of the possible positions in European Calls and Puts

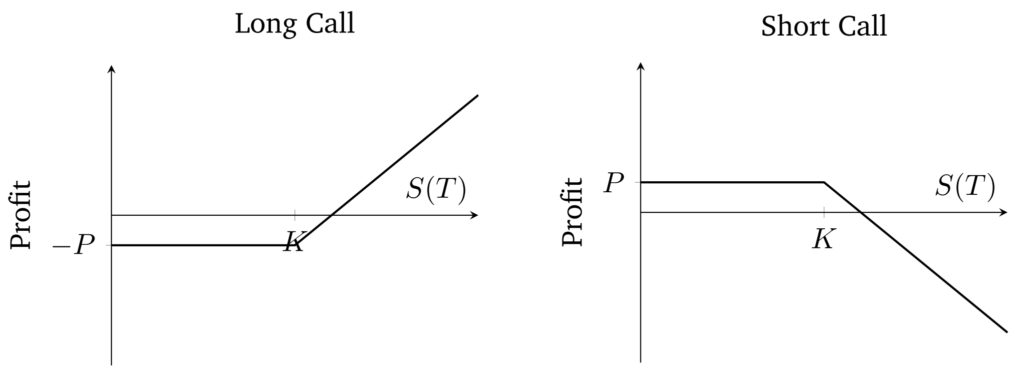



Figure 1.4.2: Payoff of a long and short position in a Call option.

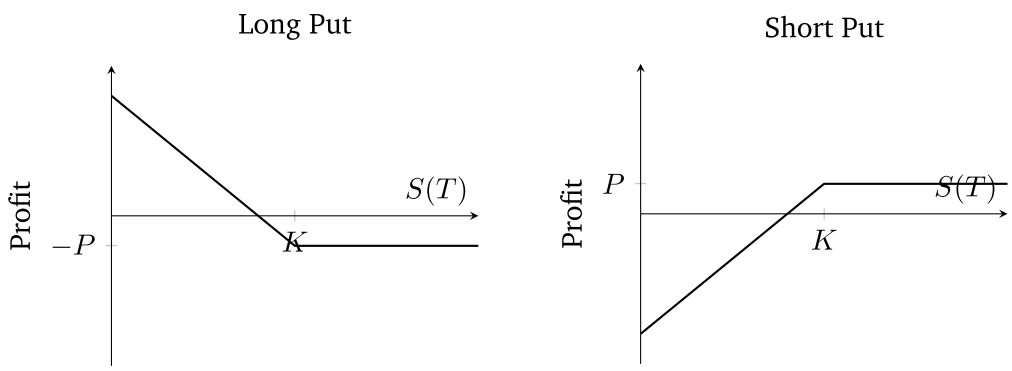



Figure 1.4.3: Payoff of a long and short position in a Put option

### 1.4.2 Relationship Between European Options and Properties

Unfortunately, with only the assumptions so far it is not possible to compute the fair price (or *premium*) of an option contracts. Nonetheless, we can still say something.

Notice that the call and put functions above share some striking similarities. From this we can prove what is known as the *put-call parity* result (Theorem 1.4.5 below).

Put-call parity relates the premiums of options with the value of the underlying stock and a certain amount of money.

The results in this section will use Market rule 1.3.4 as well as the following one:

**Market Rule 1.4.2** (Law of one price)**.** *In an arbitrage free market, two portfolios (or assets or cashflows) with identical future values must have identical present values.*

Two portfolios paying the same in every outcome at the same exact value future time must have the same price at any time prior, since otherwise arbitrage opportunities would exist in the market.

**Definition 1.4.3** (Replicating portfolio)**.** *A portfolio is called* replicating *with relation to a given contract (or portfolio) if at maturity time, the future valued of the replicating portfolio is the* exact same *as the future value of the contract (or portfolio).*

**Example 1.4.4** (On replicating portfolios)**.** *It is clear that any portfolio replicates itself.*

*Take portfolio A consisting of a long forward (set at $t = 0$ delivering $S(T)$ at $T$ for $K(0,T) = S(0)e^{rT}$. Then the portfolio B consisting of,at time $t = 0$, one stock $S$ and a loan of $£S(0)$ replicates portfolio A.*

*One can easily check that at $t = T$ both portfolios are worth $S(T) - S(0)e^{RT}$, hence replicating.*

The following results–that specify lower bounds for European puts and calls–are corollaries of the put-call parity theorem that follows later. We show these here first because the arguments used in their proof are useful.

### Lower Bound for European Call Options

Consider a call option written on an underlying stock $S$, with maturity $T$ and strike price $K$. Denote by $c_0$ the premium/fair price of the call at time $t = 0$. Recall that our generic bank account pays continuously compounded interests at rate $r$.

Consider the following two portfolios (at time $t = 0$):

*Portfolio A:* A cash amount of £$Ke^{-rT}$ and a European call option as above.
*Portfolio B:* One share $S$.

What are the balances of these portfolios at time $T$?

- **Portfolio A:**

  At time $T$, the cash has grown to $K$ (assuming it was invested at the risk-free interest rate $r$).

  Now, if underlying asset's price $S(T)$ is greater than $K$, the call option is exercised and portfolio $A$ is worth $S(T)$ (the cash amount spend in exercising the call option, and the investor receives the share that is worth $S(T)$).

  However, if $S(T) < K$, the option expires without being exercised and the portfolio is worth $K$.

  Thus, at time $T$, portfolio $A$ is worth

$$\max\{S(T), K\}$$

- **Portfolio B:**

  At time $T$, the value of Portfolio $B$ is simply $S(T)$.

Hence, at time $T$, portfolio $A$ is worth at least as much as portfolio $B$. Then, applying the Law of one price (Market Rule 1.4.2), portfolio $A$ is worth at least as much as portfolio $B$ at time $t = 0$. Hence (recalling that $c_0$ was the premium of the call option), we have

$$c + Ke^{-rT} \geq S(0) \qquad \Longrightarrow \qquad c \geq S(0) - Ke^{-rT}.$$

Moreover, since the price of a call option can only be non-negative, it follows that

$$c_0 \geq \max\{S(0) - Ke^{-rT}, 0\}.$$

In other words, if the Call option were to be sold at a price smaller than $c_0$, then an arbitrage opportunity would arise.

**Lower Bound for European Put Options**

Consider a put option written on an underlying stock $S$, with maturity $T$ and strike price $K$. Denote by $p_0$ the premium/fair price of the put at time $t = 0$. Recall that our generic bank account pays continuously compounded interests at rate $r$.

Consider the following two portfolios (at time $t = 0$):

*Portfolio A:* A European put option as above and one share $S$.
*Portfolio B:* A cash amount of £$Ke^{-rT}$.

What are the balances of these portfolios at time $T$?

- **Portfolio A:**

  If the underlying asset's price $S(T)$ is greater than $K$, the put option is not exercised and portfolio $A$ is worth $S(T)$.

  However, if $S(T) < K$ the put option is exercised and the portfolio is worth $K$ (the investor receives $K$ for selling his share).

  Thus, at time $T$, portfolio $A$ is worth

  $$\max\{S(T), K\}.$$

- **Portfolio B:**

  At time $T$, the cash has grown to $K$ (assuming it was invested at the risk-free interest rate $r$).

Therefore, at time $T$, portfolio $A$ is worth at least as much as portfolio $B$. Consequently, it is expected that portfolio $A$ is worth at least as much as portfolio $B$ at time $t = 0$. Hence

$$p_0 + S(0) \geq Ke^{-rT} \qquad \Longrightarrow \qquad p \geq Ke^{-rT} - S(0),$$

where $p_0$ denotes the price of a European put option. Taking into consideration the fact that the price of a put option cannot be negative, we have

$$p_0 \geq \max\{Ke^{-rT} - S(0), 0\}.$$

**Put-Call Parity**

Taking into consideration the last two results, we can establish a relationship between $c_0$, the price of a European call option written on a non-dividend-paying stock, and $p_0$, the price of the corresponding put option.

The put-call parity is nothing more than an equation that shows how the price of a **European** put option on a stock relates to the price of a **European** call option on the same stock (both options having the *same* strike $K$).

The proof of this relationship follows the same ideas as above. By comparing portfolios at time $T$ and then arguing, via the Law of one price (Market Rule 1.4.2), that their value must be the same at any earlier time–essentially due to the *absence of arbitrage*. If the prices were different, someone could simply buy the less expensive portfolio and sell the more expensive portfolio and make a profit without any risk.

**Theorem 1.4.5** (Put-Call parity)**.** *Consider a call and put option written on a stock $S$, with maturity $T$ and strike $K$. Denote by $c_0$ and $p_0$ the premium/fair price of the call and put at time $t = 0$ respectively. A generic bank account pays continuously compounded interest at rate $r$. Assume further that the market is arbitrage-free.*

*Then*

$$c_0 + Ke^{-rT} = p_0 + S(0).$$

*In general, let $(c_t)_{t\in[0,T]}$ and $(p_t)_{t\in[0,T]}$ denote the price of a call and put options respectively at time $t \in [0, T]$. Then, it holds for any $t \in [0, T]$ that*

$$c_t + Ke^{-r(T-t)} = p_t + S(t).$$

The proof of this theorem is based on the Law of one price (see Market rule 1.4.2) and Replicating portfolios (see Definition 1.4.3).

*Proof.* Consider the following two portfolios:

*Portfolio A:* A cash amount of £$Ke^{-rT}$ and a European call option written on a non-dividend-paying stock (the time to maturity is $T$ and the exercise price is $K$),
*Portfolio B:* A European put option written on a non-dividend-paying stock and one share (the time to maturity is $T$ and the exercise price is $K$).

It can be shown that the portfolios have identical payoffs at time $T$: $\max\{S(T), K\}$. Hence, they must have identical values at time $t = 0$ and so

$$c_0 + Ke^{-rT} = p_0 + S(0).$$

□

**Creating Synthetic Securities Using Put-Call Parity**

We can use the put-call parity equation to create *synthetic* versions of other securities, including put options, call options, stocks, and zero-coupon bonds.[3] By 'synthetic' we mean that we replicate or simulate a security by combining the features of other assets.

[3]See page 7

To this end, we construct a number of replicating portfolios below. At maturity, the values of these portfolios are the same as the security (represented on the left hand side of each equation below). Thus, the values of such a portfolio and the security must have the same value at time $t = 0$ (as a consequence of Law of one price; Market Rule 1.4.2). Note that in the equations below, a positive coefficient represents a long position, and a negative coefficient a short one.

▷ **Synthetic stock**

$$S(0) = c_0 + Ke^{-rT} - p_0.$$

A long position in the underlying is equivalent to a long position in a call option, a long position in a risk-free bond and a short position in a put option. Thus, to to create a portfolio that replicates the payoff of a share (i.e. to create a synthetic stock) one buys a call option; buys a bond with face value $K$; and sells a put option.

▷ **Synthetic bond**

$$Ke^{-rT} = S(0) + p_0 - c_0,$$

A long position in a risk-free bond with face value $K$ (a zero-coupon-bond) is equivalent to having a long position in the underlying, a long position in a put option and a short position in a call option.

▷ **Synthetic put option**

$$p_0 = c_0 + Ke^{-rT} - S(0),$$

A long position in a put option is equivalent to a long position in a call option, a long position in a risk-free bond and a short position in the underlying. The portfolio replicating the payoff of a Put option is given by: buying a call option; buying a (zero-coupon-)bond with face value $K$; and (Short) "selling" a stock.

▷ **Synthetic call option**

$$c_0 = p_0 + S(0) - Ke^{-rT},$$

A long position in a call option is equivalent to a long position in a put option, a long position in the underlying and a short position in the risk-free bond. The portfolio replicating the payoff of a Call option is given by buying a put option; buying one Stock; and selling a (zero-coupon-)bond with face value $K$.

### 1.4.3 Trading Strategies with Options for Hedging or Speculation

We define here **three** types of trading strategies. These strategies are used commonly in the market, either driven by a speculation motive or a desire to hedge. These strategies can:

1. involve a single option and a stock;
2. involve two or more options of same type (**spreads**);
3. involve both calls and puts on the same stock (**combinations**).

For the sake of simplicity, we **ignore the time value of money** (i.e. $r = 0$) in calculating the profit functions throughout. Thus, profits are calculated by subtracting the initial cost from the final payoff. In reality, it should be equal to the *present value* of the final payoff minus the initial cost.

### Trading Strategies Involving a Single Option and a Stock

These strategies seek to reduce the risk from the short position of a call option (hence the name *covered call*) or the risk of a sharp decrease in the value of a stock (hence *protective put*).

■ **Covered Call**

A trader with a short position in a call option on a stock is exposed to unlimited downside risk in the event of a big price increase, i.e. if he holds a naked call (compare with the graphic of the call function, $-\max\{S(T) - K, 0\}$). The 'covered call' strategy aims at limiting this risk. In more financial terms, this risk can be *hedged* by holding a long position in the underlying stock. See the left side of Figure 1.4.4.

■ **Protective Put**

A trader holding a stock is exposed to downside risk if the stock price decreases. The risk can be hedged by taking a long position in a put option. This position is called a 'protective put'. It is clear from the figure that the downside risk is now reduced. See right side of Figure 1.4.4.

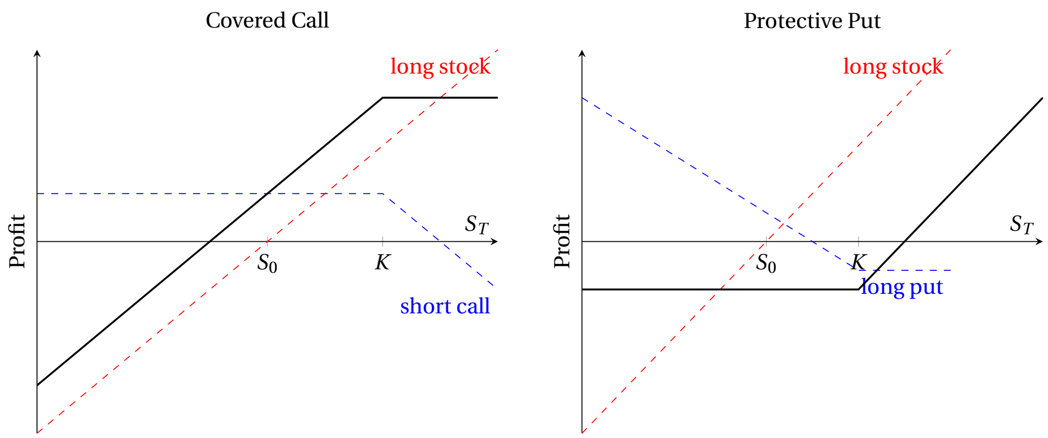



Figure 1.4.4: Payoff of Covered call on the left and Protective Put on the right

**Exercise 1.4.6.** *The reverse of a covered call and of a protective put are also possible. Draw and explain their profit diagrams.*

### Spreads: Trading Strategies Involving Two or More Options of Same Type

A **spread** is a portfolio consisting of two or more call/put options (but not a combination of both) with different positions on the same stock. We shall study the following **three** types of spreads:

- **Bull spread** - long one call with a certain strike price and short one call with higher strike price

- **Bear Spread** - long one call with a certain strike price and short one call with lower strike price
- **Butterfly Spread** - Long two calls with different strike prices and short two calls with a strike price which is the average of these two strikes

### ■ Bull Spread

A bull spread is a portfolio consisting of one long call with strike price $K_1$ and one short call with strike price $K_2$ (with $K_2 > K_1$). The payoff possibilities are summarized in following table and graphically in the left side of Figure 1.4.5.

| Pay-off Range | Long Call Pay-off | Short Call Pay-off | Total Pay-off |
|---|---|---|---|
| $S(T) \leq K_1$ | 0 | 0 | 0 |
| $K_1 \leq S(T) \leq K_2$ | $S(T) - K_1$ | 0 | $S(T) - K_1$ |
| $S(T) \geq K_2$ | $S(T) - K_1$ | $-(S(T) - K_2)$ | $K_2 - K_1$ |

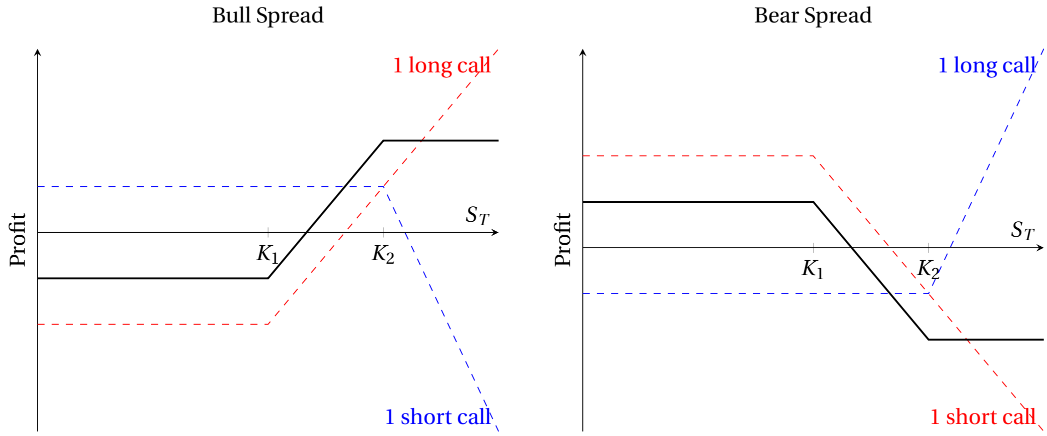



Figure 1.4.5: Bull spread on the left and Bear spread on the right

As an investment strategy, we observe that:

- the total pay-off is non-negative ($K_2 > K_1$),
- the downside risk is limited, and
- a bull spread is a useful trading strategy when the stock price increases.

**Exercise 1.4.7.** *The price of the call option with the strike price $K_2$ is less than that of the call option with strike price $K_1$. Hence a bull spread requires an initial investment. Explain why.*

### ■ Bear Spreads

A bear spread is a portfolio consisting of one long call option with strike price $K_2$ and one short call option with strike price $K_1$ (with $K_2 > K_1$). The payoff possibilities are summarized in the following table and graphically in the right side of Figure 1.4.5.

As an investment strategy, we observe that:

- the total pay-off is non-positive ($K_2 > K_1$),

| Pay-off Range | Long Call Pay-off | Short Call Pay-off | Total Pay-off |
|---|---|---|---|
| $S(T) \le K_1$ | 0 | 0 | 0 |
| $K_1 \le S(T) \le K_2$ | 0 | $-(S(T)-K_1)$ | $K_1 - S(T)$ |
| $S(T) \ge K_2$ | $S(T)-K_2$ | $-(S(T)-K_1)$ | $K_1 - K_2$ |

- the downside risk is limited, and
- a bear spread is a useful trading strategy when the stock price decreases.

**Exercise 1.4.8.** *Explain why a bear spread gives an initial income.*

### ■ Butterfly Spreads

A butterfly spread is a portfolio consisting of one long call option with strike price $K_1$, one long call option with strike price $K_3$ ($K_3 > K_1$) and two short call options with the same strike price $K_2$, where $K_2 = (K_1 + K_3)/2$. The payoff possibilities are summarized in the following table and graphically in Figure 1.4.6.

| Pay-off Range | Long Call Pay-off, $K_1$ | Long Call Pay-off, $K_3$ | 2 Short Calls Pay-off, $K_2$ | Total Pay-off |
|---|---|---|---|---|
| $S(T) \le K_1$ | 0 | 0 | 0 | 0 |
| $K_1 \le S(T) \le K_2$ | $(S(T)-K_1)$ | 0 | 0 | $S(T)-K_1$ |
| $K_2 \le S(T) \le K_3$ | $(S(T)-K_1)$ | 0 | $-2(S(T)-K_2)$ | $K_3 - S(T)$ |
| $S(T) \ge K_3$ | $S(T)-K_1$ | $S(T)-K_3$ | $-2(S(T)-K_2)$ | 0 |

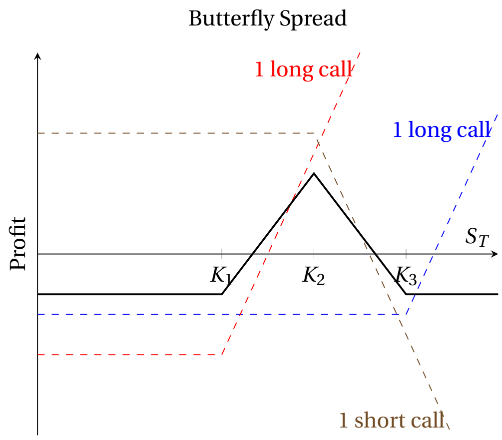



Figure 1.4.6: Payoff of Butterfly spread

As an investment strategy, the butterfly spread is a useful trading strategy when the stock price stays close to the strike price $K_2$.

**Exercise 1.4.9** (The put version of these spreads). *The analysis carried out to study the above spreads for call options can be repeated for spreads written on put options instead of call options.*

*(a) Create the above spreads by using put options and explain their profit diagrams.*

*(b) For each of them, what is the investor's perception of the market when he buys one of these spreads? Is the price of $S(T)$ expected to go up, to go down, to lie in some interval, ...?*

**Combinations - Trading Strategies Involving Calls and Puts on the Same Stock**

A **combination** is a portfolio consisting of both call and put options with the same position on the same stock. The strikes and maturities of all the options included in the portfolio are same. We shall study the following **four** types of combinations:

| | |
|---|---|
| Straddle | long one call and one put |
| Strips | long one call and two puts |
| Straps | long one put and two calls |
| Strangles | long a put and long a call |

■ **Straddle : Bottom Straddle**

A bottom straddle is combination of a long put and a long call. The payoff possibilities are summarized in the following table and graphically in left side of Figure 1.4.7.

| Pay-off Range | Call Pay-off | Put Pay-off | Total Pay-off |
|---|---|---|---|
| $S(T) \leq K$ | 0 | $K - S(T)$ | $K - S(T)$ |
| $S(T) > K$ | $S(T) - K$ | 0 | $S(T) - K$ |

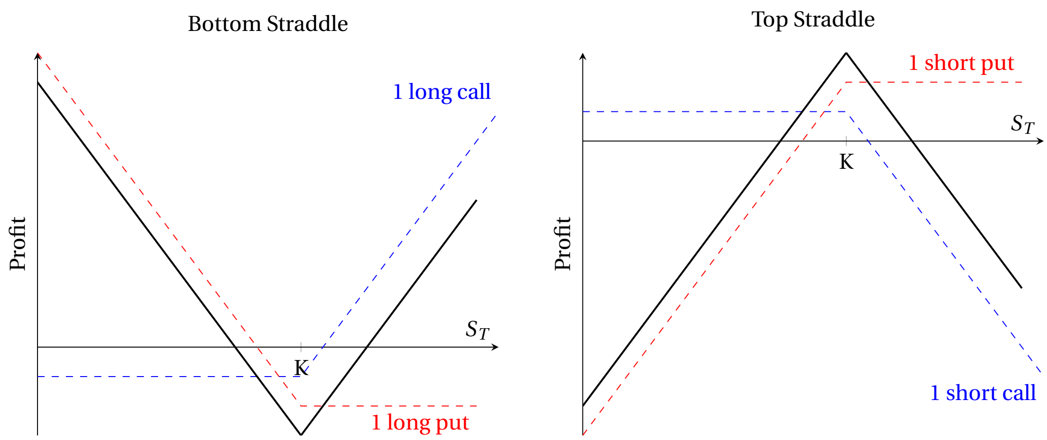



Figure 1.4.7: Bottom Straddle on the left and Top Straddle on the right

As an investment strategy, we observe that a straddle is a useful instrument when a large move in the stock price in either direction occurs.

■ **Straddle : Top Straddle**

A top straddle is a combination of a short call and a short put with same strike price and expiry date. The payoff possibilities are summarized in the following table and graphically in right side of Figure 1.4.7.

| Pay-off Range | Call Pay-off | Put Pay-off | Total Pay-off |
|---|---|---|---|
| $S(T) \leq K$ | 0 | $-(K - S(T))$ | $S(T) - K$ |
| $S(T) > K$ | $-(S(T) - K)$ | 0 | $K - S(T)$ |

As an investment strategy, the top straddle is a very risky trading strategy because a big movement in the stock price from the strike price in either direction results in an unbounded loss.

■ **Strips**

A strip is a combination of one long call and two long puts, all with same strike prices and expiry dates. The payoff possibilities are summarized in the following and graphically in left side of Figure 1.4.8.

| Pay-off Range | 1 Call Pay-off | 2 Put Pay-off | Total Pay-off |
|---|---|---|---|
| $S(T) \leq K$ | 0 | $2(K - S(T))$ | $2(K - S(T))$ |
| $S(T) > K$ | $S(T) - K$ | 0 | $S(T) - K$ |

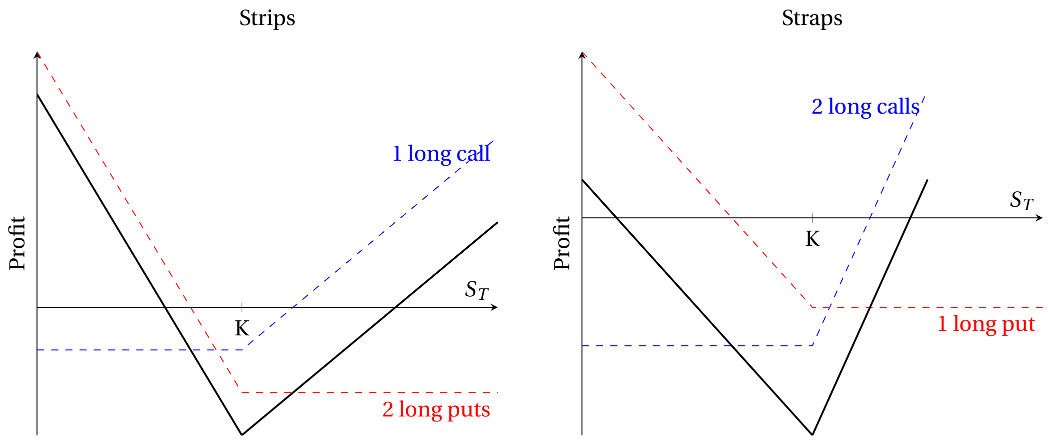



Figure 1.4.8: Payoff of a Strip on the left and a Strap on the right

As an investment strategy, a strip is a useful trading strategy when there are big movements in the stock price in either direction, and where a decrease is more likely than an increase.

■ **Straps**

A strap is a combination of one long put and two long call options with same strike prices and expiry dates. The payoff possibilities are summarized in the following table and graphically in the right side of Figure 1.4.8.

| Pay-off Range | Call Pay-off | Put Pay-off | Total Pay-off |
|---|---|---|---|
| $S(T) \leq K$ | 0 | $K - S(T)$ | $K - S(T)$ |
| $S(T) > K$ | $2(S(T) - K)$ | 0 | $2(S(T) - K)$ |

As an investment strategy, a strap is a useful trading strategy when there are expected to be large movements in stock prices in either direction, and where an increase in price is more likely than a decrease.

■ **Strangles**

A strangle is a combination of one long put at strike price $K_1$ and one long call at strike price $K_2$ ($K_2 > K_1$). The payoff possibilities are summarized in the following table and graphically in Figure 1.4.9.

A strangle is a trading strategy similar to a straddle. It is a useful when there is a big movement in the stock price from the strike price in either direction, but it is not known whether an increase (or a decrease) is more likely than a decrease (or an

| Pay-off Range | Call Pay-off | Put Pay-off | Total Pay-off |
|---|---|---|---|
| $S(T) \leq K_1$ | 0 | $K_1 - S(T)$ | $K_1 - S(T)$ |
| $K_1 < S(T) < K_2$ | 0 | 0 | 0 |
| $S(T) \geq K_2$ | $S(T) - K_2$ | 0 | $S(T) - K_2$ |

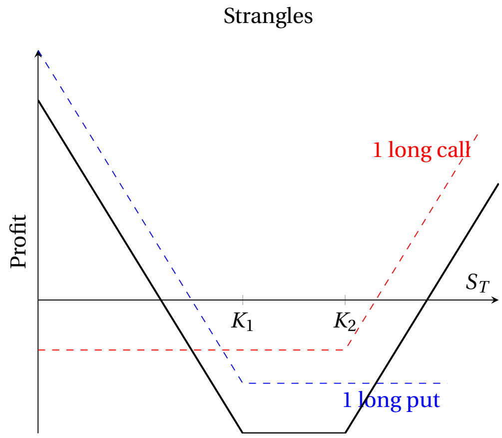



Figure 1.4.9: Payoff of a Strangle

increase). However, in the case where the stock price ends up near central value, then its downside risk is less than that of the straddle. The profit pattern of a strangle depends on the closeness of the strike prices, the further apart they are, the lesser the downside risk and hence a really big movement is needed for making profits.

The strangle is also called a bottom vertical combination. The top vertical combination is actually a sale of a bottom vertical combination and is a very risky trading strategy involving unbounded loss, if the stock price makes a big movement from the central in either direction. When the stock price remains central, then a top vertical combination gives a profit.

## 1.5 Exercises

**Exercise 1.5.1.** *An investor puts £$y_0$ in a bank account with interest rate $r$ and annual compounding. After $n$ years, the amount of money in the bank account is given by*

$$y_n = y_0(1 + r)^n.$$

*Determine, from the above equation, the expressions that yield $y_0$, $r$ and $n$ as a function of the remaining parameters.*

**Exercise 1.5.2.** *Suppose you deposit £$621$ in a bank account that pays $6\%$ compounded continuously.*

*(a) How much money will you have after $8$ years?*

*(b) Had the bank paid interest quarterly (see* (1.1.1)*), what would the amount of money in the bank account be?*

*(c) Which type of compounding would you rather have in your bank account?*

**Exercise 1.5.3.** *Suppose that a life insurance company has guaranteed a payment of £14 million to a pension fund four years from now. If the life insurance company receives* *a premium of £11 million and can invest the entire premium for four years at an annual interest rate of 6.5% (annual compounding), will it have sufficient funds from this investment to meet the £14 million obligation?*

**Exercise 1.5.4.** *A bank quotes you a (nominal) rate of interest of $i = 14\%$ per annum with quarterly compounding. What is the equivalent rate with (a) continuous compounding; (b) annual compounding (also known as effective rate of interest)?*

**Exercise 1.5.5.** *You are considering three investments: (a) one paying $7\%$ compounded quarterly; (b) one paying $7.1\%$ compounded annually; and (c) one paying $6.9\%$ compounded continuously.*

*Which investment has the highest equivalent effective annual rate of return?*

**Exercise 1.5.6.** *It is widely know that the power series, for $z \in (-1, 1)$, satisfies:*

$$\sum_{i=0}^{\infty} z^i = \frac{1}{1-z}.$$

*Let $n, m$ be natural numbers such that $0 \leq m \leq n < \infty$ and $z \in (-1, 1)$. Prove the following identities.*

$$(a) \sum_{i=n+1}^{\infty} z^i = \frac{z^{n+1}}{1-z}, \quad (b) \sum_{i=0}^{n} z^i = \frac{1-z^{n+1}}{1-z} \quad \textit{and} \quad (c) \sum_{i=m}^{n} z^i = \frac{z^m - z^{n+1}}{1-z}.$$

**Exercise 1.5.7.** *The value of a bond $B$ (at time $t = 0$) with a face value $F$, a coupon $C$ paid every year and $T$ years to maturity is given by*

$$B = C \frac{1 - \frac{1}{(1+r_A)^T}}{r_A} + \frac{F}{(1+r_A)^T} \tag{1.5.1}$$

*when the rate of interest $r_A$ per annum with annual compounding is constant for the next $T$ years.*

*Prove, in the case of a continuously compounded interest rate $r_c$, that the value of the above bond $B$ is given by*

$$B = C\,\frac{1 - e^{-r_c T}}{e^{r_c} - 1} + F e^{-r_c T}. \tag{1.5.2}$$

**Exercise 1.5.8.** *Mr. Salmond and Mr. Darling purchase bonds at the same time; each bond pays annuals coupons and a lump sum at maturity, the bond's face value is $£10{,}000$. The bank interest rate is $r = 6\%$ (compounded yearly).*

*(a) If Salmond's bond pays annual coupons at the rate of $8\%$ (of the face value) and it has $15$ years to maturity, then how much does he pay for it now?*

*(b) If Darling's bond has $10$ to maturity and he pays $£11{,}487.75$ for it, then what is the annual coupon rate*[4]*?*

**Exercise 1.5.9.** *A trader enters into a short forward contract on 100 million yen (what is a forward? short and long?). The forward exchange rate is $£0.0080$ per yen. How much does the trader gain or lose if the exchange rate at the end of the contract is (a) $£0.0074$ per yen; (b) 0.0091 per yen?*

[4]The coupon rate is defined to be $r_C := C/F$ =Coupon value/Face value.

