Mathematical Sciences

UNIVERSITY OF Southampton

Introduction to Financial Mathematics

Lecture Notes for AIMS 2019

James Vickers, Chris Howls & Gerry Kennedy

February 2019

Contents

Chapter 1

Introduction

1.1 Introduction

We shall be concerned with mathematical models for financial markets and the assets traded thereon, including "derivative" products, e.g., options and futures. Many different types of financial markets exist in many different locations, e.g.,

- Stock Markets (e.g., New York, London, Frankfurt, Hong Kong,Tokyo)
- Bond Markets: trade in government and other bonds (loans)
- Currency/ Foreign Exchange Markets: trade in currencies
- Commodity Markets: trade in oil, gold, copper, wheat etc.
- Futures and Options Markets: trade in products "derived" from others (stocks, currencies, connnodities etc)

Quantitites used to be "traded" between people in a "pit" (shouting at each other!) but deals are are now most commonly done electronically. Generically the goal is to maximise the overall profit.

The next few sections give an over-simplified description of **shares** and **share dealing**, but this will be suffice as a background for the mathematical models we shall derive. More precise definitions, descriptions and the legal processes involved, fall outside the remit of this course. However, a glossary of terms is included at the end of the notes.

1.2 A very simplified overview of shares

A **stock** (or share or equity) is used by companies to raise money, normally to finance their start-up costs or growth. It is sold at an initial price by the **board of directors** of a company to **shareholders**, who, then "own" the company. Shareholders may be individuals, corporate bodies, fundholders (for example pension funds) or even governments.

If a profit is made by the company shareholders are normally paid an annual (but variable) lump sum of cash, or **dividend**. If a company is sold, any proceeds are distributed amongst shareholders. If the company goes bust, the value of the share drops to zero and the current shareholders lose their investment. (Share ownership also carries other rights, such as votes for the selection of board chairmen etc, and there are different types of shares in the same company, but that does not concern us here.)

Shares may be bought and sold in **markets** (aka **exchanges**), for which they also have an intrinsic value. This price (determined initially each morning by **market makers** reflects the "perception" by the market about the likely future dividends and general growth of the company (both due to local and global considerations). The prices of shares are quoted at discrete time intervals, at intervals of minutes to fractions of seconds.

Since perceptions as to the value of a share are largely unquantifiable (un)educated guesses, the share price changes in an uncertain way, and it is not possible to predict its exact value at any point in the future. For that reason there is an associated **risk** in owning a share ("past performance is no guarantee of future success") and shareholders lose money if they sell a share for a price below that which they bought it for.

We shall show that it is possible to reduce the risk associated with an individual share (or other asset) by **diversification**, i.e., by possessing a collection of different financial stocks in different companies or institutions. For example, in this way it is hoped that potential losses accumulating due to the falling shares of individual companies may be cancelled out by the rising shares of companies performing better, perhaps 'in other sectors of the economy.

In this way the risk is spread ("not putting all one's eggs in the same basket"). Such a collection held by any investor in the market is called a **portfolio**.

In order to counteract the negative effects of economic recessions (sectoral, national or global slowdowns in business) on shares the portfolio may also include other financial products, such as bonds, cash in the bank, commodities, derivatives, etc.

The buying and selling of shares on an **exchange** is governed by a **contract**. The actual shares are listed on certificates of ownership which may take days or weeks to actually pass between shareholders (or may in fact never be transferred). However as soon as one has made a contract to buy shares, one is free to sell them to a third party, in accordance with rules agreed to by all those dealing on that particular exchange.

In fact it is not actually always necessary to own the shares before one sells them! This is termed **going short** on an asset. (As opposed to owning them, which is called**going long**.) The situation is resolved by allowing for **accounting periods**.

These last for a few weeks, at the end of which all promises to buy and sell shares or assets are made good. As the same shares may have been bought or sold many times by different institutions to each other, this often involves a net transfers of funds/assets between them on the **settlement date**.

More sophisticated contracts called **derivatives** exist, and methods for their fair valuation is the main object of study in the second half of the course. However before we run, we must first learn to walk.

1.3 Risky assets and returns

Let's start right from the beginning: suppose you were given £1000 and told to invest in either (or both) of two companies. Suppose also that you are told (by some supernatural being) that Megabucks Inc. shares were currently £1 each but would be worth £1.50 in one year's time, whilst shares in Bigcorp Ltd. were £2 today and would be worth £2.50 in one year.

What should you do? Well, £1000 invested in Bigcorp Ltd, would be worth £1250 in one year, whilst the same sum invested in Megabucks Inc. would be worth £1500. So anybody but a complete fool would forget completely about Bigcorp Ltd. and put the lot on Megabucks Inc. If this seems obvious, then it is worthwhile reflecting briefly on two matters:

(i) The above calculation assumes that maximising your profit is your only concern; you may have ulterior motives though. (For example, if you were a Director of Bigcorp Ltd. you might want to be seen to backing your Company).

(ii) The above calculation also assumes that you *know* exactly what will happen to the shares of these two Companies in a year's time; unfortunately such knowledge is not usually available. (Although it might be, for example if you bought a bond).

The basis of mean-variance portfolio theory is that each asset in a portfolio is **risky**. We also assume that our **utility function** (which can be thought of as our own personal "definition of happiness") is known, and in this case it simply consists of wishing to maximise our financial return.

Specifically, for the theory that we are about to consider, we introduce the following ideas

- Most assets are **risky**. These include shares, equities, futures, options, etc. etc.
- If an asset has a price $S(t_0)$ at time t_0 and $S(t)$ at some later time t then the **return** $R(t)$ on the asset is defined as

$$R(t) = \frac{S(t) - S(t_0)}{S(t_0)}.$$

  This is often expressed as a percentage.

Of course, the return depends on both the time when the investment was made, and on the time now. Normally if an asset has prices $S_1, S_2, \ldots$ (say at equally spaced time intervals) then we define

$$R_j = \frac{S_{j+1} - S_j}{S_j}$$

to be **the return at the j^{th} timestep**.

In general we only know the current value of a risky asset $S(t_0)$ and **cannot** predict $S(t)$ or $R(t)$ with any certainty. However the basis of all that follows is that we may be able to make statistical statements about these values.

1.4 Riskless assets and interest rates

Some assets are **risk free** (these include Gilts, bank interest rates, etc.). Of course, some "risk free" assets are not *really* riskless. (Anarchy in the UK, Interest rate changes, etc. etc....).

If you deposit money in a bank, they pay you a regular sum, the **interest**, for the privilege of using your money. Similarly if you borrow money from a bank, you have to pay interest to the bank to use the money they lend you (in addition to replaying the original sum). Often the money that a bank lends is the drawn from the money that is loaned to them by depositors. One way that banks make a profit is to charge a higher rate of interest to its debtors than its creditors.

Furthermore interest rates do vary according to (uncertain) economic conditions. However over the lifetime of an option (typically less than nine months) the trend of interest rates is usually known (better in some markets than others, cf Euroland vs UK), and so can be estimated. In what follows investment in banks is assumed to be a riskless way of increasing one's wealth.

Note that the above argument is not restricted to banks, and applies to all riskless assets.

1.4.1 Present values and discounting

An important tool that turns up a lot in what follows is the concept of **present value** or **discounting** of an asset. Effectively this is what one would be prepared to pay **now** to receive a guaranteed amount K at a future time T.In order to do this we first investigate in more detail the theory of interest rates.

Chapter 2

The Theory of Interest Rates

2.1 Introduction: What is interest?

Interest may be thought of as a reward paid by a borrower to a lender for the use of capital over some period of time. Depending on the precise details of the transaction, the payment of interest may not fall due until the time of repayment of the capital; on the other hand, there may be a requirement for *several* interim payments of interest to be made during the lending period.

The level of interest that a lender might expect to receive may depend upon a number of factors, such as the risk of default by the borrower, either through non-payment of interest or as a result of full or partial loss of capital, the availability of alternative lending opportunities of similar risk, the duration of the lending period, and expectations regarding the potential for decrease or increase in the purchasing power of the capital lent, due to inflation or deflation, respectively, over the lending period.

The operation of interest is most easily described within the familiar context of a savings account held in a bank. Suppose that an investor opens such an account with an initial deposit of£100.00, makes no further payments to or from the account, and finally closes the account some time later. If the investor receives a closing account balance of, say,£108.32, then 100.00 of this may be considered as a repayment of the initial capital lent to the bank and the remaining£8.32 may be considered as interest paid by the bank as a reward for the use of this capital over the duration of the account.

2.2 Simple Interest

If an investor makes an initial deposit of C into a savings account that pays simple interest at a rate of i per annum ($p.a.$), then the amount that will be paid to the investor on closure of the account some n years later is, assuming that there are no payments to or from the account in the intervening period,

$$F_n = C(1 + ni) = C + niC.$$

This represents a repayment of capital, C, and a payment of interest, niC. In the case of *simple* interest, interest credited to the account does not itself earn further interest, and so F_n grows linearly with n. This can lead to inconsistencies.

For example, suppose that the investor closes the above account after n_1 years, where $n_1 < n$, and uses the closing balance of $C(1+n_1i)$ to open a second, identical, account. If the investor closes the second account after a further n_2 years, where $n_1 + n_2 = n$, then the final closing balance will be $C(1 + n_1i)(1 + n_2i)$. Since

$$C(1 + n_1i)(1 + n_2i) = C(1 + (n_1 + n_2)i + n_1n_2i^2) = C(1 + ni) + Cn_1n_2i^2 > C(1 + ni),$$

this is greater than the closing balance of $C(1 + ni)$ that would have been received had the original account been kept open for the full n years. It would therefore be to the investors benefit to repeatedly open and close accounts in this way. In order to avoid the administrative expense associated with such frequent opening and closing of accounts, it is common commercial practice to encourage long-term investment by offering *compound* interest on savings accounts. Simple interest is, in practice, generally applied only in relation to very short-term investments.

2.3 Compound Interest

If an investor makes an initial deposit of C into a savings account that pays compound interest at a rate of i p.a., then the amount that will be paid to the investor on closure of the account some n years later is, assuming that there are no payments to or from the account in the intervening period,

$$F_n = C(1 + i)^n = C + C[(1 + i)^n - 1].$$

This represents a repayment of capital, C, and a payment of interest, $C[(1 + i)n - 1]$.

In the case of *compound* interest, interest credited to the account *does* earn further interest, and so F_n grows exponentially with n. This avoids the inconsistencies associated with simple interest.

For example, consider again the case of the investor who closes the account after n_1 years and uses the closing balance to open a second, identical, account, which in turn is closed after a further n_2 years, where $n_1 + n_2 = n$. If compound interest is earned on these accounts then the final closing balance will be

$$C(1 + i)^{n_1}(1 + i)^{n_2} = C(1 + i)^{n_1+n_2} = C(1 + i)^n,$$

and so there is no benefit to be had by repeated opening and closing of such accounts. The qualitatively different behaviours in the growth of an account balance under simple versus compound interest can lead to substantially different accumulated monetary amounts, especially after long time periods and if the interest rate is high, as is illustrated in the following tables and figure.

Period (n)	Simple	Compound
10	180.00	215.89
20	260.00	466.10
30	340.00	1,006.27
40	420.00	2,172.45
50	500.00	4,690.16
60	580.00	10,125.71
70	660.00	21,860.64
80	740.00	47,195.48

Table 2.1: Accumulation of $C = 100.00$ for n years at interest rate $i = 0.08$ p.a. under simple and under compound interest.

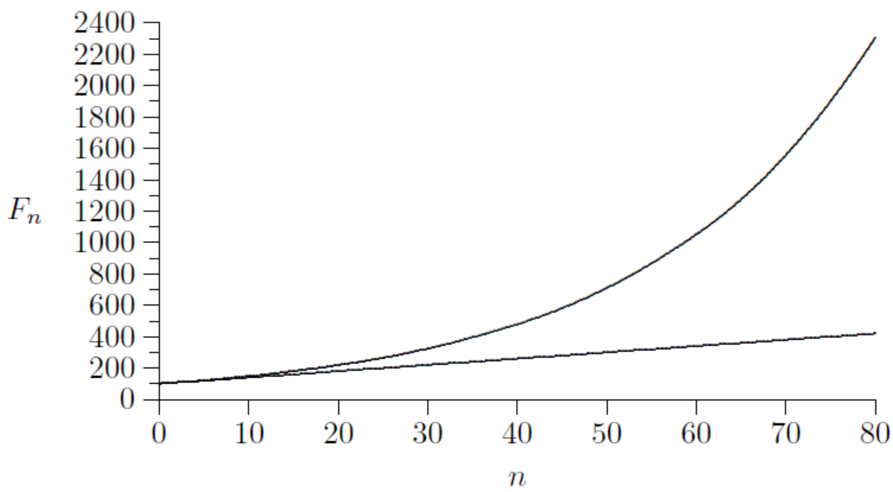

Figure 1.1. Accumulation of $C = 100.00$ for n years at interest rate $i = 0.04$ $p.a.$ under simple and under compound interest

In the next section we will develop the theory of interest rates in some generality, by considering the change in the value of an investment over time. Throughout, it will be assumed that a suitable unit of time (e.g. one year, six months, one month, one day) has been chosen, and that all time variables are quantified in relation to this unit and restricted to an appropriate time domain of relevance.