**Exercise 1.5.10.** *The current price of a stock $S(t) = £100$. Traders can borrow and lend at a rate of 5% per year (continuous compounding). Suppose that the market starts quoting a delivery price of $£101$ for a twelve month forward contract (i.e. $T - t$ =1) on the delivery of this stock. Form an arbitrage portfolio (show that there exists an arbitrage opportunity).*

**Exercise 1.5.11** (On Forward contracts)**.** *An investor enters a long forward contract on $100$ units of underlying assets $S$ and maturity $T = 4$ years. The asset $S$ pays no dividends and the spot price of one asset is $S_0 = £5$. The continuously compounded interest rate is $r = 4\%$.*

*(a) Write the formula for the fair forward delivery price $K_{0,T}$ (with $T = 4$) and calculate it.*

*(b) Consider* Portfolio A *which consists of 100 units of asset $S$. Construct a portfolio (denoted* Portfolio B*) that replicates portfolio A using a long forward contract and suitable bank deposit.*

*Show that Portfolio A and B are replicating at time $T$ and then derive/deduce the* fair forward delivery price *formula for $K_{0,T}$ which you used in part (a).*

*(c) Let $t \in [0, T]$ and denote by $V_t$ the value of the long forward contract. Derive a formula for $V_t$ in terms of $t, T, r$ and $S_t$, the latter being the price of the stock at time $t$. To do so, use the replicating portfolios from part (b) and the formula for $K_{0,T}$ (not the numerical value).*

*(d) What is the value of the forward contract at time zero? Argue using the process $V_t$ at time $t = 0$.*

*(e) Suppose that at maturity $T = 4$ years the stock price is $S_4 = £7$. Calculate the value of the forward contract for the investor at maturity $T$. Did the investor gain or loose money?*

*(f) Suppose that the investor wants to sell the forward contract at $t = 1$ year, when the stock price is $S_1 = £6$. Calculate the value $V_1$ of the forward contract for the investor.*

*(g) Suppose that the investor agrees to sell the forward contract at $t = 1$ year to a company for a price of $P_1 = £60$, when the stock price is $S_1 = £6$. How can the company construct an arbitrage?*
Hint: the company is allowed to borrow money from the bank, deposit money at the bank, buy or short-sell stocks and enter any forward contract on stocks.

**Exercise 1.5.12.** *A trader buys a European call on a share for $£4$. The stock price is $£47$ and the strike price is $£50$. Under what circumstances does the trader make a profit? Under what circumstances will the option be exercised? Draw a diagram showing the variation of the trader's profit with the stock price at the maturity of the option.*

**Exercise 1.5.13.** *Draw a diagram showing the variation of the following portfolio's profit with the price of the underlying asset: A long position in a European call with a strike price of $£45$ and a long position in a European put with a strike price of $£40$. The* *premium for the call is £4 and for the put is £3. It is assumed that both options have the same maturity. (A* STRANGLE *is a similar option trading strategy to a* STRADDLE*)*

**Exercise 1.5.14.** *A range forward contract in foreign exchange markets is designed to ensure that if the spot rate at maturity is less than $F_1$, the investor pays $F_1$ for each unit of the foreign currency; if it is between $F_1$ and $F_2$, the investor pays the spot rate; if it is greater than $F_2$, the investor pays $F_2$ for each unit of the foreign currency. Show that a long range forward contract is a combination of two options. (Draw the above diagram without taking into consideration the premiums.)*

# Chapter 2

# Review of probability

This chapter reviews some of the fundamental concepts in probability that are needed for this course. All students are assumed to have taken a probability course and this chapter is just a refresher.

A reference for a Probability course is the book "A First Course in Probability" by Sheldon and Ross[1].

Disclaimer: this Chapter has more material then what is really needed and is here as a refresher on probability. There are many small results in this chapter that are used throughout. Fundamental are the last two sections (well, three if you count the one with the exercises).

## 2.1 Summary of UoE's probability (MATH08066) course

This section contains the summary of the material taught in the School of Mathematics' Probability (MATH08066) course. This summary was created by Prof. Tibor Antal and slightly changed by me to match the this courses' notation.

### Combinatorial Analysis

*Basic principle of counting:* If in sub-experiment $i$ there are $n_i$ possible outcomes for $i = 1, \ldots, k$, then there are $n_1 n_2 \cdots n_k$ possible outcomes of the experiment.

There are $n!$ possible ways to order $n$ (distinguishable) items.

The number of ways to choose $i$ elements out of $n$ elements is "$n$ choose $i$", the binomial coefficient, which is

$$\binom{n}{i} = \frac{n!}{i!(n-i)!} \, .$$

The number of ways to distribute $n$ elements into $k$ ordered piles of sizes $n_1$, $n_2$, $\ldots, n_k$ is the multinomial coefficient

$$\binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1! n_2! \ldots n_k!} \, .$$

[1]There are many copies of it in the Noreen and Kenneth Murray Library (King's buildings).

The binomial theorem states that

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i} \; .$$

## Axioms of Probability

The sample space $S$ is the set of all possible outcomes of an experiment. An event is a subset of the sample space.

For events $A$ and $B$, we have the union $A \cup B$, the intersection $AB = A \cap B$, and the complement $A^c$. Two events are disjoint (or mutually exclusive) if $AB = \emptyset$, where $\emptyset$ is the empty set (an impossible event). The DeMorgan laws are

$$(\cup A_i)^c = \cap A_i^c \quad (\cap A_i)^c = \cup A_i^c \; .$$

Each event $A$ has a probability $\mathbb{P}(A)$ such that

- $\mathbb{P}(A) \in [0,1]$,
- $\mathbb{P}(S) = 1$,
- For disjoint events $A_i$
$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

The probability of the complement is $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

The inclusion–exclusion formula is $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(AB)$, or in general

$$\mathbb{P}(\cup_{i=1}^n A_i) = \sum_{r=1}^{n} (-1)^{r+1} \sum_{i_1 < \cdots < i_r} \mathbb{P}(E_{i_1} \cdots E_{i_r}).$$

If all outcomes are equally likely in a finite sample space, then $\mathbb{P}(A) = |A|/|S|$.

## Conditional Probability and Independence

The probability of event $A$ conditioned on event $B$ is

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)}$$

if $\mathbb{P}(B) \neq 0$. It follows that $\mathbb{P}(AB) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A)$. The multiplication rule is

$$\mathbb{P}(A_1 \cdots A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1A_2)\cdots\mathbb{P}(A_n|A_1\cdots A_{n-1}).$$

*Total probability rule:* $\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c)$. More generally, if $B_i$ for $i = 1, \ldots, k$ partition the sample space (i.e. $\cup B_i = S$ and $B_iB_j = \emptyset$ for all $i \neq j$), then

$$\mathbb{P}(A) = \sum_{i=1}^{k} \mathbb{P}(A|B_i)\mathbb{P}(B_i).$$

If we combine this with the definition of conditional probability then we obtain the *Bayes's rule*

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{j=1}^{k} \mathbb{P}(A|B_j)\mathbb{P}(B_j)} .$$

Events $A$ and $B$ are independent if $\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$. This condition is equivalent to $\mathbb{P}(A|B) = \mathbb{P}(A)$ and to $\mathbb{P}(B|A) = \mathbb{P}(B)$. Events $A_i$ for $i = 1, \dots n$ are independent if for any subset of them

$$\mathbb{P}(A_{i_1} \cdots A_{i_k}) = \mathbb{P}(A_{i_1}) \cdots \mathbb{P}(A_{i_k}).$$

## Random Variables

A random variable (RV) is a random number which depends on the outcome of an experiment. More precisely a random variable is a function $X : S \to \mathbb{R}$. Its (cumulative) distribution (function) is

$$F(x) = F_X(x) = \mathbb{P}(X \leq x).$$

$F(x)$ is non-decreasing, right-continuous, $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$. Also

$$\mathbb{P}(a < X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) = F(b) - F(a)$$

and

$$\mathbb{P}(X = x) = \mathbb{P}(X \leq x) - \lim_{y\to x^-} \mathbb{P}(X \leq y) = F(x) - \lim_{y\to x^-} F(y).$$

*Discrete RVs:* A random variable $X$ is discrete if it can take only a countable number of possible values. The (probability) mass function of $X$ is

$$p(x) = \mathbb{P}(X = x)$$

the expected (or mean) value of $X$ is

$$\mathbb{E}[X] = \sum_x x\mathbb{P}(X = x)$$

the expected value of a function of $X$ is

$$\mathbb{E}[g(X)] = \sum_x g(x)\mathbb{P}(X = x)$$

where the summation goes over all possible values of $X$. The variance of $X$ is

$$\mathrm{Var}[X] = \mathbb{E}[(X - EX)^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

Most important discrete RVs, also expressed as sums of *independent* RVs:

- Bernoulli ($p$), $\mathbb{P}(X = 1) = p, \mathbb{P}(X = 0) = 1 - p$, $\mathbb{E}[X] = p$, $\mathrm{Var}[X] = p(1 - p)$.
- binomial $(n, p) = \sum_{i=1}^{n}$ Bernoulli ($p$), $\mathbb{P}(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}, \mathbb{E}[X] = np, \mathrm{Var}[X] = np(1 - p)$.
- Poisson ($\lambda$), which is the $n \to \infty$ limit of a Binomial $(n, p)$ with $\lambda = np$, $\mathbb{P}(X = n) = e^{-\lambda}\lambda^n/n!$, $\mathbb{E}[X] = \lambda$, $\mathrm{Var}[X] = \lambda$.

- geometric ($p$): # independent Bernoulli ($p$) trials till first success, $\mathbb{P}(X = n) = (1-p)^{n-1}p$, $\mathbb{E}[X] = 1/p$, $\mathrm{Var}[X] = (1-p)/p^2$.
- negative binomial $(r, p) = \sum_{i=1}^{r}$ geometric ($p$), # independent Bernoulli ($p$) trials till the $r$-th success, $\mathbb{P}(X = n) = \binom{n-1}{r-1}(1-p)^{n-r}p^r$, $\mathbb{E}[X] = r/p$, $\mathrm{Var}[X] = r(1-p)/p^2$.
- hypergeometric $(N, m, n)$: # red balls drawn when we draw $n$ balls out of $m$ red balls and $N-m$ blue balls, $\mathbb{P}(X = i) = \binom{m}{i}\binom{N-m}{n-i}/\binom{N}{n}$, $\mathbb{E}[X] = nm/M$, $\mathrm{Var}[X] = np(1-p)[1-(n-1)/(N-1)]$, with $p = m/N$.

## Continuous random variables

A RV is continuous if its distribution function is continuous, that is if it has a density function $f : \mathbb{R} \to \mathbb{R}$ such that for any $B \subset \mathbb{R}$

$$\mathbb{P}(X \in B) = \int_B f(x)dx.$$

The distribution of $X$ is

$$F(x) = F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(x)\mathrm{d}x$$

and hence where $F(x)$ is differentiable, $f(x) = \frac{d}{dx}F(x)$. The density is normalized to one

$$F_X(\infty) = \mathbb{P}(X \leq \infty) = \int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1,$$

also

$$\mathbb{P}(a < X \leq b) = \int_a^b f(x)\mathrm{d}x$$

and for any $a \in \mathbb{R}$

$$\mathbb{P}(X = a) = \int_a^a f(x)\mathrm{d}x = 0.$$

The expected value of $X$ is

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)\mathrm{d}x$$

and the expected value of a function of $X$ is

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x.$$

The variance is the same as for discrete RVs.

Most important continuous RVs, also expressed as sums of *independent* RVs:

- uniform on [a,b], $f(x) = 1/(b-a)$ for $a < x < b$, and 0 otherwise. $\mathbb{E}[X] = (b+a)/2$, $\mathrm{Var}[X] = (b-a)^2/12$.
- normal $(\mu, \sigma^2)$, $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/2\sigma^2}$, $\mathbb{E}[X] = \mu, \mathrm{Var}[X] = \sigma^2$.
- standard normal (0,1), $f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, $\mathbb{E}[X] = 0, \mathrm{Var}[X] = 1$. Its distribution is $F(x) = \Phi(x)$ and can usually be found in tables. Also, $\mu+\sigma X$ is normal $(\mu, \sigma^2)$.

- exponential ($\lambda$), $f(x) = \lambda e^{-\lambda x}$ for $x > 0$, and 0 otherwise, $\mathbb{E}[X] = 1/\lambda$, $\mathrm{Var}[X] = 1/\lambda^2$. The exponential random variable is memoryless, that is $\mathbb{P}(X > t + s | X > t) = \mathbb{P}(X > s)$.
- gamma ($n, \lambda$) = $\sum_{i=1}^{n}$ exponential ($\lambda$), $f(x) = \lambda e^{-\lambda x} (\lambda x)^{n-1}/(n-1)!$ for $x > 0$, and 0 otherwise, $\mathbb{E}[X] = n/\lambda$, $\mathrm{Var}[X] = n/\lambda^2$.

For both discrete and continuous RVs: $\mathbb{E}(b + aX) = b + a\mathbb{E}[X]$, and $\mathrm{Var}[b + aX] = a^2 \mathrm{Var}[X]$.

DeMoivre-Laplace Theorem: $S_n$ binomial ($n, p$) RV converges to a standard normal

$$\lim_{n\to\infty} \mathbb{P}\left( a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b \right) = \Phi(b) - \Phi(a).$$

where $\Phi(x)$ is the cumulative distribution function of the standard normal RV

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy$$

### Jointly Distributed RVs

The joint distribution of $X$ and $Y$ RVs is

$$F(a, b) = F(X \leq a, Y \leq b)$$

and the marginal distributions are $F_X(a) = \mathbb{P}(X \leq a) = \mathbb{P}(X \leq a, Y \leq \infty) = F(a, \infty)$, and $F_Y(b) = F(\infty, b)$. For a rectangular region

$$\mathbb{P}(a_1 < X \leq a_2, b_1 < Y \leq b_2) = F(a_2, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, b_1).$$

$X$ and $Y$ are independent if for all $A, B \subset \mathbb{R}$, $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$.

*For discrete RVs* there is a joint mass function $p(x, y) = \mathbb{P}(X = x, Y = y)$, with marginal masses $p_X(x) = \mathbb{P}(X = x) = \sum_y p(x, y)$, and $p_Y(y) = \mathbb{P}(Y = y) = \sum_x p(x, y)$. $X$ and $Y$ are independent if $p(x, y) = p_X(x)p_Y(y)$.

*For continuous RVs* there is a joint density function $f(x, y)$, such that for any $C \subset \mathbb{R}^2$

$$\mathbb{P}((X, Y) \in C) = \iint_C f(x, y) dx\, dy.$$

The marginal densities are $f_X(x) = \int_{-\infty}^{\infty} f(x, y) \mathrm{d}y$, and $f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \mathrm{d}x$. $X$ and $Y$ are independent if $f(x, y) = f_X(x) f_Y(y)$.

Always $\mathbb{E}[\sum X_i] = \sum \mathbb{E}[X_i]$, and for independent $X_i$ variables also $\mathrm{Var}[\sum X_i] = \sum \mathrm{Var}[X_i]$.

The mass function of the sum of independent discrete variables is

$$\mathbb{P}(X + Y = a) = \sum_x \mathbb{P}(X = x)\mathbb{P}(Y = a - x)$$

The density function of the sum of independent continuous variables is

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a - y) f_Y(y) \mathrm{d}y.$$

In particular

- normal $(\mu_1, \sigma_1^2)$ + normal $(\mu_2, \sigma_2^2)$ = normal $(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$,
- $\sum_{i=1}^n$ exponential $(\lambda)$ = gamma$(n, \lambda)$,
- $\sum_{i=1}^n$ Bernoulli $(p)$ = binomial $(n, p)$,
- $\sum_{i=1}^n$ geometric $(p)$ = negative binomial $(n, p)$,
- Poisson $(\lambda_1)$ + Poisson $(\lambda_2)$ = Poisson $(\lambda_1 + \lambda_2)$.

### Moment generating function

The moment generating function of a random variable $X$ is

$$M_X(t) = M(t) = \mathbb{E}[e^{tX}]$$

for $t \in \mathbb{R}$. Note that $M(0) = 1$. The derivatives of $M(t)$ generate the moments of $X$, for example $\mathbb{E}[X] = M'(0)$, and in general

$$M^{(n)} = \mathbb{E}[X^n]$$

and hence $\mathrm{Var}[X] = M''(0) - M'(0)^2$. The moment generating function of the sum of $X, Y$ independent RVs is the product of their generating functions, that is

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

Some examples we considered:

- If $X$ is a Bernoulli $(p)$, then $M(t) = pe^t + 1 - p$.
- If $X$ is a binomial $(n, p)$, then $M(t) = (pe^t + 1 - p)^2$.
- if $X$ is a normal $(\mu, \sigma^2)$, then $M(t) = e^{t\mu + t^2\sigma^2/2}$.

### Limit theorems

*Markov Inequality*: For $X > 0$ RV, and $a > 0$

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

*Chebishev Inequality*: For $X$ RV, and $a > 0$

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{\mathrm{Var}[X]}{a^2}$$

*Weak Law of Large Numbers*: If $X_1, X_2, \ldots, X_n$ are independent, identically distributed RVs with $\mathbb{E}[X_i] = \mu$, and $\mathrm{Var}[X_i] = \sigma^2$ for all $i$, then for any $a > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq a\right) = 0$$

*Central Limit Theorem*: If $X_1, X_2, \ldots, X_n$ are independent, identically distributed RVs with $\mathbb{E}[X_i] = \mu$, and $\mathrm{Var}[X_i] = \sigma^2$ for all $i$, then

$$\lim_{n\to\infty} \mathbb{P}\left(a \leq \frac{X_1 + \cdots + X_n - \mu n}{\sigma\sqrt{n}} \leq b\right) = \Phi(b) - \Phi(a)$$

where $\Phi(x)$ is the distribution function of the standard normal RV.

To comment on the above two theorems: if we call $S_n = X_1 + \cdots + X_n$, then $\mathbb{E}[S_n] = n\mu$ and $\mathrm{Var}[S_n] = n\sigma^2$. Then the WLLN states that $(S_n - n\mu)/n$ converges to 0 in probability, or equivalently that $S_n/n$ converges to $\mu$ in probability as $n \to \infty$. The CLT states that $(S_n - \mathbb{E}[S_n])/\sqrt{\mathrm{Var}[S_n]} = (S_n - n\mu)/\sigma\sqrt{n}$ converges to a standard normal random variable in distribution.

## 2.2 Preliminary definitions (RV, Distributions, independence)

First a trial, experiment or operation is defined. The result of any trial is an outcome, which is one of a specified set of possibilities. This set is called the *sample space* and it is denoted by $\Omega$. Therefore, every *element* $\omega \in \Omega$ is a possible outcome of the trial (experiment). For example, we can think of the trial as the tossing of a coin, for which the sample space would be $\Omega = \{\text{head, tail}\}$, or we can think of the life span of a light bulb and then $\Omega = [0, \infty)$ . An *event* $A$ is any subset of $\Omega$, which is indicated by writing $A \subseteq \Omega$.

Let us now suppose that an experiment, whose sample space is $\Omega$, is repeatedly performed under exactly the same conditions. For each event $A$, where $A \subset \Omega$, an integer $n(A)$ is defined to be the number of times that the event $A$ occurred after the first $n$ repetitions of the experiment. Then, the probability of the event $A$, $\mathbb{P}(A)$, is defined by

$$\mathbb{P}(A) = \lim_{n\to\infty} \frac{n(A)}{n} = \lim_{\#\text{runs} \to\infty} \frac{\#\text{ favorable runs}}{\#\text{ runs}}$$

known as the *relative frequency* definition of probability.

### 2.2.1 Probability measures

In order to define the concept of a probability measure, first we must define the concept of $\sigma-$algebra.

A $\sigma-$algebra $\mathcal{F}$ is a collection of subsets of the sample space $\Omega$ such that the following properties hold

- The sample space belongs to $\mathcal{F}$, i.e.$\Omega \in \mathcal{F}$.
- If $A$ is an element of $\mathcal{F}$ then the complement of $A$,(i.e. $A\cup A^C = \Omega$ and $A\cap A^C = \emptyset$) , denoted by $A^c = \Omega/A$ is also in $\mathcal{F}$.
- For any collection of elements of $\mathcal{F}$, say $\{A_i\}_{i\geq 1} \subset \mathcal{F}$, the union of this subsets is also an element in $\mathcal{F}$. In other words, $\bigcup_{i\geq 1} A_i \in \mathcal{F}$.

Any element of $\mathcal{F}$ is called an **event**. For every event $A$, where $A \subset \Omega$, it is assumed that the number $\mathbb{P}(A)$ is defined and satisfies the following axioms:

1) $\mathbb{P}(\Omega) = 1$.

2) $0 \leq \mathbb{P}(A) \leq 1$, for every $A \in \mathcal{F}$.

3) For any sequence of mutually disjoint events $\{A_i\}_{i\geq 1} \subset \mathcal{F}$

$$\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

where $A_i \bigcap A_j = \emptyset$ and $i \neq j$.

Axiom 3 states that the probability of at least one of the events of any sequence of mutually disjoint events occurring is just the sum of their respective probabilities. Suppose $\Omega$ is the sample space, $\mathcal{F}$ is a $\sigma-$algebra of events of $\Omega$ and $\mathbb{P}$ is a probability measure. The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is called a 'probability space'. By taking into consideration the three axioms, it is easy to prove the following proposition

**Proposition 2.2.1.** *For any event $A \in \mathcal{F}$ $\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$, where $A \subset \Omega$ and $A^C$ is its complement .*

*Proof.* $1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^C) = \mathbb{P}(A) + \mathbb{P}(A^C) \quad \Rightarrow \quad \mathbb{P}(A^C) = 1 - \mathbb{P}(A).$ □

**Proposition 2.2.2.** *If $A \subset B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$.*

*Proof.* $A \subset B \quad \Rightarrow \quad B = A \cup (A^C \cap B) \quad \Rightarrow \quad \mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(A^C \cap B).$ □

**Proposition 2.2.3.** $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

*Proof.* Draw the Venn Diagrams!

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \cup (A^C \cap B)) = \mathbb{P}(A) + \mathbb{P}(A^C \cap B).$$

By taking into consideration that $B = (A \cap B) \cup (A^C \cap B)$ and axiom 3, we obtain

$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A^C \cap B) \quad \Rightarrow \quad \mathbb{P}(A^C \cap B) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

□

**Example 2.2.4.** *Let us calculate the probability that any of the events $A$, $B$ and $C$ occurs:*

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}((A \cup B) \cup C) = \mathbb{P}(A \cup B) + \mathbb{P}(C) - \mathbb{P}((A \cup B) \cap C).$$

*By taking into consideration the preceding equation and that*

$$\mathbb{P}((A \cup B) \cap C) = \mathbb{P}((A \cap C) \cup (B \cap C))$$

*(try to draw the corresponding Venn diagram), we obtain that*

$$\begin{aligned}
\mathbb{P}(A \cup B \cup C) &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) + \mathbb{P}(C) - \mathbb{P}((A \cap C) \cup (B \cap C)) \\
&= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) \\
&+ \mathbb{P}((A \cap C) \cap (B \cap C)) \\
&= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) \\
&+ \mathbb{P}(A \cap B \cap C).
\end{aligned}$$

Another important concept in probability theory is the notion of **conditional probability**. It is often the case that additional information concerning the result of an experiment is available. As a result, we can "reduce" the sample space to all the events that have a common intersection with the event defining such "an information".

Therefore we must consider a new probability measure $\mathbb{P}^\star$ this is called a conditional probability. If the 'information' known to happen can be described by an event $B \in \mathcal{F}$ then such a probability measure $\mathbb{P}^\star$ is given by.

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \qquad \text{if } \mathbb{P}(B) > 0.$$

**Example 2.2.5.** *A coin is flipped twice. The sample space Ω has four points, Ω = { (H, H), (H, T), (T, H), (T, T)}, and all of them are equally likely to occur. What is the probability that both flips result in heads, given that the first flip does?*

*If A = {(H, H)} denotes the event that both flips result in heads and B = { (H, H), (H, T)} the event that the first flip results in heads. Then, the conditional probability is given by*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{(H,H)\} \cap \{(H,H),(H,T)\})}{\mathbb{P}(\{(H,H),(H,T)\})} = \frac{\frac{1}{4}}{\frac{2}{4}} = \frac{1}{2}.$$

### 2.2.2 Random variables

Fix a sample space Ω and a collection of random events $\mathcal{F}$. A **random variable** is a real-valued function defined on the sample space Ω whose outcomes can be fully described by our collection of random events $\mathcal{F}$.

**Example 2.2.6.** *Let us consider the case where a fair coin is flipped 3 times. Let $X$ denote the number of heads appearing. In this case*

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

*and if the events $\mathcal{F}$ are the possible subsets of Ω (i.e. the 'power set' of Ω) then we can know if $X = 2$ is described as a set that belongs to $\mathcal{F}$.*

*Then, $X$ is a random variable taking the values 0, 1, 2 and 3. Since the value of the random variable is determined by the outcome of the experiment, probabilities are assigned to its possible values. Thus, we calculate:*

$$\begin{aligned}
\mathbb{P}(X = 0) &= \mathbb{P}(\{(T,T,T)\}) = \frac{1}{8};\\
\mathbb{P}(X = 1) &= \mathbb{P}(\{(T,T,H),(T,H,T),(H,T,T)\}) = \frac{3}{8};\\
\mathbb{P}(X = 2) &= \mathbb{P}(\{(T,H,H),(H,H,T),(H,T,H)\}) = \frac{3}{8};\\
\mathbb{P}(X = 3) &= \mathbb{P}(\{(H,H,H)\}) = \frac{1}{8}.
\end{aligned}$$

*Note that*

$$\mathbb{P}(\bigcup_{i=0}^{3}[X = i]) = \sum_{i=0}^{3} \mathbb{P}(X = i) = 1$$

*since we take into consideration all the possible values of the random variable.*

**Example 2.2.7.** *Let us now consider the case where a series of independent trials, consisting of the flipping of a coin, takes place until either a head ($H$) occurs or the coin is flipped $n$ times. Let us assume that $\mathbb{P}(H) = p$ and $\mathbb{P}(T) = 1 - p$ (the case of not having a fair coin, $p \neq 0.5$, is included). Then, $X$, which denotes the number of times the coin is flipped, is a random variable taking on one of the values 1, 2, 3, . . ., $n$ with respective* *probabilities:*

$$
\begin{aligned}
\mathbb{P}(X=1) &= \mathbb{P}(\{H\}) = p \\
\mathbb{P}(X=2) &= \mathbb{P}(\{T,H\}) = (1-p)p \\
\mathbb{P}(X=3) &= \mathbb{P}(\{T,T,H\}) = (1-p)^2 p \\
&\vdots \\
\mathbb{P}(X=n-1) &= \mathbb{P}(\{\underbrace{T,T,\ldots,T}_{n-2},H\}) = (1-p)^{n-2}p \\
\mathbb{P}(X=n) &= \mathbb{P}(\{\underbrace{T,T,\ldots,T}_{n-1},T\},\{\underbrace{T,T,\ldots,T}_{n-1},H\}) = (1-p)^{n-1}
\end{aligned}
$$

*Note that*

$$\mathbb{P}(\bigcup_{i=1}^{n}\{X=i\}) = \sum_{i=1}^{n}\mathbb{P}(X=i) = 1$$

*since*

$$
\begin{aligned}
\sum_{i=1}^{n}\mathbb{P}(X=i) &= \sum_{i=1}^{n-1}(1-p)^{i-1}p + (1-p)^{n-1} \\
&= p\frac{1-(1-p)^{n-1}}{1-(1-p)} + (1-p)^{n-1} = 1.
\end{aligned}
$$

The two previous examples refer to **discrete random variables**, i.e. random variables whose set of possible values is either finite or countably infinite. However, there also exist random variables whose set of possible values is uncountable.

An important family of random variables that are not discrete, is the family of **continuous random variables**. They will be defined in the next section.

### 2.2.3 Distributions

As before, we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The **cumulative distribution function (c.d.f.)**, or more simply the *distribution function* $F$, of the random variable $X$ is defined by

$$F_X(b) = \mathbb{P}(X \leq b), \qquad \forall\, b \in \mathbb{R}.$$

In other words, $F_X(b)$ denotes the probability that the random variable $X$ takes on a value that is less or equal to $b$.

Important properties of the cumulative distribution function are the following:

1) $\mathcal{F}_X$ is a non-decreasing function (i.e. $\forall a < b \Rightarrow F_X(a) \leq F_X(b)$).

2) $\lim_{b\to\infty} F_X(b) = 1$ and $\lim_{b\to-\infty} F_X(b) = 0$.

3) $F_X$ is right continuous (i.e. for any real $b$ and any decreasing sequence $b_n$, $n \geq 1$, that converges to b, we obtain $\lim_{n\to\infty} F_X(b_n) = F_X(b)$).

**Example 2.2.8.** *Let us assume that the distribution function of the variable $X$ is given by*

$$F_X(x) = \begin{cases} 0, & x < 0, \\ \frac{x}{2}, & 0 \leq x < 1, \\ \frac{2}{3}, & 1 \leq x < 2, \\ \frac{11}{12}, & 2 \leq x < 3, \\ 1, & x \geq 3. \end{cases}$$

*By drawing the graph of $F(x)$, the properties of the distribution function can be seen. Check the calculation of the following probabilities:*

$$(a) \quad \mathbb{P}(X < 3) = \frac{11}{12} \quad \textit{and} \quad \mathbb{P}(X = 3) = \frac{1}{12}.$$

$$(b) \quad \mathbb{P}(X < 2.5) = \frac{11}{12} \quad \textit{and} \quad \mathbb{P}(X = 2.5) = 0.$$

$$(c) \quad \mathbb{P}(X > \frac{1}{2}) = 1 - \mathbb{P}(X \leq \frac{1}{2}) = 1 - F(\frac{1}{2}) = 1 - \frac{1}{4} = \frac{3}{4}.$$

$$(d) \quad \mathbb{P}(2 < X \leq 4) = F(4) - F(2) = \frac{1}{12}.$$

The expectation or the mean of a *discrete random variable $X$*, denoted $\mathbb{E}[X]$, is defined by

$$\mathbb{E}[X] = \sum_x xp(x)$$

where the sum is over all possible values $x$ of $X$ and $p(x) = \mathbb{P}(X = x)$ denotes its corresponding **probability density function** (or **probability mass function**). For example, the mean for $X$ = "the sum of the numbers when two dice are thrown" is given by

$$\mathbb{E}[X] = \sum_{x=2}^{12} x\mathbb{P}(X = x) = 2\frac{1}{36} + 3\frac{2}{36} + \ldots + 12\frac{1}{36} = 7.$$

**Remark 2.2.9.** *Moreover, the expectation of a real-valued function $g(X) : \Omega \to \mathbb{R}$, where $X$ is a discrete random variable that takes one of the values $x_k$, $k \geq 1$, with respective probabilities $p(x_k) = \mathbb{P}(X = x_i)$, is given by*

$$\mathbb{E}[g(X)] = \sum_i g(x_i)p(x_i)$$

**Example 2.2.10.** *Let $X$ denote a discrete random variable such that $\mathbb{P}(X = 0)$ = 0.1, $\mathbb{P}(X = 1)$ = 0.1, $\mathbb{P}(X = 2)$ = 0.6, $\mathbb{P}(X = 3)$ = 0.1 and $\mathbb{P}(X = 4)$ = 0.1 (Note that $\sum_{k=0}^{4} \mathbb{P}(X = k)$ = 1). Then*

$$\mathbb{E}[X^2] = \sum_{k=0}^{4} k^2\mathbb{P}(X = k) = 0^2 0.1 + 1^2 0.1 + 2^2 0.6 + 3^2 0.1 + 4^2 0.1 = 5$$

As a result, the following properties are obtained:

(1) $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$ for any real constants $a$ and $b$ and any random variables $X$ and $Y$.

(2) If $X = c$, where $c \in \mathbb{R}$, i.e. $X(\omega)$ is constant, for every $\omega \in \Omega$, then $\mathbb{E}[X] = c$.

A random variable $X$ is called *continuous* if

- there exists a non-negative function $f$, defined for all $x \in \mathbb{R}$, such that

$$\mathbb{P}(X \in A) = \int_A f(x)dx$$

for any (measurable) set $A$ of real numbers.

- Such a function $f$ must have an integral on $\mathbb{R}$ equal to the probability of the sample space $\mathbb{P}(\Omega) = 1$.

$$\int_{\mathbb{R}} f(x)dx = 1.$$

The function $f$ is called the **probability density function (p.d.f.)** of the random variable $X$. Note that the p.d.f. plays the role that is played for discrete random variables by the probability distribution $\mathbb{P}(X = x)$. Consequently, $f$ must satisfy:

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx. \tag{2.2.1}$$

In particular, notice that, if $X$ is a continuous random variable, 2.2.1 implies that $\mathbb{P}(X = a) = 0$ for any $a \in \mathbb{R}$.

The corresponding definition of the expected value for a continuous random variable $X$ is given by

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)dx$$

Moreover, the expectation of a real-valued function $g(X) : \Omega \to \mathbb{R}$, where $X$ is a continuous random variable with probability density function $f(x)$, is given by

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x$$

We define by $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ the (vector) space of functions such that $\mathbb{E}[X^2] < \infty$. The **variance** of a random variable $X \in \mathcal{L}^2(\Omega, \mathcal{F}\,\mathbb{P})$, denoted by $Var[X]$, with mean $\mu = \mathbb{E}[X]$ is given by

$$\text{Var}[X] = \mathbb{E}[(X - \mu)^2].$$

Some useful properties of the variance of a random variable $X \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$:

$$\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2, \tag{2.2.2}$$

$$\text{Var}[aX + b] = a^2\text{Var}[X]. \tag{2.2.3}$$

### 2.2.4 Independence

Two events $A$ and $B$ are said to be independent if:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Taking into consideration the formula for the conditional probability, the following result is obtained for the case of two independent events $A$ and $B$:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(B)} = \mathbb{P}(A)$$

(i.e. the occurrence of the event $B$ does not affect the probability of the occurrence for the event $A$!)

If two events $A$ and $B$ are independent, then $A$ and $B^C$ are also independent (*hint:* $A = (A \cap B) \cup (A \cap B^C)$).

The number $\mathbb{E}[X^p]$ ($p > 0$) is called the $p$-th moment of $X$. The family of all random variables whose $p$-th moment is finite is denoted by $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$.
Two discrete random variables are said to be *independent* if, for all $x$ and $y$ in their respective ranges, the events $X = x$ and $Y = y$ are independent:

$$\mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

Just as in the case of real-valued random variables, a **vector**-valued random variable is a function $X : \Omega \to \mathbb{R}^d$.

### 2.2.5 Joint distribution

And the definition of the **joint distribution** is given by

$$F_{X,Y}(x, y) = \mathbb{P}(X \leq x, Y \leq y),$$

And the joint probability density is a function $f(x, y)$ such that

$$F_{X,Y}(x, y) = \int_A \int_B f_{X,Y}(x, y)\, dxdy$$

Two continuous random variables, $X$ and $Y$ are *independent* if and only if for 'any' $A$, $B \subset \mathbb{R}$

$$\mathbb{P}(X \in A Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B),$$

Moreover, $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x, y) = f_X(x) f_Y(y).$$

Moreover, $X$ and $Y$ are two independent random variables then

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] \quad \text{and} \quad \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$$

(which holds no matter if they are continuous or discrete).

The **covariance** of two random variables $X$ and $Y$, denoted by $Cov(X, Y)$, with corresponding expected values $\mu$ and $\nu$ is given by

$$Cov(X, Y) = \mathbb{E}[(X - \mu)(Y - \nu)]$$

Note that in case of two independent random variables $X$ and $Y$ their covariance is equal to zero. However, the opposite is not necessarily true!

## 2.3 Binomial Distribution

Let us consider a trial or an experiment having two possible outcomes, usually they are called "success" and "failure", with corresponding probabilities

$$\mathbb{P}(X = \text{"success"}) = p \quad \text{and} \quad \mathbb{P}(X = \text{"failure"}) = 1 - p,$$

where $0 \leq p \leq 1$ of course.

Now, let us consider the case where $n$ of these independent trials are to be performed. Then, the discrete random variable $X$ = "number of successes that occur in the $n$ trials" is called **binomial** with parameters $(n, p)$. The corresponding probability mass function is given by

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \text{where } k = 0, 1, \ldots, n.$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad \text{(known as the binomial coefficient)}$$

denotes the different sequences of the $n$ outcomes leading to $k$ successes and $n-k$ failures. For example, consider the case of $n = 4$ and $k = 2$, i.e. four trials and two successes, then there are $\binom{4}{2} = 6$ ways in which the four trials can result in two successes:

$$(s, s, f, f), (s, f, s, f), (s, f, f, s), (f, s, s, f), (f, s, f, s), (f, f, s, s).$$

Then, the probability of having two successes (when $p = 0.4$) is given by

$$\mathbb{P}(X = 2) = \binom{4}{2} 0.4^2 0.6^2 = 0.3456. \quad \square$$

Note that by the binomial theorem

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

the probabilities add up to 1

$$\sum_{k=0}^{\infty} \mathbb{P}(X = k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = [p + (1-p)]^n = 1.$$

### Properties of Binomial Random variables

By taking into consideration that

$$k\binom{n}{k} = k\frac{n!}{k!(n-k)!} = n\frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!} = n\binom{n-1}{k-1}$$

the $q$-th moment can be calculated

$$\begin{aligned}
\mathbb{E}[X^q] &= \sum_{k=0}^{n} k^q \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^{n} k^q \binom{n}{k} p^k (1-p)^{n-k} \\
&= np \sum_{k=1}^{n} k^{q-1} \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} \\
&= np \sum_{j=0}^{n-1} (j+1)^{q-1} \binom{n-1}{j} p^j (1-p)^{(n-1)-j} = np\mathbb{E}[(Y+1)^{q-1}].
\end{aligned}$$

where $Y$ is a binomial random variable with parameters $n-1$ and $p$. Thus, for $q=1$, we obtain

$$\mathbb{E}[X] = np\mathbb{E}[(Y+1)^0] = np.$$

Similarly,

$$\mathbb{E}[X^2] = np\mathbb{E}[Y+1] = np[(n-1)p+1]$$

and since $\mathbb{E}[X] = np$ we obtain

$$\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = np[(n-1)p+1] - n^2p^2 = np(1-p)$$

Another interesting property concerning the relationship of consecutive probabilities:

$$\frac{\mathbb{P}(X=k+1)}{\mathbb{P}(X=k)} = \frac{\frac{n!}{(k+1)!(n-k-1)!}p^{k+1}(1-p)^{n-k-1}}{\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}} = \frac{(n-k)p}{(k+1)(1-p)}$$

Thus,

$$\mathbb{P}(X=k+1) = \frac{p}{1-p}\frac{n-k}{k+1}\mathbb{P}(X=k).$$

Another interesting way of calculating the mean and the variance of a binomial random variable $X$ is to consider that

$$X = \sum_{k=1}^{n} X_k, \qquad (k \text{ independent trials}),$$

where $X_k$ is defined to equal 1 if the trial $k$ is a "success" (with probability $p$) and to equal 0 otherwise. Then,

$$\mathbb{E}[X] = \mathbb{E}[\sum_{k=1}^{n} X_k] = \sum_{k=1}^{n}\mathbb{E}[X_k] = n[p\times 1 + (1-p)\times 0] = np$$

Similarly, for each $X_k$,

$$\text{Var}[X_k] = \mathbb{E}[(X_k - \mathbb{E}[X_k])^2] = (1-p)^2\times p + (0-p)^2\times(1-p) = p(1-p)$$

and since

$$\text{Var}[\sum_{k=1}^{n} X_k] = \sum_{k=1}^{n}\text{Var}[X_k]$$

due to independence, we obtain

$$\text{Var}[X] = \text{Var}[\sum_{k=1}^{n} X_k] = np(1-p).$$

## 2.4 Normal Random Variables

A continuous random variable $X$ is said to be **normally distributed** with parameters $\mu \in \mathbb{R}$ and $\sigma^2 > 0$, written $X \sim \mathcal{N}(\mu, \sigma)$, if the probability density function of $X$ is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty < x < \infty. \tag{2.4.1}$$

The density function is a bell-shaped curve that is symmetric about $\mu$.

**Proposition 2.4.1.** *The function $f$ defined in* (2.4.1) *is a probability density function.*

*Proof - do this to see a couple of cool math tricks.* In order to prove that $f(x)$ is indeed a probability density function, the following result should hold:

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = 1$$

By setting $y = (x-\mu)/\sigma$, the integral under consideration becomes

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy$$

Now, let

$$I = \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy$$

then

$$I^2 = \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}dz = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}e^{-\frac{z^2}{2}}dydz$$

Consequently

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{z^2+y^2}{2}}dydz$$

and by introducing polar coordinates ($z = r\cos\theta$, $y = r\sin\theta$ and $dzdy = rd\theta dr$), we obtain

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{z^2+y^2}{2}}dydz = \int_{0}^{\infty}\int_{0}^{2\pi} e^{-r^2/2}rd\theta dr = 2\pi\int_{0}^{\infty} re^{-r^2/2}dr = 2\pi$$

Therefore, $I = \sqrt{2\pi}$ and thus

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \frac{1}{\sqrt{2\pi}}I = \frac{1}{\sqrt{2\pi}}\sqrt{2\pi} = 1.$$

□

For completeness we also state the expected value and standard deviation of a normal distribution.

**Lemma 2.4.2.** *Take $\mu \in \mathbb{R}$ and $\sigma \in (0,\infty)$. Let $X$ be a Normally distributed random variable such that $X \sim \mathcal{N}(\mu,\sigma^2)$. Then, $\mathbb{E}[X] = \mu$ and Var$[X] = \sigma^2$.*

*Proof.* This proof is simple computation of the involved quantities. We start with the expected value.

$$\begin{aligned}\mathbb{E}[X] &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} xe^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-\mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx + \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\mu e^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} ye^{-\frac{y^2}{2\sigma^2}}dy + \mu\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 0 + \mu\times 1 = \mu\end{aligned}$$

and

$$\mathrm{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x$$

By making the substitution $y = (x-\mu)/\sigma$, we obtain

$$\begin{aligned}\mathrm{Var}[X] &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} \mathrm{d}y = \frac{\sigma^2}{\sqrt{2\pi}} \Big( - y e^{-\frac{y^2}{2}} |_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \mathrm{d}y \Big) \\ &= \sigma^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} = \sigma^2.\end{aligned}$$

$\square$

## Other properties

Let $X$ be a normally distributed with parameters $\mu$ and $\sigma^2$, i.e. $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Z = (X-\mu)/\sigma$ is also normally distributed with zero mean and unit variance, in other words $Z \sim \mathcal{N}(0,1)$.

The new random variable $Z$ is called *standard normal distribution* and the corresponding cumulative distribution function is denoted by $\Phi(x)$ (or sometimes by $N(x)$):

$$\Phi : \mathbb{R} \to \mathbb{R}, \qquad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} \mathrm{d}y, \qquad x \in \mathbb{R}. \tag{2.4.2}$$

### Symmetry properties

As can be seen from (2.4.1), the density function of a standard normal distribution is an *even function* and hence for any odd function $g$ it easily follows that $\mathbb{E}[g(X)] = 0$ when $X \sim \mathcal{N}(0,1)$. Can you convince yourself why such is true?

There are other symmetry properties of the normal distribution. Those stated next are fundamental for our course and will be used again and again. It is strongly recommended to the student to prove the next result.

**Proposition 2.4.3.** *Let $\phi$ and $\Phi$ be the standard normal density function* (2.4.1) *(with $\mu = 0$ and $\sigma = 1$) and distribution function* (2.4.2) *respectively.*

*Then for any $x \in \mathbb{R}$, for any $\lambda \in \mathbb{R}$ the following two identities hold*

$$\Phi(-x) = 1 - \Phi(x), \qquad \textit{and} \qquad \Phi(\lambda x) = 1 - \Phi(-\lambda x).$$

*Proof.* The first statement is rather easy and can be proved directly using that $\phi$ is an even function. Let $x \in \mathbb{R}$, then

$$\Phi(-x) = \int_{-\infty}^{-x} \phi(y) \mathrm{d}y = \int_{-\infty}^{+\infty} \phi(y) \mathrm{d}y - \int_{-x}^{+\infty} \phi(y) \mathrm{d}y.$$

The first integral in the RHS is equal to one. For the second integral one uses the change of variables formula with $u = -y$. This yields

$$\Phi(-x) = 1 - \int_{x}^{-\infty} \phi(-u)(-1) \mathrm{d}u = 1 - \int_{-\infty}^{x} \phi(u) \mathrm{d}u = 1 - \Phi(x).$$

The second is slightly more complicated and is left for the assignment sheets O_O! $\square$

## 2.5 Limit theorems (LLN and CLT) and transformations

### 2.5.1 Laws of large numbers

An important result in probability theory, known as the DeMoivre-Laplace limit theorem, states that when $n$ is large, a binomial random variable with parameters $n$ and $p$ will have approximately the same distribution as a normal random variable with the same mean and variance as the binomial.

**Theorem 2.5.1** (The DeMoivre-Laplace limit theorem). *If $X_n$ denotes the number of successes that occur when $n$ independent trials, each resulting in a success with probability $p$, are performed then, for any $a < b$,*

$$\lim_{n\to\infty} \mathbb{P}\left[a \leq \frac{X_n - np}{\sqrt{np(1-p)}} \leq b\right] = \Phi(b) - \Phi(a).$$

The above theorem is only a special case of the famous central limit theorem. The latter is one of the most remarkable results in probability theory. Amongst other things, it help us explain the fact that the empirical frequencies of so many natural populations exhibit bell-shaped (i.e. normal) curves. In its simplest form the central limit theorem is as follows.

**Theorem 2.5.2** (The central limit theorem (CLT)). *Let $X_1, X_2, \cdots, X_n, \cdot$ be a sequence of independent and identically distributed random variables each having finite mean $\mu$ and variance $\sigma^2$. Then, the distribution of the RV*

$$Z = \frac{1}{\sigma\sqrt{n}}\left(\sum_{k=1}^{n} X_k - n\mu\right)$$

*converges to the standard normal distribution as $n \to \infty$. That is, for any $a \in (-\infty, \infty)$,*

$$\lim_{n\to\infty} \mathbb{P}[Z \leq a] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}}\, dx.$$

Note that this theorem holds for *any* distribution of the $X_i$'s as long as $\mathbb{E}[X] < \infty$, Var$[X] < \infty$; herein lies its power.

### 2.5.2 Transformation of random variables and their density functions

**Theorem 2.5.3.** *Let $X$ be a continuous random variable having probability density function $f_X$. Suppose that $g(x)$ is a strictly monotone (increasing or decreasing), differentiable (and thus continuous) function of $x$. Then, the random variable $Y$ defined by $Y = g(X)$ has a probability density function given by*

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y))|\frac{d}{dy}g^{-1}(y)|, & \text{if } y = g(x) \text{ for some } x, \\ 0, & \text{if } y \neq g(x) \text{ for all } x, \end{cases}$$

*where $g^{-1}(y)$ is defined to equal that value of $x$ such that $g(x) = y$.*

*Proof.* Suppose that $g(x)$ is an increasing function and let $y \in \text{Range}(g)$ then there exists $x$ such that $y = g(x)$. Then, with $Y = g(X)$,

$$F_Y(y) = \mathbb{P}[g(X) \le y] = \mathbb{P}[X \le g^{-1}(y)] = F_X\big(g^{-1}(y)\big).$$

The $pdf$ of $Y$ is obtained by differentiating the distribution function $F_Y(y)$ with respect to $y$. Since $F_Y$ written as a function of $F_X$ involves a composition of functions, one must use the chain rule. Let $y \in \text{Range}(g)$, then

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}\left[F_X\big(g^{-1}(y)\big)\right] \\ &= \left(\frac{d}{dx}F_X(x)\right)\Big|_{x=(g^{-1}(y))} \cdot \frac{d}{dy}\Big(g^{-1}(y)\Big) = f_X\Big(g^{-1}(y)\Big) \cdot \frac{d}{dy}\Big(g^{-1}(y)\Big). \end{aligned}$$

□

A common application example of the above result is the derivation of the distribution of a log-normal from a normal one.

**Example 2.5.4** (The log-normal distribution). *A random variable $Y$ is said to have a lognormal distribution if the random variable $X$ defined as $X = \ln Y$ has a normal distribution with mean $\mu$ and variance $\sigma^2$.*

*Goal: Derive the probability density function $Y$. We denote it by $f_Y(\cdot)$.*

*From the above result, we want to compute the probability density function of $Y$, knowing that the pdf of the RV $X$, denoted $f_X$, is given by* (2.4.1).

*Define for any $x \in \mathbb{R}$ the function $g(x) = e^x$. We have that $Y = e^X = g(X)$ and the map $x \mapsto g(x) = e^x$ is a strictly increasing differentiable one. Therefore, we fulfill all the conditions of the above theorem. Note that for any $y \in (0, \infty)$ we have $g^{-1}(y) = \ln(y)$. Hence for any $y \in (0, \infty)$ we have*

$$f_Y(y) = f_X\Big(g^{-1}(y)\Big) \cdot \frac{d}{dy}\Big(g^{-1}(y)\Big) = f_X\Big(\ln(y)\Big) \cdot \frac{1}{y} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}}.$$

## 2.6 Exercises

### Discrete random variables

**Exercise 2.6.1.** *Let $A$ and $B$ be events with some probability.*

*(a) Which order relation, $\le$, $=$ or $\ge$, should be placed in the below boxes.*

*Can you convince your colleagues?*

$$\mathbb{P}(A \cap B) \boxed{\phantom{\le}} \min\{\mathbb{P}(A), \mathbb{P}(B)\}; \qquad \mathbb{P}(A \cup B) \boxed{\phantom{\le}} \max\{\mathbb{P}(A), \mathbb{P}(B)\}.$$

*(b) Assume $\mathbb{P}(A) = \frac{3}{4}$ and $\mathbb{P}(B) = \frac{1}{3}$.*

*1) Show that $\frac{1}{12} \le \mathbb{P}(A \cap B) \le \frac{1}{3}$.*

*2) Find corresponding upper and lower bounds for $\mathbb{P}(A \cup B)$.*

**Exercise 2.6.2.** *For a nonnegative integer-valued random variable $X$, show that*

$$\mathbb{E}[X] = \sum_{j=1}^{\infty} \mathbb{P}[X \geq j].$$

**Exercise 2.6.3.** *For a nonnegative integer-valued random variable $X$, show that*

$$\sum_{k=0}^{\infty} k\mathbb{P}[X > k] = \frac{1}{2}(\mathbb{E}[X^2] - \mathbb{E}[X]).$$

**Exercise 2.6.4** (Central limit theorem)**.** *Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables for which $\mathbb{E}[\frac{1}{X_1}]$ exists. Show that, if $m \leq n$, then $\mathbb{E}[\frac{S_m}{S_n}] = \frac{m}{n}$, where $S_n = X_1 + X_2 + \ldots + X_n$.*

*Tip: prove that $\mathbb{E}[\frac{X_1}{S_n}] = \frac{1}{n}$.*

## Continuous random variables [VERY IMPORTANT SECTION]

**Exercise 2.6.5.** *Let $X$ be a discrete random variable. Show that if $\mathbb{E}[X^2] = 0$ then $\mathbb{P}(X = 0) = 1$. Furthermore, show that if $Var[X] = 0$ then there exists $c \in \mathbb{R}$ such that $\mathbb{P}(X = c) = 1$, i.e. $X$ is almost surely constant.*

**Exercise 2.6.6.** *Let $X$ be a random variable that takes on values between 0 and $c$ (i.e. $\mathbb{P}[0 \leq X \leq c] = 1$). Show that: (a) $\mathbb{E}[X^2] \leq c\mathbb{E}[X]$. (b) Var$[X] \leq \frac{c^2}{4}$.*

**Exercise 2.6.7.** *Show that the continuous random variables $X$ and $-X$ have the same distribution function if and only if $f_X(x) = f_X(-x)$ for all $x \in \mathbb{R}$.*

**Exercise 2.6.8** (Normal distribution)**.** *Let $X$ be a normal random variable. Prove that for any $\alpha \in \mathbb{R} \setminus \{0\}$ and $\beta \in \mathbb{R}$, the random variable $\alpha X + \beta$ is also normally distributed with parameters $\alpha\mathbb{E}[X] + \beta$ and $\alpha^2$Var$[X]$.*

**Exercise 2.6.9** (Normal distribution)**.** *Let $X$ be a normally distributed random variable with mean $\mu$ and variance $\sigma^2$ and $g : \mathbb{R} \to \mathbb{R}$ is a differentiable and bounded function (i.e., $|g(x)| \leq K$ for all $x \in \mathbb{R}$, where $K$ is a constant). Show that $\mathbb{E}[(X - \mu)g(X)] = \sigma^2\mathbb{E}[g'(X)]$.*

**Exercise 2.6.10** (Log-Normal)**.** *A random variable $Y$ is said to have a lognormal distribution if the random variable $X = \ln Y$ has a normal distribution with mean $\mu$ and variance $\sigma^2$. Derive the probability density function for $Y$ and show that $\mathbb{E}[Y^n] = e^{n\mu + n^2\frac{\sigma^2}{2}}$.*

**Exercise 2.6.11.** *Let $\phi$ and $\Phi$ be the standard normal density and distribution functions respectively. Show that:*

*(a) $\Phi(\lambda x) = 1 - \Phi(-\lambda x)$*

*(b) $\mathbb{E}[e^{\mu+\sigma Z}\mathbb{1}_{\{Z>-d\}}] = e^{\mu+\frac{\sigma^2}{2}}\Phi(d + \sigma)$, where $Z \sim \mathcal{N}(0, 1)$ (standard normal distribution) and $\mathbb{1}$ denotes the indicator function of an event:*

$$\mathbb{1}_{\{Z>-d\}} = \begin{cases} 1, & \textit{when } Z > -d, \\ 0, & \textit{otherwise.} \end{cases}$$

# Chapter 3

# Risk-neutral pricing

In this chapter we build on the material of Chapter 1 and assume this time a model for the stock process. The main themes here are the so-called *pricing tree models*, risk-neutral pricing and risk free portfolios.

In this section we define the notion of *arbitrage* and *portfolio* within the context of a financial market that consists of a risky asset $S$ and a riskless asset "bond" (or bank account) between *two points in time* (over a single time period). We denote the passing of the time, in periods of length $\Delta t > 0$, as $n = 0, 1, \ldots$ and the passing of time in hours/days/months/years by the letter $t_n = n\Delta t$.

## 3.1 The one-period model

The market consists of a stock and a bank account over a single period, $n \in \{0, 1\}$. The bank account pays interests discretely at a rate $r$ over the time period.

The stock price is a stochastic process[1] $(S_n)_{n=0,1}$ where $S_0$ is a positive real number and $S_1$ is a random variable of Bernoulli type: in fact, we will use the Bernoulli RV $Z \sim \text{Ber}(p_u)$ namely $Z : \Omega \to \{u, d\}$ ($u, d \in \mathbb{R}$ are the possible outcomes) and $\mathbb{P}[Z = u] = p_u$, $\mathbb{P}[Z = d] = p_d$ such that $p_u + p_d = 1$ to help define $S_1$, see Figure 3.1.1.

We always assume that the price of the stock today is known as are the positive constants $d, u, p_u, p_d$. We assume that $d < u$ and, of course, that $p_d + p_u = 1$. We refer to the above described market model as the **one-period model** or the **one-period binomial model**.

For this market, we make assumptions as in Chapter 1 (see Section 1.2).

### Portfolios and arbitrage

Given the above described market, an investor can always deposit/borrow money from the bank as well as buying/selling stocks. We formalize this in the next definition.

**Definition 3.1.1** (Trading strategy or portfolio in the one-period model)**.** *Let* $x_0, \Delta_0 \in \mathbb{R}$*. The value* $x_0$ *denotes the number of units of money in the "bank account" or money*

[1] A stochastic process is roughly a random variable that depends on a time variable.

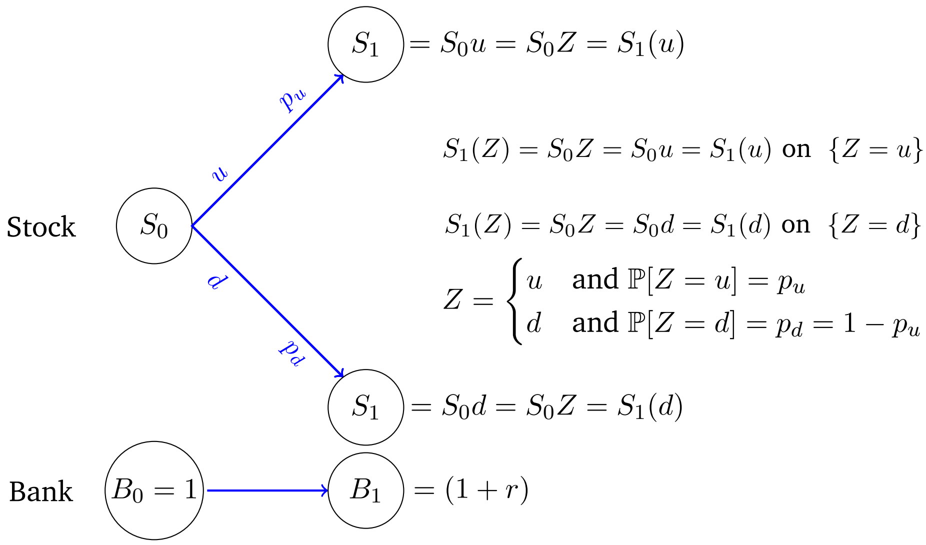



Figure 3.1.1: Evolution of Stock price and Bank deposit in the one-period model.

*invested in a riskless bond at time $t = 0$ (or period $n = 0$). The value $\Delta_0$ denotes the number of units in the stock at time $t = 0$ (or period $n = 0$).*

*The vector $h := (x_0, \Delta_0) \in \mathbb{R}^2$ is called* portfolio *or* strategy *or* trading strategy.

For example, take the portfolio $h = (0.25, -2)$. This means that the investor as short sold two units of stock and has $0.25$ units of money deposited in the bank at time $t = 0$.

**Exercise 3.1.2.** *Can you quickly compute how much money the portfolio is worth at $n = 1$?*

**Definition 3.1.3** (Value process or Wealth process of a strategy)**.** *Suppose $h = (x_0, \Delta_0)$ is a portfolio. We define the* ***value process/wealth process*** *of portolio $h$ to be*

$$V_n^h := x_0 B_n + \Delta_0 S_n, \qquad n = 0, 1.$$

Writing the wealth process explicitly at $n = 0$ and $n = 1$ (beggining and end of the period) we have:

$$V_0^h = x_0 + \Delta_0 S_0 \qquad \text{and} \qquad V_1^h(Z) = x_0(1 + r) + \Delta_0 S_1(Z).$$

Note that the wealth process is deterministic (a non-random number) at time zero and a random variable at the end of the period. $V_1^h(Z)$ depends on the future value of the stock $S_1(Z)$, in fact, $V_1^h(Z)$ is still Bernoulli distributed but its mean and standard deviation have changed (not those of $Z$).

**Exercise 3.1.4.** *Is $V_n^h$ at $n = 0$ random or deterministic? and at $n = 1$? If so, compute $\mathbb{E}[V_n^h]$ and Var$[V_n^h]$ (recall that $Z$ is Bernoulli distributed).*

As in Chapter 1, everyone wants to make a profit by trading and hence we define Arbitrage next.

**Definition 3.1.5** (Arbitrage)**.** *Take $h = (x_0, \Delta_0) \in \mathbb{R}^2$ to be a portfolio and let $V^h$ be the associated wealth process. We say $h$ is an* ***arbitrage*** *portfolio if the following conditions hold,* *i)* $V_0^h = 0$.

*ii)* $V_1^h \geq 0$ *with probability* $1$ *(in other words* $\mathbb{P}[V_1^h \geq 0] = 1$*) and* $\mathbb{P}[V_1^h > 0] > 0$.

*If no such portfolio* $h$ *exists, we say that the market* ***free of arbitrage****.*

How can one interpret the above definition? Notice that in our setting this is equivalent to have $V_1^h(Z) \geq 0$ and $V_1^h(Z) > 0$ on at least one of the events $\{\omega : Z(\omega) = u\}$ or $\{\omega : Z(\omega) = d\}$.

An arbitrage portfolio is thus basically a deterministic money making machine, and we interpret the existence of an arbitrage portfolio as equivalent to a serious case of mispricing on the market. It is now natural to investigate when a given market model is arbitrage free, i.e. when there are no arbitrage portfolios.

**Proposition 3.1.6.** *Assume* $d < u$. *The 1-period model is* ***free of arbitrage*** *if and only if* $d \leq 1 + r \leq u$.

*Proof.* Check notes in the lecture and the exercises of Assignment 2. $\square$

Proposition 3.1.6 states simply that in order to have a market free of arbitrage the return of the "bank account" *cannot dominate* the return of the stock and vice versa.

## The risk-neutral measure

The condition $d \leq 1+r \leq u$ in Proposition 3.1.6 is actually very powerful. Since $1+r$ lies between $u$ and $d$, by convex combination[2], there exists we can write

$$\exists q_u, q_d \in [0,1] \quad \text{s.th.} \quad 1 + r = q_u \cdot u + q_d \cdot d \quad \text{and where } q_u + q_d = 1.$$

**Example 3.1.7.** *How to compute* $q_u, q_d$*? The above statement reveals two equations with two unknowns. In other words, we solve the following* $2 \times 2$*-system*

$$\begin{cases} q_u + q_d = 1 \\ q_u \cdot u + q_d \cdot d = 1 + r \end{cases} \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 1 \\ u & d \end{bmatrix} \begin{bmatrix} q_u \\ q_d \end{bmatrix} = \begin{bmatrix} 1 \\ 1+r \end{bmatrix}.$$

*The matrix is invertible as long as* $\det(\textit{matrix}) = u - d \neq 0$. *By assumption* $d < u$ *and hence the system is solvable with solutions given by*

$$q_u = \frac{(1+r) - d}{u - d} \qquad \textit{and} \qquad q_d = \frac{u - (1+r)}{u - d}.$$

Moving away from the computations of $q_u$ and $q_d$ itself. These weights $q_u$ and $q_d$ can be interpreted as a new probability measure $\mathbb{Q}$ on $\Omega$ (from a Bernoulli distribution) with the property that $\mathbb{Q}[Z = u] = q_u$ and $\mathbb{Q}[Z = d] = q_d$. This new probability measure $\mathbb{Q}$ seems to have appeared quite casually without true implications.

[2]This operation is called *convex combination*. Given $N$ numbers $\{x_i\}_{i=1,\cdots,N}$ and $N$ weights $\{w_i\}_{i=1,\cdots,N}$ such that $\sum_{i=1}^N w_i = 1$ then $\hat{x} := \sum_{i=1}^N w_i x_i$ is the convex combination of the $x_i$'s according to the weights $w_i$. If you think on it for a second, the expected value $\mathbb{E}[X]$ of a discrete rv $X$ is a convex combination on the possible values $X$ can take (do write it down to see).

Let us compute the expected price of the Stock,

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[S_1] &= S_1(u)\mathbb{Q}[Z=u] + S_1(d)\mathbb{Q}[Z=d] \\ &= S_0 u \cdot q_u + S_0 d \cdot q_d \\ &= S_0\left(1+r\right).\end{aligned}$$

In other words, $S_0 = (1+r)^{-1}\mathbb{E}^{\mathbb{Q}}\left[S_1\right]$. The value today of the stock is the expected discounted value tomorrow under the probability measure $\mathbb{Q}$. This just stated rule is very intuitive for economists.