2.4 Accumulation and Discount Functions

The *accumulation function*, $u(s,t)$, is defined as the value at time t of an investment that has unit value at time s. The *discount function*, $v(s,t)$, is defined as the value at time s of an investment that has unit value at time t. These definitions are applicable for all values of the time arguments, s and t, and, since

$$u(s,t) = v(t,s),$$

they are conceptually equivalent. Indeed, the theory of interest rates could be developed solely in terms of the accumulation function or solely in terms of the discount function, without reference to the other. However, it is customary to employ both functions, $u(s,t)$ and $v(s,t)$, but to restrict the domain to $s \leqslant t$, and this ordering of the time arguments will generally be implicitly understood. For a savings account, we may interpret these functions as follows: $u(s,t)$ represents the closing balance at time t of an account that has unit opening balance at time s; $v(s,t)$ represents the opening balance at time s of an account that has unit closing balance at time t. In each case, we assume that there are no payments to or from the account in the intervening period (s,t). The actions of $u(s,t)$ and of $v(s,t)$ may be represented diagrammatically as follows:

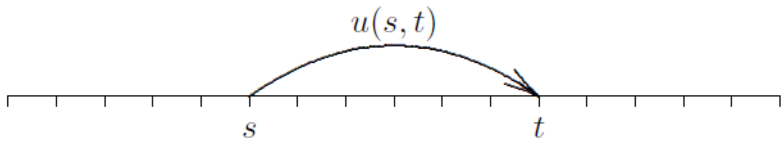

Figure 2.1. Action of $u(s,t)$.

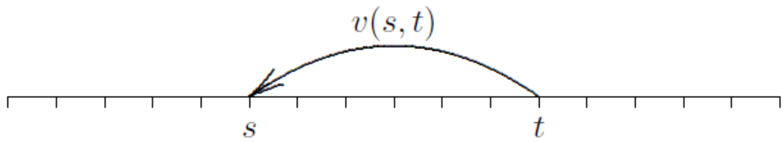

Figure 2.2. Action of $v(s,t)$.

We will speak of accumulating a cash flow from time s to time t under the action of $u(s,t)$, and of discounting a cash flow from time t to time s under the action of $v(s,t)$. The underlying investment market will generally be assumed to satisfy the following principles:

- *Scaleability.* The accumulated value at time t of a cash flow C at time s is $Cu(s,t)$.
- *Rationality.* $u(s,s) = 1$.
- *Consistency.* For $r < s < t$, $u(r,t) = u(r,s)u(s,t)$, which may be represented as follows:

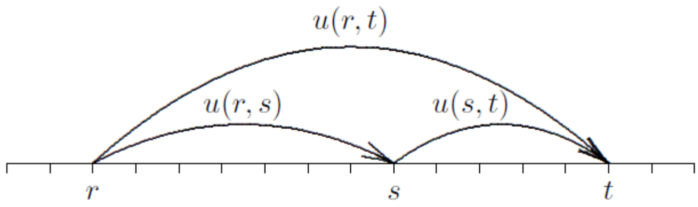

Figure 2.3. Illustration of the consistency principle for the accumulation function.

Note that, when rationality holds, the consistency principle extends to the assertion that, for $r \leqslant s \leqslant t$, $u(r,t) = u(r,s)u(s,t)$.

In practice, these principles may not *strictly* hold true: scaleability may be violated if there are fixed dealing expenses, or if better terms are offered for investments of larger amount; rationality may be broken if there are any dealing expenses, whether fixed or variable; consistency may likewise be inexact if account is taken of dealing expenses, the effects of taxation, or other factors. However, generally, deviations from these principles are small, and assuming them to hold provides us with a tractable theory which serves as a good approximation to reality.

Example 2.1. Determine whether each of the following functional forms for $u(s,t)$ satisfies the principles of rationality and of consistency:

(i) $u(s,t) = st$;

(ii) $u(s,t) = \frac{t}{s}$;

(iii) $u(s,t) = 1 + i(t-s)$, where i is a constant;

(iv) $u(s,t) = (1+i)^{t-s}$, where i is a constant.

Solution.

(i) $u(s,s) = s^2 \neq 1$, so not rational. Also $u(r,t) = rt$, but $u(r,s)u(s,t) = rs^2t \neq rt$, so not consistent.

(ii) $u(s,s) = \frac{s}{s} = 1$, so rational. $u(r,t) = \frac{t}{r}$, and $u(r,s)u(s,t) = \frac{s}{r}\frac{t}{s} = \frac{t}{r}$, so consistent.

(iii) $u(s,s) = 1 + i(s-s) = 1$, so rational. $u(r,t) = 1 + i(t-r)$, but

$$u(r,s)u(s,t) = [1 + i(s-r)][1 + i(t-s)] = 1 + i(t-r) + i^2(s-r)(t-s) \neq 1 + i(t-r),$$

so not consistent. This is the case of simple interest.

(iv) $u(s,s) = (1+i)^0 = 1$, so rational. $u(r,t) = (1+i)^{t-r}$, and

$$u(r,s)u(s,t) = (1+i)^{s-r}(1+i)^{t-s} = (1+i)^{t-r},$$

so consistent. This is the case of compound interest.

Example 2.2. Let time be measured in years, and suppose that, for all times s and t such that $s \leqslant t$,

$$u(s,t) = e^{0.03(t-s)}.$$

(i) Show that the principles of rationality and consistency hold.

(ii) Find the accumulated value after 20 years of an investment of 1,000.00 at any time.

Solution.

(i) $u(s,s) = e^{0.03(s-s)} = e^0 = 1$, so rationality holds.

$$u(r,s)u(s,t) = e^{0.03(s-r)}e^{0.03(t-s)} = e^{0.03(t-r)} = u(r,t),$$

so consistency holds.

(ii) $1000.00u(s,s+20) = 1000.00e^{0.03(s+20-s)} = 1000.00e^{0.6} = 1000.001.8221188 = 1822.19.$

As a consequence of scaleability, it is easily seen that $u(s,t)$ and $v(s,t)$ must satisfy the relationship

$$v(s,t)u(s,t) = 1,$$

or, equivalently,

$$v(s,t) = \frac{1}{u(s,t)}.$$

It immediately follows that the principles of scaleability, rationality, and consistency, although stated initially in terms of $u(s,t)$, have corresponding expressions in terms of $v(s,t)$:

- *Scaleability.* The discounted value at time s of a cash flow C at time t is $Cv(s,t)$.
- *Rationality.* $v(s,s) = 1$.
- *Consistency.* For $r < s < t$, $v(r,t) = v(r,s)v(s,t)$, which may be represented as follows:

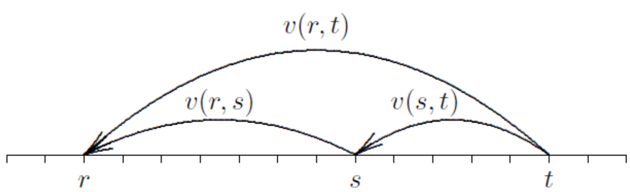

Figure 2.4. Illustration of the consistency principle for the discount function.

Again, when rationality holds, the consistency principle extends to the assertion that, for $r \leqslant s \leqslant t$, $v(r,t) = v(r,s)v(s,t)$. As a consequence of consistency, the s and t-dependences of each of $u(s,t)$ and $v(s,t)$ may be separated, as follows:

$$u(s,t) = \frac{u(r_0,t)}{u(r_0,s)}$$

and

$$v(s,t) = \frac{v(r_0,t)}{v(r_0,s)},$$

where r_0 is some 'reference time' $r_0 < s < t$. Note these expressions are *independent* of the chosen arbitrary reference time r. (Prove this!) By rationality, such separability extends to $r \leqslant s \leqslant t$.

Often, $r_0 = 0$, (i.e. the present), is chosen as the reference time, resulting in the variable-separated expressions

$$u(s,t) = \frac{u(0,t)}{u(0,s)} = \frac{u(t)}{u(s)}$$

where $u(t)$ is defined to be $u(0,t)$. Similarly

$$v(s,t) = \frac{v(0,t)}{v(0,s)} = \frac{v(t)}{v(s)},$$

where $v(t)$ is defined as $v(0,t)$.

Definition: The value at time 0 of a cash flow C at time t is called the *present value* of that cash flow, whether obtained by accumulating from time $t \leqslant 0$ or by discounting from time $t \geqslant 0$.

2.5 Effective and Nominal Interest Rates and Discount Rates

The accumulation may be written as

$$u(s,t) = 1 + i(s,t)$$

This expression may be used to define the *effective rate of interest* for the period (s,t), by

$$i(s,t) := u(s,t) - 1,$$

Similarly we may write

$$v(s,t) = 1 - d(s,t),$$

and use this to define the *effective discount rate* for the period (s,t), according to

$$d(s,t) := 1 - v(s,t).$$

Since $v(s,t) = 1/u(s,t)$ the effective rates of interest and discount for the period (s,t) are not independent of each other, but are related via the constraint

$$[1 - d(s,t)][1 + i(s,t)] = 1,$$

from which follow the equivalent expressions

$$d(s,t) = \frac{i(s,t)}{1 + i(s,t)} = \frac{i(s,t)}{u(s,t)} = v(s,t)i(s,t),$$

which expresses $d(s,t)$ as the discounted value at time s of $i(s,t)$ at time t. Similarly

$$i(s,t) = \frac{d(s,t)}{1 - d(s,t)} = \frac{d(s,t)}{v(s,t)} = u(s,t)d(s,t),$$

which expresses $i(s,t)$ as the accumulated value at time t of $d(s,t)$ at time s. Thus, payment of interest $i(s,t)$ at time t is equivalent to payment of discount $d(s,t)$ at time s.

The accumulation and discount functions may also be written as

$$u(t, t+h) = 1 + hi_h(t)$$

and

$$v(t-h, t) = 1 - hd_h(t),$$

respectively. These expressions serve to define the *nominal rate of interest per unit time* for the period $(t, t+h)$, which is therefore given by

$$i_h(t) := \frac{u(t, t+h) - 1}{h},$$

and the *nominal rate of discount per unit time* for the period $(t-h, t)$,by

$$d_h(t) := \frac{1 - v(t-h, t)}{h}.$$

Note that $hi_h(t) = i(t, t+h)$ is the effective rate of interest for the period $(t, t+h)$, and $hd_h(t) = d(t-h, t)$ is the effective rate of discount for the period $(t-h, t)$. A nominal rate is therefore nothing more than the effective rate over the relevant period expressed as a rate per unit time. Indeed, a nominal rate can be thought of as the simple rate per unit time, which, when applied over the given period, is equivalent to the effective rate over that period. In the particular case that $h = 1$, the nominal and effective rates are identical, and we write

$$i(t) = i_1(t) = i(t, t+1)$$

and

$$d(t) = d_1(t) = d(t-1, t).$$

As with the corresponding effective rates, the nominal rates of interest and of discount per unit time over a given period are not independent. In fact, if the relationship for the effective rates is applied over the interval $(t-h, t)$ we obtain

$$d_h(t) = \frac{i_h(t-h)}{1 + hi_h(t-h)} = \frac{i_h(t-h)}{u(t-h, t)} = v(t-h, t)i_h(t-h),$$

while if this relationship is applied over the interval $(t, t+h)$ we obtain

$$i_h(t) = \frac{d_h(t+h)}{1 - hd_h(t+h)} = \frac{d_h(t+h)}{v}(t, t+h) = u(t, t+h)d_h(t+h).$$

If $i_h(t)$ and $d_h(t)$ are independent of t, we simply write i_h and d_h, respectively. If, further, $h = 1$, and they are independent of t we simply write i and d instead of $i(t)$ and $d(t)$, respectively.

The definitions of effective and nominal rates of interest and discount, and their relationships, are illustrated in the following diagram:

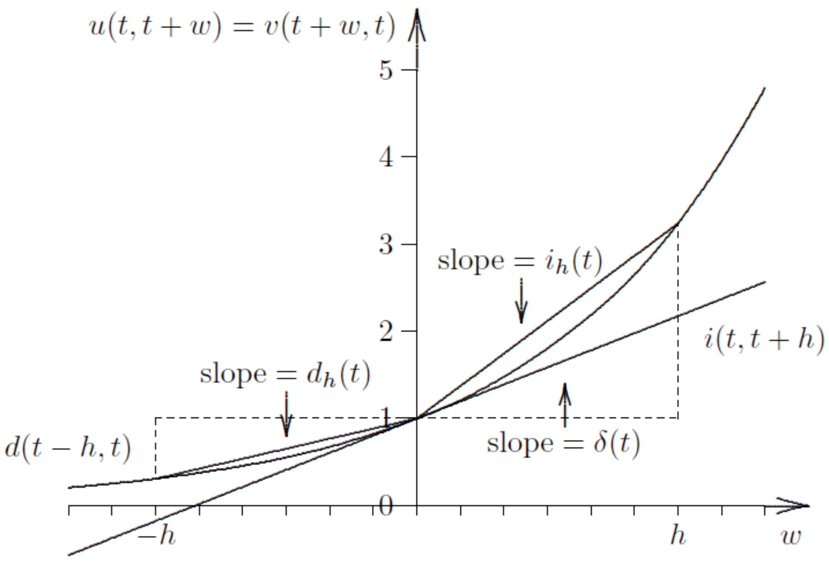

Figure 2.5. Illustration of effective and nominal rates of interest and discount, and force of interest.

Example 2.3. On a particular day, the nominal rates of interest per annum quoted for short-term deposits are as follows: Term: Nominal Rate of Interest (%)
1 day: 4.75
2 days: 4.70
7 days: 4.50
1 month: 4.35
3 months: 4.25

Find the accumulated value of an investment at this time of 10,000.00 for each of these terms.

Solution.
1 day: $10,000.00(1 + \frac{1}{365} \times 0.0475) = 10,001.30$
2 days: $10,000.00(1 + \frac{2}{365} \times 0.0470) = 10,002.58$
7 days: $10,000.00(1 + \frac{7}{365} \times 0.0450) = 10,008.63$

1 month: $10,000.00(1+\frac{1}{12}\times 0.0435) = 10,036.25$
3 months: $10,000.00(1+\frac{1}{4}\times 0.0425) = 10,106.25.$

This example raises a few conundrums. First, if the 10,000.00 is invested for 1 day at the given 1 day rate, and this is repeated for a further 1 day, the accumulated value will be found to be greater than that obtained if the 10,000.00 is invested for 2 days at the 2 day rate. Is this a failure of consistency? What can you learn by comparing the 1 day rate with the 2 day rate? Second, the given nominal rates clearly decrease as the term increases, yet inspection of Figure 2.5 suggests that $i_h(t)$ is an increasing function of h. How can this apparent contradiction be resolved?