Why do we call $\mathbb{Q}$ Risk-neutral? Let us go back to the above equation:

$$S_0(1+r) = \mathbb{E}^{\mathbb{Q}}[S_1] = S_0\mathbb{E}^{\mathbb{Q}}[Z].$$

On the left hand side, we have the wealth generated by depositing $S_0$ in the bank and letting interests accumulate. On the right hand side we have $S_0\mathbb{E}^{\mathbb{Q}}[Z]$ which is how much we expect to gain in the stock market by having a long one-unit position in the stock.

Because there is an equality, this means that what we expect to gain in the stock market is the same as depositing the money in the bank account. And hence, in terms of an investment, we are indifferent between the bank account or the stock. We are neutral.

In practice, we can either make money (stock goes up) or lose it (stock goes down) but in terms of "expected" gains, the investments are the same under $\mathbb{Q}$.

For completeness, notice that under $\mathbb{P}$

$$\mathbb{E}^{\mathbb{P}}[S_1] = S_0 u \cdot p_u + S_0 d \cdot p_d \neq \mathbb{E}^{\mathbb{Q}}[S_1] \quad \text{in general.}$$

**Definition 3.1.8.** *A probability measure $\mathbb{Q}$ is called the **martingale measure, risk neutral measure** or even **risk adjusted measure** if the following condition is satisfied*

$$S_0 = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}\left[S_1\right].$$

We can now state Proposition 3.1.6 in an alternative way.

**Proposition 3.1.9.** *Consider the 1-period market model. Then the following statements are equivalent.*

*i) The market model is arbitrage free*

*ii) There exists a risk neutral measure $\mathbb{Q}$.*

*iii) $d \leq 1+r \leq u$.*

And for the one-period (and for the multi-period in the sections below) market model it is very easy to compute the risk-neutral probabilities.

**Proposition 3.1.10.** *For the 1-period 'binomial market model' the probability measure $\mathbb{Q}$ defined by*

$$q_u = \frac{(1+r)-d}{u-d} \qquad \textit{and} \qquad q_d = \frac{u-(1+r)}{u-d},$$

*is the **unique risk-neutral measure** $\mathbb{Q}$.*

**Exercise 3.1.11** (Continuous compounding)**.** *Rewrite this whole section under the condition that the bank account pays continuously compounded interests with interest rate $r$ (over the period's time length).*

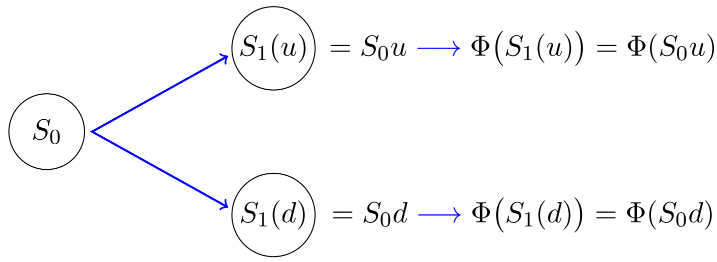



Figure 3.2.1: Contingent Claim $\Phi(S_1)$

## 3.2 Contingent claims

Let us now work under the assumption that the market is indeed arbitrage free. Like in Chapter 1 we now introduce the derivative products (e.g. call and put options) and aim at finding the prices for them. In this chapter, we will be able to go further than just the Put-Call parity formula of Chapter 1 and indeed we can compute the premiums or fair prices of such contracts.

**Definition 3.2.1** (Contingent claim)**.** *A **contingent claim** (or financial derivative) is any stochastic variable $X$ of the form $X = \Phi(S_1)$, where $S_1$ the stochastic variable representing the stock given above, see Figure 3.2.1. The real valued function $\Phi$ is called the contract function.*

**Example 3.2.2.** *A typical example of a contingent claim is the European call option on the stock $S$ at the end of the period with strike price $K$. Assume that $S_0d \leq K \leq S_0u$ and*

$$\Phi(S_1) = \Phi\big(S_1(Z)\big) = \big(S_1(Z) - K\big)^+ = (S_0Z - K)^+.$$

*If $S_1 > K$ then we use the option, pay $K$ to get the stock and then sell the stock on the market for $S_0u$; thus making a net profit of $S_0u - K$. If $S_1 < K$ then the option is not exercised. This allows us to write the payoff of the option as*

$$\Phi\left(S_1\right) = \begin{cases} (S_0u - K) & \text{if } Z = u, \\ 0 & \text{if } Z = d \end{cases}.$$

The main question is the following: What is the fair price (to be charged at time $t = 0$) of $\Phi\left(S_1\right)$? In what sense is the price *fair*?

**Definition 3.2.3** (Price process of a claim)**.** *Denote by $\Pi\left(n; \Phi(S_1)\right)$ the price operator associated with the claim $\Phi(S_1)$ which returns, for $n \in \{0, 1\}$, the price of the claim at period $n$.*

- When $n = 1$, then we must have $\Pi\big(1; \Phi(S_1)\big) = \Phi\big(S_1\big)$.
- When $n = 0$, ... ?

From Chapter 1, we have bounds for the price of the call options as well as the put-call parity (see Section 1.4.2 and Theorem 1.4.5: $c_0 + Ke^{-rT} = p_0 + S_0$). Namely, that

$$\Pi\left(0; \Phi\left(S_1\right)\right) \geq \max\{S_0 - K(1+r)^{-1}, 0\} \qquad \text{and} \qquad \Pi\left(0; \Phi\left(S_1\right)\right) \leq S_0.$$

We have assumed that the market is *arbitrage free* or there is *absence of arbitrage*. Hence, we know that we cannot make money out of nothing, nonetheless let us see what can be achieved on the market.

**Definition 3.2.4** (Reachability and replicating portfolio). *A contingent claim $X$ is said to be* ***reachable*** *or replicable if there exists a portfolio $h$ such that*

$$V_1^h = X,$$

*with probability 1. In that case we say that the portfolio $h$ is a* ***hedging*** *portfolio or a* ***replicating*** *portfolio.*

**If a certain claim $X$ is reachable with a replicating portfolio $h$, then, from a financial point of view, there is no difference between holding the claim and holding the portfolio. No matter what happens on the stock market, the value of the claim at period $n = 1$ will be exactly equal to the value of the portfolio at period $n = 1$. Thus the price of the claim should equal the market value of the portfolio, and we have reached the principle from the Law of one price (Market Rule 1.4.2) - Two assets or portfolios with identical future cash flows must have identical present values.**

### 3.2.1 Pricing principle and risk-neutral valuation

If a claim $\Phi$ is replicable with a replicating or hedging portfolio $h = (x, y)$ then its *fair price*, denoted by $\Pi(n; \Phi)$, is given by the solution of

$$\Pi(n; \Phi) = V_n^h, \quad n = 0, 1.$$

How to find $\Pi(1; \Phi)$? Let us start by collecting what we know.

#1) By no-arbitrage arguments, the price $\Pi(1; \Phi)$ of the claim $\Phi$ at maturity must be the payoff of the claim itself. This means that for any possible outcome of $S_1$ (or indeed the r.v. $Z$) one must have $\Pi(1; \Phi) = \Phi$.

#2) Given a portfolio $h = (x, y)$, the wealth process is expressed by $V_n^h = xB_n + yS_n$ for $n = 0, 1$. Hence at period $n = 1$ we have

$$V_1^h = x(1+r) + yS_0Z = \begin{cases} x(1+r) + yS_0u & \text{,if } Z = u \\ x(1+r) + yS_0d & \text{,if } Z = d \end{cases}.$$

#3) Since the portfolio $h$ is the *replicating portfolio*, we have by definition that $V_1^h = \Phi(S_1)$. And hence that $V_1^h = \Phi(S_1) = \Pi(1; \Phi)$.

*The conclusion* from the three points above, is that we return to our argument with the Law of one price (Market Rule 1.4.2). More concretely, since $V_1^h$ and $\Pi(1; \Phi)$ have the same payoff at time $n = 1$ (they are replicating), then they must also have the same present values (by the Law of One-price), hence $V_0^h = \Pi(0; \Phi)$.

Writing the equality explicitly we have a $2 \times 2$-system of equations from which we can determine $x$ and $y$,

$$V_1^h = \Phi = \Pi(1; \Phi) \Leftrightarrow \begin{cases} x(1+r) + yS_0u = \Phi(S_0u) & \text{, if } Z = u \\ x(1+r) + yS_0d = \Phi(S_0d) & \text{, if } Z = d \end{cases}.$$

Writing the system in matrix form

$$\begin{bmatrix} 1+r & S_0u \\ 1+r & S_0d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \Phi(S_0u) \\ \Phi(S_0d) \end{bmatrix}, \qquad \text{where } \det(\text{matrix}) = (1+r)S_0(d-u) \neq 0.$$

The matrix is invertible because of the no-arbitrage condition of Proposition 3.1.6, where $d < u$. The solution $h^\star = (x^\star, y^\star)$ is easily computable and given by

$$x^\star = \frac{1}{1+r}\frac{u\Phi\big(S_0 d\big) - d\Phi\big(S_0 u\big)}{u-d}, \tag{3.2.1}$$

$$y^\star = \frac{1}{S_0}\frac{\Phi\big(S_0 u\big) - \Phi\big(S_0 d\big)}{u-d}. \tag{3.2.2}$$

With $h^\star = (x^\star, y^\star)$, the price of the option at $t = 0$ is simply $V_0^{h^\star}$ that can be explicitly computed:

$$\begin{aligned}
\Pi(0;\Phi) &= V_0^{h^\star} = x^\star + S_0 y^\star \\
&= \cdots \qquad \text{(do these computations; its a nice exercise)} \\
&= \frac{1}{1+r}\Big(\Phi\big(S_0 u\big)q_u + \Phi\big(S_0 d\big)q_d\Big) \\
&= \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}\left[\Phi\left(S_1\left(Z\right)\right)\right].
\end{aligned}$$

Notice that one does not see $\mathbb{Q}$ anywhere in (3.2.1)-(3.2.2). Nonetheless, when one computes $x^\star + S_0 y^\star$ the constants (almost) magically align and the risk neutral probabilities appear.

We can neatly summarize the above calculations in the following main theorem that tells us that we do not really need to compute prices through the portfolios but just through the discounted value to the expected payoff under the risk-neutral measure.

**Theorem 3.2.5.** *If the binomial model is free of arbitrage then any claim $\Phi$ (in the sense of Definition 3.2.4) is replicable and the* no-arbitrage price *of $\Phi$ is given by*

$$\Pi(0;\Phi) = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}\left[\Phi\left(S_1\right)\right] = \frac{1}{1+r}\Big(\Phi\big(S_1(u)\big)q_u + \Phi\big(S_1(d)\big)q_d\Big).$$

*Where $\mathbb{Q}$ is the risk-neutral probability measure that is uniquely determined by the relation*

$$S_0 = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}\left[S_1\right] \quad \textit{where} \quad q_u = \frac{(1+r)-d}{u-d} \quad \& \quad q_d = \frac{u-(1+r)}{u-d}.$$

*The replicating portfolio is $h^\star = (x^\star, y^\star)$ given by (3.2.1) and (3.2.2). The r.v. $y^\star$ in (3.2.2) is called the $\Delta-$hedge/delta hedge.*

*Moreover, the claim's price $\Pi(0;\Phi)$ at time $t = 0$ is exactly the execution cost of the optimal portfolio $h^\star$ given by $V_0^{h^\star} = x^\star + y^\star S_0$.*

The main economic moral of this result can be summarized as follows.

- The only role played by the objective probabilities ($p_u$ and $p_d$) is that they determine which events are possible and which are impossible. In more abstract probabilistic terminology they thus determine the class of equivalent probability measures. (This will, unfortunately, not be covered in this course.)
- When we compute the arbitrage free price of a financial derivative we carry out the computations as if we live in a risk neutral world (under the measure $\mathbb{Q}$).
- This does not mean that we live (or believe that we live) in a risk neutral world.
- The valuation formula holds for all investors, regardless of their attitude towards risk, as long as they prefer more deterministic money to less

### 3.2.2 An example on the pricing principle and risk-neutral valuation

Take a one-period model as follows: $S_0 = 100$, $u = 1.2$, $d = 0.8$, $p_u = 0.6$, $p_d = 0.4$ and $r = 0$. We want to price an European Call option with $K = 110$.

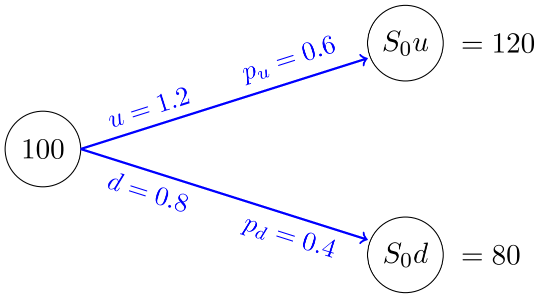



Figure 3.2.2: The one-period binomial model for this example

**Step 0: Build the market tree.** This has been done in Figure 3.2.2.

**Step 1: Verify the no-arbitrage condition.** We start by verifying the no-arbitrage condition: $0.8 = d < u = 1.2$ ✓ and $0.8 = d \le 1 = (1+r) \le u = 1.2$ ✓. And we now follow just like in Theorem 3.2.5.

**Step 2: Compute the risk-neutral measure** $\mathbb{Q}$ through its defining equation

$$\frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}[S_1] = S_0 \Leftrightarrow (1+r)S_0 = q_u S_1(u) + q_d S_1(d)$$
$$\Leftrightarrow 100(1+r) = q_u \cdot 120 + (1-q_u)\cdot 80,$$
$$\Rightarrow \quad q_u = \frac{(1+r)-d}{u-d} = \frac{1-0.8}{1.2-0.8} = \frac{1}{2} \quad \text{and} \quad q_d = \frac{u-(1+r)}{u-d} = \frac{1}{2}.$$

Note that computing the discounted expected value of the Stock $S_1$ under the physical measure $\mathbb{P}$ leads to $(1+r)^{-1}\mathbb{E}^{\mathbb{P}}[S_1] = 120 \cdot 0.6 + 80 \cdot 0.4 = 104 \neq 100$. The market is risk averse.

**Step 3: Compute the price under $\mathbb{Q}$.** Take $K = 110$ and $\Phi(S_1) = (S_1 - K)^+$ a European call option we can compute the risk-neutral price of the option by applying Proposition 3.2.5.

$$c_0 = \Pi(0; \Phi(S_1)) = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}[\Phi(S_1)] = \left[\overset{Z=u}{\Phi(120)} \cdot \frac{1}{2} + \overset{Z=d}{\Phi(80)}\frac{1}{2}\right] = \frac{10}{2} + 0 = 5.$$

For completeness, note that the computations done under $\mathbb{P}$ lead to different results

$$c_0 = \Pi(0; \Phi(S_1)) = \frac{1}{1+r}\mathbb{E}^{\mathbb{P}}[\Phi(S_1)] = \cdots = 6.$$

**Step 4: Compute the optimal portfolio or hedging portfolio**. Computing the hedging portfolio $h^\star = (x^\star, y^\star)$. If the risk-neutral pricing is correct, then (3.2.1) and (3.2.2) replicate the payoff of the call option $\Phi$. We compute $h^\star$

$$h^\star = \begin{cases} x^\star & = \frac{1.2\cdot 0 - 0.8\cdot 10}{1.2-0.8} = -20 \\ y^\star & = \frac{1}{100}\cdot\frac{10}{1.2-0.8} = \frac{1}{4}. \end{cases}$$

one borrows $|x^\star| = 20$ from the bank (due to the minus sign) and invests this money in $\frac{1}{4}$ of a share at time $t = 0$.

**(i)** $V_0^{h^\star} = -20 \cdot 1 + \frac{1}{4} \cdot 100 = 5 \equiv \Pi\left(0; \Phi(S_1)\right)$. This is the option's premium.

**(ii)** $V_1^{h^\star} = -20 \cdot 1 + \frac{1}{4} \cdot 120 = 10$, if $S_1 = 120$ or $Z = u$.

**(iii)** $V_1^{h^\star} = -20 \cdot 1 + \frac{1}{4} \cdot 80 = 0$, if $S_1 = 80$ or $Z = d$.

Point (ii) and (iii) prove that $h^\star$ is the replicating portfolio.

If anyone pays $\pounds 6$ for the call option then we can make arbitrage! ($\pounds 1$ of gain for sure). Indeed, we sell a call at $\pounds 6$ and invest $\pounds 5$ on the hedging portfolio $h^\star$ constructed above, we have $V_1^{h^\star} = \Phi\left(S_1\right)$ and we pay back the investor on one's bank account we have $\pounds 1$.

**Remark 3.2.6.** *Assuming that the transaction must happen, suppose that $S_1 = 120$ then, as we have $\frac{1}{4}$ of the stock and the holder of the option will buy a unit of stock from us, we short sell '$\frac{3}{4}$' of the stock we sell it to the holder of the option for $K = 110$, that allows to use $\pounds 90$ to buy back '$\frac{3}{4}$' of the stock and 'replace' the stock we short sell, we still have a debt of $\pounds 20$ that we meet with the remaining cash from the call option transaction, ensuring that $\pounds 1$ is on our bank account.*

### 3.2.3 The continuous-time version

In this section we re-write Theorem 3.2.5 when the bank account pays continuously. The underlying assumption is that the bank account now pays interests at a continuous rate of $r$ and, one finally sees the importance of the period's time length $\Delta t$.

The bank account deposit now grows at a rate $e^{r\Delta t}$ instead of rate $(1 + r)$.

**Proposition 3.2.7.** *Assume the (1-period) binomial model is arbitrage-free, i.e.*

$$\text{if } d < u \quad \text{and} \quad 0 < d \leq e^{r\Delta t} \leq u.$$

*Then any claim $\Phi(S_1)$ is replicable and the risk-neutral price $\Pi\left(0; \Phi\right)$ of the contingent claim with payoff function $\Phi$ is given by*

$$\begin{aligned} \Pi\big(0; \Phi(S_1)\big) &= e^{-r\Delta t} \mathbb{E}^{\mathbb{Q}}\big[\Phi\left(S_1\right)\big] \\ &= e^{-r\Delta t}\big[\Phi\left(S_1(u)\right) q_u + \Phi\left(S_1(d)\right) q_d\big]. \end{aligned} \tag{3.2.3}$$

*Furthermore, the probabilities $q_u$ and $q_d$ are uniquely determined by the equation*

$$S_0 = e^{-r\Delta t} \mathbb{E}^{\mathbb{Q}}\left[S_1\right],$$

*where $q_u$ and $q_d$ are explicitly given by*

$$q_u = \frac{e^{r\Delta t} - d}{u - d} \qquad \& \qquad q_d = \frac{u - e^{r\Delta t}}{u - d}.$$

*The replicating portfolio $h^\star := (x^\star, y^\star)$ is given by*

$$x^\star = e^{-r\Delta t} \frac{u\Phi\left(S_1(d)\right) - d\Phi\left(S_1(u)\right)}{u - d} \quad \& \quad y^\star = \frac{1}{S_0} \frac{\Phi\left(S_1(u)\right) - \Phi\left(S_1(d)\right)}{u - d}. \tag{3.2.4}$$

*Moreover, the claim's price $\Pi\left(0; \Phi\right)$ at time $t = 0$ is exactly the execution cost of the optimal portfolio $h^\star$ given by $V_0^{h^\star} = x^\star + y^\star S_0$.*

The next section contains an illustrative example of the previous theorem.

### 3.2.4 Risk-neutral pricing and risk-free portfolio pricing

Theorem 3.2.5 or Theorem 3.2.7 state that *Risk-neutral pricing* works by computing the risk neutral probability measure $\mathbb{Q}$ and compute the value of the contingent claim backwards through the time period. This is reading the theorems from top to bottom.

On the other hand, reading the theorems from bottom to top, gives rise to the so-called *risk-free portfolio pricing* argument. Risk-free portfolio pricing argument focuses on constructing a portfolio such that one can *adjust* the portfolio (by buying/selling stock) at every trading time in order to ensure that there is no uncertainty about the value of the portfolio at the *maturity* of the contingent claim $n = 1$.

The next example illustrates the use of Theorem 3.2.7 as well as both pricing arguments described above.

**Example 3.2.8.** *Take the one-period model, $n \in \{0, 1\}$, where the period has time length $\Delta t = 0.25$ (a period of 3 months over the year). Suppose the bank account pays interest continuously with rate $r = 0.12$.*

*Suppose that in our market model we have a stock with a spot price of $S_0 = £20$ which will either rise at the end of the period to $£22$ or fall to $£18$. See Figure 3.2.3.*

*Let there be a European call option written on $S_1$ with strike $K = 21$.*

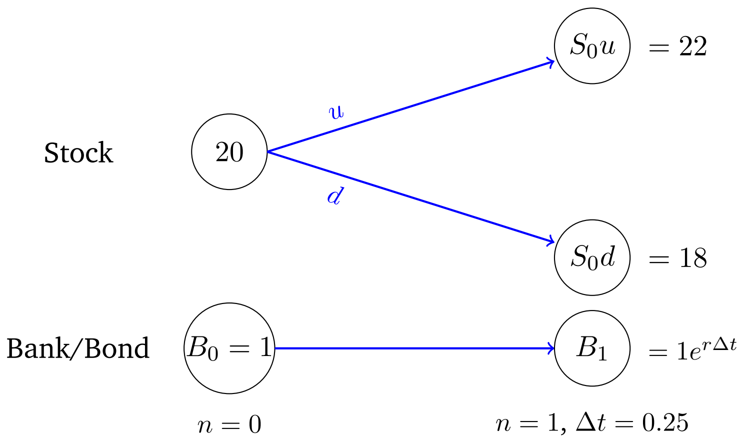



Figure 3.2.3: The Stock market and the Bank account

How to price the call option above? We start with the usual method, the Risk-Neutral measure one.

#### Pricing under the risk-Neutral measure

From the example setup above, it is easy to compute $u$ and $d$. Namely,

$$\begin{aligned} 22 = S_1(u) = S_0 u = 20 \cdot u \quad &\Rightarrow \quad u = 1.1 \\ 18 = S_1(d) = S_0 d = 20 \cdot d \quad &\Rightarrow \quad d = 0.9. \end{aligned}$$

We have then, $0.9 = d < u = 1.1$✓. For the no-arbitrage condition, we first compute $e^{r\Delta t}$ and verify that $0.9 = d < e^{r\Delta t} = \exp\{0.12 \times 0.25\} = 1.03045 < u = 1.1$✓.

The risk-neutral probabilities follow easily: $q_u = (e^{r\Delta t} - d)/(u-d) = 0.6523$. And hence, the call option's price can be computed from the usual equation

$$\Pi\big(0;\Phi(S_1)\big) = e^{-0.12\cdot\frac{1}{4}}\left(0.6523\cdot(22-21)^+ + 0.3477\cdot(18-21)^+\right) = 0.633.$$

The optimal portfolio

$$y^\star = \frac{1}{S_0}\frac{\Phi\left(S_1(u)\right) - \Phi\left(S_1(d)\right)}{u-d} = \frac{1}{20}\frac{1-0}{1.1-0.9} = \frac{1}{4}.$$

**Pricing through riskless portfolios**

Consider the following investment. A portfolio $h = (0,\Delta)$ (here $\Delta \geq 0$, i.e. a long position in the stock) and a short position (of one unit) in the European call. This investment, say $I_n$ for $n = 0, 1$, is worth

$$\begin{aligned} &\text{at time } n=0: \quad I_0 := 0 + \Delta\cdot S_0 - \Pi\big(0;\Phi(S_1)\big),\\ &\text{at time } n=1: \quad I_1 := I_1(S_1) = 0 + \Delta\cdot S_1 - \Phi(S_1). \end{aligned}$$

We now want to find the number $\Delta$ (acquired at $n=0$) that makes the investment $I_1$ have the same value independently of the value of the stock. In other words, the $\Delta$ such that $I_1(S_1(d)) = I_1(S_1(u))$. Writing the equation we have

$$\begin{aligned} &I_1\big(S_1(u)\big) = I_1\big(S_1(d)\big)\\ \Leftrightarrow\ &\Delta\cdot S_1(u) - \Phi\big(S_1(u)\big) = \Delta\cdot S_1(d) - \Phi\big(S_1(d)\big)\\ &\Leftrightarrow \quad 22\Delta - (22-21)^+ = 18\Delta - (18-21)^+ \ \Leftrightarrow\ \Delta = 0.25. \end{aligned}$$

With this specific value $\Delta$, we know that by construction $I_1(S_1(d)) = I_1(S_1(u)) = 4.5$.

Recall the Law of one price (Market Rule 1.4.2). If two investments have the same future value, then their present value must also be the same.

In this case, since $I_1$ always takes the same value, we can compare it with a deposit in the bank account. Namely, a deposit at time $t=0$ of value $e^{-r\Delta t}I_1$ will pay the exact same at the end of the period as the investment $I_1$. Then (Law of one price - Market Rule 1.4.2) at time $t=0$ the investment and the deposit must have equal values:

$$I_0 = e^{-r\Delta t}I_1 \quad \Leftrightarrow \quad 20\Delta - \Pi\left(0;\Phi\right) = e^{-r\Delta t}I_1.$$

Given the numerical values we computed above, the equation can be solved for $\Pi\left(0;\Phi\right)$,

$$\begin{aligned} \Pi\left(0;\Phi\right) &= 20\Delta - e^{-r\Delta t}I_1\\ &= 20\times 0.25 - e^{-0.12\times 0.25}\times 4.5 = 5 - 4.367 = £0.633. \end{aligned}$$

The name risk-less portfolio derives from the fact that with the particular $\Delta = 0.25$, the investment $I_1$ behaves just like money in a bank account because it has no risk in it.

## 3.3 Multi-period model

In this section we extend the one-period model to a multi-dimensional one. In terms of added difficulty, there isn't much. Indeed, for multi-period models one just repeats for each period the ideas of the one-period model.

### 3.3.1 The market, the involved quantities and no-arbitrage

The multi-period binomial model is a discrete-time model with, say $N \in \mathbb{N}$ periods, and a running time index $n = 0$ to $n = N$. The periods have a fixed time length of $\Delta t$ and the total time of the $N$ periods is simply $T = N\Delta t$. Let $r \in [0, \infty)$ be period's interest rate.

As before, we have two underlying assets, a bond with price process $(B_n)_{n=0,\cdots,N}$ and a stock with price process $(S_n)_{n=0,1\cdots,N}$. We assume that the return on the bond $(B_n)_{n=0,\cdots,N}$ to be deterministic and constant over the trading stages $n = 0, \cdots, N-1$.

$$\text{Bank/Bond} = \begin{cases} B_{n+1} & = (1+r)B_n, \qquad \Rightarrow B_{n+1} = (1+r)^{n+1} \\ B_0 & = 1 \end{cases},$$

$$\text{Stock} = \begin{cases} S_{n+1} & = S_n \cdot Z_n, \\ S_0 & = 1. \end{cases}$$

where the collection of r.v. $Z_0, \ldots, Z_{N-1}$ are i.i.d. Bernoulli r.v. and $\mathbb{P}[Z_i = u] = p_u$ and $\mathbb{P}[Z_i = d] = p_d$ for any $i \in \{0, \cdots, N-1\}$.

We can illustrate the dynamics of the stock by means of a tree (see figure). In what

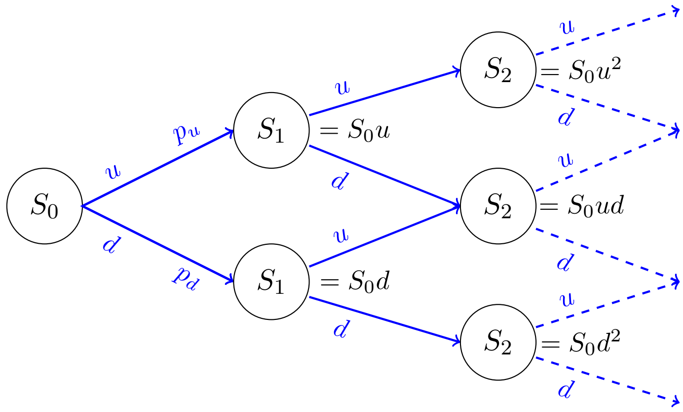



Figure 3.3.1: Stock prices evolution in the multi-period model.

follows we revisit and adapt to the multi-period case the concepts of portfolio, wealth process, arbitrage. And recover as well the conditions for no-arbitrage.

#### Portfolios and arbitrage

Given that it is possible to trade in this market we define what are portfolios, wealth processes and self-financing strategies.

**Definition 3.3.1** (Portfolio strategy and Wealth process). *A **portfolio strategy** is a stochastic process*

$$\{h_n = (x_n, y_n) \,:\, n = 1, \ldots, N\}$$

*such that $(h_n)$ is a function of the stock prices $S_0, S_1, \ldots, S_{N-1}$. For a given portfolio strategy $h$ we set $h_0 = h_1$ by convention.*

*The **value/wealth process** corresponding to it by*

$$V_n^h = x_n B_n + y_n \cdot S_n = x_n(1+r)^n + y_n \cdot S_n \qquad \textit{for } n = 0, \cdots, N$$

Here $x_n$ is the amount of money we invest in the bank account at time $n-1$ and we keep it until time $n$, similarly $y_n$ is the number of shares that we buy at time $n-1$ and keep until time $n$. We allow the portfolio strategy to depend on all the information we have from the evolution of the market up to time $n$. We are, however, not allowed to look into the future.

**Definition 3.3.2** (Self-financing strategy). *A portfolio strategy $h_n = (x_n, y_n)$ is said to be a **self-financing** strategy if $x_n(1+r) + y_n \cdot S_n = x_{n+1}(1+r) + y_{n+1} \cdot S_n$, for all $n = 1, \ldots, N$. In other words,*

$$(x_{n+1} - x_n)(1+r) + (y_{n+1} - y_n)S_n = 0,$$

*i.e. there is no "extra" capital into the portfolio at any time $n = 1, \ldots, N$.*

A self-financing strategy simply means that if one wishes to deposit money in the bank account, this extra money has to come from the stocks (either by selling of short-selling). The same goes for the stock, in order to purchase more stocks, the money has to be borrowed from the bank or be withdraw money from our deposit.

In a multi-period model we assume all portfolio strategies to be **self-financing**. We extend the definition of arbitrage to the multi-period case.

**Definition 3.3.3** (Arbitrage). *Given an $N$-period binomial model, an **arbitrage** possibility is a self-financing portfolio with the following properties:*

$$V_0^h = 0, \qquad \mathbb{P}[V_N^h \geq 0] = 1, \qquad \text{and} \qquad \mathbb{P}[V_N^h > 0] > 0.$$

*If no such portfolio $h$ exists with those properties, we say that the market **free of arbitrage**.*

As in the one-period case, we assume throughout that $d < u$ and we can state a no-arbitrage as well as the existence of the Risk-neutral probability $\mathbb{Q}$ result for the model just like before.

**Theorem 3.3.4.** *Let $d < u$. The model is arbitrage-free if and only if $d \leq 1+r \leq u$.*

*Moreover, under the above condition there exists a unique risk-neutral measure $\mathbb{Q}$ whose defining probabilities $q_d$ and $q_u$ are uniquely defined by the relation*

$$s = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}[S_{n+1} | S_n = s],$$

*and are explicitly computed through*

$$q_u = \frac{(1+r) - d}{u - d} \qquad \text{and} \qquad q_d = \frac{u - (1+r)}{u - d}.$$

### Contingent claims

As before a **contingent claim** in the multi-period model is an easy extension of the concept from before.

**Definition 3.3.5** (Contingent claim). *A **contingent claim** (or financial derivative) is any stochastic variable $X$ of the form $X = \Phi(S_N)$, where $S_N$ the stochastic variable representing the stock given above. The real valued function $\Phi$ is called the contract function.*

### 3.3.2 The pricing mechanism in the multi-dimensional model

How does the pricing work? The pricing mechanist for a claim is to repeat the pricing theorems of the previous section.