Example 2.4. On a particular day, the commercial discount rate on 3 month Treasury bills is quoted as 5.375% p.a. Obtain the price of such bills per 100.00 nominal.

Solution. The information given is that $d_h(t) = 0.05375$ for $h = 0.25$. The discounted value of 100.00 due in 3 months is therefore

$$100.00(1 - 0.25 \times 0.05375) = 98.656.$$

2.6 Force of Interest

By the consistency principle,

$$u(0,t) = u(0,s)u(s,t).$$

on taking logs it follows that

$$\ln u(0,t) = \ln(0,s) + \ln u(s,t),$$

and therefore that

$$\frac{\partial}{\partial t}(\ln u(0,t)) = \frac{\partial}{\partial t}(\ln u(s,t)).$$

Thus,

$$\frac{\partial}{\partial t}(\ln u(s,t)) := \delta(t)$$

is independent of s and is a function solely of t, called the *force of interest per unit time* at time t.

Integrating with respect to time we get:

$$\int_s^t \delta(w)dw = \int_s^t \frac{\partial}{\partial w}(\ln u(s,w))\,dw = [\ln u(s,w)]_{w=s}^{w=t} = \ln u(s,t) - \ln u(s,s) = \ln u(s,t)$$

Taking the exponential gives:

$$u(s,t) = e^{\int_s^t \delta(w)dw}$$

Since $v(s,t) = 1/u(s,t)$, it follows that

$$v(s,t) = e^{-\int_s^t \delta(w)dw}.$$

Thus, both the accumulation and discount functions, each of which is a function of two variables, may be simply expressed in terms of the force of interest function, which is a function of a single variable. From the mathematical point of view, the force of interest function plays a central role in the theory of interest rates, but in financial transactions nominal and effective rates are often used instead.

The earlier expressions for effective and nominal rates of interest and discount may be recast in terms of the force of interest function, as follows:

$$\begin{aligned}
i(s,t) &= u(s,t) - 1 = e^{\int_s^t \delta(w)dw} - 1, \\
d(s,t) &= 1 - v(s,t) = 1 - e^{\int_s^t \delta(w)dw}, \\
i_h(t) &= \frac{u(t,t+h)-1}{h} = \frac{e^{\int_t^{t+h} \delta(w)dw} - 1}{h}, \\
d_h(t) &= \frac{1 - v(t-h,t)}{h} = \frac{1 - e^{-\int_{t-h}^{t} \delta(w)dw}}{h}
\end{aligned}$$

Note that by l'Hôptital's rule

$$\lim_{h\to 0+} i_h(t) = \lim_{h\to 0+}\left(\frac{e^{\int_t^{t+h} \delta(w)dw} - 1}{h}\right) = \lim_{h\to 0+} \delta(t-h) = \delta(t)$$

Thus, the force of interest per unit time at time t may be regarded as an instantaneous nominal rate of interest per unit time at time t, and this is often used as an alternative definition of this function. In the same vein, the force of interest per unit time at time t is also sometimes referred to as the continuously compounded instantaneous rate of interest (or discount) per unit time at time t.

The interest $i(s,t)$ can be shown to be equivalent to a *continuous* payment of interest at rate $\delta(w)$ throughout the interval $s \leqslant w \leqslant t$. This follows from:

$$\begin{aligned}
\int_s^t \delta(w)u(w,t)dw &= \int_s^t \left(\delta(w)e^{\int_w^t \delta(z)dz}\right) dw && (2.1)\\
&= \int_s^t -\frac{d}{dw}\left(e^{\int_w^t \delta(z)dz}\right) dw && (2.2)\\
&= \left[-e^{\int_w^t \delta(z)dz}\right]_{w=s}^{w=t} && (2.3)\\
&= e^{\int_z^t \delta(z)dz} - 1 && (2.4)\\
&= u(s,t) - 1 && (2.5)\\
&= i(s,t) && (2.6)
\end{aligned}$$

Example 2.5. During a certain year, the force of interest per annum earned by bank deposits is given by

$$\delta(t) = \begin{cases} 0.03 + 0.06t & 0 \leqslant t \leqslant \frac{1}{3} \\ 0.05 & \frac{1}{3} \leqslant t \leqslant \frac{2}{3} \\ 0.07 - 0.03t & \frac{2}{3} \leqslant t \leqslant 1 \end{cases}$$

where t denotes the fraction of the year completed. A woman opens an account with a deposit of £50,000 at the start of the year, makes a further deposit of £25,000 halfway through the year, and closes the account at the end of the year, there being no other payments either to or from the account during its operation.

(i) Determine the closing balance of the account.

(ii) Find the effective yield per annum achieved by the woman in the operation of this account. [Note that the yield is the equivalent constant effective interest rate per annum.]

(iii) Obtain the nominal interest rate per annum and the nominal discount rate per annum achieved on deposits of term six months commencing halfway through the year.

Solution. Use one year as the time unit. We require the values of the integrals $u(0, \frac{1}{2}) = \exp\left(\int_0^{\frac{1}{2}} \delta(t)dt\right)$ and $u(\frac{1}{2}, 1) = \exp\left(\int_{\frac{1}{2}}^1 \delta(t)dt\right)$. For the first integral:

$$\int_0^{\frac{1}{2}} \delta(t)dt = \int_0^{\frac{1}{3}} (0.03 + 0.06t)dt + \int_{\frac{1}{3}}^{\frac{1}{2}} 0.05dt \tag{2.7}$$

$$= \left[0.03t + 0.03t^2\right]_0^{\frac{1}{3}} + [0.05t]_{\frac{1}{3}}^{\frac{1}{2}} \tag{2.8}$$

$$= 0.01 + 0.00333333 + 0.025 - 0.01666667 = 0.021666667 \tag{2.9}$$

Hence

$$u(0, \frac{1}{2}) = e^{0.021666667} = 1.021903093.$$

For the second integral:

$$\int_{\frac{1}{2}}^1 \delta(t)dt = \int_{\frac{1}{2}}^{\frac{2}{3}} (0.05)dt + \int_{\frac{2}{3}}^1 (0.07 - 0.03t)dt \tag{2.10}$$

$$= [0.05t]_{\frac{1}{2}}^{\frac{2}{3}} + \left[0.07t - 0.03t^2\right]_{\frac{2}{3}}^1 \tag{2.11}$$

$$= 0.0333333 - 0.025 + 0.07 - 0.015 - (0.04666667 - 0.00666667) = 0.02333333 \tag{2.12}$$

Hence

$$u(\frac{1}{2}, 1) = e^{0.02333333} = 1.023607685.$$

The closing balance of the account is therefore

$$\left[50,000u(0, \frac{1}{2}) + 25,000\right] u(\frac{1}{2}, 1) = [50,000 \times 1.021903093 + 25,000] \times 1.023607685 = 77,891.59$$

(ii) Let i be the effective yield per annum. Then i satisfies the equation

$$50,000(1+i) + 25,000(1+i)^{\frac{1}{2}} = 77,891.59$$

Setting $x = (1+i)^{\frac{1}{2}}$ this gives the quadratic equation

$$50,000x^2 + 25,000x - 77,891.59 = 0$$

which has **positive** root 1.022922543, so that $i = x^2 - 1 = 0.046370528$ or 4.6370528%.

(iii) We have

$$i_{\frac{1}{2}}(\frac{1}{2}) = \frac{u(\frac{1}{2},1)-1}{\frac{1}{2}} = 2(1.023607685 - 1) = 0.04721537$$

or 4.721537%

Also

$$d_{\frac{1}{2}}(1) = \frac{1 - v(\frac{1}{2},1)}{\frac{1}{2}} = 2\left(1 - \frac{1}{1.023607685}\right) = 0.046126431$$

or 4.61264312%

2.7 Annuities and some actuarial notation

An annuity is a *regular* stream of payments. If the payments are of equal amount, the annuity is said to be *level*; otherwise, it is said to be increasing, decreasing, or, more generally, *varying*. If the pattern of the regular stream of payments commences immediately, the annuity is said to be *immediate*; otherwise, it is said to be *deferred*. If the payments are not contingent on survivorship or death, the annuity is said to be an *annuity-certain*; otherwise, it is said to be a *life annuity*. If the stream of payments continues indefinitely, the annuity is referred to as a *perpetuity*. In this subsection we will look at the properties of some standard annuities-certain.

2.7.1 Immediate Level Annuities

Consider a series of n payments, each of amount 1, at times 1, 2, 3, . . . , n, as illustrated in Figure 3.2. Such an annuity is sometimes said to have *payments in arrear*.

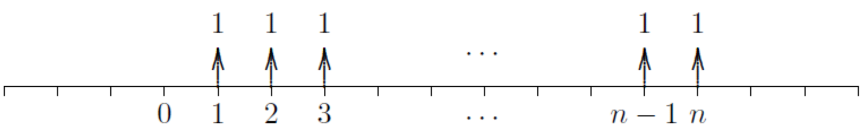

Figure 3.2. An immediate level annuity.

The *present value* of this payment stream is denoted by $a_{\overline{n}|}$. If $i = 0$ then $a_{\overline{n}|} = n$, while if $i \neq 0$ then

$$a_{\overline{n}|} = v + v^2 + v^3 + \cdots + v^n = v\left(\frac{1-v^n}{1-v}\right) = \frac{1-v^n}{i}.$$

On the other hand the *accumulated value* of this payment stream is denoted by $s_{\overline{n}|}$. If $i = 0$ then $s_{\overline{n}|} = n$, while if $i \neq 0$ then

$$s_{\overline{n}|} = 1 + u + u^2 + u^3 + \cdots + u^{n-1} = \frac{1 - u^n}{1 - u} = \frac{u^n - 1}{i}$$

Clearly,

$$s_{\overline{n}|} = u^n a_{\overline{n}|}$$

so that $s_{\overline{n}|}$ is the present value accumulated forward an amount n units of time.

Next, consider a series of n payments, each of amount 1, at times 0, 1, 2, . . . , $n - 1$, as illustrated in Figure 3.3. Such an annuity is sometimes said to have payments in advance, and is referred to as an *annuity-due*.

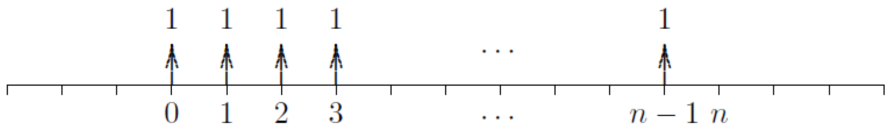

Figure 3.3. An immediate level annuity-due.

The *present value* of this payment stream is denoted by $\ddot{a}_{\overline{n}|}$. If $i = 0$ then $\ddot{a}_{\overline{n}|} = n$, while if $i \neq 0$ then

$$\ddot{a}_{\overline{n}|} = 1 + v + v^2 + v^3 + \ + v^{n-1} = \frac{1 - v^n}{1 - v} = \frac{1 - v^n}{d},$$

The *accumulated value* of this payment stream at time n is denoted by $\ddot{s}_{\overline{n}|}$. If $i = 0$ then $\ddot{s}_{\overline{n}|} = n$, while if $i \neq 0$ then

$$\ddot{s}_{\overline{n}|} = u + u^2 + u^3 + \ + u^n = u\left(\frac{1 - u^n}{1 - u}\right) = \frac{u^n - 1}{d}$$

The present and accumulated values are related by

$$\ddot{s}_{\overline{n}|} = u^n \ddot{a}_{\overline{n}|}, \qquad \ddot{a}_{\overline{n}|} = v^n \ddot{s}_{\overline{n}|}.$$

Finally consider an annuity payable *continuously* at the rate of 1 per time unit over the interval $(0, n)$.

The *present value* of this payment stream is denoted by $\bar{a}_{\overline{n}|}$. If $i = 0$ then $\bar{a}_{\overline{n}|} = n$, while if $i \neq 0$ then

$$\bar{a}_{\overline{n}|} = \int_0^n v^t dt = \int_0^n e^{-\delta t} dt = \left[\frac{e^{-\delta t}}{-\delta}\right]_0^n = \frac{1 - v^n}{\delta}.$$

Similarly the *accumulated value* of this payment stream is denoted by $\bar{s}_{\overline{n}|}$. If $i = 0$ then $\bar{s}_{\overline{n}|} = n$, while if $i \neq 0$ then

$$\bar{s}_{\overline{n}|} = \int_0^n u^t dt = \int_0^n e^{\delta t} dt = \left[\frac{e^{\delta t}}{\delta}\right]_0^n = \frac{u^n - 1}{\delta}.$$

Clearly

$$\bar{s}_{\overline{n}|} = u^n \bar{a}_{\overline{n}|}, \quad \text{and} \quad \bar{a}_{\overline{n}|} = v^n \bar{s}_{\overline{n}|}$$

There are certain relationships between these various annuity functions. For instance,

$$\begin{aligned} i a_{\overline{n}|} &= d \ddot{a}_{\overline{n}|} = \delta \bar{a}_{\overline{n}|}, \\ i s_{\overline{n}|} &= d \ddot{s}_{\overline{n}|} = \delta \bar{s}_{\overline{n}|}. \end{aligned}$$

The functions $a_{\overline{n}|}$, $\ddot{a}_{\overline{n}|}$, $\bar{a}_{\overline{n}|}$, $s_{\overline{n}|}$, $\ddot{s}_{\overline{n}|}$, and $\bar{s}_{\overline{n}|}$ are all increasing functions of n. If n is infinite then the annuity is called a perpetuity, and, if $i > 0$, has present value

$$\begin{aligned} a_{\overline{\infty}|} &= \frac{1}{i}, \\ \ddot{a}_{\overline{\infty}|} &= \frac{1}{d}, \\ \bar{a}_{\overline{\infty}|} &= \frac{1}{\delta}, \end{aligned}$$

as appropriate.