**Proposition 3.3.6.** *The price of a contingent claim is computed in the multi-period model by applying Theorem 3.2.5 or Theorem 3.2.7, backwards in time to each branch of the tree.*

*Given a claim $\Phi$ written on $S$ we denote, for any $n \in \{0, \cdots, N\}$*

$$\Pi(n; \Phi)\Big|_{S_n}$$

*as the price of the option $\Phi$ at period $n$ when the stock takes the value $S_n$.*

Step-wise we have

i) Build the tree for the stock

ii) Verify there is no arbitrage & compute $q_u$ and $q_d$.

iii) Compute the pay-off of $\Phi$ at the last period

iv) Apply the **risk-neutral** pricing theorem in the correct context (continuous or discrete compounding), backwards, over all the branches of the tree until $t = 0$ is reached.

Let us see an example for clarity.

**Example 3.3.7.** *Take a 2-period binomial model ($N = 2$) with $d = 0.8$, $u = 1.2$, $S_0 = 50$, $r = 5\%$ (continuously compounded). Each period has time length of $\Delta t = 1$ year and the total time the model covers is $T = N \times \Delta t = 2$, two years. Take an European call option written on the value of the stock at $n = 2$, denoted $S_2$, with strike $K = 52$.*

***Step 0: Build the market tree.*** *This has been done in Figure 3.3.2 next.*

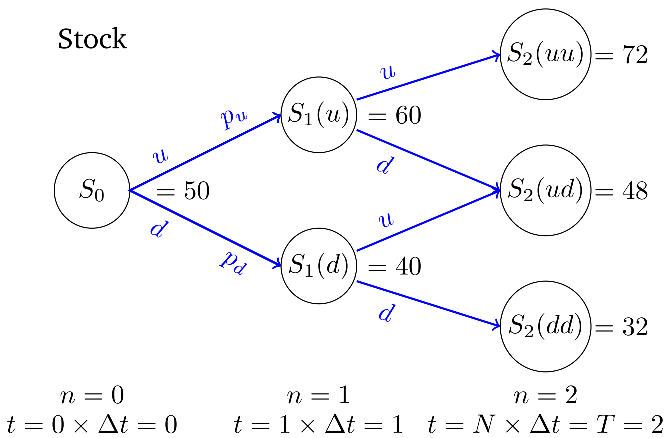


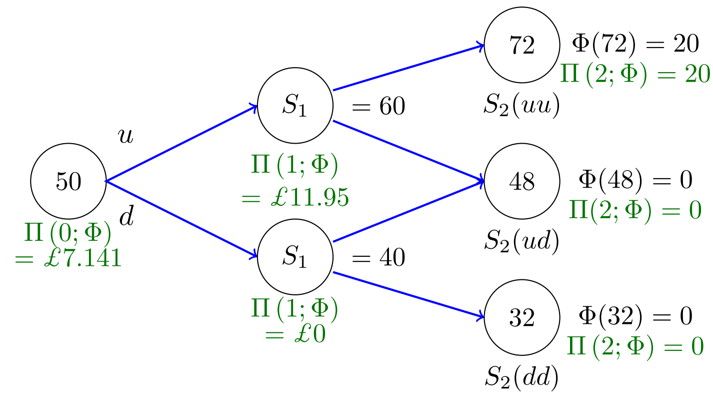

Figure 3.3.2: Evolution of the stock in the 2-period model of Example 3.3.7.

***Step 1: Verify the no-arbitrage condition.*** *we start with $0.8 = d < u = 1.2$ ✓. For the no-arbitrage $0.8 = d < e^{0.05} = 1.057 < u = 1.2$ ✓. We conclude that the market is free of arbitrage.*

*Step 2: Compute the risk-neutral measure* $\mathbb{Q}$. *By following the formulas:* $q_u = 0.6282$ *and* $q_d = 0.3718$.

*Step 3: Compute the price under* $\mathbb{Q}$. *We now compute the price of the call option in each of the nodes of the tree. For this, we go back to Figure 3.3.2 and write next to it the payoff of the Call option at period* $n = 2$*; we compute* $\Phi(S_2(uu)), \Phi(S_2(ud))$ *and* $\Phi(S_2(dd))$.

*The payoff of the call option is simply* $\Phi(S_2) = (S_2 - K)^+$ *and at period* $n = N = 2$ *we also have* $\Pi\big(0; \Phi(S_2)\big) = \Phi(S_2) = (S_2 - K)^+$.




*To compute the Call option prices for period* $n = 1$ *and* $n = 0$ *we simply use* (3.2.3) *at each node of the tree.*

- *At node* $n = 1$*, for* $S_1(u) = 60$

$$\Pi(1;\Phi)\Big|_{S_1(u)} = e^{-0.05}\left[0.6282 \cdot 20 + 0.3718 \cdot 0\right] = 11.951.$$

- *At node* $n = 1$*, for* $S_1(d) = 40$

$$\Pi(1;\Phi)\Big|_{S_1(d)} = e^{-0.05}\left[0.6282 \cdot 0 + 0.3718 \cdot 0\right] = 0.$$

- *At node* $n = 0$*, for* $S_0 = 50$*. Here notice that we are using the prices of the call at* $n = 1$ *and not its prices at* $n = 2$*. we have then*

$$\Pi(0;\Phi)\Big|_{S_0} = e^{-0.05}\left[11.951 \cdot 0.6282 + 0 \cdot 0.3718\right] = £7.141.$$

### Algebraic pricing formulas

From an algebraic point of view, given a European type claim $\Phi(S_N)$ the pricing formula (3.2.3) can be repeated backwards in time to yield a very "clean" formula that needs to be calculated only *once*.

**Proposition 3.3.8.** *Consider a* $N$*-period binomial market model and on it an European type claim* $\Phi(S_N)$*. If the market is arbitrage free then*

$$\Pi\left(0; \Phi\left(S_N\right)\right) = \left(e^{-r\Delta t}\right)^N \sum_{k=0}^{N} \binom{N}{k} q_u^k q_d^{N-k} \Phi\left(S_0 u^k d^{N-k}\right).$$

*In particular, for the 2-period model the* risk-neutral price *formula of the European type contingent claim is given by*

$$\Pi\left(0;\Phi(S_2)\right) = e^{-r\cdot 2\Delta t}\left(q_u^2\underbrace{\Phi\left(S_0u^2\right)}_{\Phi(S_2(u,u))}+2q_uq_d\underbrace{\Phi\Big(S_0ud\Big)}_{S_2(u,d)}+q_d^2\underbrace{\Phi\Big(S_0d^2\Big)}_{S_2(d,d)}\right).$$

*Proof.* This proof is a direct iteration of the *Risk-neutral pricing* Theorem 3.2.7. □

**Exercise 3.3.9.** *Write explicitly the pricing formula for the 3-period model.*

One can also write a similar statement based on Theorem 3.2.5 when interests are paid periodically at the end of each period.

**Example 3.3.10.** *Continuing Example 3.3.7. Instead of computing the prices of the call option along the tree we can calculate it directly using the 2-period algebraic formula of the previous proposition.*

*We have then*

$$\begin{aligned}\Pi\left(0;\Phi(S_2)\right) &= e^{-r\cdot 2\Delta t}\left(q_u^2\Phi\left(S_0u^2\right)+2q_uq_d\Phi\Big(S_0ud\Big)+q_d^2\Phi\Big(S_0d^2\Big)\right)\\ &= e^{-0.05\times 2}\Big(q_u^2\times 20+2q_uq_d\times 0+q_d^2\times 0\Big)\\ &= \cdots = £7.141.\end{aligned}$$

*Try to compute the price of the European Put option.*

### The Put-Call parity formula

As in Chapter 1 we can state (and prove in the same way) a version of the Put-Call parity formula (see Theorem 1.4.5). We do not prove the result, nonetheless we will use it.

**Theorem 3.3.11** (Put-Call parity)**.** *Consider the $N$-period binomial model where the bank account pays continuously at rate $r$ and each period has time length $\Delta t$. Assume that the market is arbitrage-free.*

*Consider a European call and put option written on a stock $S$, with maturity in period $N$ (or at time $T = N\times\Delta t$) and strike $K$. For $n = 0,\cdots,N$ we denote by $\Pi(n;Call)$ and $\Pi(n;Put)$ the risk-neutral price of the call and put option at period $n$ respectively.*

*Then at time $n = 0$ or $t = 0$ it holds that*

$$\Pi(0;Call) + Ke^{-r\times N\times\Delta t} = \Pi(0;Put) + S_0.$$

*In general, for any $n\in\{0,\cdots,N\}$ it holds that*

$$\Pi(n;Call) + Ke^{-r\times(N-n)\times\Delta t} = \Pi(n;Put) + S_n. \tag{3.3.1}$$

Note that $Ke^{-r\times(N-n)\times\Delta t}$ corresponds to $Ke^{-r\times(T-t)}$ in the formula of Chapter 1, where $T = N\times\Delta t$ and $t = n\times\Delta t$.

**Important note:** pay attention to the interest rate term. Mistakes tend to appear in the interest rate term when one is computing the put-call parity in the intermediate periods as one confuses the time to maturity $(T-t)$ with just the maturity $T$!

**Exercise 3.3.12.** *Take the market model from Example 3.3.7.*

a) *Compute the prices of an European Put option with strike $K = 110$.*

b) *Go to Example 3.3.7 and collect the prices of the European Call option priced there. Verify that the Put-Call parity formula* (3.3.1) *holds at each node of the tree*[3] *(see Figure 3.3.2).*

## 3.4 American options

We briefly discussed American type options in Chapter 1. Assume we are trading in an $N$-period binomial model containing a stock $S$ with prices $(S_n)_{n=0,1,\cdots,N}$. Recall the call function defined as $\Phi(x) = \max\{x - K, 0\} = (x - K)^+$.

A European call option gives the holder the right, not the obligation, to buy the underlying asset $S$ at maturity time $T$ for the strike price $K$.

The **American option** works just like the European one but, it can be exercised at **any time**. Exercising the American option before the maturity time, is called an **early-exercise**.

**Example 3.4.1.**

- *European call $\Phi(S_N) = (S_N - K)^+$, can be exercised only at the maturity period $N$ (or time $T$).*
- *American call $\Phi(S_\cdot) = (S_\cdot - K)^+$ can be exercised at* any *trading time.*

*Here $S_\cdot$ means $S(t)$ for any $t \in [0, T]$ (using the language of Chapter 1) or $S_n$ at period $n \in \{0, 1, \ldots, N\}$ (using the language of this chapter).*

Intuitively, the American call option should be more expensive than the European call since it may give us more possibilities of making a profit.

To avoid possible confusions later on, we define the *intrinsic value* of an American option.

**Definition 3.4.2** (Intrinsinc value of an American option)**.** *The **intrinsic value** of an American option at period $n$ is the market value of the option (its payoff) if one exercises the option at period $n$.*

The intrinsic value *is usually different* from the fair price of the option and the concept of intrinsic value of a European option does not make sense.

### 3.4.1 Pricing the American option

The pricing mechanism for the American option works recursively backwards through the tree as before, but **it does so by comparing the expected fair price of the American option if it was an European one** and **its intrinsic value**!

[3]Note that in this exercise, at period $n = 0$ the discounting factor is $e^{-r\times 2\times\Delta t} = e^{-2r}$ but at period $n = 1$ the discounting factor is $e^{-r\times(2-1)\times\Delta t} = e^{-r}$. Confusing the amount periods in the discounting factor is a known common mistake.

If the intrinsic value is higher then one would exercise the option at that tree node and one call this **early-exercise**.

Step-wise it works as follows:

▹ at period $N$ (or time $T$) the price of the option is exactly its pay-off (as it is in the case of a European option): $\Pi(N;\Phi)\big|_{S_N} = \Phi(S_N)$.

▹ at period $n < N$ the price of the American option is the maximum between

- its expected future gain given by the *risk-neutral* price formula (as if an early-exercise would not be possible, i.e. an European type option)

$$e^{-r\Delta t}\big(q_u\Pi\left(n+1;\Phi\right)\big|_{S_n} + q_d\Pi\left(n+1;\Phi\right)\big|_{S_n}\big),$$

- and its intrinsic value given by the pay-off from an early exercise $\Phi\left(S_n\right)$.

in other words: take any $n = 0, \cdots, N-1$ and let $S_n = s$, then the price of the american option at period $n$ is given by

$$\Pi(n;\Phi)\big|_{S_n=s}$$
$$:= \max\left\{ \underbrace{\Phi(s)}_{\text{intrinsic value}}, \underbrace{e^{-r\Delta t}\Big(q_u\Pi\left(n+1;\Phi\right)\big|_{S_{n+1}=us} + q_d\Pi\left(n+1;\Phi\right)\big|_{S_{n+1}=ds}\Big)}_{\text{expected future gain in the option given } S_n} \right\}.$$

In the case of the intrinsic value at period $n$ is less than the future expected value of the option (given by the risk-neutral price), then *I expect to make more money if I wait* and hence I will not exercise early the option. If it had been reversed (intrinsic value higher than expected future payoff), then one would exercise early the option.

Note that, if there is no early-exercise of the American call, then the price of the American call has to be the price of the European call.

We continue with the market setup of Example 3.3.7 and we now discuss an example with the pricing of an American Call option.

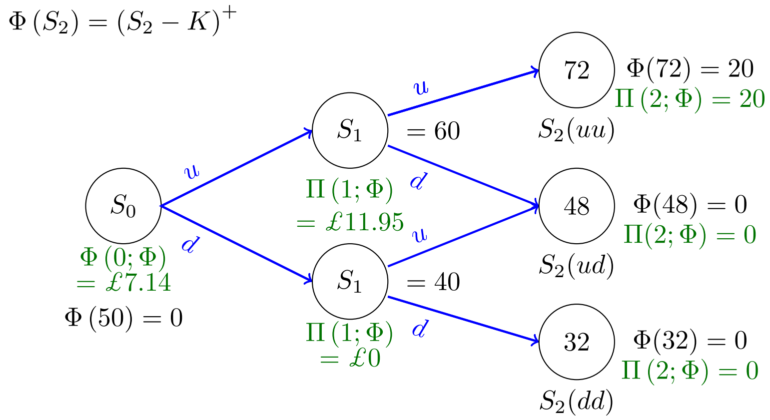
**Example 3.4.3** (American call option)**.** *Take the market setup of Example 3.3.7. The Stock price tree can be see in Figure 3.3.2.*

*We now compute the prices of the American call option backwards through the tree nodes.*

- *Price at $S_1(u) = 60$ at $(n = 1)$*
  - *Early exercise/payoff/Intrinsic value* $\Phi\left(60\right) = 8$
  - *The risk-neutral value* $e^{-0.05}\left[q_u \cdot 20 + q_d \cdot 0\right] = 11.958.$
  - *Then* $\Pi\left(1;\Phi\right)\big|_{S_1(u)} := \max\left\{8, 11.958\right\} = 11.958 \quad \Rightarrow$ *No early exercise!*
- *Price at $S_1(d) = 40$ at $(n = 1)$*
  - *Early exercise/payoff/Intrinsic value* $\Phi\left(40\right) = 0.$
  - *The risk-neutral value* $e^{-0.05}\left[q_u \cdot 0 + q_d \cdot 0\right] = 0$
  - *Then* $\Pi\left(1;\Phi\right)\big|_{S_1(d)} := \max\left\{0, 0\right\} = 0 \quad \Rightarrow$ *No early exercise!*

- <u>*Price at* $S_0 = 50$ *at* $(n = 0)$</u>

  - *Early exercise/payoff/Intrinsic value* $\Phi(50) = 0$.
  - *The risk-neutral value* $e^{-0.05}[q_u \cdot 11.985 + q_d \cdot 0] = 7.140$.
  - *Then* $\Pi(0;\Phi)\big|_{S_0} := \max\{0, 7.140\} = 7.140$ $\Rightarrow$ *No early exercise!*




*The investor does not exercise the option at any earlier time (he does make a profit from having the stock earlier). Thus the price of the American option is, in this case, equal to the price of a European call option.*

**Remark 3.4.4** (Further remarks:)**.**
- *If the option would not be exercised in period* $n$ *then its price would be exactly the price as if we are pricing an European option with maturity equal to the next "trading period".*
- *If the option was exercised at certain period* $n \le N$ *then its price would be exactly its payoff at* $\Pi(n;\Phi) = \Phi(S_n)$.
- *The price of an American option at period* $n$ *is the* $\max\{$*Payoff now*, *Price of the equivalent option having the next period as maturity.*$\}$.

### 3.4.2 An example - the American put option

We continue with the market setup of Example 3.3.7 and we now discuss the pricing of the American Put option $\Phi(S_\cdot) = (K - S_\cdot)^+$.

**Example 3.4.5** (The European Put option)**.** *We continue working on the market model of Example 3.3.7. Take the* 2 *period binomial model with* $d = 0.8$, $u = 1.2$, $S_0 = 50$, $r = 5\%$ *(continuously compounded)* $T = 2$, *and* $\Delta t = 1$ *year.*

*Take an European put option with strike* $K = 52$.

*We can copy* ***Steps 0, 1 and 2*** *from Example 3.3.7. The tree can be seen in Figure 3.3.2.*

***Step 3: Compute the prices under*** $\mathbb{Q}$***.***

*Prices and payoff must match at* $n = N$. *Using the pricing formulae:*

- *For* $S_1(u) = 60$ $(n = 1)$

$$\Pi(1;\Phi)\big|_{S_1(u)} = e^{-0.05}[0.6282 \cdot 0 + 0.3718 \cdot 4] = £1.415.$$

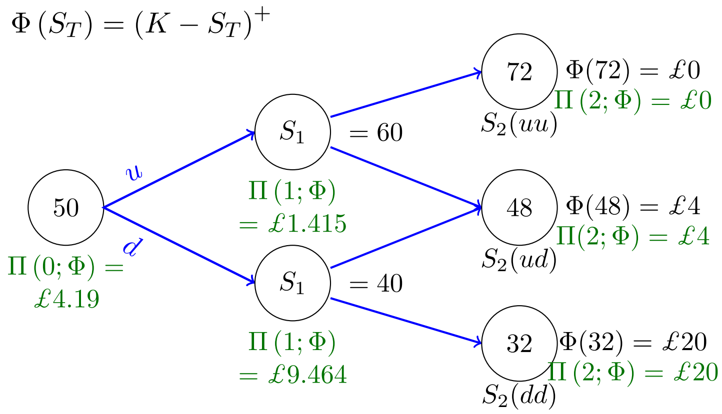



Figure 3.4.1: European put option's value along $n = 0, 1, 2$.

- *For* $S_1(d) = 40$ *(*$n = 1$*)*

$$\Pi(1;\Phi)\big|_{S_1(d)} = e^{-0.05}\left[0.6282 \cdot 4 + 0.3718 \cdot 20\right] = £9.464.$$

- *For* $S_0 = 50$ *(*$n = 0$*)*

$$\Pi(0;\Phi)\big|_{S_0} = e^{-0.05}\left[1.415 \cdot 0.6282 + 9.464 \cdot 0.3718\right] = £4.19.$$

The next example goes through the prices of the American put, and the previous example serves as a comparison to see where the early-exercise optionality makes a difference.

**Example 3.4.6** (American Put option)**.** *We continue working on the market model of Example 3.3.7. Take the* 2 *period binomial model with* $d = 0.8$, $u = 1.2$, $S_0 = 50$, $r = 5\%$ *(continuously compounded)* $T = 2$, *and* $\Delta t = 1$ *year.*

*Take an American put option with strike* $K = 52$.

*We can copy* ***Steps 0, 1 and 2*** *from Example 3.3.7. The tree can be seen in Figure 3.3.2.*

***Step 3: Compute the prices under*** $\mathbb{Q}$**.**

- *Price at* $S_1(u) = 60$ $(t = 1)$.
  - *Early exercise* $\Phi(60) = 0$
  - *The risk-neutral value* $e^{-0.05}\left[q_u \cdot 0 + q_d \cdot 4\right] = £1.415$.
  - *Then* $\Pi(1,\Phi)\big|_{S_1(u)} := \max\{0, 1.415\} = £1.415. \quad \Rightarrow$ *No early exercise!*

- *Price at* $S_1(d) = 40$ $(t = 1)$.
  - *Early exercise/payoff* $\Phi(40) = £12$.
  - *The risk-neutral value* $e^{-0.05}\left[q_u \cdot 4 + q_d \cdot 20\right] = £9.436$
  - *Then* $\Pi(1,\Phi)\big|_{S_1(d)} := \max\{12, 9.463\} = £12 \quad \Rightarrow$ *Early exercise!*

- *Price at* $S_0 = 50$ $(t = 0)$

- *Early exercise/payoff* $\Phi(50) = £2$.
- *The risk-neutral value* $e^{-0.05}[q_u \cdot 1.415 + q_d \cdot 12] = £5.089$.
- *Then* $\Pi(0,\Phi)\big|_{S_0} := \max\{2, 5.089\} = £5.089 \quad \Rightarrow$ *No early exercise!*

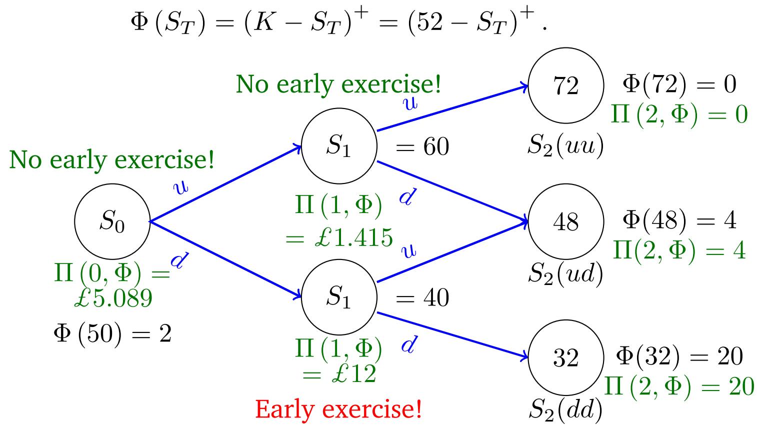



Figure 3.4.2: American put option's value along $n = 0, 1, 2$.

***Here is a trick question:*** *How about the put-call parity? Do you think it holds for the American Call and Put?*

## 3.5 Volatility and model calibration

Note: this section is based in Chapter 11 (*Binomial Trees*) of J. Hull's book.

In the real world, one important quantitative property of a price is its *volatility*. The volatility of a stock, denoted by $\sigma$, is a measure of our uncertainty about the returns of a stock[4]. Stocks typically have a volatility between 15% and 60%.

In the early time of mathematical finance, statistical testing led traders and researchers to believe that in a small time interval, $\Delta t$, that $\sigma\sqrt{\Delta t}$ is approximately equal to the standard deviation of the returns/percentage change in the stock price at that time. Suppose that $\sigma = 0.3$ or 30% per annum and the current stock price is £50. The standard deviation of the percentage change in the stock price in one week is approximately

$$30 \times \sqrt{\frac{1}{52}} \approx 4.16\%.$$

Therefore, a one-standard deviation move in the stock price over one week is $50 \times 0.0416\% \approx £2.08$. These above ideas, led practitioners and researchers to propose another model (more realistic) to simulate the evolution of the stock.

The returns of a stock are modeled over time as (Black-Scholes model)

$$\frac{\Delta S_t}{S_t} = \frac{S_{t+\Delta t} - S_t}{S_t} = \underbrace{\mu \Delta t}_{\text{drift or trend}} + \underbrace{\sigma\sqrt{\Delta t}\varepsilon}_{\text{volatility}}, \quad \text{with } \varepsilon \sim \mathcal{N}(0,1).$$

The equation above shows that our uncertainty about a future stock price, as measured by its standard deviation, increases or decreases (approximately), with the square root of how far ahead we are looking.

[4]Recall that returns are $(S_{n+1} - S_n)/S_n$: percentage change in stock price from $n$ to $n+1$

Converting the above model to the binomial tree language, this means that for $t_k = k \cdot \Delta t$ and $k = 0, \ldots, N$ (or $\Delta t = \frac{T}{N}$)

$$S(t_{k+1}) = S(t_k) + \mu S(t_k)\Delta t + \sigma S(t_k)\sqrt{\Delta t}\varepsilon, \qquad \varepsilon \sim \mathcal{N}(0,1) \tag{3.5.1}$$
$$= S(t_k)\Big(1 + \mu\Delta t + \sigma\sqrt{\Delta t}\varepsilon\Big)$$

Our bridge between these new ideas and those of the previous sections is that the binomial tree model is built to be an approximation of this (harder to analyze) model.

Lastly, volatility is a measure of how uncertain we are about future stock prices movements. As volatility increases, the chance that the stock will do very well or very poorly increases as well. For the owner of a stock, these two outcomes tend to offset each other. However, this is not the case for the owner of a option. The owner of a call benefits from price increases but has limited downside risk in the event of price decreases because the most that he or she can lose is the price of the option (premium). The owner of a put also benefits since he or she has limited downside risk in the event of price increases. Hence, the value of both calls and puts increases as volatility increases.

### 3.5.1 How to match volatility with the one-period model?

Let us recall the one-period model in Figure 3.5.1.

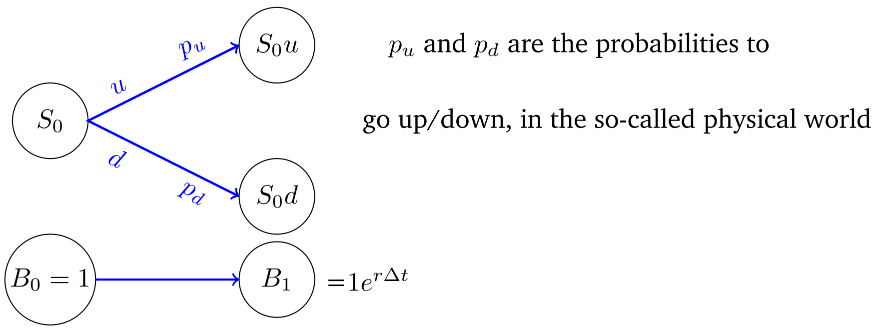



Figure 3.5.1: Example $N = 1$

The stock price either moves up by a proportional amount $u$ or moves down by a proportional amount $d$. The probability of an upward movement *in the real world* is $p_u$ ($p_d$ for the downward movement). The time step is of length $\Delta t = T/N$, and the evolution of the stock price is given by (compare with Figure 3.1.1)

$$S_{n+1} = S_n \cdot Z, \qquad Z \text{ is a rv s.th. } \mathbb{P}(Z = u) = p_u, \mathbb{P}(Z = d) = p_d. \tag{3.5.2}$$

Notation wise, we will use the notation $S(t_n)$ to identify stock prices of model (3.5.1) at time $t_n$ and $S_n$ to identify stock prices of model (3.5.2) at period $n$ that corresponds to time $t_n = n \times \Delta t$.

Our goal is to have the binomial model approximate the possible path of the stock price, in a way coherent with what has been observed, see Figure 3.5.2. For this, the values $u$ and $d$ must be appropriately chosen so that the binomial model can mimic (approximate) the stock evolution.

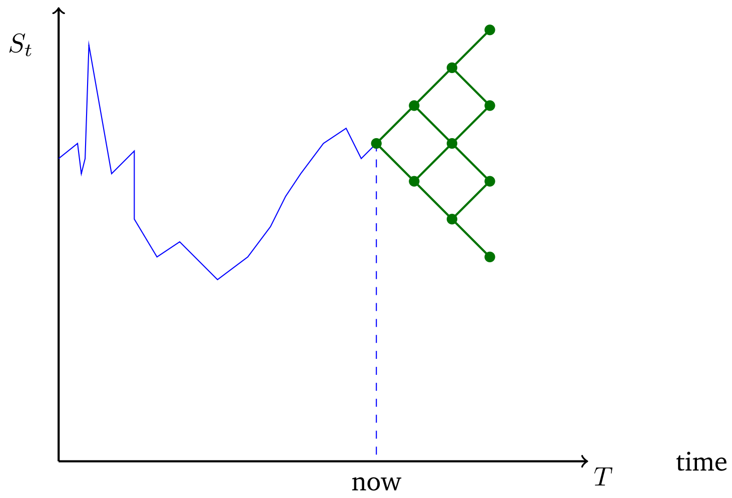



Figure 3.5.2: Using a binomial tree to model future price of a stock.

**The goal is,** assuming that $\sigma$ and $\mu$ have been computed from real world data, we want to choose $u$ and $d$ such that $S_{n+1}$ from (3.5.2) is as close as possible (under some optimizing criteria) to $S(t_{n+1})$ in (3.5.1).

**The goal is achieved by** matching the first two moments (Expectation and Variance) of both models. In other words, choose $u$, $d$, $p_u$ and $p_d = (1 - p_u)$ such that

$$i)\ \mathbb{E}^{\mathbb{P}}[S_1] = \mathbb{E}^{\mathbb{P}}[S(t_1)] \qquad \text{and} \qquad ii)\ \mathrm{Var}^{\mathbb{P}}[S_1] = \mathrm{Var}^{\mathbb{P}}[S(t_1)],$$

*where $S_1$ comes form the Binomial model and $S(t_1)$ comes from the Black-Scholes model.*

### Matching the expectations

We start with $i)$ $\mathbb{E}^{\mathbb{P}}[S_1] = \mathbb{E}^{\mathbb{P}}[S(t_1)]$. Where $t_0 = 0$ and $t_1 = \Delta t$ and $S(0) = S_0$.

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}[S_1] &= \mathbb{E}^{\mathbb{P}}[S_0 Z] = S_0(u \cdot p_u + d \cdot (1 - p_u)), \\
\mathbb{E}^{\mathbb{P}}[S(t_1)] &= \mathbb{E}^{\mathbb{P}}[S(t_0) + \mu S(t_0)\Delta t + \sigma S(t_0)\sqrt{\Delta t}\varepsilon] \\
&= S_0 + \mu S_0 \Delta t + \sigma S_0 \sqrt{\Delta t}\, \mathbb{E}^{\mathbb{P}}[\varepsilon] \\
&= S_0(1 + \mu \Delta t) \approx S_0(1 + e^{\mu \Delta t}),
\end{aligned}$$

where in the last line we used the approximation that for $x << 1$ it holds that

$$1 + x \approx e^x.$$

Given your knowledge in the Taylor, this approximation simply says that $e^x$ is approximated by the first two terms of its Taylor series. (Can you write down the Taylor Series for $e^x$ around point $x_0 = 0$?)

Since we are interested in having matching the moments we have

$$\mathbb{E}^{\mathbb{P}}[S_1] = \mathbb{E}^{\mathbb{P}}[S(t_1)] \Leftrightarrow S_0(u \cdot p_u + d \cdot (1 - p_u)) = S_0(1 + e^{\mu \Delta t}) \quad \Rightarrow \quad p_u = \frac{e^{\mu \Delta t} - d}{u - d}.$$

**Matching the variances**

We go now to $ii)$ $\text{Var}^{\mathbb{P}}[S_1] = \text{Var}^{\mathbb{P}}[S(t_1)]$.

Recall from Section 2.1 or Section 2.3 that for $X \sim \text{Bernoulli}(p_u)$ taking values $u$ with probability $p_u$ and $d$ with probability $p_d$ we have $\text{Var}[X] = p_u(1-p_u)(u-d)^2$ (please verify this yourself). And from Section 2.1 or Section 2.4, that $X \sim \mathcal{N}(0,1)$ we have $\text{Var}[X] = 1$.

$$\begin{aligned}\text{Var}^{\mathbb{P}}[S_1] &= \text{Var}^{\mathbb{P}}[S_0 Z] = S_0^2 p_u(1-p_u)(u-d)^2,\\ \text{Var}^{\mathbb{P}}[S(t_1)] &= \text{Var}^{\mathbb{P}}[S(t_0) + \mu S(t_0)\Delta t + \sigma S(t_0)\sqrt{\Delta t}\varepsilon]\\ &= \sigma^2 S_0^2 \Delta t\, \text{Var}^{\mathbb{P}}[\varepsilon] = \sigma^2 S_0^2 \Delta t.\end{aligned}$$

Recalling that $p_u = e^{\mu\Delta t} - d/(u-d)$ and that we are interested in having matching variances from both models we have then

$$\begin{aligned}\text{Var}^{\mathbb{P}}[S_1] = \text{Var}^{\mathbb{P}}[S(t_1)] \Leftrightarrow &\quad S_0^2 p_u(1-p_u)(u-d)^2 = \sigma^2 S_0^2 \Delta t\\ \Rightarrow &\quad p_u(1-p_u)(u-d)^2 = \sigma^2 \Delta t\\ \Rightarrow &\quad e^{\mu\Delta t}(u+d) - ud - e^{2\mu\Delta t} = \sigma^2 \Delta t. \end{aligned} \tag{3.5.3}$$

Since $\Delta t << 1$, one can use again the Taylor expansions trick on the exponential and ignore all higher powers of $\Delta t$, i.e. $(\Delta t)^n$ with $n > 1$. This leads to the following solution to the above equation

**Lemma 3.5.1.**

$$u = e^{\sigma\sqrt{\Delta t}} \qquad \textit{and} \qquad d = e^{-\sigma\sqrt{\Delta t}}.$$

*Note that* $u \cdot d = 1$.