The functions $a_{\overline{n}|}$, $\ddot{a}_{\overline{n}|}$, and $\bar{a}_{\overline{n}|}$ are decreasing functions of i, whereas the functions $s_{\overline{n}|}$, $\ddot{s}_{\overline{n}|}$, and $\bar{s}_{\overline{n}|}$ are increasing functions of i.

Example 3.3. A mortgage loan of 150,000 is to be repaid over 20 years by level monthly payments in arrear, calculated on the basis of an effective interest rate of 0.5% per month. Determine the amount of each monthly payment.

Solution. Let one month be the time unit, and let X be the amount of each monthly payment. At an effective interest rate of $i = 0.005$ per time unit, X satisfies the condition that the present value of the interest payments is the same as the value of the loan. This gives:

$$150,000 = X a_{\overline{240}|} = X \frac{1 - v^{240}}{i} = X \frac{1 - (1.005)^{-240}}{0.005} = 139.58077X,$$

whence

$$X = \frac{150,000}{139.58077} = 1074.6466 = 1074.65 \quad \text{(to 2 decimal places)}.$$

2.8 Nominal Interest and Discount Rates

In the case of **constant** force of interest, δ, the nominal interest rate per unit time and the nominal discount rate per unit time for any time interval of length h time units are given by

$$i_h = \frac{e^{\delta h} - 1}{h} = \frac{(1+i)^h - 1}{h}$$

and

$$d_h = \frac{1 - e^{-\delta h}}{h} = \frac{1 - (1-d)^h}{h},$$

respectively, where we have used the relationships $e^{\delta} = 1 + i$ and $e^{-\delta} = 1 - d$.

If $h = \frac{1}{p}$, where p is a positive integer, so that h is a simple fraction of a time unit, then it is customary to write $i^{(p)}$ instead of $i_{\frac{1}{p}}$, and $d^{(p)}$ instead of $d_{\frac{1}{p}}$, and to refer to $i^{(p)}$ and $d^{(p)}$ as, respectively, the nominal rate of interest per unit time and the nominal rate of discount per unit time, convertible p-thly (or payable p-thly).Thus,

$$i^{(p)} = p\left[(1+i)^{\frac{1}{p}} - 1\right]$$

and

$$d^{(p)} = p\left[1 - (1-d)^{\frac{1}{p}}\right].$$

These equations may be rewritten as, respectively,

$$1 + i = \left(1 + \frac{i^{(p)}}{p}\right)^p,$$

which simply states that a single accumulation over a time interval of unit length at an effective interest rate of i per time unit is equivalent to p successive accumulations over adjacent subintervals each of length $\frac{1}{p}$ at an effective interest rate of $\frac{i^{(p)}}{p}$ per subinterval, and

$$1 - d = \left(1 - \frac{d^{(p)}}{p}\right)^p,$$

which simply states that a single discounting over a time interval of unit length at an effective discount rate of d per time unit is equivalent to p successive discountings over adjacent subintervals each of length $\frac{1}{p}$ at an effective discount rate of $\frac{d^{(p)}}{p}$ per subinterval.

Both of these equivalences follow, of course, from the consistency property, which implies

$$u(0,1) = u\left(0, \frac{1}{p}\right) u\left(\frac{1}{p}, \frac{2}{p}\right) u\left(\frac{2}{p}, \frac{3}{p}\right) \cdots u\left(\frac{p-1}{p}, 1\right)$$

and

$$v(0,1) = v\left(0, \frac{1}{p}\right) v\left(\frac{1}{p}, \frac{2}{p}\right) v\left(\frac{2}{p}, \frac{3}{p}\right) \cdots v\left(\frac{p-1}{p}, 1\right).$$

The total accumulated value at time 1 of the p interest instalments, each of $\frac{i^{(p)}}{p}$ at times $\frac{1}{p}, \frac{2}{p}, \frac{3}{p}, \ldots \frac{p-1}{p}, 1$ must therefore equal i, and so $i^{(p)}$ may be interpreted as the *total* amount of interest, payable in equal instalments at the end of each such subinterval of length $\frac{1}{p}$, that is equivalent to a single payment of interest of amount i at the end of a unit time interval.

Likewise, the total discounted value at time 0 of the p discount instalments, each of $\frac{d^{(p)}}{p}$ at times $0, \frac{1}{p}, \frac{2}{p}, \frac{3}{p}, \ldots \frac{p-1}{p}$, must therefore equal d, and so $d^{(p)}$ may be interpreted as that total amount of discount, payable in equal instalments at the start of each subinterval of length $\frac{1}{p}$, that is equivalent to a single payment of discount of amount d at the start of a unit time interval.

The nominal rates of interest and of discount are, of course, related via the constraint

$$\left[1-\frac{d^{(p)}}{p}\right]\left[1+\frac{i^{(p)}}{p}\right]=1$$

From which follows:

$$d^{(p)}=\frac{i^{(p)}}{1+\frac{i^{(p)}}{p}}$$

which expresses $d^{(p)}$ as the discounted value of $i^{(p)}$, and

$$i^{(p)}=\frac{d^{(p)}}{1-\frac{d^{(p)}}{p}}$$

which expresses $i^{(p)}$ as the accumulated value of $d^{(p)}$.

Example 4.1. If $i^{(12)}=0.06$ determine i, δ, d, and $d^{(12)}$.

Solution. We have

$$1+i=\left(1+\frac{i^{(12)}}{12}\right)^{12}=(1.005)^{12}=1.0616778,$$

so

$$i=(1+i)-1=1.0616778-1=0.0616778,$$

$$\delta=\ln(1+i)=\ln(1.0616778)=0.0598505,$$

$$d=\frac{i}{1+i}=\frac{0.0616778}{1.0616778}=0.0580946,$$

and

$$d^{(12)}=\frac{i^{(12)}}{1+\frac{i^{(12)}}{12}}=\frac{0.06}{1.005}=0.0597015.$$

Example 4.2. If $\delta=0.04$, find the values of

i, $i^{(4)}$, $i^{(365)}$, d, and $d^{(12)}$ and $d^{(52)}$.

Solution. We first find i using

$$1+i=e^{\delta}$$

The value of d is then given by

$$1-d=e^{-\delta}=\frac{1}{1+i}$$

The remaining values may be found using the formulae

$$i^{(p)}=p\left[(1+i)^{\frac{1}{p}}-1\right]$$

and

$$d^{(p)}=p\left[1-(1-d)^{\frac{1}{p}}\right].$$

Chapter 3

Investment Appraisal and Investment Performance

In this chapter we will apply *discounted cash flow techniques* to the valuation and appraisal of investment or business projects, and to the measurement of investment performance.

3.1 Net Present Value

Let us consider a general investment or business project, involving a series of discrete net cash flows, $\{c_{t_j}\}$ at times $\{t_j, j = 1, 2, ..., n\}$, and a continuous net cash flow at rate $\rho(t)$ throughout the relevant t-domain, $[0, T]$, where T may be infinite. The present value of this stream of net cash flows at a constant effective rate of interest of i per time unit is called the *net present value* of the investment or business project at rate i and is expressible as

$$NPV(i) = \sum_{j=1}^{n} c_{t_j}(1+i)^{-t_j} + \int_0^T \rho(t)(1+i)^{-t} dt.$$

Note the net present value is calculating the value at $t = 0$ of cash flows that take place in the future which is the reason for the negative powers of $(1 + i)$.

Note that if instead of using i we were to use the *force of interest* δ then the expression for the net present value could be written

$$NPV = \sum_{j=1}^{n} c_{t_j} e^{-t_j \delta} + \int_0^T \rho(t) e^{-t\delta} dt.$$

If instead we want to find the vale at some future time $T < \infty$ we can calculate the *net accumulated value* at time T of this stream of net cash flows at a constant effective rate of i per unit time. This is expressible as

$$NAV(i) = \sum_{j=1}^{n} c_{t_j}(1+i)^{T-t_j} + \int_0^T \rho(t)(1+i)^{T-t} dt$$

Note the $NAV(i) = (1+i)^T NPV(i)$ which expresses the fact that the net accumulated value is nothing but the net present value accumulated T years into the future. Equivalently the net present value is nothing but the net accumulated value at time T discounted T years to the present.

If an investor may borrow or lend money at rate i then the net accumulated value may be regarded as the *net profit* at time T arising from the investment or business project: if the net cash flows associated with the project are accumulated in an account in which interest is credited or debited at rate i, according to whether the balance in the account is positive or negative, respectively, then the balance at time T will be $NAV(i)$. In such a case, the investment or business project will be profitable if and only if

$$NAV(i) > 0,$$

or, equivalently, if and only if

$$NPV(i) > 0.$$

Definition: The *yield* (or internal rate of return) per unit time for the investment or business project is the unique solution $i^* > -1$ of the equation $NPV(i) = 0$. i.e. the yield is the "break-even" value of i for which the project neither makes nor loses money. We call the equation $NPV(i) = 0$ the *equation of value* for the project. Typically $NPV(i)$ is a monotonically decreasing function of i^* (see figure)

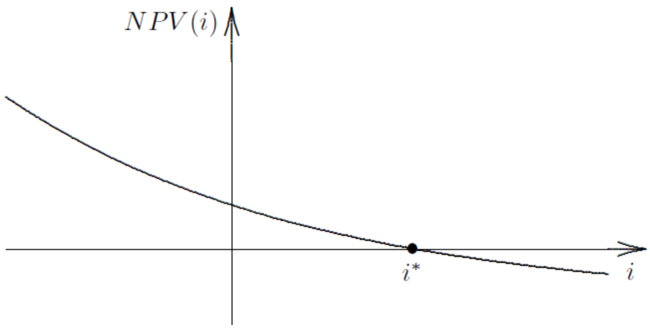

Figure 5.1. Criterion for profitability: $NPV(i) > 0$ or, equivalently, $i < i^*$.

So that a project is profitable if the common lending and borrowing rate i is less than the yield.

Example 5.1. An investor is considering an investment project which requires an immediate payment of 15,000 and a further payment of 5,000 three years from now, in return for receipt of 25,000 in seven years time. The investor may borrow or lend money at an effective rate of 3% per annum. Determine whether or not this investment project is profitable, and find the profit or loss when the project ends in seven years time.

Solution. Choose one year as the unit of time. Then since

$$NPV(i) = -15,000 - 5,000(1+i)^{-3} + 25,000(1+i)^{-7},$$

we easily find $NPV(0.03) = 751.58$, and therefore the net profit when the project ends in seven years time is

$$NAV(0.03) = (1.03)^7 NPV(0.03) = 924.35.$$

3.2 Comparison of Investment Projects

Consider two investment or business projects, A and B. The net present value functions for these two projects will be denoted by $NPV_A(i)$ and $NPV_B(i)$, respectively, and the corresponding yields will be denoted by i_A^* and i_B^* respectively.

An investor, who may borrow or lend money at a constant effective interest rate of i per unit time, wishes to determine which of the two projects is the more favourable to select for investment. It might be thought that such an investor should always preferentially select the project with the higher yield, and, in fact, this course of action should be followed if the investor has sufficient money to fund either project without need for borrowing. In such a case, the yield on either project may be interpreted as that rate of interest that would be required to be earned on a comparable deposit account into which the investor pays amounts corresponding to the negative net cash flows at the corresponding times, and from which the investor receives amounts corresponding to the positive net cash flows at the corresponding times, the final such inflow to or outflow from the account resulting in a zero balance. Thus, if $i_A^* > i_B^*$ such an investor should choose project A in preference to project B.

If, however, the investor has no such funds, but must use the available facility for borrowing and lending at rate i, then an appropriate criterion would be to preferentially select the project with the higher net present value at this rate. Thus, if $NPV_A(i) > NPV_B(i)$, such an investor should choose project A in preference to project B.

It is important to note that these two criteria are applicable to two different investor circumstances, and are not interchangeable. For instance, it may be the case that $i_A^* > i_B^*$, but that, at the given borrowing and lending rate, i, $NPV_A(i) < NPV_B(i)$, because the graphs of the two net present value functions cross, as illustrated in the figure below.

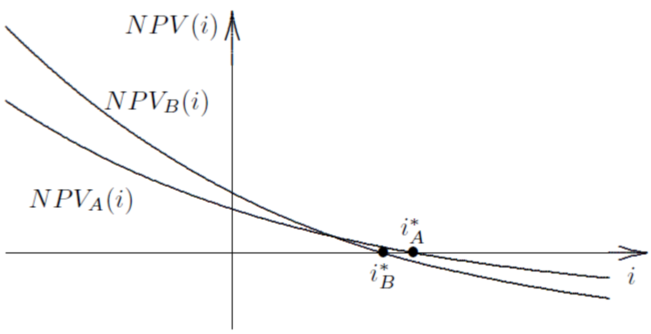

Figure 5.2. Crossover of $NPV(i)$ functions: $i_A^* > i_B^*$ but $NPV_A(i) < NPV_B(i)$.

Example 5.2. An investor, who has no available funds, but may borrow or lend money at an effective interest rate of 3.75% per annum, is considering whether to invest in either or both of the following projects:

A For a purchase price of 100,000, the investor will receive 7,550 per annum payable monthly in arrear for 20 years;

B For a purchase price of 90,000, the investor will receive 4,000 per annum payable annually in arrear for 15 years, and a return of the purchase price at the end of this period.

Would you advise the investor to invest in either project, and, if so, which would be the more profitable?

Solution. For project A,

$$NPV_A(0.0375) = -100,000 + 7,550\left[\frac{1-(1.0375)^{-20}}{12\left((1.0375)^{\frac{1}{12}}-1\right)}\right] = 6,707.61.$$

For project B,

$$NPV_B(0.0375) = -90,000 + 4,000\left[\frac{1-(1.0375)^{-15}}{1.075-1}\right] + 90,000(1.0375)^{-15} = 7,072.06.$$

Thus, at the given borrowing and lending rate, $i = 0.0375$, the advice would be that both projects are profitable, but that project B is the more profitable.