*Proof.* Let us verify that the above $u$ and $d$ are indeed approximate solutions to (3.5.3). Ignoring the terms $|\Delta t|^n$ order $n > 1$ we have: $e^{2\mu\Delta t} \approx 1 + 2\mu\Delta t$ and

$$u = e^{\sigma\sqrt{\Delta t}} \approx 1 + \sigma\sqrt{\Delta t} + \sigma^2\Delta t \quad \text{and} \quad d = e^{-\sigma\sqrt{\Delta t}} \approx 1 - \sigma\sqrt{\Delta t} + \sigma^2\Delta t.$$

Coming back (3.5.3), replacing the approximations and ignoring (due to the products) all $\Delta t$ terms with power strictly greater then one, we have

$$\begin{aligned}&\frac{1}{S_0}\text{Var}^{\mathbb{P}}[S(t_1)]\\ &\quad = e^{\mu\Delta t}(u+d) - ud - e^{2\mu\Delta t}\\ &\quad \approx \Big(1 + \mu\Delta t\Big)\Big(2 + 2\sigma^2\Delta t\Big) - \Big(1 + 2\sigma^2\Delta t\Big) - \Big(1 + 2\mu\Delta t\Big) + o(\Delta t)\\ &\quad = 2 + 2\sigma^2\Delta t + 2\mu\Delta t - \big(1 + 2\sigma^2\Delta t\big) - \big(1 + 2\mu\Delta t\big) + o(\Delta t)\\ &\quad = 2\sigma^2\Delta t + o(\Delta t) \approx \frac{1}{S_0}\text{Var}^{\mathbb{P}}[S_1].\end{aligned}$$

The variances are then very close and as an approximation for small enough $\Delta t$ one can say that they match. □

**And under the risk-neutral measure $\mathbb{Q}$?**

The pricing mechanisms designed in the previous subsections all used the risk-neutral probability measure $\mathbb{Q}$; in fact, we saw that pricing and the physical measure $\mathbb{P}$ doesn't lead to the same type of results.

The computations we have done so far, to find the values for $p_u, p_d, u$ and $d$, we done using $\mathbb{P}$. What happens under $\mathbb{Q}$?

Under $\mathbb{Q}$, the probability $q_u$ is defined through

$$S_0 = e^{-r\Delta t}\mathbb{E}^{\mathbb{Q}}[S_1] \quad \text{where} \quad q_u = \frac{e^{r\Delta t} - d}{u - d}.$$

Comparing the expectations under the respective measures

$$\mathbb{E}^{\mathbb{P}}[S(t_1)] = \cdots = S_0(1 + \mu\Delta t) \approx S_0 e^{\mu\Delta t} \neq \mathbb{E}^{\mathbb{Q}}[S_1] = S_0 e^{r\Delta t}$$

Hence, $\mathbb{E}^{\mathbb{Q}}[S_1] \neq \mathbb{E}^{\mathbb{P}}[S(t_1)]$ unless $r = \mu$.

Comparing now the variances,

$$\begin{aligned} \text{Var}^{\mathbb{Q}}[S_1] &= \text{Var}^{\mathbb{Q}}[S_0 Z] = S_0^2 q_u(1 - q_u)(u - d)^2 \\ &= \cdots = S_0^2\sigma^2\Delta t, \\ \text{Var}^{\mathbb{P}}[S(t_1)] &= \cdots = S_0^2\sigma^2\Delta t. \end{aligned}$$

The conclusion is that if $u$ and $d$ are properly calibrated (i.e. $u = e^{\sigma\sqrt{\Delta t}}$ and $d = e^{-\sigma\sqrt{\Delta t}}$) then the binomial model (for $\Delta t$ small enough) and under $\mathbb{Q}$ does reproduce the model as pricing is concerned.

This analysis shows that when we move from the real world to the risk-neutral world the expected return on the stock changes but its volatility remains the same (at least in the limit as $\Delta t$ tends to zero). This is an illustration of an important general result known as *Girsanov's theorem* (which we unfortunately will not discuss). When we move from a world with one set of risk preferences to a world with another risk preferences, the expected growth rates in variables change but their volatilities remain the same. Moving from one set of risk preferences to another is sometimes referred to as *changing the (probability) measure*.

In practice, an analyst can expect to get only a very rough approximation to an option price by assuming that stock price movements during the life of the option consist of one or two binomial steps.

When binomial trees are used in option pricing, the life of the derivative is typically divided into 30 or more time steps of length $\Delta t$. In each time step there is a binomial stock price movement. With 30 time steps this means that 31 terminal stock prices and $2^{30}$, or about 1 billion, possible stock price paths are considered.

**Using the calibrated binomial model to compute prices**

Take a $N$-periods binomial model, each period with time-length $\Delta t$ and let $\Phi(S_N)$ be an European type claim. The bank account pays continuously interest rates at rate $r$. **Assume the market has volatility** $\sigma$. Then the price of $\Phi$ is given exactly as the previous sections, through Proposition 3.2.7 for the one-period models or like Proposition 3.3.6 for the multi-period ones.

In other words, simply compute $u$ and $d$ as in Lemma 3.5.1 and then proceed with pricing through Proposition 3.2.7 for the one-period models or like Proposition 3.3.6 for the multi-period ones.

Compute $q_u$ through

$$q_u = \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}} \qquad \text{and} \qquad q_d = \frac{e^{\sigma\sqrt{\Delta t}} - e^{r\Delta t}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}} = 1 - q_u.$$

and compute the prices at $n = 0$ (or $t = 0$) of an European type with the formula from Proposition 3.3.8, e.g.

$$\Pi\left(0; \Phi\right)\big|_{S_0} = \left(e^{-r\Delta t}\right)^N \sum_{k=0}^{N} \binom{N}{k} q_u^k\, q_d^{N-k}\, \Phi\Big(S_0 u^k d^{N-k}\Big). \tag{3.5.4}$$

### 3.5.2 The continuous-time limit of the multi-period binomial model

The main motivation for this section is the following question: If the discrete-time (binomial) model does indeed reproduce reality for $\Delta t$ small enough and can calibrate the market's volatility, can we take limits and compute the continuous-time case? What happens to the pricing formulas (3.5.4) when we $\Delta t \to 0$?

#### The setup

The market's continuous time interest rate is a constant $r$ and is of volatility $\sigma$. Let $\Phi\big(S(T)\big)$ be an European Call or put Option with maturity $T$ written on the stock $(S(t))_{t\in[0,T]}$.

Divide the time interval $[0, T]$ into $N$ sub-intervals with time-step $\Delta t := \frac{T}{N}$. Define the time points $t_n^N := n \times \Delta t$, $n \in \{0, 1, \dots, N\}$.

For all $n \in \{0, 1, \dots, N\}$, define the factors $u_N := e^{\sigma\sqrt{\Delta t}}$ and $d_N := e^{-\sigma\sqrt{\Delta t}}$ and the probabilities $q_{u,N}$ and $q_{d,N}$

$$q_{u,N} := \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}, \qquad \text{and} \qquad q_{d,N} := 1 - q_{u,N}.$$

#### The main results

**Theorem 3.5.2** (Continuous time limits of Call and Put option prices in the Black-Scholes market)**.** *Define $d_+$ and $d_-$ as follows*

$$d_+ = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} \qquad \textit{and} \qquad d_- = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}},$$

*(note that $d_- = d_+ - \sigma\sqrt{T}$) and let the map $\mathbb{R} \ni x \mapsto N(x) \in [0, 1]$ denote the cumulative Normal distribution function.*

*Then the continuous time limit (as $\Delta t \to 0$ or $N \to \infty$) of* (3.5.4) *for a Call or Put option are as follows.*

▷ *Let* $\Phi$ *be the* Call *option function with strike* $K$*. Then*

$$c_0 := \lim_{\Delta t \to 0} \Pi\left(0; \Phi\right)\big|_{S_0} = S_0 N\big(d_+\big) - K e^{-rT} N\big(d_-\big),$$

*where* $\Pi(0;\Phi)|_{S_0}$ *is given by* (3.5.4).

▷ *Let* $\Phi$ *be the* Put *option function with strike* $K$*. Then*

$$p_0 := \lim_{\Delta t \to 0} \Pi\left(0; \Phi\right)\big|_{S_0} = K e^{-rT} N\left(-d_-\right) - S_0 N\left(-d_+\right),$$

*where* $\Pi(0;\Phi)|_{S_0}$ *is given by* (3.5.4).

*Tip: try to solve Exercise 3.6.6.*

In practice, when given specific values for the involved parameters, the value of $N(\cdot)$ must then be found in the usual tables for the standard Normal cumulative probability function.

The above result is the so-called classical **Black-Scholes** European call/put option valuation formula. It is an asymptotic limit of option prices in a sequence of binomial type models (also known as Cox-Ross-Rubinstein model) with a special choice of parameters. In the continuous-time Black-Scholes model the dynamics of the (stochastic) stock price process $S(t)$ are modeled by a *geometric Brownian motion*, which implies that the asset returns $\ln S(t)$ are normally distributed, that will be studied in the next chapter.

We leave the proof of theorem 3.5.2 to the end of the section. The proof is highly educational.

The general case of the above result, i.e. a price formula for any point in time $t$ is stated in the next result that we present without proof. The missing proof can be obtained from that of the above result.

**Theorem 3.5.3** (Continuous time limits of Call and Put option prices General case)**.** *Let* $t \in [0,T]$*. Define* $d_+$ *and* $d_-$ *as follows*

$$d_\pm = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r \pm \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}},$$

*and let the map* $\mathbb{R} \ni x \mapsto N(x) \in [0,1]$ *denote the cumulative Normal distribution function.*

*Denote by* $C^E\left(t, S_t\right) := \Pi\left(t; \mathit{Call}(T,K)\right)$ *and* $P^E\left(t; S_t\right) := \Pi\left(t, \mathit{Put}(T,K)\right)$ *the prices of a Call option and a Put option of European type at a given time* $t$ *for the stock spot price* $S_t$*.*

*Then the continuous time limit (as* $\Delta t \to 0$ *or* $N \to \infty$*) of* (3.5.4) *for a Call or Put option are as follows.*

▷ *Let* $\Phi$ *be the* Call *option function with strike* $K$*. Then*

$$C^E\left(t, S_t\right) = \lim_{\Delta t \to 0} \Pi\left(t, \mathit{Call}(T,K)\right) = S_t N(d_+(t)) - K e^{-r(T-t)} N(d_-(t)).$$

▷ *Let* $\Phi$ *be the* Put *option function with strike* $K$*. Then*

$$P^E\left(t, S_t\right) = \lim_{\Delta t \to 0} \Pi\left(t, \mathit{Put}(T,K)\right) = K e^{-r(T-t)} N(-d_-(t)) - S_t N(-d_+(t)).$$

**Remark 3.5.4** (Put-Call parity). *One can verify that the put-call parity formula of Chapter 1 also holds for the above defined price formulas.*

We now prove Theorem 3.5.2.

*Proof of Theorem 3.5.2.* We show only the Call option case. Denote by $c_0$ the price of the Call option at time $t = 0$. The price formula $\Pi(0;\Phi)|_{S_0}$ is given by (3.5.4),

$$c_0 = \Pi^N(0,\Phi)\Big|_{S_0} = e^{-r\Delta t N}\sum_{k=0}^{N}\binom{N}{k} q_{u,N}^k q_{d,N}^{N-k}\left(S_0 u_N^k d_N^{N-k} - K\right)^+$$
$$= e^{-rT}\sum_{k=a_N}^{N}\binom{N}{k} q_{u,N}^k q_{d,N}^{N-k}\left(S_0 u_N^k d_N^{N-k} - K\right),$$

for some time point $a_n = \min\{k \in \mathbb{N}_0 \text{ such that } S_0 u_N^k d_N^{N-k} \geq K\}$. It follows then that

$$c_0 = S_0\sum_{k=a_N}^{N}\binom{N}{k}\left(\frac{q_{u,N}u}{e^{r\Delta t}}\right)^k\left(\frac{q_{d,N}d}{e^{r\Delta t}}\right)^{N-k} - Ke^{-rT}\sum_{k=a_N}^{N}\binom{N}{k}\left(\frac{q_{u,N}}{e^{r\Delta t}}\right)^k\left(\frac{q_{d,N}}{e^{r\Delta t}}\right)^{N-k}.$$

Having in mind the properties of the Binomial distribution with we observe that

$$\sum_{k=a_N}^{N}\binom{N}{k}\left(\frac{q_{u,N}u}{e^r\Delta t}\right)^k\left(\frac{q_{d,N}d}{e^{r\Delta t}}\right)^{N-k} = \mathbb{P}\left(a_N \leqslant Y_N \leqslant N\right),$$

where $Y_N \sim$ Binomial $\left(N, q_{u,N}\cdot e^{\sigma\sqrt{\Delta t}-r\Delta t}\right)$.

The Central limit theorem, Theorem 2.5.2, implies

$$\mathbb{P}\left[\frac{a_N - \mathbb{E}[Y_N]}{\sqrt{Var[Y_N]}} \leq \frac{Y_N - \mathbb{E}[Y_N]}{\sqrt{Var[Y_N]}} \leq \frac{N - \mathbb{E}[Y_N]}{\sqrt{Var[Y_N]}}\right] \longrightarrow^{N\to\infty} N(d_-) - N(d_+),$$

where $d_- = \frac{\ln\left(\frac{S_0}{K}\right)+(r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$ and $d_+ = \frac{\ln\left(\frac{S_0}{K}\right)+(r+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$. Compare with $d_+$ and $d_-$. □

## 3.6 Exercises

**Exercise 3.6.1.** *The current price of a stock is £30. It is known that the stock price will either increase by 8% or decrease by 10% during each 2-month period for next 4 months. The yearly risk-free rate of interest is 5% (cont. compounding).*

*Find the price of a European-type derivative written on this non-dividend paying stock that pays*[5] $f(S_T) = \left(\max\{S_T - 30, 0\}\right)^2$, *where* $S_T$ *is the stock price at maturity (4 months). Start by pointing out how many periods the model has (i.e.* $T$ *and* $\Delta t$*).*

*If the derivative is American-style, then should it be exercised early?*

**Exercise 3.6.2.** *Using the Binomial model of the previous exercise, compute the price of an European type option with pay-off function* $\Phi(x) = \left(\max(30 - x, 0)\right)^2$. *The claim is* $\Phi(S_2) := [\max(30 - S_2, 0)]^2$.

[5]For clarity $f(x) = (\max\{x, 0\})^2$.

**Exercise 3.6.3** (Question 1 of Aug/2014 FiM Exam). *Let us assume that today's stock price is £$48$ and it is expected either to increase by 5% or to decrease by 6% every month over the next two 3-months periods (two-period Binomial tree). Also suppose that the continuously compounded risk-free rate of interest is 7% per annum.*

(a) *Find the price of a 6-month European call option with strike price £$50$, written on a non-dividend paying stock.* ***[6 Marks]***

(b) *Find the price of a 6-month European put option with strike price £$50$, written on a non-dividend paying stock.* ***[6 Marks]***

(c) *Verify if put-call parity relationship holds using the prices calculated in (a) and (b).* ***[3 Marks]***

(d) *If the call and put options considered above were American instead of European, then in each case verify if an early exercise is optimal.* ***[10 Marks]***

**Exercise 3.6.4** (European/American option; volatility; - Question 1 of FMI 2013 Exam).

1. *Calculate the price of a one-year* **European call** *option on a non-dividend-paying stock when the current stock price is $\pounds 10$, the strike price is $\pounds 11$, the continuously compounded risk-free rate is 2% per annum and the volatility is 30 % per annum:*

   (a) *Using a two-period binomial tree ($\Delta t = \frac{1}{2}(T-t) = \frac{1}{2}$).* *[4 marks]*

   (b) *Using a three-period binomial tree ($\Delta t = \frac{1}{3}(T-t) = \frac{1}{3}$).* *[6 marks]*

2. *Calculate the price of a one-year* **American put** *option on a non-dividend-paying stock when the current stock price is $\pounds 10$, the strike price is $\pounds 11$, the continuously compounded risk-free rate is 2% per annum and the volatility is 30 % per annum:*

   (a) *Using a two-period binomial tree ($\Delta t = \frac{1}{2}(T-t) = \frac{1}{2}$).* *[7 marks]*

   (b) *Using a three-period binomial tree ($\Delta t = \frac{1}{3}(T-t) = \frac{1}{3}$).* *[8 marks]*

**Exercise 3.6.5** (Multi-period Binomial model; American option; Volatility). *Using a four-period binomial tree, calculate the price of a fourth-month American put option on a non-dividend-paying stock when the current stock price is $\pounds 100$, the strike price is $\pounds 100$, the risk-free rate is 5% per annum and the volatility is 25 % per annum ($\Delta t = 1/12$).*

**Exercise 3.6.6** (Pricing using limit if binomial model). *Suppose that the stock price is £$52$, risk-free rate is 12% per annum and volatility is 30% per annum. Find the price of a 3-month European call option on a non-dividend paying stock with strike price £$50$.*

# Chapter 4

# Stochastic analysis

In Chapter 3 the stock price was modeled in a discrete time market with its price evolving according to a binomial tree. Given the starting price $S_0$ the possible subsequence prices for the stock are given recursively through the model

$$S_{n+1} = S_n \cdot Z, \quad \text{where } Z \sim \text{Ber}(p_u) \text{ and } Z \text{ takes values in } Z \in \{u, d\}.$$

Given its $N$ periods, the models covers only a finite amount of time points in the time interval $[0, T]$, namely the points $t_n = n \times \Delta t$ for $n = 0, \cdots, N$ and where $\Delta t := T/N$.

In this chapter, stock prices will be modeled using a continuous time stochastic processes defined for all time points $t \in [0, T]$. More precisely, we will work with the continuous time limit version of the process in (3.5.1), the so-called geometric Brownian motion, which is given for any $t \in [0, T]$ by

$$S_t = S_0 \exp\left\{(\mu - \frac{\sigma^2}{2})t + \sigma W_t\right\}, \qquad \text{where } W_t \sim \mathcal{N}(0, t).$$

Here $S$ is a stochastic process and so is $W$. The latter in fact, is known has the Brownian motion while $S$ is the geometric Brownian motion. We will start by discussing $W$ and its properties to then try to reach a theory that allows us to repeat the pricing arguments of Chapter 3 and find prices for options. The use of the geometric Brownian motion as a model for stock prices dates back to the seventies and is due to Black & Scholes.

## 4.1 The basics

In recent years, use of the theory of stochastic differential equations in the area of financial studies has significantly increased. The success of the Black-Scholes formula, which price an European call option under the assumption of a lognormally distributed underlying asset prices, has initiated the introduction of Ito's calculus into the daily life of market practitioners.

In general, not every subset of $\Omega$ (*sample space*) is observable or likely to occur. Therefore, let $\mathcal{F}$ denote the set of all observable events for a single trial. This set must include $\Omega$ (the certain event), the impossible set $\emptyset$ (the empty set) and for every event $A$, its complement $A^C$. Furthermore, given two events $A$ and $B$, the union $A \cup B$ and the intersection $A \cap B$ also belong to $\mathcal{F}$. Thus, $\mathcal{F}$ is an *algebra of events*. If we create a family $\mathcal{F}$ that contains all the subsets of $\Omega$ that are observable and has the following properties:

(1) $\emptyset \in \mathcal{F}$, where $\emptyset$ is the empty set.

(2) If $A \in \mathcal{F}$, then the complement of it $A^C \in \mathcal{F}$.

(3) $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ when $A_i \in \mathcal{F}$ for $i \geq 1$.

then the family is called a *sigma-algebra*, a special case of an *algebra of events*. Moreover, the triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *probability space*, where $\Omega$ is the space of *elementary events* $\omega$ ("*market situations*", in the present context);

$\mathcal{F}$ is some *sigma-algebra* of subsets of $\Omega$ (the set of " *observable market events*");

$\mathbb{P}$ is a probability (or *probability measure*) on $\mathcal{F}$.

The probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is equipped with a *filtration*: a nondecreasing family $\{\mathcal{F}_t; t \geq 0\}$ of sub-sigma-fields of $\mathcal{F} : \mathcal{F}_t \subseteq \mathcal{F}_s \subseteq \mathcal{F}_\infty$ for all $0 \leq t < s < \infty$. We set $\mathcal{F}_\infty = \sigma(\bigcup_{t \geq 0} \mathcal{F}_t)$.

A *stochastic process* is a collection of random variables $X = \{X_t; 0 \leq t < \infty\}$ on the *sample space* $(\Omega, \mathcal{F})$, which takes values in a second measurable space $(\Omega', \mathcal{Q})$, called the *state space*. For a fixed sample point $\omega \in \Omega$, the function $t \mapsto X_t(\omega); t \geq 0$ is the *sample path* (realization, trajectory) of the process $X$ associated with $\omega$. It provides the mathematical model for a random experiment whose outcome can be observed continuously in time (e.g. the trajectory for the prices of a stock subjected to the random disturbances represented by Brownian motion). It may be useful to think of $t$ as "time" (particularly when we apply the theory of stochastic calculus to financial studies) and each $\omega$ as an individual "particle" or "experiment". It may be also helpful to consider that the picture $X_t(\omega)$ represents the position (or result) at time $t$ of the particle (or experiment) $\omega$. Furthermore, we may write $X(t, \omega)$ instead of $X_t(\omega)$, to regard the process as a function of two variables

$$\begin{aligned} X : [0, T] \times \Omega &\to \mathbb{R}^n \\ (t, \omega) &\mapsto X(t, \omega). \end{aligned}$$

Moreover, it may be useful to think of $F_t$ as the set of events *observable through time* $t$. In other words, $\mathcal{F}_t$ is the *"information"* on the market situation that is available to an observer up to time $t$ inclusive. Therefore, it is natural to assume that $S(t)$ is $\mathcal{F}_t$-measurable, or that (using a more descriptive language) prices are formed on the basis of the developments observable on the market up to time $t$. We say they are *adapted* to the flitration.

## 4.2 The Brownian motion

Brownian motion is the name given to the irregular movement of pollen, suspended in water, observed by the Scottish botanist Robert Brown. The range of applications of Brownian motion covers areas such as physics, biology, economics and many more. The first quantitative work on Brownian motion is due to Bachelier (1900), but Einstein (1905) derived the transition density for it. A rigorous treatment of Brownian motion began with N. Wiener (1923), who provided the first existence proof.

**Definition 4.2.1** (The Wiener process or Brownian motion)**.** *Consider the stochastic process $(W_t)_{t\geq 0}$. The process $W$ is called a* Brownian motion *or the* Wiener process *if it satisfies the following points.*

1. $W_0 = 0$ *with probability* $1$.
2. *It is a* Gaussian process*. That is, for all $0 \leq t < s$ the random variable $W_s - W_t$ is normally distributed with mean zero and variance $s - t$.*
3. *It is an independent increments process, i.e. for any $r < s \leq t < u$ then $W_u - W_t$ and $W_s - W_r$ are independent random variables.*
4. *The sample paths of the process are almost surely continuous. That is, the functions from $[0, \infty)$ to $\mathbb{R}$ given by $t \mapsto W_t(\omega)$ are almost surely continuous.*

At this point, we strongly suggest the reader to try Exercise 4.7.1.

Due to the random behavior of the Brownian motion even though for fixed $\omega \in \Omega$ the paths are continuous, i.e. the functions $t \mapsto W_t(\omega)$ are continuous, they are not differentiable. Worst yet, the path function is no-where differentiable - it is not differentiable at any point.

Add picture of Brownian Motion

**Example 4.2.2** (Is a given process a Brownian motion?)**.** *Let $(W_t)_{t\geq 0}$ be a Brownian motion and let $c \in (0, \infty)$.*

*Define the a new process $B_t := cW_{c^{-2}t}$. Is $(B_t)_{t\geq 0}$ a Brownian motion?*

*To solve this problem we need only to verify the defining properties of the Brownian motion.*

1) *at $t = 0$, $B_0 = cW_{c^{-2}0} = cW_0 = 0$ ✓*

2) *Goal: show that $B$ has independent increments, i.e. $B_u - B_t \perp\!\!\!\perp B_s - B_r$ (for $r < s \leq t < u$) using that $W$ has independent increments.*

   *This one is easy since if $r < s \leq t < u$ then one also has $\frac{r}{c^2} < \frac{s}{c^2} \leq \frac{t}{c^2} < \frac{u}{c^2}$; using that the increments of $W$ are independent for the times $\frac{r}{c^2} < \frac{s}{c^2} \leq \frac{t}{c^2} < \frac{u}{c^2}$ yields that the increments of $B$ are also independent (simply the way $B$ is defined). ✓*

3) *Take $s < t$. Goal: Show $B_t - B_s \sim \mathcal{N}(0, t - s)$.*

   *We have*

$$B_t - B_s = cW_{c^{-2}t} - cW_{c^{-2}s} = c\big(W_{c^{-2}t} - W_{c^{-2}s}\big).$$

   *Notice now that $W_{c^{-2}t} - W_{c^{-2}s} \sim \mathcal{N}(0, c^{-2}t - c^{-2}s)$ (this follows from property iii) for the BM $W$). Moreover, from the properties of the BM again*

$$\sqrt{c^{-2}}(W_t - W_s) = c^{-1}(W_t - W_s) \sim \mathcal{N}(0, c^{-2}t - c^{-2}s) = \mathcal{N}\big(0, c^{-2}(t - s)\big).$$

   *Bringing all together, we have*

$$B_t - B_s = c\big(W_{c^{-2}t} - W_{c^{-2}s}\big) \sim cc^{-1}(W_t - W_s) \sim \mathcal{N}(0, t - s).✓$$

4) *Fix $\omega$. Using the definition of $\tilde{W}$, we have $t \mapsto \tilde{W}_t(\omega) = cW_{c^{-2}t}(\omega)$. Since we know that the BM $t \mapsto W_t(\omega)$ is a continuous function of $t$, then for any $\alpha > 0$ so is $t \mapsto W_{\alpha t}$. Take $\alpha = c^{-2}$ and the argument follows.* ✓

   *Think of it this way, one the $\omega$ is fixed, then $W_t(\omega)$ is a simple function of $t$; if this is confusing name it $g(t) := W_t(\omega)$. Now, say that $g$ is a continuous function, what can you say about the continuity of the function $g(c^{-2}t)$? The constant changes nothing, right? Composition of continuous functions is still continuous.*

## 4.3 Stochastic Integration

The big theoretical development in this chapter is the development of the theory flowing naturally from the continuous time limit results of the Binomial model, see Section 3.5.2.

In a market where stocks can be bought and sold, we saw in Chapter 3 that if $\Delta_0$ is the number of units of stocks one has in the portfolio at time $t = 0$ then the market value of such a portfolio is simply $\Delta_0 S_0$.

We give now the broad idea of what we are about to do. In this chapter we model the evolution of stocks in continuous time; let $(S_t)_{t\in[0,T]}$ be the said stock price process. Let us denote $(\pi_t)_{t\in[0,T]}$ the number of units of stocks held at time $t$, then the value of the portfolio over the time interval $[0,t]$ is (recall that we go from discrete time to continuous time)

$$\text{Value of the portfolio} = \int_0^t \pi_u \mathrm{d}S_u.$$

Note here the essential requirement that at time $t$ the strategy $\pi_t$ cannot have access to information (on $S_t$) *posterior* to time $t$. The process $\pi$ cannot know the future, we call this property *adaptness* or we say $\pi$ is *adapted* to the filtration.

This is all fine, but assume for an instant that $S$ is the Brownian motion $W$? What is the meaning of $\int f(u)dW_u$ for some map $f$? Were $u \mapsto W_u$ a differentiable function this integral would be very easy to compute. In fact, one would use the ideas of *Line Integrals*[1] and write

$$\int f(u)\mathrm{d}W_u = \int f(W_u)|\frac{\mathrm{d}W_u}{\mathrm{d}u}|\mathrm{d}u.$$

The point is that $\frac{\mathrm{d}W_u}{\mathrm{d}u}$ **does not exist!** And hence what is written above is pure rubbish. The other implication is that to go further a theory to stochastic integration must be introduced.

In the next pages we give superficial ideas on how stochastic integration works. The core theory is left for another course (*Probability, Measure & Finance (MATH10024)*). For reference's sake, we suggest Chapter 4 of [Bjo09].

### 4.3.1 Simple integrands

Just like for Riemann integrals, we start by defining the integral for "simple" integrands.

[1]Don't remember? Check wikipedia: https://en.wikipedia.org/wiki/Line_integral

**Definition 4.3.1** (Simple process). *The stochastic process* $(g_t)_{t\geq 0}$ *is called a* simple processes *if it is of the following form*

$$g_t(\omega) = \sum_{i=0}^{\infty} g_i(\omega)\mathbb{1}_{[t_i,t_{i+1})}(t), \tag{4.3.1}$$

*where* $g_i$ *are are random variables that are adapted to the information up to time* $t_i$ *and* $0 = t_1 < t_2 < \cdots$ *are real numbers.*

Note that $g_t$ is the same as step function where at each step one has a random variable.

Let us define, for real numbers $a, b$ the number $a \wedge b := \min\{a, b\}$. The *stochastic integral for a simple process* is then

$$X_t = \int_0^t g_s \mathrm{d}W_s := \sum_{i=0}^{\infty} g_i \times (W_{t_{i+1}\wedge t} - W_{t_i \wedge t}).$$

Note that for any *finite* time $t$, the sum will be finite as well. So everything is well-defined.

**Example 4.3.2** (Integrating constants). *Let* $a \in \mathbb{R}$ *and assume* $g_t = a$ *for any* $t \in [0, T]$. *Then direct calculations show* $I := \int_0^t a\mathrm{d}W_u = a(W_t - W_0) = aW_t$.

*Moreover,* $I \sim \mathcal{N}(0, a^2 t)$.

The key property for the whole stochastic integration theory is the so called *Itô Isometry*, which allows to pass from stochastic integrals (which we are trying to build) to Riemann/Lebesgue integrals (which we know how to deal with).

**Lemma 4.3.3** (The Itô isometry for simple processes). *Take* $T \in (0, \infty)$ *and let* $(g_s)_{s\in[0,T]}$ *be a simple process. Then*

$$\mathbb{E}\left[\left(\int_0^T g_s \mathrm{d}W s\right)^2\right] = \mathbb{E}\left[\int_0^T |g_s|^2 \mathrm{d}s\right].$$

*Proof.* Using the definition of Itô stochastic integral for simple processes we get

$$I := \mathbb{E}\left[\left(\int_0^T g_s \mathrm{d}W s\right)^2\right] = \mathbb{E}\left[\left(\sum_{i=1}^{\infty} g_i \Delta W_i\right)^2\right] = \sum_{i,j=1}^{\infty} \mathbb{E}\Big[g_i g_j \Delta W_i \Delta W_j\Big],$$

where we define $\Delta W_i := W_{t_{i+1}\wedge T} - W_{t_i \wedge T}$.

If $i < j$ then $g_i g_j \Delta W_i$ is independent of $\Delta W_j$. This is because $g_i g_j \Delta W_i$ is $t_j$-information adapted and by definition of Brownian motion, the process $W_{t_{j+1}} - W_{t_j}$ has no "memory" (is independent) of the random variables before time $t_j$.

Hence

$$\mathbb{E}\left[g_i g_j \Delta W_i \Delta W_j\right] = \mathbb{E}\left[g_i g_j \Delta W_i\right] \mathbb{E}\left[\Delta W_j\right] = 0,$$

given that $\Delta W_j$ is normally distributed with mean $0$.