Note that, in this example, the yield on project A is the solution of the following equation of value at time 0:

$$NPV_A(i) = -100,000 + 7,550\left[\frac{1-(1+i)^{-20}}{12\left((1+i)^{\frac{1}{12}}-1\right)}\right] = 0$$

Since $NPV_A(0.045) = 219.40$ and $NPV_A(0.05) = -3,733.08$ linear interpolation gives

$$i_A^* \approx 0.045 + \left(\frac{219.40-0}{219.40-(-3,733.08)}\right)(0.05-0.045) = 0.04527468$$

Similarly the yield on project B is given by the equation of value

$$NPV_B(i) = -90,000 + 4,000\left[\frac{1-(1+i)^{-15}}{1+i-1}\right] = 0$$

Solving this gives $i_b^* \approx 0.44444444$.

Thus, in this example, the investor should preferentially select project B, even though the yield on project A is higher than that on project B.

3.3 Different Rates for Borrowing and Lending

In practice, it is improbable that an investor considering the financing of an investment or business project will be able to borrow or lend money at the same rate of interest per unit time, i. Rather, it is likely that the interest rate per unit time at which such an investor may borrow money, i_b, will be higher than that at which money may be lent, i_ℓ, that is, $i_b > i_\ell$, this differential reflecting, among other factors, market margins and the credit-worthiness of the investor.

In such circumstances, the concepts of net present value and yield are no longer generally applicable, and it is necessary to proceed from first principles, the net cash flows associated with an investment or business project being accumulated at a rate of interest that varies according to whether or not the accumulated balance is positive or negative

Example 5.5. An electricity generation company is considering an investment project involving the construction, operation, and eventual decommissioning of a power station. Construction is expected to take 3 years, and to cost £450 million, payable in equal instalments at the start of each year. Once constructed, the power station is expected to enjoy 25 years of operational life, to produce an annual output of 3 billion kWh of energy, and to incur an annual operating cost of £150 million, both of these annual figures being spread uniformly throughout the year. At the end of its operational life, it will be necessary to decommission the power station, a process that is expected to take 2 years and to cost £150 million, payable in equal instalments at the start of each year. The company has no spare funds to finance the project, but may borrow the construction costs from its bankers, who charge an effective annual rate of interest of 6% on borrowings and pay an effective annual rate of interest of 41 2% on deposits.

(i) Determine the minimum price (in pence per kWh) at which electricity must be sold if this project is to just break even, and find the length of time that must elapse before the company will have repaid its bank indebtedness at this minimum price, assuming that borrowings may be reduced by repayment at any time.

(ii) Identify some of the factors that the company should consider before embarking on this project.

Solution.

(i) Let P be the minimum price (in £ per kWh) at which electricity must be sold if the project is just to break even. Let $t = 0$ be the time at which the power station starts producing electricity and let T be the time at which the initial bank load has been repaid. Then accumulating the cost of building the power station forward to $t = 0$ must match the income produced up until time t discounted back to $t = 0$ (we need to compare expenditure and income at the **same** time). Thus P and T must satisfy the equation:

$$150(u_b + u_b^2 + u_b^3) = (3000P - 150) \int_{t=0}^{T} u_b^{-t} dt \qquad (*)$$

where $u_b = (1 + i_b) = 1.06$.

Once the debt has been repaid the power station starts making an operating profit but this must be used to pay for the cost of decommissioning. We continuously accumulate this profit forward to $t = 25$ at a rate of 0.045 and compare this with the cost of decommissioning discounted back to $t = 25$. This gives the equation

$$75(1 + u_\ell^{-1}) = (3000P - 150) \int_{t=T}^{25} u_\ell^t dt \qquad (**)$$

where $u_\ell = (1 + i_\ell) = 1.045$.

We know the LHS of both equation and eliminating P by dividing one equation by the other gives:

$$3.448874054 = \frac{\int_0^T (1.06)^{-t} dt}{\int_T^{25} (1.045)^t dt}$$

Solving this equation for T numerically gives $T = 21.6894$ years.Inserting this value of T into (*) and solving for P gives $P = 0.063704$ or 6.37p per kWh.

(ii) Factors that should be considered might be as follows: inflation of construction cost estimates, or delays in construction; inflation of annual operating cost estimate; the effect of competition, market capacity, and market demand on the price at which electricity may be sold; inflation or potential underestimation of decommissioning cost; length of operating life, particularly at full generation output; and environmental factors, if relevant.

Chapter 4

Measures of risk

4.1 Introduction

Up until now we have looked at situations where the future behaviour of an investment is known for certain and modelled using the accumulation function $u(s,t)$ or the force of interest δ. However in reality financial outcomes are *uncertain* and the best one can do is to use a statistical or stochastic model to predict future behaviour.

In the next chapter we will examine the use of a *portfolio* of shares to reduce the risk of an investment. Most mathematical investment theories of investment use *variance of return* as the measure of risk. However, it is not obvious that variance necessarily corresponds to investors' perception of risk and other measures have been proposed as being more appropriate. We start by looking at the variance of return and then briefly review other measures of investment risk.

4.2 Variance of return

Variance of return is defined as

$$\int_{-\infty}^{+\infty} (\mu - x)^2 f(x) dx \tag{4.1}$$

where μ is the mean return at the end of the chosen period and $f(x)$ is the probability density function of the return.

Variance has the advantage over most other measures in that it is mathematically tractable and the mean-variance framework discussed below leads to elegant solutions for optimal portfolios selection. This ease of use should not be lightly disregarded and to justify using a more complicated measure it would have to be shown that

- it was both more theoretically correct, and

- that it would lead to significantly different choices of optimal portfolios than the use of variance.

In fact the use of mean-variance theory has been shown to give a good approximation to several other proposed methodologies. Mean-variance portfolio theory can be shown to lead to optimum portfolios if returns can be assumed to be normally distributed.

4.3 Semi-variance of return

The main argument against the use of variance as a measure of risk is that most investors do not dislike uncertainty of returns as such; rather they dislike the possibility of low returns. One measure that seeks to quantify this view is downside semi-variance. This is defined as

$$\int_{-\infty}^{\mu} (\mu - x)^2 f(x)dx \tag{4.2}$$

Semi-variance is not so easy to handle mathematically and it takes no account of variability above the mean.

Exercise for student: if returns on assets are symmetrically distributed about the mean show that semi-variance is proportional to variance.

4.4 Shortfall probabilities

A shortfall probability measures the probability of returns falling below a certain level. In its simplest form it can be expressed as a probability:

$$\int_{-\infty}^{L} f(x)dx \tag{4.3}$$

where L is the benchmark level chosen by the user. Alternatively the risk measure can be expressed as the expected shortfall below a certain level:

$$E[\text{Max}(L - x, 0)] = \int_{-\infty}^{L} (L - x)f(x)dx \tag{4.4}$$

where L is again the benchmark level chosen by the user. Note the absence of the square on the bracket.

- The benchmark level can be expressed as the return on a benchmark fund (a fund that is typical of a sector, e.g, Asia ex-Japan) if this is more appropriate than an absolute level.

- Any of the risk measures discussed here can be expressed as measures of the risk relative to a suitable benchmark which may be an index (e.g., HSI, FTSE, DJ, NASDAQ etc), a median fund or some level of inflation.

- Downside risk measures have also been proposed based on an increasing function, say, $g(L-x)$, rather than the linear function $(L-x)$ itself in the integral above.

$$E[\text{Max}(g(L-x),0)] = \int_{-\infty}^{L} g(L-x)f(x)dx \tag{4.5}$$

  We shall focus here on either the linear function or the simple probability definition. Which of these we use will be made clear in the context.

- Shortfall measures are useful for monitoring a fund's exposure to risk because the expected underperformance relative to a benchmark is a concept that is apparently easy to understand.

- As with semi-variance, however, no attention is paid to the distribution of outperformance of the benchmark.

4.5 Value at Risk

Value at Risk (VaR) is a measure of how the market value of an asset or of a portfolio of assets is likely to decrease over a certain time period (usually over 1 day or 10 days) under usual conditions. It generalises the likelihood of underperforming by providing a statistical measure of downside risk.

VaR has three parameters:

- The time horizon to be analyzed - the **holding period**. The typical holding period is 1 day, although 10 days are used, for example, to compute capital requirements under the European Capital Adequacy Directive (CAD). For some problems, even a holding period of 1 year is appropriate.

- The confidence level at which the estimate is made. Popular confidence levels usually are 99% and 95%.

- The unit of the currency which will be used to denominate the value at risk (VaR).

The VaR is the minimum amount at risk to be lost from an investment (under "normal" market conditions) over a given holding period, at a particular confidence level. As such, it is the converse of shortfall probability, in that it represents the amount to be lost with a given probability, rather than the probability of a given amount to be lost.

Note that VaR cannot anticipate changes in the composition of the portfolio during the day. Instead, it reflects the riskiness of the portfolio based on the portfolio's current composition.

Example:

Suppose that the market value in US dollars of a portfolio is known today, but its market value tomorrow is not known. The holder of the portfolio might report that its portfolio has a 1-day VaR of \$2 million at the 5% level. This implies that (provided usual conditions will prevail over the 1 day) the holder expects that, with a probability of 95%, the value of its portfolio will decrease by at most \$2 million during 1 day:

$$P(\text{loss} \leq 2\text{million}) = 0.95.$$

In other words, there is a 5% probability that the value of its portfolio will decrease by 2 million or more during 1 day:

$$P(\text{loss} \geq 2\text{million}) = 0.05.$$

The key thing to note is that the target level (5% in the above example) is the given parameter here; the output from the calculation (\$2 million in the above example) is the maximum amount at risk (the value at risk) for that confidence level.

- VaR assesses the potential losses on a portfolio over a given future time period with a given degree of confidence. It can be measured either in absolute terms or relative to a benchmark. Again, VaR is based on assumptions that may not be immediately apparent.

- In particular, it is frequently calculated assuming a normal distribution of returns. If the distribution of returns is "fat-tailed", or skewed, tracking error (with its focus on the standard deviations of returns) may be misleading.

- Unfortunately, portfolios exposed to credit risk, systematic bias or derivatives may exhibit non-normal distributions. The usefulness of VaR in these situations depends on modelling skewed or fat-tailed distributions of returns, either in the form of statistical distributions (such as the Gumbel, Frechet or Weibull distributions) or via Monte Carlo simulations.

- However, the further one gets out into the"tails" of the distributions, the more lacking the data material and, hence, the more arbitrary the choice of the underlying probability becomes.

VaR is now a highly controversial topic. Some blame the assumptions underlying the VaR for the poor risk rating attached to the assets that led to the credit crunch. Others blame the incorrect choice of underlying probability distributions of the returns on these assets. Some blame the investors in these assets for not understanding just what is meant by VaR in the first place!

4.6 Example of Risk Measure Calculations

4.6.1 Discrete Probabilities

This example is based on discrete probabilities, but is easy to generalise to continuous distributions.

Suppose there are three possible future scenarios for a particular asset:

State	Return	Probability
1	10%	0.5
2	20%	0.3
3	50%	0.2

The above three possible different measures of investment risk are then determined as follows:

- **The mean return** μ:

$$\mu = 0.1 \times 0.5 + 0.2 \times 0.3 + 0.5 \times 0.2 = 0.21 = 21\%$$

- **The variance of the return**:

  Let R_i be the return on the i^{th} asset with probability p_i. Then the variance of the return is

$$\sum_{i=1}^{3}(\mu - R_i)^2 p_i = (0.1-0.21)^2 \times 0.5 + (0.2-0.21)^2 \times 0.3 + (0.5-0.21)^2 \times 0.2 = 0.0229 = 2.29\%$$

- **The semi variance of the return**:

  With the same notation, the semi variance of the return is

$$\sum_{i=1,\ (R_i<\mu)}^{3} (\mu - R_i)^2 p_i = (0.1 - 0.21)^2 \times 0.5 + (0.2 - 0.21)^2 \times 0.3 = 0.00608 = 0.608\%$$

- **The shortfall probability**:

  In words, the shortfall probability is the probability that the return will fall below a certain level L. If we set $L = 25\%$, say, then for discrete probabilities the shortfall probability is

$$\sum_{i=1,\ (R_i<L)}^{3} p_i = 0.5 + 0.3 = 0.8 = 80\%$$

Whereas if we were to set $L = 15\%$, then the shortfall probability would reduce to 50%.

4.6.2 Continuous Probabilities

Let's suppose that an investment is **assumed** to generate a return of R where:

$$R = 200,000 - 100,000N[1,1]$$

where $N[1,1]$ is a random variable (with mean 1 and variance 1). We can calculate the following:

- **Mean return:**

$$\begin{aligned}\bar{R} &= \langle R \rangle \\ &= \langle 200,000 - 100,000N[1,1] \rangle \\ &= \langle 200,000 \rangle - 100,000 \langle N[1,1] \rangle \\ &= 200,000 - 100,000 \times 1 \\ &= 100,000.\end{aligned}$$

- **Variance:** We use the result that if X is a random variable and a and b are constants then:$\sigma^2_{a+bX} = b^2\sigma^2_X$. In our case $a = 10^5$ and $\sigma^2_X = 1$ so that

$$\mathrm{var}(200,000 - 100,000N[1,1]) = 100,000^2 \times \mathrm{var}(N[1,1]) = 10^{10} \times 1 = 10^{10}.$$

- **Downside Semi-Variance:** Here the probability distribution is distributed symmetrically about the mean, since the random element is normal. (This will not always be the case). Therefore the downside semi-variance = upside semi variance = 1/2 the variance = 5×10^9.