If $j < i$ then the same argument leads to $\mathbb{E}\left[g_i g_j \Delta W_i \Delta W_j\right] = 0$. Finally if $i = j$ we note that $g_i^2$ is a $t_i$-information adapted random variable and again by definition of the Brownian motion the process $\Delta W_i$ has no "memory" (is independent) of the random variables before time $t_i$.

We have then

$$\mathbb{E}\left[g_i^2(\Delta W_i)^2\right] = \mathbb{E}\left[g_i^2\right]\mathbb{E}\left[(\Delta W_i)^2\right] = \mathbb{E}\left[g_i^2\right](t_{i+1}\wedge T - t_i \wedge T),$$

since $\Delta W_i$ is normally distributed with mean $0$, variance $t_{i+1}\wedge T - t_i\wedge T$. Hence

$$I = \sum_{i=1}^{\infty}\mathbb{E}\left[g_i^2\right]\times\left(t_{i+1}\wedge T - t_i\wedge T\right) = \int_0^T \mathbb{E}\left[|g_t|^2\right]\mathrm{d}t.$$

□

**Example 4.3.4.** *Let $g_t(\omega) = t$ for any $t$, in other words a deterministic function. Then,*

$$\mathbb{E}\Big[\Big(\int_0^T g_u \mathrm{d}W_u\Big)^2\Big] = \mathbb{E}\Big[\Big(\int_0^T u\mathrm{d}W_u\Big)^2\Big] = \mathbb{E}\Big[\int_0^T u^2\mathrm{d}u\Big] = \frac{u^3}{3}\Big|_0^T = \frac{T^3}{3}.$$

### 4.3.2 The general case

The results of the Itô Isometry are revealing. They essentially hint that one will be able to define the stochastic integral for any adapted integrand process as long as the process is square-integrable.

**Definition 4.3.5** (Space of square integrable adapted processes). *A stochastic process $g$ belongs to the class $L^2(\Omega\times[0,T])$ if it is information adapted and if*

$$\|g\|_{L^2(\Omega\times[0,T])} := \mathbb{E}\left[\int_0^T |g_s|^2\mathrm{d}s\right]^{\frac{1}{2}} < \infty.$$

*Notationwise we omit the $\Omega$ when we write $L^2(\Omega\times[0,T])$ and refer to it as simply $L^2([0,T])$.*

Note that a simple process will surely belong to the class $L^2([0,T])$ but a process in the $L^2([0,T])$ class is not necessarily a simple process.

The general definition of stochastic integral follows through a usual approximation proceadure. To define a stochastic integral for a general $g\in L^2([0,T])$ we create an approximating sequence of simple processes $g^n$ such that $g^n\to g$ in the norm of $L^2(\Omega\times[0,T])$.

Indeed, a general Itô integrable processes $(g_s)_{s\in[0,T]}$ can be approximated by taking a partition of $[0,T]$ given by $0=t_0<t_1<\cdots<t_n=T$ with $t_i-t_{i-1}=\tau_n := T/n$ and defining the simple process

$$g_t^n = \sum_{i=0}^{n-1} g_{t_i}\mathbb{1}_{[t_i,t_{i+1})}(t)$$

It is important to note that we have chosen the value of the process at the left-hand point of the interval $[t_i,t_{i+1})$. With stochastic integrals (unlike classical Riemann–Stieltjes integrals) this is very important and with different choice we could obtain different notions of stochastic integrals. It can then be shown that

$$\mathbb{E}\left[\int_0^T |g_s - g_s^n|^2 ds\right] \to 0 \text{ as } n\to\infty.$$

We can define

$$\int_0^T g_s \mathrm{d}Ws := \lim_{n\to\infty} \int_0^T g_s^n \mathrm{d}Ws,$$

where the limit is understood in the sense of the $L^2([0,T])$ norm defined above.

In this framework we can prove the Itô isometry for general $L^2([0,T])$ processes.

**Lemma 4.3.6** (The Itô isometry). *For any process $(g_s)_{s\in[0,T]} \in L^2[0,T]$ we have*

$$\mathbb{E}\left[\left(\int_0^T g_s \mathrm{d}Ws\right)^2\right] = \mathbb{E}\left[\int_0^T |g_s|^2 \mathrm{d}s\right].$$

**Corollary 4.3.7.** *For any two $L^2([0,T])$ processes $X$ and $Y$ we have*

$$\mathbb{E}\left[\left(\int_0^T X_s \mathrm{d}W_s\right)\left(\int_0^T Y_s \mathrm{d}W_s\right)\right] = \mathbb{E}\left[\int_0^T X_s Y_s \mathrm{d}s\right].$$

*In particular, if $Y_t = 1$ for all $t \in [0,T]$, then $\forall 0 \le a \le b \le T$*

$$\mathbb{E}\left[\left(\int_a^b X_s \mathrm{d}W_s\right)(W_b - W_a)\right] = \mathbb{E}\left[\int_a^b X_s \mathrm{d}s\right].$$

*We leave the easy proof of this corollary for the exercises.*

Here are some important properties of Itô integrals that are very similar to those of standard integration, except the last one which is very important.

**Theorem 4.3.8** (Properties of the Itô integral). *Let $(g_t)_{t\in[0,T]}$ and $(f_t)_{t\in[0,T]}$ be processes in $L^2([0,T])$ and let $0 \le S < U < T$. Then:*

1. *We have*

$$\int_S^T g_t \mathrm{d}W_t = \int_S^U g_t \mathrm{d}W_t + \int_U^T g_t \mathrm{d}W_t \qquad \textit{almost surely.}$$

2. *For any $\alpha \in \mathbb{R}$, a constant, we have*

$$\int_0^T (\alpha g_t + f_t)\,\mathrm{d}W_t = \alpha \int_0^T g_t \mathrm{d}W_t + \int_0^T f_t \mathrm{d}W_t \qquad \textit{almost surely.}$$

3. *For any fixed $t \in [0,T]$ the stochastic process $X$ given by the integral $X_t := \int_0^t g_s \mathrm{d}W_s$ is a square-integrable information adapted process such that*

$$\mathbb{E}\left[\int_0^t g_s \mathrm{d}W_s\right] = 0 \quad \textit{for any } t \in [0,T].$$

**Sort of an example**

Having defined a stochastic integral is of course nice but how to work with it? having to go through a simple process approximation is copmlicated and inconvinient. Let us see it once.

**Example 4.3.9.** *Let us calculate*

$$\int_0^T W_s \mathrm{d}W s.$$

*We approximate* $(W_s)_{s\in[0,T]}$ *by the simple process*

$$g_s^n := \sum_{i=0}^{n-1} W_{t_i} \mathbb{1}_{[t_i,t_{i+1})}(s),$$

*where* $t_i = iT/n$. *We will always sum from* $0$ *to* $n-1$. *Let* $\Delta W_i := W_{t_{i+1}} - W_{t_i}$. *Then the integral is the limit, as* $n \to \infty$, *of*

$$\sum_i W_{t_i} \Delta W_i.$$

*We begin by observing that*

$$W_T^2 = \sum_i \left(W_{t_{i+1}}^2 - W_{t_i}^2\right) = \sum_i \left(W_{t_{i+1}} - W_{t_i}\right)^2 + 2\sum_i W_{t_i}\left(W_{t_{t+1}} - W_{t_i}\right). \quad (4.3.2)$$

*This is simply a consequence of the identity, true for* $a, b \in \mathbb{R}$,

$$(a-b)b = \frac{1}{2}\left(a^2 - b^2 - (a-b)^2\right).$$

*Now we wish to take the limit, as* $n \to \infty$, *in* (4.3.2).

*It can be shown that*

$$\sum_{i=0}^{n-1} \left(W_{t_{i+1}} - W_{t_i}\right)^2 \to T \quad \textit{in} \quad L^2(\Omega).$$

*We know that*

$$\sum_{i=0}^{n-1} W_{t_i}\left(W_{t_{i+1}} - W_{t_i}\right) = \sum_{i=0}^{n-1} W_{t_i}\Delta W_i \to \int_0^T W_s \mathrm{d}W s \quad \textit{in} \quad L^2(\Omega).$$

*Thus, in the limit,* (4.3.2) *becomes*

$$W_T^2 = T + 2\int_0^T W_s \mathrm{d}W s.$$

This example shows that it is a lot of work to calculate something as simple as

$$\int_0^T W_s \mathrm{d}W s = \frac{1}{2}W_T^2 - \frac{1}{2}T. \quad (4.3.3)$$

Recall that in "classical" calculus, where $f : [0,T] \to \mathbb{R}$ is differentiable, we have the Fundamental theorem of calculus telling us that

$$\int_a^b f'(s)\mathrm{d}s = f(b) - f(a), \quad a, b \in [0,T].$$

Clearly, if (4.3.3) is correct (and it is), in Itô calculus things work differently. But how? The answer is given by the Itô formula below.

## 4.4 Stochastic differential equations

Stochastic integrals are important for any theory of *stochastic calculus* since they can meaningfully defined. It is very common though to use a new notation for expressions such as

$$X_t = x_0 + \int_0^t g_u \mathrm{d}Wu, \qquad \text{note that } X_0 = x_0 \in \mathbb{R}.$$

The shorthand comes from "differentiating" the above equation and is

$$dX_t = g(t)\mathrm{d}W_t, \qquad \text{and } X_0 = x_0 \qquad \text{for every } t \in [0, T]$$

Think of $\mathrm{d}W_t$ as being a normal random variable with mean zero and variance $dt$. These two forms for the stochastic integral are meant to be equivalent. The latter form resembles the form of an ordinary differential equation. Pushing this idea further, given two functions $f$ and $h$ the following equation in differential form

$$dX_t = f(t, X_t)\mathrm{d}t + h(t, X_t)\mathrm{d}W_t, \qquad X_0 = x_0 \quad \text{for every } t \in [0, T],$$

known as a **stochastic differential equation (SDE)** , can be rewritten as

$$X_t = x_0 + \int_0^t f(u, X_u)\mathrm{d}u + \int_0^T h(u, X_u)\mathrm{d}W_u, \qquad \text{for any } t \in [0, T].$$

**Example 4.4.1** (Brownian motion with drift)**.** *Consider the following the simple stochastic differential equation:*

$$dX_t = \mu\mathrm{d}t + \sigma\mathrm{d}W_t, \qquad X_0 = 0 \quad \textit{for every } t \in [0, T],$$

*where $\mu$ and $\sigma$ are positive constants. Then,*

$$\begin{aligned} X_T = \int_0^T \mu\mathrm{d}t + \int_0^T \sigma\mathrm{d}W_t &= \mu T + \sigma(W_T - W_0) \\ &= \mu T + \sigma W_T. \end{aligned}$$

*In other words, $X_T$ is a normal random variable with mean $\mu T$ and variance $\sigma^2 T$. Seen as a stochastic process $(X_t)_{t\in[0,T]}$ has a distribution at every time point $t \in [0, T]$ given by $X_t \sim \mathcal{N}(\mu t, \sigma^2 t)$.*

### 4.4.1 Itô processes and Itô's formula

Given the general form that an SDE can take, a process that is defined as the sum of a standard integral plus an Itô one we define first an Itô process then the Itô formula.

**Definition 4.4.2** (Itô process)**.** *Take $t \in (0, \infty)$ and let $(W_t)_{t\in[0,T]}$ be a Brownian motion. An* Itô process *is a stochastic process $(X_t)_{t\in[0,T]}$ of the form*

$$X_t = x_0 + \int_0^t U_s \mathrm{d}s + \int_0^t V_s \mathrm{d}W_s \tag{4.4.1}$$

*where $x_0 \in \mathbb{R}$ and the processes $(U_t)_{t\in[0,T]}$ and $(V_t)_{t\in[0,T]}$ are adapted and square-integrable processes.*

Recall that writing (4.4.1) is equivalent to write $\mathrm{d}X_t = U_t\mathrm{d}t + V_t\mathrm{d}W_t$, $X_0 = x_0$ in the shorthand differential notation.

**The Itô formula or chain rule for stochastic processes**

Before we go into the main result, let us go over an example from classic analysis. Take three functions, $f : [0,T] \times \mathbb{R} \to \mathbb{R}$, $g : [0,T] \to \mathbb{R}$ and $h(t) := f\big(t, g(t)\big)$. Let us compute $\frac{\mathrm{d}}{\mathrm{d}t}h(t)$.

Since is the composition of functions, the rule we use here is the standard chain. There is just one little adjustment, in that the variation in $t$ of $h$ comes from the variation of $g$ and also from the $T$ component in $f$. Taking that into account we have

$$\frac{\mathrm{d}}{\mathrm{d}t}h(t) = \big(\partial_t f\big)\big(t, g(t)\big) + \big(\partial_x f\big)\big(t, g(t)\big)\frac{\mathrm{d}}{\mathrm{d}t}g(t).$$

Since we want to compare with Brownian motion, we cannot take derivatives but deal only with the infinitesimal changes. The above can then be written as (we are passing the $\mathrm{d}t$ term on the left hand side to the right hand side; we drop some parenthesis for easiness of presentation),

$$\mathrm{d}h(t) = \partial_t f\big(t, g(t)\big)\mathrm{d}t + \partial_x f\big(t, g(t)\big)\mathrm{d}g(t). \tag{4.4.2}$$

Can the same thing be done for stochastic processes? In other words, when $g(t)$ is a stochastic process and not a function? If you quickly compare (4.4.2) with (4.4.3) below, you will see a fundamental difference: the second derivative term! It appear there exactly because the Brownian motion in not differentiable and hence a correction to (4.4.2) is needed.

We have then the following important result.

**Theorem 4.4.3** (The Itô formula). *Let $(X_t)_{t\in[0,T]}$ be an Itô process given by*

$$\mathrm{d}X_t = U_t\mathrm{d}t + V_t\mathrm{d}W_t, \quad X_0 = x_0.$$

*Let $f \in C^{1,2}\left([0,T] \times \mathbb{R}, \mathbb{R}\right)$ and define the new process $Y_t := f(t, X_t)$, $t \in [0,T]$.*

*Then $Y$ is again an Itô process and satisfies the following stochastic differential equation*

$$\mathrm{d}Y_t = \frac{\partial f}{\partial t}(t, X_t)\mathrm{d}t + \frac{\partial f}{\partial x}(t, X_t)\mathrm{d}X_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X_t)\langle \mathrm{d}X_t, \mathrm{d}X_t\rangle, \tag{4.4.3}$$

*where $Y_0 = f(0, x_0)$ and $\langle \mathrm{d}X_t, \mathrm{d}X_t\rangle = V_t^2\mathrm{d}t$ and is the quadratic variation of $X$.*

*Fully replacing the equation for $\mathrm{d}X$, we have further*

$$\mathrm{d}Y_t = \Big[\frac{\partial f}{\partial t}(t, X_t) + U_t\frac{\partial f}{\partial x}(t, X_t) + V_t^2\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X_t)\Big]\mathrm{d}t + \Big[V_t\frac{\partial f}{\partial x}(t, X_t)\Big]\mathrm{d}W_t$$

*which reads equivalently in integral form*

$$\begin{aligned} Y_t = f(0, x_0) &+ \int_0^t \Big[V_s\frac{\partial f}{\partial x}(s, X_s)\Big]\mathrm{d}W_s \\ &+ \int_0^t \Big[\frac{\partial f}{\partial t}(s, X_s) + U_s\frac{\partial f}{\partial x}(s, X_s) + V_s^2\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s, X_s)\Big]\mathrm{d}s, \quad t \in [0,T]. \end{aligned}$$

Formally, we can write the quadratic variation terms $\langle \mathrm{d}X_t, \mathrm{d}X_t\rangle = \mathrm{d}X_t \cdot \mathrm{d}X_t$, where we calculate according to the following rules, $\mathrm{d}t \cdot \mathrm{d}t = 0$, $\mathrm{d}t \cdot \mathrm{d}W_t = \mathrm{d}W_t \cdot \mathrm{d}t = 0$ and $\mathrm{d}W_t \cdot \mathrm{d}W_t = \mathrm{d}t$ cleanly summarized in the below table

| Product | $\mathrm{d}t$ | $\mathrm{d}W_t$ |
|---|---|---|
| $\mathrm{d}t$ | 0 | 0 |
| $\mathrm{d}W_t$ | 0 | $\mathrm{d}t$ |

*Proof.* We give here a proof using intuitive ideas that are not 100% right but get us to the right place.

A "naive" Taylor series expansion of $f(X)$, completely disregarding the nature of $X$ and treating $\mathrm{d}X$ as a small increment in $X$, yields

$$f(X + \mathrm{d}X) = f(X) + \frac{\mathrm{d}f}{\mathrm{d}X}\mathrm{d}X + \frac{1}{2}\frac{\mathrm{d}^2 f}{\mathrm{d}X^2}\mathrm{d}X^2$$

ignoring higher-order terms.

Moreover, one can argue that $f(X + \mathrm{d}X) - f(X)$ is just the "change in" $f$, thus

$$\mathrm{d}f = \frac{\mathrm{d}f}{\mathrm{d}X}\mathrm{d}X + \frac{1}{2}\frac{\mathrm{d}^2 f}{\mathrm{d}X^2}\mathrm{d}X^2$$

By taking into consideration that

$$\mathrm{d}X_t = U_t\mathrm{d}t + V_t\mathrm{d}W_t$$

and also that $(\mathrm{d}W_t)^2 = \mathrm{d}t$, the desired result (Itô's formula) is obtained.

Although this approach lacks any rigor, it does give the correct result. □

**Example 4.4.4** (Example 4.3.9 again). *Suppose we go back to Example 4.3.9 and try to compute using Itô's formula* $\int W_s \mathrm{d}W_s$.

*Take the function* $f(t, x) := x^2$ *and let us apply Itô's formula to* $Y_t = W_t^2 = f(t, W_t)$ *given by* (4.4.3).

*We first identify how we can go from this little problem to the use of* (4.4.3). *The process* $X$ *there is* $\mathrm{d}X = U\mathrm{d}t + V\mathrm{d}W$ *and* $Y_t = f(t, X_t)$. *Here* $X$ *is simply the Brownian motion.*

*What is the SDE for the Brownian motion? This one is easy, simply take* $U = 0$, $V = 1$ *and* $X_0 = 0$. *We have*

$$dX_t = 0\mathrm{d}t + 1\mathrm{d}Wt, \quad X_0 = 0. \quad \Rightarrow \quad X_t = 0 + \int_0^t 1\mathrm{d}W_s = W_t \quad \checkmark$$

*The function* $f(t, x)$ *is given here,* $f(t, x) = x^2$. *We need only to compute the required derivatives*

$$\partial_t f = 0, \qquad \partial_x f = 2x, \qquad \partial_{xx} f = 2.$$

*Replacing these terms in formula* (4.4.3) *we have*

$$
\begin{aligned}
Y_t = W_t^2 = f(0,x_0) &+ \int_0^t \Big[V_s \frac{\partial f}{\partial x}(s,X_s)\Big]\mathrm{d}W_s \\
&\qquad + \int_0^t \Big[\frac{\partial f}{\partial t}(s,X_s) + U_s \frac{\partial f}{\partial x}(s,X_s) + V_s^2 \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s,X_s)\Big]\mathrm{d}s \\
&= f(0,0) + \int_0^t \Big[1 \times 2W_s\Big]\mathrm{d}W_s + \int_0^t \Big[0 + 0 \times 2W_s + 1^2 \times \frac{1}{2} \times 2\Big]\mathrm{d}s \\
&= 0^2 + 2\int_0^t W_s \mathrm{d}W_s + \int_0^t 1\mathrm{d}s \\
&= 2\int_0^t W_s \mathrm{d}W_s + t \\
&\Leftrightarrow \quad \int_0^t W_s \mathrm{d}W_s = \frac{W_t^2}{2} - \frac{t}{2}.
\end{aligned}
$$

*We found the same result as in Example 4.3.9, this time using the Itô formula.*

From this example we see that the Itô formula for the general case simplifies greatly in the case of the Brownian motion. We state, for hat case, an explicit corollary

**Corollary 4.4.5** (The Itô formula for the Brownian motion)**.** *Let* $(W_t)_{t\in[0,T]}$ *be the standard Brownian motion, let* $f \in C^{1,2}([0,T]\times\mathbb{R},\mathbb{R})$ *and define the new stochastic process* $Y_t := f(t,W_t)$ *for any* $t \in [0,T]$*.*

*Then* $Y$ *is again an Itô process and satisfies the following stochastic differential equation*

$$
\mathrm{d}Y_t = \Big[\frac{\partial f}{\partial t}(t,W_t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,W_t)\Big]\mathrm{d}t + \Big[\frac{\partial f}{\partial x}(t,W_t)\Big]\mathrm{d}W_t, \quad Y_0 = f(0,0) \tag{4.4.4}
$$

*which reads equivalently in integral form for any* $t \in [0,T]$

$$
Y_t = f(0,0) + \int_0^t \Big[\frac{\partial f}{\partial x}(s,W_s)\Big]\mathrm{d}W_s + \int_0^t \Big[\frac{\partial f}{\partial t}(s,W_s) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s,W_s)\Big]\mathrm{d}s.
$$

Here are a couple of examples to test the applications of Itô's formula:

**Example 4.4.6** (Examples from the Lecture notes of "Stochastic Analysis in Finance" (MATH11154) by Prof. Gyongy)**.** *These follow from simple and direct application of the Itô formula. Use it to verify the following inequalities*

1. $de^{X_t} = e^{X t}(a_t + \frac{1}{2}b_t^2)dt + e^{X t}b_t dW_t$*, where the Itô process is* $dX_t = a_t dt + b_t dW_t$
2. $dW_t^p = pW_t^{p-1}dW_t + \frac{p(p-1)}{2}W_t^{p-2}dt$*; for any real number* $p \geq 2$*;*
3. *dsin* $(W_t) = \cos(W_t)dW_t - \frac{1}{2}\sin(W_t)dt$*;*
4. $d(e^{t/2}\sin(W_t)) = e^{t/2}d\sin(W_t) + \sin(W_t)de^{t/2} = e^{t/2}\cos(W_t)dW_t$.

### 4.4.2 The Geometric Brownian Motion

Let us consider now the following stochastic differential equation, where the drift and randomness scale with the value of the random variable $S_t$ for any $t \in [0,T]$

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \sigma S_t \mathrm{d}W_t, \qquad \text{and } S_0 = S_0 \in \mathbb{R},$$

where $\mu$ and $\sigma$ are positive constants.

Note that if $S$ starts out positive it can never go negative, the closer that $S$ gets to zero the smaller the increments $\mathrm{d}S$.

#### Examples with the Geometric Brownian motion

Let $(W_t)_{t\geq 0}$ be a Brownian motion and let $(S_t)_{t\geq 0}$ denote the geometric Brownian motion given by the stochastic process

$$S_t = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right\}, \qquad S_0 \in (0,\infty),\ \sigma \in (0,\infty),\ \mu \in \mathbb{R}.$$

Consider, for any $t \geq 0$ the process $Y_t := \ln S_t$.

Let us use Itô's formula to show that Stochastic differential equation for $X$ is given by

$$\mathrm{d}Y_t = (\mu - \frac{\sigma^2}{2})\mathrm{d}t + \sigma \mathrm{d}W_t \quad \text{and hence} \quad Y_t = Y_0 + \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t,$$

and hence deduce (by taking exponentials) that

$$S_t = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right\}.$$

The SDE for the geometric Brownian motion given by (see definition above)

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \sigma S_t \mathrm{d}W_t, \qquad S_0 = S_0.$$

Since $Y_t := \ln S_t = f\big(t, S_t\big)$ with $f(t,x) := \ln x$.

Computing the derivatives appearing in the Itô formula

$$\partial_t f = 0, \qquad \partial_x f = \frac{1}{x}, \qquad \partial_{xx} f = -\frac{1}{x^2}.$$

To apply Itô's formula we identify $X, U$ and $V$ there with $S, \mu S$ and $\sigma S$ here. We have then by direct substitution with all the terms we computed

$$\begin{aligned}\mathrm{d}Y_t &= \Big[0 + \mu S_t \frac{1}{S_t} + \frac{1}{2}\big(\sigma S_t\big)^2 \times \Big(-\frac{1}{(S_t)^2}\Big)\Big]\mathrm{d}t + \Big[\sigma S_t \frac{1}{S_t}\Big]\mathrm{d}W_t \\ &= (\mu - \frac{1}{2}\sigma^2)\mathrm{d}t + \sigma \mathrm{d}W_t.\end{aligned}$$

Integrating the SDE above from $t = 0$ to $t = t$ yields

$$Y_t = \ln(S_0) + \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t.$$

**Some properties of the gBm**

Using Itô's formula, one can obtain the equation that describes $S(t)$:

$$S(t) = S(0)e^{(\mu-\frac{\sigma^2}{2})t+\sigma W_t}.$$

By taking into account that

$$\begin{aligned}\mathbb{E}[e^{\sigma W_t}] &= \frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\infty} e^{\sigma x}e^{-\frac{x^2}{2t}}dx = \frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\infty} e^{-\frac{(x-\sigma t)^2}{2t}}e^{\frac{\sigma^2}{2}t}dx \\ &= e^{\frac{\sigma^2}{2}t}\frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\infty} e^{-\frac{z^2}{2t}}dz = e^{\frac{\sigma^2}{2}t}\end{aligned}$$

we obtain

$$\begin{aligned}\mathbb{E}[S(t)] &= \mathbb{E}\Big[S_0\exp\big\{(\mu-\frac{\sigma^2}{2})t+\sigma W_t\big\}\Big] \\ &= S_0 e^{(\mu-\frac{\sigma^2}{2})t}\mathbb{E}\big[e^{\sigma W_t}\big] = S_0e^{\mu t}.\end{aligned}$$

Properties of $S(t)$:

$$\begin{cases}\mu < \frac{\sigma^2}{2} & \Leftrightarrow \lim_{t\to\infty} S(t) = 0 \text{ a.s.} \\ \mu = \frac{\sigma^2}{2} & \Leftrightarrow \limsup_{t\to\infty} S(t) = \infty \text{ and } \liminf_{t\to\infty} S(t) = 0 \text{ a.s.} \\ \mu > \frac{\sigma^2}{2} & \Leftrightarrow \lim_{t\to\infty} S(t) = \infty \text{ a.s.}\end{cases}$$

A variable has a lognormal distribution if the natural logarithm of the variable is normally distributed.

Thus, in case where the stock price process, $S(t)$, is assumed to follow a geometric Brownian motion, then $\ln\frac{S(t)}{S(0)}$ (the returns process) is normally distributed with mean $(\mu-\frac{\sigma^2}{2})t$ and variance $\sigma^2 t$. The standard deviation of the stock price is $\sigma\sqrt{t}$ and thus it is proportional to the square root of **how far ahead we are looking**. A random variable that has a lognormal distribution can take any value between zero and infinity.

## 4.5 The Black-Scholes market and the pricing of options

We first introduce the market and its model and then describe what is an investment in this market. With these ingredients we argue just like in Section 3.2.4 to find the price function (a function of time and the value of the stock) of a given European type contract like a Put or a Call option.

### 4.5.1 Modeling the assets and the Black-Scholes assumptions

The evolution of the risk free asset, the bank account, is given for $t \in [0, T]$ by

$$dB_t = rB_t dt, \quad B_0 = 1.$$

This equation's solution is $B_t = e^{rt}$.

The choice for the stock process, has alluded before, is the one taken by Black and Scholes which is givne by the Geometric Brownian motion,

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \;\; t \geq s \;\text{ with }\; S_0 = S$$

for some $\mu \in \mathbb{R}$ (called the *drift*) and $\sigma \in \mathbb{R}$ (called the *volatility*).

Why is the Black-Scholes market model a reasonable model?

1. The risky asset price cannot become negative, recall that $S_t = S_0 \exp\{(\mu - \sigma^2/2)t + \sigma W_t\}$.
2. It captures the phenomenon of noise in the evolution of the risky asset.
3. The stochastic differential equation is easy to solve, as we have seen.

What are the main drawbacks?

1. Empirical data show that the log returns of many assets are not normally distributed.
2. The volatility parameter $\sigma$ is constant but in practice the volatility of risky instruments is observed to be variable and random.

**The Black-Scholes market assumptions**

Just like in Section 1.2, we now overview the assumptions associated with the Black-Scholes market model.

- *The price of the underlying asset follows a geometric Brownian motion*: This assumption implies that stock returns are normally distributed and leads to explicit solutions. However, in reality stock prices can **jump** and that immediately invalidates the assumption under consideration since a geometric Brownian motion has continuous sample paths. Moreover, this assumption is contradictory to the empirical work that has been done previously (known as the problem of "**fat tails**").
- *The risk-free interest rate is a known function of time or a constant*: This restriction is to help us find explicit solutions again. In reality the rate $r$ is not known in advance and is itself stochastic.
- *There are no dividends on the underlying asset*: An explicit solution can be obtained in case that the underlying asset pays cash dividends.
- *There are no transaction costs on the underlying asset or short-sale restrictions*: In reality there are short-sale restrictions and transactions costs and thus delta hedging is sometimes an expensive business. This assumption simplifies the construction of a suitable hedging strategy.
- *Delta hedging is done continuously*: This is an impossibility. Hedging must be done in discrete time. Often the time between rehedges will depend on the level of transaction costs in the market for the underlying asset; the lower the costs, the more frequent the rehedging.

- *There are no arbitrage opportunities*: In reality of course there are arbitrage opportunities and a lot of investors make a lot of money finding them.
- *The volatility $\sigma$ of the underlying asset is a known function of time or a constant*: This is one assumption in Black-Scholes that is clearly not true and can alter the option price significantly. Note also that volatility is the only parameter in the Black-Scholes pricing formulae that cannot be directly observed. Many analysts estimate the volatility parameter by using historical data. An alternative approach involves what is called *implied volatility*.

Implied volatilities can be used to monitor the market's opinion about the volatility of a particular stock. Analysts often calculate implied volatilities from actively traded options on a certain stock and then use them to calculate the price of a less actively traded option on the same stock.

### 4.5.2 SDEs and some models used in finance

(This small section is loosely based on a similar section found in the Lecture notes of "Stochastic Analysis in Finance" (MATH11154) by Prof. Gyongy.])

Before addressing the question of how to price of option within the Black-Scholes model, let us highlight a few other models that are also used. We consider here some examples of Stochastic differential Equations (SDEs), just like the one for the gBM above, that are often used in applications.

**Black-Scholes SDE**: Let $\mu$ and $\sigma$ be some constants. We look for a solution of

$$dX_t = \mu X_t dt + \sigma X_t dW_t$$

with initial value $X_0 = X$, where $X$ is an $\mathcal{F}_0$-adapted random variable.

We have seen above that for any constants $\lambda$ and $\sigma$ the process $S_t := e^{\lambda t + \sigma W_t}$ has the stochastic differential

$$dS_t = S_t(\lambda + \frac{1}{2}\sigma^2)dt + S_t \sigma dW_t$$

Taking here $\lambda := \mu - \frac{1}{2}\sigma^2$ we see that the process $S_t := e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}$ is a solution of the the above SDE with initial value $S_0 = 1$. Hence $X_t := X S_t = X e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}$ is a solution with initial value $X_0 = X$. This process is usually called the the "geometric Brownian motion" (gBM).

**Orstein-Uhlenbeck process:** Let $\alpha > 0$ and $\sigma$ be constants. Consider

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

with initial condition $X_0$.

We can solve this equation by the method of integrating factor as follows

$$d(e^{\alpha t} X_t) = (X_t e^{\alpha} \alpha - e^{\alpha t} \alpha X_t)dt + e^{\alpha t} \sigma dW_t = e^{\alpha t} \sigma dW_t.$$

Hence

$$e^{\alpha t} X_t = X_0 + \sigma \int_0^t e^{\alpha s} dW_s,$$

which gives

$$X_t = e^{-\alpha t}(X_0 + \sigma \int_0^t e^{\alpha s} dW_s) \; .$$

Using Itô's formula we can verify that this process is indeed a solution to the SDE.