- **Shortfall probability where the shortfall level is 50,000:** This is the probability P such that

$$\begin{aligned}P(R < 50,000) &= P(200,000 - 100,000N[1,1] < 50,000) \\ &= P(N[1,1] > 1.5) \\ &= P(N[0,1] + 1 > 1.5) \\ &= P(N[0,1] > 0.5) \\ &= 1 - \Phi(0.5) \\ &= 1 - 0.691462 \\ &= 0.308538.\end{aligned}$$

- **Value at Risk at the 10% level :** The VaR (not to be confused with the variance, Var, which is written with lower case "r") at the $x\%$ level is the value $t > 0$ such that the probability that the return is less than $-t$, (i.e., the loss on the return exceeds t) is x. So here we want the value t such that

$$\begin{aligned} P(R \leq -t) &= 0.1 \\ \Rightarrow P(200,000 - 100,000N[1,1] \leq -t) &= 0.1 \\ \Rightarrow P(N[1,1] \geq 2 + t/100,000) &= 0.1 \\ \Rightarrow P(N[0,1] \geq 1 + t/100,000) &= 0.1 \\ \Rightarrow \Phi(1 + t/100,000) &= 1 - 0.1 = 0.9 \\ \Rightarrow 1 + t/100,000 &= 1.28155 \\ \Rightarrow t &= 28,155. \end{aligned}$$

  plus change. Hence the VaR of the investment with return R at the 10% level is 28,155, i.e., there is a 10% chance that the asset will fall by more than 28,155 in value.

Note that this result is purely based on the **assumption** that the returns are distributed according to the stated distribution (here normal). It could be that the probability of large losses or gains is actually much higher (fat tails). It could be that the distribution is in fact actually completely different from the assumed one. Even if the assumed distribution is correct, there is no guarantee that you will, or will not, lose more or less than the VaR, since life can only be lived once! Just because there is only a (small) 10% chance that you will lose more than the VaR, it's doesn't mean to say that you won't!

It is possible to assume different probability distributions for returns (at least on exam papers!) and work out the corresponding measures of risk.

Chapter 5

Portfolio Theory

5.1 Introduction

In portfolio analysis we seek to answer some or all of the following questions:

- What are the basic principles underlying portfolio choice?
- How can investors structure their choices?
- Given a number of assets with known properties, what is the best group of assets to hold and in what proportions should they be held?
- What makes some portfolios "better" than others?
- How can we determine the composition of "better" portfolios?
- What portfolio best meets the needs of a given investor?

Mean-variance portfolio theory, sometimes called **modern portfolio theory (MPT)**, specifies a method for an investor to construct a portfolio that gives the maximum return for a specified risk, or the minimum risk for a specified return. However, the theory relies on some strong and limiting assumptions about the properties of portfolios that are important to investors and so it is also vital to understand the limitations of the theory. For example, in the form described here the theory ignores the investor's liabilities although it is possible to extend the analysis to include them.

The application of the mean-variance framework to portfolio selection falls conceptually into two parts:

- First, the definition of the properties of the portfolios available to the investor - **the opportunity set**.
- Second, the determination of how the investor chooses one out of all the **feasible portfolios** in the opportunity set.

5.2 Specification of the opportunity set

As the number of potential assets available to a given investor continually increases, the investor has an ever-expanding universe of portfolio choices, called the **opportunity set**). In specifying the opportunity set it is necessary to make some assumptions about how investors make decisions. Then the properties of portfolios can be specified in terms of relevant characteristics. It is assumed that investors select their portfolios on the basis of the expected return and the variance of that return over a single time horizon. Thus all the relevant properties of a portfolio can be specified with just two numbers - the mean return and the variance of the return. The variance (or standard deviation) is known as the risk of the portfolio.

To calculate the mean and variance of return for a portfolio it is necessary to know the expected return on each individual security and also the variance/covariance matrix for the available universe of securities.

5.2.1 Efficient portfolios

Two familiar assumptions about investor behaviour help with the concept of defining an efficient portfolio. The assumptions are:

- **Investors are never satiated.** At given level of risk, they will always prefer a portfolio with a higher return to one with a lower return.
- **Investors dislike risk.** For a given level of return they will always prefer a portfolio with lower variance to one with higher variance.

Once the set of efficient portfolios has been identified all others can be ignored.

A portfolio is inefficient if the investor can find another portfolio with the same expected return and lower variance, or the same variance and higher expected return. A portfolio is efficient if the investor cannot find a better one in the sense that it has both a higher expected return and a lower variance. However, an investor may be able to rank efficient portfolios by using a so-called utility function (see full notes for details on this).

5.2.2 Preliminaries

Now we will consider what happens to a portfolio. We shall assume (eventually) that:

- We have N risky assets, call them S_i, $(i = 1, 2, .., N)$.
- An **expected return** is associated with each risky asset: this expected return is constant for each asset.

- A **variance of return** is associated with each asset: this variance is constant for each asset. The variance measures how much the return on an asset differs from the expected return.
- A **covariance** is defined between each pair of two risky assets: this covariance is constant. (Often this assumption is distinctly dubious; we have to assume something however, so for the moment it'll have to do). The covariance measures how returns on assets move together.
- We may also have (usually just one) riskless asset (call it S_0 if we need it).

Assume that the ith asset has a return R_i. which is known at various times t_k $(k = 1, .., K)$. Then we define the **expected return** on R_i to be

$$E[R_i] = \bar{R}_i = \frac{1}{K}\sum_{k=1}^{K} R_i(t_k).$$

(How often we actually know the relevant values of the return depends upon the asset; in a typical case say we might know the return every month and calculate the expected return over 2 years so that $K = 24$.)

The **variance** of R_i is defined in the normal way by

$$\sigma_i^2 = E[(R_i(t_k) - \bar{R}_i)^2] = \frac{1}{K}\sum_{k=1}^{K}(R_i(t_k) - \bar{R}_i)^2.$$

Of course, the variance measures how far individual returns are from the mean return. In general, the higher the variance of a return, the higher the risk associated with it.

The **covariance** σ_{ij} between two risky assets i and j should be large if the two assets are likely to change "in harmony". We therefore define

$$\sigma_{ij} = E[(R_i - \bar{R}_i)(R_j - \bar{R}_j)] = \frac{1}{K}\sum_{k=1}^{K}(R_i(t_k) - \bar{R}_i)(R_j(t_k) - \bar{R}_j).$$

As there are two indices i, j, it is convenient to represent the covariance as a matrix

$$\sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}$$

- From the definitions above [1]

$$\sigma_{ij} = \sigma_{ji}$$

[1] Reminder of properties of covariances: Let X, Y, V, W be real random variables and a, b, c, d constants (or values of deterministic functions), then:

$$\mathrm{cov}(X, a) = 0, \qquad \mathrm{cov}(X, X) = \mathrm{var}X, \qquad \mathrm{cov}(X, Y) = \mathrm{cov}(Y, X),$$

$$\mathrm{cov}(aX + bY, cW + dV) = ac\ \mathrm{cov}(X, W) + ad\ \mathrm{cov}(X, V) + bc\ \mathrm{cov}(Y, W) + bd\ \mathrm{cov}(Y, V)$$

$$\sigma_{ii} = \sigma_i^2$$

so that the covariance matrix is symmetric about its leading diagonal.

- Note that (off diagonal) covariances σ_{ij}, $(i \neq j)$ may be positive, zero, or negative.
- The maximum value that σ_{ij} can take is given by $\sqrt{\sigma_i^2 \sigma_j^2} = \sigma_i \sigma_j$, while the minimum value is $-\sqrt{\sigma_i^2 \sigma_j^2} = -\sigma_i \sigma_j$.
- If the covariance is close to this value, then the two assets move together. If the covariance is 0, then they have no correlation and are essentially "independent"; if the covariance is close to its minimum value $- \mid \sigma_i \sigma_j \mid$ then the assets move in opposite directions to each other (see problem 1, sheet 1).
- For practical purposes, it is often easier to use the **correlation**; we define the correlation ρ_{ij} between S_i and S_j by

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

  (by definition ρ_{ij} therefore always lies between -1 and 1).
- An idea of the behaviour of an investment can easily be visualised by constructing the **risk-reward diagram**; we simply plot $\bar{R}$ against σ for given portfolios. Then increasing return corresponds to points higher on the y-axis, whilst increasing risk corresponds to increased values on the x-axis.

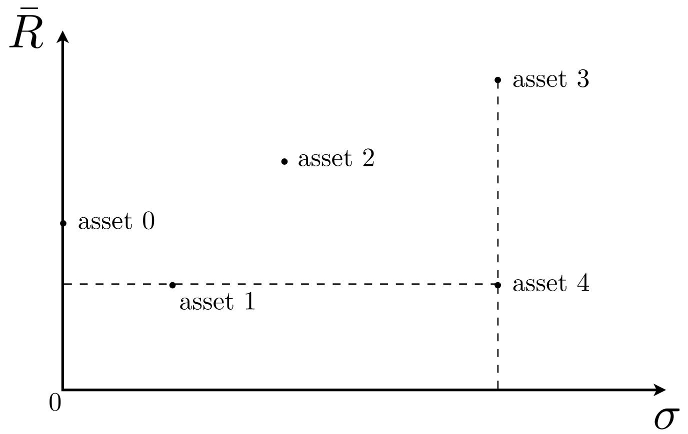

Figure 5.1: The expected returns and associated risks of five different assets plotted on a risk-reward diagram. Asset 0 is a riskless (zero risk) asset with a positive return (until recently: money in the bank). Asset 1 is, relative to assets 2-4, a low-risk, low-return asset. It has the same return as asset 4, but is less risky. Asset 2 has intermediate risk and return. Asset 3 is a high-risk, high-return asset, but is expected to return more than asset 4 for the same amount of risk. A rational investor would choose asset 3 over asset 4. However whether an investor choses assets 0,2 or 3 would depend on the amount of risk that each investor would wish to accept.

- Technical note: the correlations (or equivalently, the covariances) cannot take just *any* values between -1 and 1. (For example, how could we have three shares that were all perfectly negatively correlated with each other?.) If we are "making up" correlations (for an example, say) then the condition that share price histories exist with such specified correlations is that the correlation matrix is positive definite (i.e. that all its eigenvalues are non-negative).

5.2.3 Constructing a Portfolio

Now that we've completed the definitions, let us consider what happens when we hold a number of risky assets. To keep things simple at first, let us suppose that we have a portfolio Π composed of just two assets S_1 and S_2. The two assets are not of identical relative importance; for we have invested λ_1 of our total wealth in S_1 and λ_2 in S_2. We'll also assume for the present that $\lambda_1 + \lambda_2 = 1$ and λ_1 and λ_2 are both greater than 0 (thus prohibiting "short selling", which will be explained later).

We first note that

$$\bar{R}_\pi = E[\lambda_1 R_1 + \lambda_1 R_2] = \lambda_1 \bar{R}_1 + \lambda_2 \bar{R}_2$$

since expected values add and constant factors can be taken out.

Variances, on the other hand, do NOT add. We have

$$\begin{aligned}
\sigma_\pi^2 &= E[(R_\pi - \bar{R}_\pi)^2] \\
&= E[(\lambda_1 R_1 + \lambda_2 R_2 - \lambda_1 \bar{R}_1 - \lambda_2 \bar{R}_2)^2] \\
&= E[(\lambda_1 (R_1 - \bar{R}_1) + \lambda_2 (R_2 - \bar{R}_2))^2] \\
&= E[\lambda_1^2 (R_1 - \bar{R}_1)^2 + 2\lambda_1\lambda_2 (R_1 - \bar{R}_1)(R_2 - \bar{R}_2) + \lambda_2^2 (R_2 - \bar{R}_2)^2] \\
&= \lambda_1^2 \sigma_1^2 + 2\lambda_1\lambda_2\sigma_{12} + \lambda_2^2\sigma_2^2 \\
&= \lambda_1^2 \sigma_1^2 + 2\lambda_1\lambda_2\rho_{12}\sigma_1\sigma_2 + \lambda_2^2\sigma_2^2.
\end{aligned}$$

We conclude from this that if $\rho_{12} > 0$, then the variance of the portfolio is greater than the weighted sum of the individual variances - so the risk is higher. If, on the other hand, $\rho_{12} < 0$ then the risk is lower (the assets can "hedge" each other to some extent).

This calculation was fairly easy for 2 assets; but in general we wish to perform it for the case of N assets. So suppose our portfolio Π is given by

$$\Pi = \sum_{i=1}^{N} \lambda_i S_i, \quad \text{with} \quad \left(\sum_{i=1}^{N} \lambda_1 = 1 \right).$$

By the linearity of expectation we have, as before,

$$\bar{R}_\pi = E[R_\pi] = \sum_{i=1}^{N} \lambda_i E[R_i] = \sum_{i=1}^{N} \lambda_i \bar{R}_i$$

whilst for the variance we have (by simply generalising the previous result)

$$\sigma_\pi^2 = \sum_{i=1}^{N} \lambda_i^2 \sigma_i^2 + \sum_{i=1}^{N} \sum_{j=1 (j \neq i)}^{N} \rho_{ij} \lambda_i \lambda_j \sigma_i \sigma_j. \tag{5.1}$$

It's worth examining (5.1) a little more closely. If all the assets move independently and so all the covariances are zero, then all the ρ_{ij} are zero and thus

$$\sigma_\pi^2 = \sum_{i=1}^{N} \lambda_i^2 \sigma_i^2.$$

Suppose that $\lambda_i = 1/N$, so that the same amount of each asset is held. Then

$$\sigma_\pi^2 = \frac{1}{N} \left(\frac{1}{N} \sum_{i=1}^{N} \sigma_i^2 \right)$$

or

$$\sigma_\pi^2 = \frac{1}{N} E[\sigma_i^2].$$

This clearly tends to zero as N tends to infinity, showing that the variance or risk associated with a portfolio of uncorrelated assets decays like the number of assets.

This seems to argue for diversification within a portfolio; but sadly in the real world assets are almost never uncorrelated. For most assets, the covariances and correlations are positive. The risk cannot therefore be made to tend to zero. Nevertheless, the risk of the portfolio may be much less than the variance of an individual asset: with $\rho_{ij} \neq 0$, assume once again that $\lambda_i = 1/N$ for all i. Then

$$\sigma_\pi^2 = \sum_{i=1}^{N} \lambda_i^2 \sigma_i^2 + \sum_{i=1}^{N} \sum_{j=1 (j \neq i)}^{N} \rho_{ij} \lambda_i \lambda_j \sigma_i \sigma_j$$

and, assuming again that $\lambda_i = 1/N$, we have

$$\sigma_\pi^2 = \frac{1}{N} \sum_{i=1}^{N} \frac{\sigma_i^2}{N} + \sum_{i=1}^{N} \sum_{j=1 (j \neq i)}^{N} \frac{1}{N^2} \rho_{ij} \sigma_i \sigma_j$$

so that

$$\sigma_\pi^2 = \frac{1}{N} E[\sigma_i^2] + \frac{(N-1)}{N} \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1 (j \neq i)}^{N} \rho_{ij} \sigma_i \sigma_j.$$

Now the double sum has $N(N-1)$ terms, so that $1/(N(N-1))$ times it is none other than the average of $\rho_{ij}\sigma_i\sigma_j$ which is the average of σ_{ij}. Thus

$$\sigma_\pi^2 = \frac{1}{N} (average \;\; variance) + \frac{(N-1)}{N} (average \;\; covariance).$$

Now as the number of assets tends to infinity, we see that the portfolio variance tends to the average covariance. The individual risk of assets can thus be diversified away, but the risk arising from the covariances cannot.

5.2.4 Interpretations of a Two-Asset Portfolio

For simplicity, let's return again to the 2-asset portfolio and look at it a little more carefully. We will consider what happens to the risk-reward curve for various values of the correlation ρ_{12}.

CASE 1: $\rho_{12} = 1$.

Here the assets are perfectly correlated, and, from our previous expressions for the two-asset case, we have

$$\bar{R}_\Pi = \lambda_1 \bar{R}_1 + \lambda_2 \bar{R}_2, \qquad \sigma_\Pi = \lambda_1 \sigma_1 + \lambda_2 \sigma_2.$$

(since with the absence of correlations σ_π^2 is an exact square).

Now suppose we examine the effect of varying the λs. To do this, let $\lambda_1 = t$ and $\lambda_2 = 1 - t$ where $t \in [0, 1]$. Then

$$\bar{R}_\Pi = t\bar{R}_1 + (1-t)\bar{R}_2, \qquad \sigma_\Pi = t\sigma_1 + (1-t)\sigma_2.$$

Eliminating t, we find that

$$\bar{R}_\Pi = \frac{\sigma_\Pi(\bar{R}_1 - \bar{R}_2) + \sigma_1 \bar{R}_2 - \sigma_2 \bar{R}_1}{\sigma_1 - \sigma_2}$$

and thus traces out a straight line (of slope $(\bar{R}_1 - \bar{R}_2)/(\sigma_1 - \sigma_2)$) in the risk-reward curve as shown in figure 5.2. The conclusion is obvious; the return and risk of the portfolio is just the weighted average of the return and risk on each individual asset. So diversification achieves nothing for perfectly correlated assets (as we might expect). We can alter λ_1 and λ_2 as we wish, but all that we find is that proportionally more risk leads to proportionally more reward.

CASE 2: $\rho_{12} = -1$.

Now the assets are perfectly negatively correlated, and, from our previous expressions we have

$$\bar{R}_\Pi = \lambda_1 \bar{R}_1 + \lambda_2 \bar{R}_2, \qquad \sigma_\Pi = \mid \lambda_1 \sigma_1 - \lambda_2 \sigma_2 \mid .$$

Let us again examine the effect of varying the λs. We set $\lambda_1 = t$ and $\lambda_2 = 1 - t$ again and note immediately that when

$$\lambda_1 = \frac{\sigma_2}{\sigma_1 + \sigma_2}$$

σ_Π, and hence the risk, is zero! (And the corresponding reward is easily worked out to be $(\sigma_2 \bar{R}_1 + \sigma_1 \bar{R}_2)/(\sigma_1 + \sigma_2)$.) This means that for two perfectly negatively correlated assets we can construct a completely risk free portfolio by choosing λ_1 in this way. This is "hedging" in its purest form. Eliminating t as usual, we find that the risk-reward curve is now two straight lines, specifically

$$\bar{R}_\Pi = \sigma_\Pi \left(\frac{\bar{R}_2 - \bar{R}_1}{\sigma_1 + \sigma_2} \right) + \left(\frac{\sigma_1 \bar{R}_2 + \sigma_2 \bar{R}_1}{\sigma_1 + \sigma_2} \right) \qquad (t < \sigma_2/(\sigma_1 + \sigma_2))$$

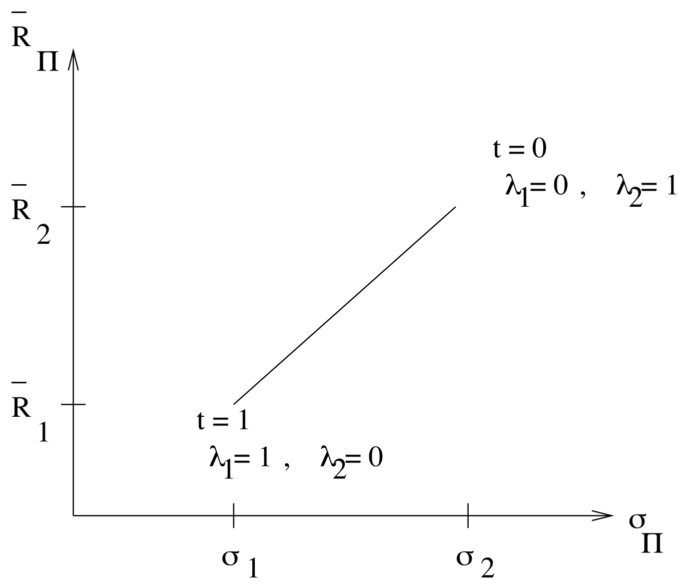

Figure 5.2: Risk-reward diagram for two perfectly correlated assets

$$\bar{R}_\Pi = \sigma_\Pi \left(\frac{\bar{R}_1 - \bar{R}_2}{\sigma_1 + \sigma_2} \right) + \left(\frac{\sigma_1 \bar{R}_2 + \sigma_2 \bar{R}_1}{\sigma_1 + \sigma_2} \right) \qquad (t > \sigma_2/(\sigma_1 + \sigma_2)).$$

The risk-reward curve for $t \in [0, 1]$ is shown in figure 5.3.

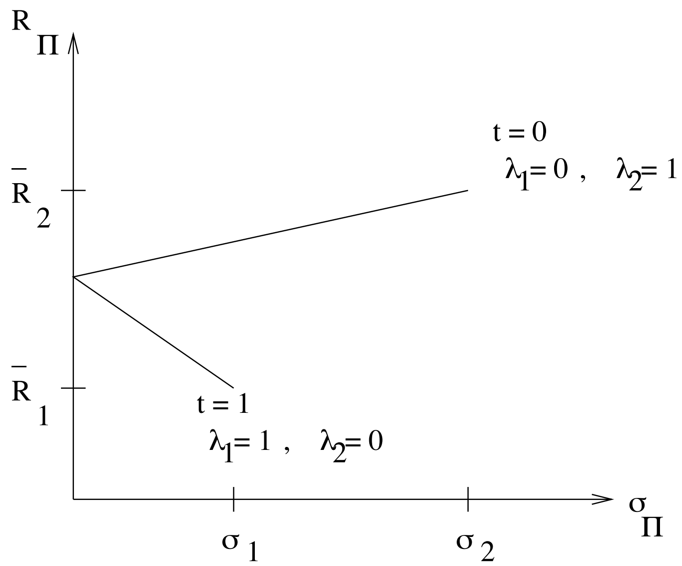

Figure 5.3: Risk-reward diagram for two perfectly negatively correlated assets

CASE 3: $\rho_{12} = 0$**.**

Now the assets have no correlation and move completely independently. Proceeding as before, we have

$$\bar{R}_\Pi = t\bar{R}_1 + (1-t)\bar{R}_2, \qquad \sigma_\Pi^2 = t^2\sigma_1^2 + (1-t)^2\sigma_2^2.$$

This curve is now more complicated to draw, (in general σ_Π^2 is given by a quadratic in $\bar{R}_\Pi$; see figure 5.4 for a typical curve) but we note immediately that

$$\frac{\partial \sigma_\Pi^2}{\partial t} = 2t\sigma_1^2 - 2(1-t)\sigma_2^2$$

and thus σ_Π has a minimum when

$$t = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

(which is between 0 and 1 for all values of σ_1 and σ_2). This means that when $\rho_{12} = 0$ then there is ALWAYS a combination of the two assets that has a risk smaller than each one on its own - and of course this is the most desirable portfolio to hold.

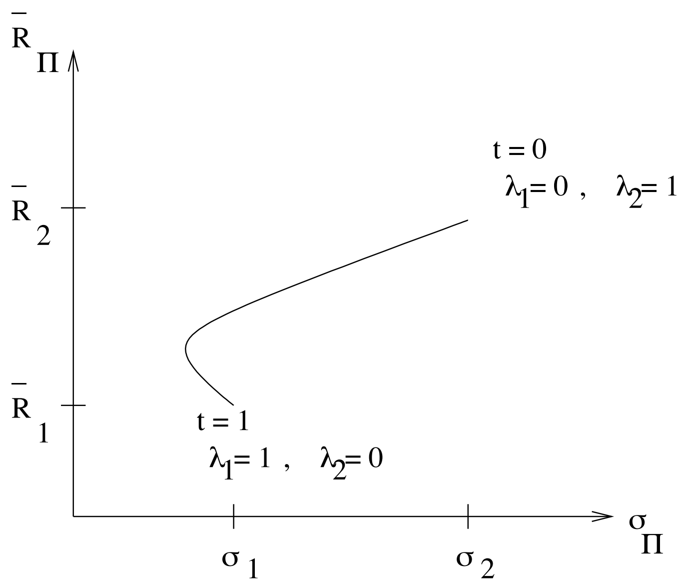

Figure 5.4: Risk-reward diagram for two typical uncorrelated assets

GENERAL CASE: $-1 \leqslant \rho_{12} \leqslant 1$.

How about the general case? Well, whether or not a minimum σ_Π exists depends on the values of σ_1, σ_2 and the correlation ρ_{12}. As we saw above, a minimum always exists for $\rho_{12} \leqslant 0$ and MAY exist for larger values of ρ_{12} (depending on the variances). (The exact condition turns out to be that there is a minimum whenever $\rho_{12} < \sigma_2/\sigma_1$, but this is not very important) One thing is for sure though: if we we look at the risk-reward curve for various values of the correlation, then the risk-reward curve ALWAYS lies between the $\rho_{12} = -1$ and $\rho_{12} = 1$ cases consider above. (See figure 5.5).

It is also possible (though tedious) to show that the risk-reward curve is CONVEX. (Here convexity means that if any 2 points on the curve are joined with a straight line then the line lies to the right of the curve.) We also note that any point on the risk-reward curve is itself an asset with an expected return and risk. This completes our initial examination of the 2-asset case.

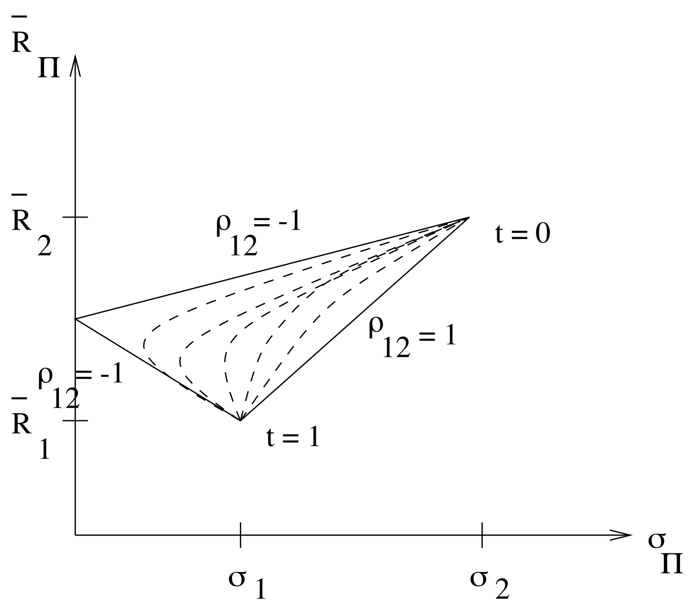

Figure 5.5: Risk-reward curves for various ρ_{12} between 1 and -1 (broken lines)

5.2.5 A Portfolio of N Assets and the Efficient Frontier

What can we say about general portfolios with N assets? We know that the lower the correlation coefficient between assets, the higher the payoff from diversification. But what does the risk-reward space look like? With more than 2 assets, it's fairly obvious that we won't just get a curve (as for the two asset case). Instead we get a REGION, which is usually known as the **Portfolio possibilities** region. In figure 5.6 the following experiment was carried out: we took 4 assets (actually with respective $\bar{R}_i$ and σ_i equal to 4, 5, 7, 8 and 5, 7, 8 11 and correlations $\rho_{12} = 0.3$, $\rho_{13} = 0.5$, $\rho_{14} = -0.6$, $\rho_{23} = 0.9$, $\rho_{24} = 0.5$, $\rho_{34} = 0.3$) (note: the ρ_{ij} *do* lead to a positive definite, and thus realisable correlation matrix) and randomly generated 5000 portfolios (by generating 4 random $\lambda_i \geqslant 0$ whose sum was one).

Its clear that

- the portfolio possibility space forms a connected region in $(\sigma, \bar{R})$ space
- that the portfolio possibility space is bounded to the left by a convex curve and
- an investor wishing to maximise return for a given risk, or minimise risk for a given return would be interested ONLY in the curve bounding the "top left" part of the portfolio possibility space.
- The set of portfolios lying along this curve is called the **efficient frontier**.

More carefully defined, the efficient frontier is the boundary of the portfolio possibilities set that satisfies $d\bar{R}/d\sigma > 0$ and $d\sigma/d\bar{R} > 0$.