**Vasicek's interest rate model:** The instantaneous interest rate $r_t$, in the Vasicek model for interest rate satisfies the SDE

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

where $a, b, \sigma$ and $r_0$, the initial value, are positive constants. One can use Itô's formula to see that the below is the solution to the SDE

$$r_t = a(1 - e^{-bt}) + r_0 e^{-bt} + e^{-bt}\sigma \int_0^t e^{bs} dW_s.$$

A quick analysis show that $r_t$ is a normal random variable with mean

$$\mathbb{E}[r_t] = a(1 - e^{-bt}) + r_0 e^{-bt}$$

and variance

$$\text{Var}(r_t) = e^{-2bt}\sigma^2 \int_0^t e^{2bs} ds = \sigma^2 e^{-2bt}(e^{2bt} - 1)/2b = \sigma^2(1 - e^{-2bt})/2b.$$

Notice that for $t \to \infty$

$$\mathbb{E}[r_t] \to a, \quad \text{Var}(r_t) \to \sigma^2/2b.$$

**Cox-Ingersoll-Ross model** The evolution of the instantaneous interest rate $r_t$ is modeled by the SDE

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

where $a, b, \sigma$ are positive constants. One can show that for every given initial condition $r_0 > 0$ this equation has aunique solution $(r_t)_{t\in[0,T]}$, and it takes only non-negative values.

### 4.5.3 The pricing problem and the main result

Recall from the previous chapters that an *European option* contract gives the buyer the right but not the obligation to buy (or sell) a certain asset at a future time $T$ (the exercise time or maturity) for a price $K$ (the strike price). The option's payoff can be for example $(S_T - K)^+$ for an European call option or $(K - S_T)^+$ for an European put option.

Let us denote by $v(t, s)$ the value of the option at time $t \in [0, T)$ when the price of the risky asset at that time is known to be $s$, i.e. $S_t = s$. The aim of this section is to find a way of determining $v(t, s)$, in other words, the price function of the option contract.

#### The main result

As seen in Chapter 3, in particular Section 3.2.4, the pricing of options can be done in different ways. The first one analysed then was the risk-neutral measure pricing and the same concept follows here. The approach is simple and uses the following rule:

*The fair value of an option is the present value of the expected payoff at maturity under the risk-neutral measure for the price of the underlying asset.* (just like in the previous chapter of with the binomial trees.)

$$\text{option's value at time } t = e^{-r(T-t)}\mathbb{E}[\text{ final payoff }]. \tag{4.5.1}$$

Using the above we essentially obtain the same results as in Theorem 3.5.2.

**Theorem 4.5.1** (Price of Puts and Calls in a Black-Scholes market)**.** *Let $(S_t)_{t\in[0,T]}$ be a geometric Brownian motion with volatility $\sigma \in (0,\infty)$ and drift $r \in \mathbb{R}$. A bank account is available paying interests continuously at rate $r$. Let $t \in [0,T]$. Define $d_+$ and $d_-$ as follows*

$$d_\pm = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r \pm \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}, \qquad \textit{and } d_- = d_+ - \sigma\sqrt{T-t}.$$

*Let the map $N : \mathbb{R} \mapsto [0,1]$ denote the cumulative Normal distribution function.*

*Denote by $C(t,S_t;T,K) := \Pi\left(t; \mathit{Call}(T,K)\right)|_{S_t}$ and $P(t,S_t;T,K) := \Pi\left(t; \mathit{Put}(T,K)\right)|_{S_t}$ the prices of a Call option and a Put option of European type (with maturity $T$ and strike $K$) at a given time $t$ for the stock spot price $S_t$.*

*Then the Black-Scholes formulae for the prices, at time $t$, of a European call and Put are given by*

$$\begin{aligned} C(t,S_t;T,K) &= e^{-r(T-t)}\mathbb{E}\big[(S_T-K)^+\big] \\ &= S_t N\big(d_+(t)\big) - Ke^{-r(T-t)}N\big(d_-(t)\big). \end{aligned}$$

*and*

$$\begin{aligned} P(t,S_t;T,K) &= e^{-r(T-t)}\mathbb{E}\big[(K-S_T)^+\big] \\ &= Ke^{-r(T-t)}N\big(-d_-(t)\big) - S_t N\big(-d_+(t)\big). \end{aligned}$$

*Proof.* This proof is rather straightforward. One "only" needs to compute the expectations. □

One may doubt, by looking at the divisor term $T-t$ in the formulas that, $C(T,S_T;T,K) = (S_T-K)^+$. In order to see that the Black-Scholes formulae have properties that are consistent with options payoff at $t=T$, one needs to look at what $d_+$ and $d_-$ as $T-t \to 0$.

If $S_t > K$, then both tend to $+\infty$ and thus $N(d_-) = N(d_+) = 1$ and $N(-d_-) = N(-d_+) = 0$. Consequently, the price of the call is given by $C(T,S_T;T,K) = S_T - K$ and the price of the put is $P(T,S_T;T,K) = 0$. If $S_t < K$, then both $d_+$ and $d_-$ tend to $-\infty$ and thus $N(d_+) = N(d_-) = 0$ and $N(-d_+) = N(-d_-) = 1$. This means that $C(T,S_T;T,K) = 0$ and and $P(T,S_T;T,K) = K - S_T$.

To provide an interpretation for the terms in the Black-Scholes formula for a European call option, we note that it can be written as follows

$$C(t,S_t;T,K) = e^{-r(T-t)}[S(t)e^{r(T-t)}N(d_1) - KN(d_2)].$$

The expression $N(d_-)$ is the probability that the option will be exercised in a risk-neutral world, so that $KN(d_-)$ is the strike price times the probability that the strike price will be paid.

The expression $S_t e^{r(T-t)} N(d_+)$ is the expected value of a random variable that equals $S_T$ at $t = T$, if $S_T > K$, and is zero otherwise (in the risk-neutral world).

**Example 4.5.2** (The Delta hedge in the Black-Scholes market)**.** *Calculate the* delta hedge *of a European call option on a non-dividend-paying stock using the Black-Scholes formulae by direct computations. Moreover, compute the* delta hedge *a European put option on the same setup using the put-call parity.*

*Using the notation used in the* revisions *section of the Workshop sheets, proceed step-by-step in the following way (Tip: start by writting down the relevant formulae):*

*i) Find $N'(x)$;*

*ii) Show that $Ke^{-r(T-t)}N'(d_-) = S_t N'(d_+)$;*

*(Suggestion: Use that $d_- = d_+ - \sigma\sqrt{T-t}$.)*

*iii) Calculate $\frac{\partial d_+}{\partial S_t}$ and $\frac{\partial d_-}{\partial S_t}$;*

*iv) Using Black-Scholes formula of the price $C(t, S_t; T, K)$ of a European call option on a non-dividend-paying stock (with maturity $T$ and strike $K$), show that*

$$\frac{\partial}{\partial S_t}\big(C(t, S_t; T, K)\big) = N(d_+).$$

*v) Show that the prices given by the Black-Scholes formulae for the European call and put option satisfy the put-call parity;*

*(Suggestion: recall that $N(x) = 1 - N(-x)$.)*

*vi) Using the put call parity, find the value of* delta hedge *of the European put option, i.e. using the put-call parity find an expression for $\frac{\partial}{\partial S_t}\big(P(t, S_t; T, K)\big)$.*

*We now follow the steps proposed above.*

*Recall that Black-Scholes formulae for the prices, at time $t$, of a European call option $C$ on a non-dividend-paying stock and a European put option $P$ on a non-dividend-paying stock are given by*

$$C(t, S_t; T, K) = S_t N(d_+) - Ke^{-r(T-t)}N(d_-)$$

*and*

$$P(t, S_t; T, K) = Ke^{-r(T-t)}N(-d_-) - S_t N(-d_+)$$

*where*

$$d_+ = \frac{\ln(\frac{S_t}{K}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}, \quad d_- = \frac{\ln(\frac{S_t}{K}) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} = d_+ - \sigma\sqrt{T-t},$$

*and $N(x)$ is the cumulative probability distribution function for a random variable that is normally distributed with a mean of $0$ and a variance of $1$.*

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

*ii) Now,*

$$\begin{aligned}
N'(d_-) = N'(d_+ - \sigma\sqrt{T-t}) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(d_+ - \sigma\sqrt{T-t})^2} \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(d_+^2 + \sigma^2(T-t) - 2d_+\sigma\sqrt{T-t})} \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}d_+^2} e^{-\frac{1}{2}\sigma^2(T-t) + \ln(\frac{S_t}{K}) + (r + \frac{\sigma^2}{2})(T-t)} \\
&= N'(d_+) e^{\ln(\frac{S_t}{K})} e^{r(T-t)} = N'(d_+) \frac{S_t}{K} e^{r(T-t)}
\end{aligned}$$

*which gives*

$$Ke^{-r(T-t)} N'(d_-) = S_t N'(d_+)$$

*iii) Also,*

$$\frac{\partial d_+}{\partial S_t} = \frac{\partial d_-}{\partial S_t} = \frac{1}{S_t \sigma \sqrt{T-t}}$$

*iv) Thus*

$$\begin{aligned}
\frac{\partial}{\partial S_t}\Big(C(t, S_t; T, K)\Big) &= N(d_+) + S_t N'(d_+) \frac{\partial d_+}{\partial S_t} - Ke^{-r(T-t)} N'(d_-) \frac{\partial d_-}{\partial S_t} \\
&= N(d_+) + S_t N'(d_+) \frac{\partial d_+}{\partial S_t} - S_t N'(d_+) \frac{\partial d_+}{\partial S_t} = N(d_+).
\end{aligned}$$

*v) We now see that*

$$\begin{aligned}
P(t, S_t; T, K) + S_t &= Ke^{-r(T-t)} N(-d_-) - S_t N(-d_+) + S_t \\
&= Ke^{-r(T-t)}(1 - N(d_-)) - S_t(1 - N(d_+)) + S_t \\
&= Ke^{-r(T-t)} - Ke^{-r(T-t)} N(d_-) + S_t N(d_+) \\
&= Ke^{-r(T-t)} + C(t, S_t; T, K).
\end{aligned}$$

*An alternative method to solve this problem, is to a) write down the Put-call parity formula; then b) replace the formulas for the prices; c) re-organize and d) use the identity $N(x) = 1 - N(-x)$.*

*vi) From the above put-call parity, we have $\frac{\partial}{\partial S_t}\Big(P(t, S_t; T, K)\Big) + 1 = \frac{\partial}{\partial S_t}\Big(C(t, S_t; T, K)\Big)$ which gives*

$$\frac{\partial}{\partial S_t}\Big(P(t, S_t; T, K)\Big) = N(d_+) - 1 = -N(-d_+).$$

## 4.6 Hedging options and the Black-Scholes PDEs

In this section we now try to price a call option (same arguments go for the Put option) using the ideas of block *Pricing through riskless portfolios* in Section 3.2.4. The call option has maturity $T$, strike $K$, is written on the stock $(S_t)_{t\in[0,T]}$ which is modeled by a geometric Brownian motion (drift $\mu$ and volatility $\sigma$) and available is also a Bank account paying interests continuously at rate $r$.

Let us shorten the notation of the last section and let $f(t, s) := C(t, s; T, K)$ denote the price of the call option (maturity $T$, strike $K$) at time $t$ when the stock price if $S_t = s$.

### Portfolio and investment

Let $I_t$ denote the value at time $t$ of an **Investment of one long option position and a short position in some quantity $\Delta$ of the underlying asset** $S$ (compare with Section 3.2.4):

$$I(t) = f(t, S_t) - \Delta S_t,$$

note that the portfolio strategy is described as *self-financing*, i.e. there is no inflow or outflow of money. Moreover, let us recall that the price of the underlying asset follows a geometric Brownian motion:

$$\begin{aligned} \mathrm{d}S_t &= \mu S_t \mathrm{d}t + \sigma S_t \mathrm{d}W_t, \qquad \text{for every } t \in [0, T], \\ \mathrm{d}B_t &= r B_t \mathrm{d}t, \qquad \text{for every } t \in [0, T], \end{aligned}$$

where $T$ denotes the maturity of the option contract.

### The pricing mechanism

The main question is how the value of the investment $I(t)$ changes in time from time $t$ to time $t + dt$. It should be fairly clear that the changed in the investment's value are due partly to the changes in the option value and partly to the changes in the price of the underlying asset. This translates in to

$$\mathrm{d}I(t) = \mathrm{d}f(t, S_t) - \Delta \mathrm{d}S_t.$$

Notice that $\Delta$ is the position in the stock at time $t$ and that does not changes during the infinitesimal time step $\mathrm{d}t$ (check the example for calculating $\Delta$ in discrete-time modeling in Section 3.2.4).

Ito's lemma yields

$$\mathrm{d}f(t, S_t) = \Big[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial s}\mu S_t + \frac{1}{2}\frac{\partial^2 f}{\partial s^2}\sigma^2 S_t^2\Big]\mathrm{d}t + \Big[\frac{\partial f}{\partial s}\sigma S_t\Big]\mathrm{d}W_t$$

Thus, the investment $I$ changes by

$$\begin{aligned} \mathrm{d}I(t) &= \mathrm{d}f(t, S_t) - \Delta \mathrm{d}S_t \\ &= \Big[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial s}\mu S_t + \frac{1}{2}\frac{\partial^2 f}{\partial s^2}\sigma^2 S_t^2\Big]\mathrm{d}t + \Big[\frac{\partial f}{\partial s}\sigma S_t\Big]\mathrm{d}W_t - \Big(\mu \Delta S_t \mathrm{d}t + \sigma \Delta S_t \mathrm{d}W_t\Big) \\ &= \Big[\frac{\partial f}{\partial t} + \Big(\frac{\partial f}{\partial s} - \Delta\Big)\mu S_t + \frac{1}{2}\frac{\partial^2 f}{\partial s^2}\sigma^2 S_t^2\Big]\mathrm{d}t + \Big[\Big(\frac{\partial f}{\partial s} - \Delta\Big)\sigma S_t\Big]\mathrm{d}W_t. \end{aligned}$$

The right hand side of the above equation contains two integral types, the Riemann integral and an Itô integral (one deterministic and one random). The random terms are the risk in our portfolio and in order to eliminate the risk[2] (Delta hedging in continuous time), we can carefully select an appropriate $\Delta$.

In this case, if we choose

$$\Delta = \frac{\partial f}{\partial s}(t, S_t) \tag{4.6.1}$$

[2]Any reduction in randomness is generally termed *hedging*. The perfect elimination of risk, by exploiting correlation between two instruments (in this case an option and its underlying asset) is generally called *delta hedging*.

then the, after some cancellations, the value of the investment $I$ changes by the amount

$$\mathrm{d}I(t) = \left[\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial s^2}\sigma^2 S_t^2\right]\mathrm{d}t$$

and thus we have a perfect elimination of risk. The change $\mathrm{d}I$ is completely *riskless*. Consequently, a completely risk-free change $\mathrm{d}I$ in the investment value $I$ is expected to be the same as the growth we would get if we put the equivalent amount of cash in a risk-free interest-bearing bank account. Thus,

$$\mathrm{d}I(t) = rI(t)\mathrm{d}t.$$

This is an example of the *no-arbitrage* principle and is very very similar to the ideas presented in Section 3.2.4.

Delta hedging is also an example of a *dynamic hedging* strategy. From one time-step to the next, the quantity $\frac{\partial f}{\partial s}$ changes since it is, like $f$, a function of the ever-changing variables $S$ and $t$. This means that the perfect hedge must be continually rebalanced.

### 4.6.1 The Black-Scholes partial differential equation (PDE)

Continuing from before. Writing the equations on the left and right for the riskless portfolio we obtain

$$\mathrm{d}I(t) = rI(t)\mathrm{d}t \quad \Leftrightarrow \quad \left[\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial s^2}\sigma^2 S_t^2\right]\mathrm{d}t = r\left[f(t, S_t) - \frac{\partial f}{\partial s}S_t\right]\mathrm{d}t.$$

The above equation result in the following PDE (let $S_t = x$)

$$\frac{\partial f}{\partial t}(t,x) + rx\frac{\partial f}{\partial s}(t,x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial s^2}(t,x) - rf(t,x) = 0.$$

This is the **Black-Scholes equation**. The equation was first written down in 1969, but a few years passed, with Fischer Black and Myron Scholes justifying the model, before it was published. The derivation of the equation was finally published in 1973. The Black-Scholes equation is a *linear parabolic partial differential equation* and it is related to the heat or diffusion equation of mechanics.

The Black-Scholes equation contains all the obvious variables and parameters such as the price of the underlying asset, time and volatility but there is no mention of the drift rate $\mu$. Any dependence on the drift dropped out at the same time as we eliminated the risk component of the portfolio. The economic argument for this is that since we can perfectly hedge the option with the underlying asset, we should not be rewarded for taking unnecessary risk. Notice that only the risk-free rate of return is in the equation. This means that if the two parties agree on the volatility of an asset, they will agree on the value of its derivatives *even if they have differing estimates of the drift*.

Another way of looking at the hedging argument is to ask what happens if we hold a portfolio consisting of just the stock, in a quantity $\Delta$, and cash. If $\Delta$ is the partial derivative of some option value then such a portfolio will yield an amount at expiry that is simply that option's payoff. In other words, we can use the same Black-Scholes argument to **replicate** an option just by buying and selling the underlying asset. This leads to the idea of a **complete market**. In a complete market an option can be replicated with the underlying asset, thus making options redundant. However, in real life many things conspire to make markets incomplete such as transaction costs.

#### The well-posed problem and boundary conditions

The Black-Scholes equation is derived without taking into consideration what kind of option we are valuing, whether it is for example a call or a put, or what is the strike price and the maturity. These points are dealt with by the *final condition*. We should therefore specify the option value $f(T, S_T)$ as a function of the price of the underlying asset $S_T$ at the expiry date $T$.

For example, the final payoff of a call option is given by

$$f(T, S_T) = \max\{\, S(T) - K; 0\,\}.$$

The Black-Scholes PDE for the price of a call option reads then

$$\frac{\partial f}{\partial t} + rx\frac{\partial f}{\partial s} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial s^2} - rf = 0, \qquad \forall (t,x) \in [0,T) \times (0,\infty) \tag{4.6.2}$$

$$f(T,x) = \max\{\, x - K; 0\,\}, \quad \forall x \in (0,\infty). \tag{4.6.3}$$

This equation has simple solutions for calls, puts and some other contracts. There several ways of obtaining these solutions.

### 4.6.2 Using SDEs to solve PDEs

In the previous section we started from a stochastic model and finished by identifying a partial differential equation (PDE) for the price of the option. This is indeed a product of a deeper result in stochastic analysis. Certain types of PDEs can be solved by using stochastic analysis, the so-called Feynman-Kac formula.

For our particular case, if one would take the Black-Scholes PDE (4.6.2)-(4.6.3) then would would find a formula similar to (4.5.1).

**Theorem 4.6.1** (Feynman-Kac formula)**.** *Assume that the function $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ is a solution to the following boundary value problem PDE*

$$\frac{\partial f}{\partial t}(t,x) + \mu(t,x)\frac{\partial f}{\partial x}(t,x) + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 f}{\partial x^2}(t,x) - rf(t,x) = 0, \tag{4.6.4}$$

$$f(T,x) = \Phi(x). \tag{4.6.5}$$

*For any $(t,x) \in [0,T] \times \mathbb{R}$, define the stochastic process $(X_s)_{s\in[t,T]}$ as the solution to the SDE*

$$\mathrm{d}X_s = \mu(s, X_s)\mathrm{d}s + \sigma(s, X_s)\mathrm{d}W_s, \ \ \forall\, s \in [t,T], \qquad X_t = x. \tag{4.6.6}$$

*Assume that the stochastic process $\big(e^{-rs}\sigma(s,X_s)\frac{\partial f}{\partial x}(s,X_s)\big)_{s\in[t,T]} \in L^2([0,T]\times\mathbb{R})$.*

*Then the solution $f$ of* (4.6.4)-(4.6.5) *can be expressed as*

$$f(t,x) = e^{-r(T-t)}\mathbb{E}_{(t,x)}[\, \Phi(X_T)\,] \qquad \forall\,(t,x) \in [0, T \times \mathbb{R}.$$

This result may not look particularly fancy, but it is link in which Monte-Carlo simulation techniques are used to numerically approximate partial differential equations. This is particularly important in high-dimensional problems.

*Proof.* The proof is rather straightforward and is based on a direct application of Itô's formula (see Theorem 4.4.3).

Define the process $(Y_s)_{s\in[t,T]}$ as $Y_s = e^{-rs}f(s,X_s)$ where $X$ is given by (4.6.6). Applying Itô's formula to $Y$, i.e. computing $\mathrm{d}Y_s$ gives

$$\begin{aligned}
\mathrm{d}Y_s &= \mathrm{d}\Big(e^{-rs}f(s,X_s)\Big)\\
&= (-r)e^{-rs}f\mathrm{d}s + e^{-rs}\partial_t f\mathrm{d}s + e^{-rs}\partial_x f\mathrm{d}X_s + \frac{1}{2}e^{-rs}\partial_{xx}f\big(\mathrm{d}X_s\big)^2\\
&= e^{-rs}\big[-rf + \partial_t f + \mu\partial_x f + \frac{1}{2}\sigma^2\partial_{xx}f\big]\mathrm{d}s + e^{-rs}\big[\sigma\partial_x f\big]\mathrm{d}W_s
\end{aligned}$$

Using the equality given by (4.6.4) (the PDE) we see that the $\mathrm{d}s$ term disappears completely leaving

$$\mathrm{d}Y_s = \mathrm{d}\Big(e^{-rs}f(s,X_s)\Big) = e^{-rs}\big[\ \sigma(s,X_s)\partial_x f(s,X_s)\ \big]\mathrm{d}W_s.$$

Integrating both sides from $s=t$ to $s=T$ gives

$$\begin{aligned}
\big[e^{-rs}f(s,X_s)\big]\Big|_{s=t}^{s=T} &= \int_t^T e^{-ru}\sigma(u,X_u)\partial_x f(u,X_u)\mathrm{d}W_u\\
\Leftrightarrow e^{-rt}f(t,X_t) &= e^{-rT}f(T,X_T) - \int_t^T e^{-ru}\sigma(u,X_u)\partial_x f(u,X_u)\mathrm{d}W_u,\\
\Leftrightarrow f(t,X_t) &= e^{-r(T-t)}f(T,X_T) - \int_t^T e^{-r(u-t)}\sigma(u,X_u)\partial_x f(u,X_u)\mathrm{d}W_u.
\end{aligned}$$

Taking expectations $\mathbb{E}_{(t,x)}[\cdot]$ on both sides (recall that the process $X$ starts at time $t$ in position $x$; this is the meaning of the subscript $(t,x)$ in the expectation sign),

$$f(t,X_t) = e^{-r(T-t)}\mathbb{E}_{(t,x)}\big[f(T,X_T)\big],$$

where the expectation of the stochastic integral disappears due to the properties of the stochastic integral, see property 3 in Theorem 4.3.8 since by assumption we have $\big(e^{-rs}\sigma(s,X_s)\partial_x f(s,X_s)\big)_{s\in[t,T]} \in L^2([0,T]\times\mathbb{R})$. □

## 4.7 Exercises

### Brownian Motion

**Exercise 4.7.1** (Brownian Motion). *Show that if $(W_t)_{t\geq 0}$ is a Wiener process then so is*

*i) the process given by $\bar{W}_t = W_{t+r} - W_r$ for any fixed $r \geq 0$.*

*ii) the process given by $\tilde{W}_t = cW_{c^{-2}t}$ for any fixed $c > 0$.*

*Lastly verify that $\mathbb{E}[W_tW_s] = \min\{t,s\}$.*

*Suggestion: use the independence of increments and that for $X$, $Y$ independent rv's then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.*

**Exercise 4.7.2** (Brownian Motion). *Let $Z$ be a standard normal variate and define $X_t := \sqrt{t}Z$. Is $X := (X_t)_{\{t\geq 0\}}$ a Brownian motion?*

**Exercise 4.7.3** (Brownian Motion). *Let $(W_t)_{t\geq 0}$ denote a standard Brownian motion and define $Y_t := tW_{1/t}$ for all $t > 0$ and $Y_0 := 0$. Then for any $0 < s < t < \infty$, show that*

*(a)* $\mathbb{E}[Y_t - Y_s] = 0$,

*(b)* $\mathbb{E}[Y_t Y_s] = s$,

*(c)* *Var*$[Y_t - Y_s] = t - s$.

**Exercise 4.7.4** (Covariance of the Brownian Motion). *Let $(W_t)_{t\geq 0}$ denote a standard Brownian motion.*

*Show that the covariance Cov$(W_s, W_t) = s$, where $0 < s < t < \infty$.*

**Exercise 4.7.5** (Geometric Brownian motion). *Suppose that the stock price at time $T$ is given by the RV*

$$S_T = S_0 \exp\{(r - \frac{\sigma^2}{2})T + \sigma W_T\}, \qquad \text{where } W_T \sim \mathcal{N}(0, T).$$

*Prove the probability that a European call option will be exercised is $N(d_-)$, in other words prove that*

$$\mathbb{P}[S_T > K] = N(d_-),$$

*where $N$ is the CDF of the $\mathcal{N}(0, T)$ distribution.*

## Itô calculus

**Exercise 4.7.6** (Itô's Isometry). *Use Itô's isometry and the formula $2ab = (a+b)^2 - a^2 - b^2$ to show that*

$$\mathbb{E}\left[\left(\int_0^T X_s \mathrm{d}W_s\right)\left(\int_0^T Y_s \mathrm{d}W_s\right)\right] = \mathbb{E}\left[\int_0^T X_s Y_s ds\right].$$

**Exercise 4.7.7** (Itô Formula). *Take $T \in (0, \infty)$ and $(t, x) \in [0, T] \times \mathbb{R}$ and define the function $f(t, x) := t\cos(x)$.*

*Use Itô's formula to calculate $df(t, W_t)$.*

**Exercise 4.7.8** (Itô Formula). *Use Itô's formula to prove that for any $t \geq 0$*

$$\int_0^t s \mathrm{d}W_s = tW_t - \int_0^t W_s \mathrm{d}s.$$

**Exercise 4.7.9** (Itô Formula). *Use Itô's formula to show that*

$$\int_0^t W_t^2 \mathrm{d}W_t = \frac{1}{3}W_t^3 - \int_0^t W_t \mathrm{d}t$$

*and*

$$\int_0^T W_t^3 \mathrm{d}W_t = \frac{1}{4}W_T^4 - \frac{3}{2}\int_0^T W_t^2 dt.$$

**Exercise 4.7.10** (Itô Formula). *Let $(W_t)_{t\geq 0}$ be a Brownian motion and define for any $t \geq 0$ the stochastic process $X_t = e^{ct + \alpha W_t}$ for some constants $c$ and $\alpha$.*

*Prove that*

$$dX_t = \left(c + \frac{1}{2}\alpha^2\right)X_t dt + \alpha X_t \mathrm{d}W_t.$$

**Exercise 4.7.11** (Solve an SDE). *Solve the stochastic differential equation*

$$dX_t = -X_t dt + e^{-t}\mathrm{d}W_t, \qquad X_0 = x_0,$$

*by proceeding in the following way:*

*i) apply Itô's formula to $f(t, X_t)$ with $f(t,x) := e^t x$;*

*ii) integrate the resulting SDE and obtain an explicit expression for $f(t, X_t)$;*

*iii) from the expression for $f(t, X_t)$ deduce that for $X_t$.*

## Pricing in a Black-Scholes market

**Exercise 4.7.12** (Price of a put option on a Black-Scholes market). *Find the price of a 6-month European put option on a non-dividend paying stock when the stock price is £69, strike price is £70, risk-free rate of interest is 5% per annum and volatility is 35% per annum.*

*Find the price of the call option with the same characteristics.*

*Show that the Put-Call parity holds.*

**Exercise 4.7.13** (Price of a call option in continuous time, Math10003 2014 Exam question). *For a fixed $T > 0$, let $(W_t)_{t\in[0,T]}$ denote the standard Brownian motion (as in the question above). Suppose that the evolution of the stock price $(S_t)_{t\in[0,T]}$ is as follows:*

$$S_t = S_0 \exp\Big\{\big(r - \frac{\sigma^2}{2}\big)t + \sigma W_t\Big\}, \qquad \textit{for any } t \in [0,T],$$

*where $r$ is the risk-free rate of interest, $\sigma$ is the stock's volatility and $S_0 > 0$ is today's stock price.*

*Show that the price of the European call option with pay-off $\Phi(S_T) := \max(S_T - K, 0)$, with maturity $T$, strike price $K$ and written on the non-dividend paying stock $S$ is given by (you only need to prove the last formula on the right)*

$$\Pi\Big(0; \textit{Call}(T,K)\Big) = e^{-rT}\mathbb{E}[\Phi(S_T)] = S_0 N(d_+) - Ke^{-rT}N\big(d_-\big),$$

*where*

$$d_\pm = \frac{\ln\big(\frac{S_0}{K}\big) + \big(r \pm \frac{\sigma^2}{2}\big)T}{\sigma\sqrt{T}},$$

*and $N$ denotes the cdf of the standard Normal distribution.*

***Suggestions:*** *$\triangleright$ Define $\mathbb{1}_{S_T>K}$ as the indicator function of the event $S_T > K$, use it to rewrite $\Phi(S_T)$ inside the expectation and then compute the said expectation;*

*$\triangleright$ Recall that the cdf $N$ satisfies the identity $N(x) = 1 - N(-x)\ \forall x \in \mathbb{R}$;*

**Exercise 4.7.14** (Delta hedge on the Black-Scholes, Past FiM/FMI exam question). *Calculate the* delta hedge *of a European call option on a non-dividend-paying stock using the Black-Scholes formulae by direct computations. Moreover, compute the* delta hedge *a European put option on the same setup using the put-call parity. Using the notation used in the* revisions *section of the Workshop sheets, proceed step-by-step in the following way (Tip: start by writting down the relevant formulae):*

*i) Find $N'(x)$;*

*ii) Show that* $Ke^{-r(T-t)}N'(d_-) = S_t N'(d_+)$*;*

*(Suggestion: Use that* $d_- = d_+ - \sigma\sqrt{T-t}$*.)*

*iii) Calculate* $\frac{\partial d_+}{\partial S_t}$ *and* $\frac{\partial d_-}{\partial S_t}$*;*

*iv) Using Black-Scholes formula of the price* $C(t, S_t; T, K)$ *of a European call option on a non-dividend-paying stock (with maturity* $T$ *and strike* $K$*), show that* $\frac{\partial}{\partial S_t}\big(C(t, S_t; T, K)\big) = N(d_+)$*;*

*v) Show that the prices given by the Black-Scholes formulae for the European call and put option satisfy the put-call parity;*

*(Suggestion: recall that* $N(x) = 1 - N(-x)$*.)*

*vi) Using the put call parity, find the value of* delta hedge *of the European put option, i.e. using the put-call parity find an expression for* $\frac{\partial}{\partial S_t}\big(P(t, S_t; T, K)\big)$.

**Exercise 4.7.15** (Pricing the All of nothing option)**.** *Take* $T \in (0, \infty)$ *and consider a Black-Scholes market for a stock* $(S_t)_{t\in[0,T]}$*. Consider the* all or nothing option *with the payoff at time* $T$ *defined by*

$$g(S) = \begin{cases} P & \text{if} \quad S \geq K, \\ 0 & \text{if} \quad S < K, \end{cases}$$

*where* $P > 0$ *and* $K > 0$ *are given constants.*

*Assume that we work in the Black–Scholes framework and compute the price* $\Pi\big(0; g(S_T)\big)$ *of the contract* $g(S_T)$*, i.e. compute an expression for the right hand side of*

$$\Pi\big(0; g(S_T)\big) = e^{-rT}\mathbb{E}[\, g(S_T)\,].$$

# Bibliography


[BR98] Baxter, Martin and Rennie, Andrew. *Financial Calculus: An Introduction to Derivative Pricing*. Cambridge University Press 1998

[Bjo09] Björk, Tomas. *Arbitrage theory in continuous time*. 3rd Edition, Oxford University Press 2009

[Gar13] Garrett, S. J.. *An introduction to the mathematics of finance*. Pearson Education India, 2006

[Hul06] Hull, John C. *Options, Futures and Other Derivatives*. Elsevier/Butterworth Heinemann, 2013

[Ros08] Ross, Sheldon. *A First Course in Probability*. Pearson, 8th Edition, 2008

[Shr04a] Shreve, Steven E.. *Stochastic calculus for finance. I*. Springer-Verlags 2004

[Shr04b] Shreve, Steven E.. *Stochastic calculus for finance. II*. Springer-Verlag 2004