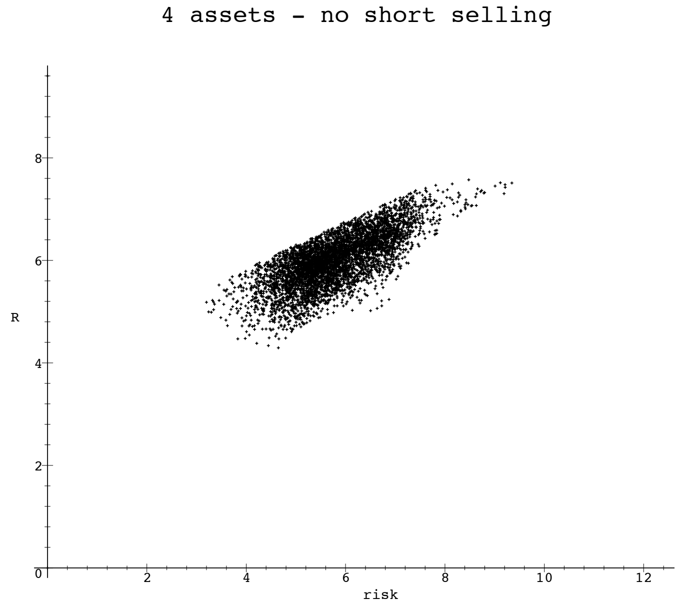

Figure 5.6: The (numerically generated) portfolio possibilities region for 4 assets

5.3 Including Short Selling

Short selling is popular in general, but although we would like to extend our theory to include short selling, we have not wasted our time in studying the case when no short selling is allowed. The reason for this is that most institutional investors do not short sell. Many are forbidden by law to short sell, and others just don't do it.

In any case, it's pretty easy to include short selling in our theory. It's easy to see that short selling just corresponds to have *less than zero* of a particular asset. For 2 assets, we still have $\lambda_1 + \lambda_2 = 1$, but now the λ_i may be either postive or negative numbers if we allow short selling. We still have

$$\bar{R}_\Pi = \lambda_1 \bar{R}_1 + \lambda_2 \bar{R}_2, \qquad \sigma_\Pi^2 = \lambda_1^2 \sigma_1^2 + 2\rho_{12}\lambda_1\lambda_2\sigma_1\sigma_2 + \lambda_2^2\sigma_2^2$$

and we may now reconsider some of our earlier risk-reward diagrams.

CASE 1: $\rho_{12} = 1$**.**

Now the assets are perfectly correlated, and our diagram is as before, but with the straight line extended (see figure 5.7). With $\lambda_1 = t$ etc. as usual, the extensions of the lines correspond to $t < 0$ (short selling asset 1 and investing the proceeds in the riskier asset 2) and $t > 1$ (short selling asset 2 and investing the profits in the less-risky asset 1). Not much has changed though; both risk and reward still increase exactly in step with each other, as we'd expect.

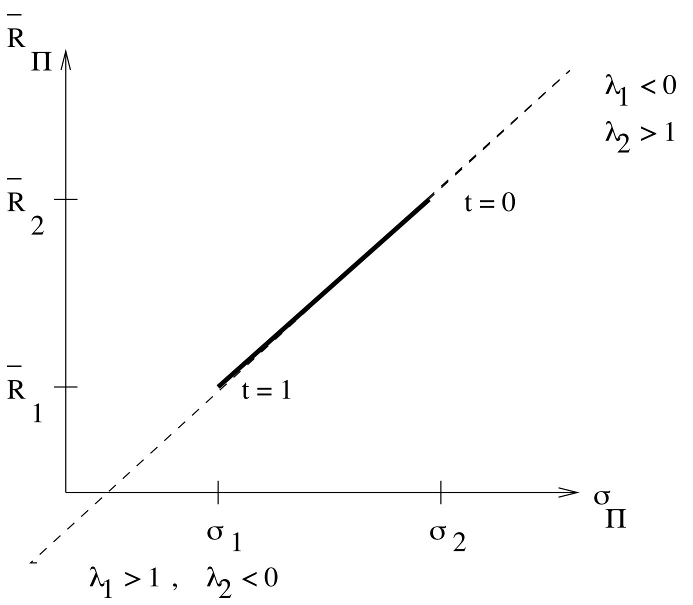

Figure 5.7: Risk-reward diagram for two perfectly correlated assets with short selling

CASE 2: $\rho_{12} = -1$**.**

Now the assets are perfectly negatively correlated. Once again, the diagram is as before but with the two previous straight lines extended (figure 5.8): perfect hedging is still possible.

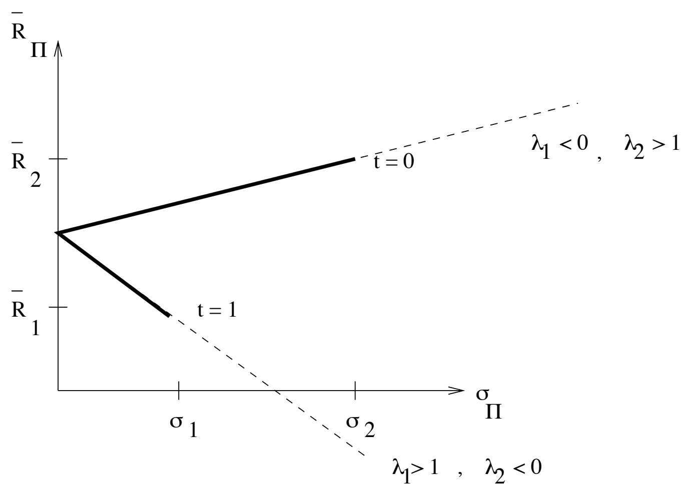

Figure 5.8: Risk-reward diagram for two perfectly negatively correlated assets with short selling

GENERAL CASE: $-1 \leqslant \rho_{12} \leqslant 1$**.**

The risk-reward curve for the general 2-asset case when short sales are allowed is very predictable. Once again, it is exactly the same as the case when no short sales are allowed, except that the curves are extended in each direction.

5.3.1 A portfolio with N assets where Short Selling is Allowed

For an N-asset portfolio, we can still generate the portfolio possibilities region as before. The only difference now is that it will be larger than before as there are more possible values of the portfolio proportions. To produce the portfolio possibilities region shown in figure 5.9 5000 random portfolios were generated; short selling was allowed in that t was allowed to take values greater than -1. (Once again, we used respective $\bar{R}_i$ and σ_i equal to 4, 5, 7, 8 and 5, 7, 8 and 11 and correlations $\rho_{12} = 0.3$, $\rho_{13} = 0.5$, $\rho_{14} = -0.6$, $\rho_{23} = 0.9$, $\rho_{24} = 0.5$, $\rho_{34} = 0.3$, just as in figure 5.6). As before, we are only interested in the part of the bounding curve (the efficient frontier) with $d\bar{R}/d\sigma > 0$ and $d\sigma/d\bar{R} > 0$. The efficient frontier again starts at the minimum variance portfolio, and is again concave since each individual two-portfolio case is. The main difference between this and the case where no short selling is allowed is that now there are portfolios whose return is arbitrarily large: this is not surprising since with short selling it is possible to sell assets with low returns and buy ones with high returns using the resulting funds. The efficient frontier thus has no bound above (though of course such high-return portfolios inevitably involve higher risk).

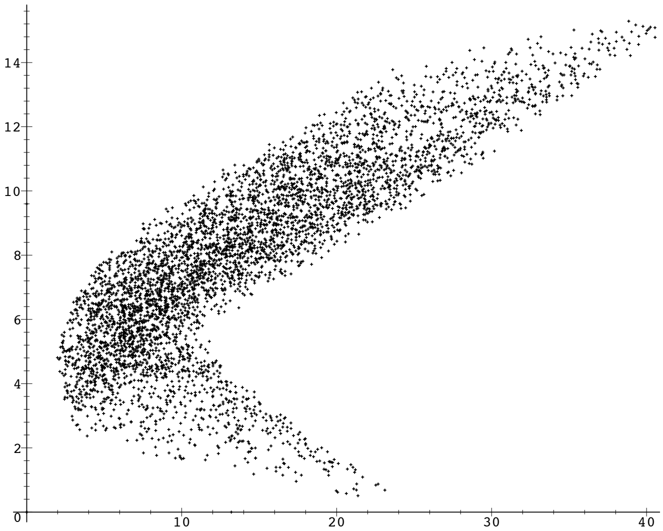

Figure 5.9: The (numerically generated) portfolio possibilities region for 4 assets with short selling allowed

5.4 Including Riskless Asset(s)

Our portfolio might also include some riskless assets. These may be bank deposits or Government bonds, for example. We have assumed that *all* riskless assets attract the same rate of interest, so without loss of generality we shall suppose that we have one riskless asset, say S_0, with return R_0. What sort of risks are associated with this asset? By definition, none.

Such an asset has zero variance and zero correlation with any of the risky assets, so that

$$\bar{R}_0 = R_0$$

and

$$\sigma_0 = \rho_{0i} = 0.$$

Let us now consider a portfolio where a fraction λ_0 of the total portfolio is invested in the riskless asset and the remainder $\lambda_p = 1 - \lambda_0$ in a single risky asset (or a portfolio of many risky assets). The portfolio is thus

$$\Pi = \lambda_0 S_0 + \lambda_p S_p.$$

The risk-reward structure of this portfolio is given by

$$\bar{R}_\Pi = E[\lambda_0 R_0 + \lambda_p R_p] = \lambda_0 \bar{R}_0 + \lambda_p \bar{R}_p$$

and

$$\sigma_\Pi^2 = \lambda_p^2 \sigma_p^2$$

(since the riskless asset has no variance). Thus

$$\sigma_\Pi = \mid \lambda_p \mid \sigma_p,$$

(the plus or minus sign depending on whether λ_p is greater than or less than zero; the latter possibly applying as usual when short selling is allowed). This makes the relationship between the risk and the reward particularly simple, for we have

$$\bar{R}_\Pi = \left(1 - \frac{\sigma_\Pi}{\sigma_p}\right) \bar{R}_0 + \frac{\sigma_\Pi}{\sigma_p} \bar{R}_p \qquad (\lambda_p > 0)$$

$$\bar{R}_\Pi = \left(1 + \frac{\sigma_\Pi}{\sigma_p}\right) \bar{R}_0 - \frac{\sigma_\Pi}{\sigma_p} \bar{R}_p \qquad (\lambda_p < 0).$$

The consequent risk-reward diagram (a pair of straight lines, of course) is shown in figure 5.10.

5.5 Optimal Portfolio, Capital Market Line and Market Price of Risk

If investors have homogeneous expectations, then they are all faced by the same efficient frontier of risky securities.

If in addition they are all subject to the same risk-free rate of interest, then the calculations above show that the efficient frontier collapses to the straight line in $(\bar{R}, \sigma)$ space which passes through the risk-free rate of return R_0 on the $\bar{R}$-axis.

Consequently there is a unique point that simultaneously:

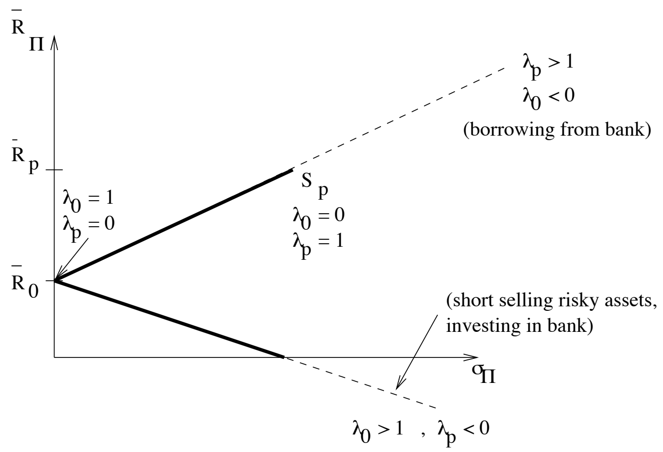

Figure 5.10: Risk-reward diagram for portfolio of riskless and risky assets (short selling allowed)

- lies within the opportunity set of the N risky assets;
- lies within the efficient frontier of a portfolio of N riskless and one riskless asset;
- maximises the slope of the line of possible riskless/risky portfolios as discussed above.

A rational investor should therefore *always* choose to invest in a combination of the riskless asset S_0 and the portfolio of risky assets at the tangent point along the line of maximal slope. This line is called the **Capital Market Line** (CML).

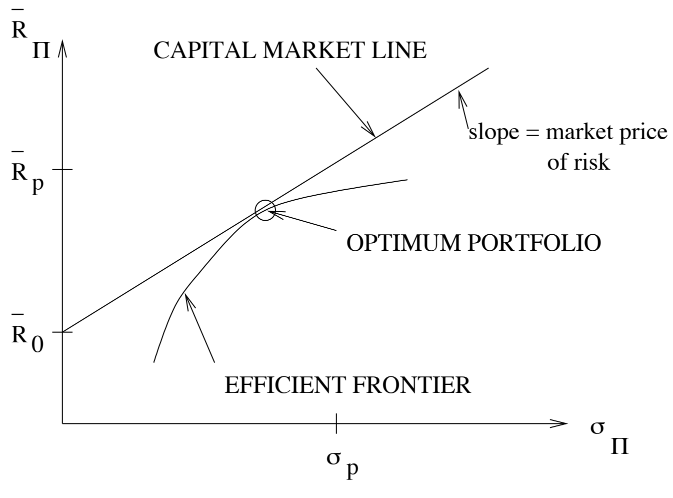

Figure 5.11: Optimal portfolio and capital market line

Why would a rational investor do this?

- Any other combined riskless-risky portfolio will lie on a line within the risky opportunity set and thus for each given risk will have a lower return than a combined portfolio on the CML.

- No combined riskless/risky portfolio with a higher rate of return per risk is possible since this straight line will lie outside the opportunity set of risky assets.

Now a key assumption of the Capital Asset Pricing Model (CAPM see full notes for details) is that all investors in the market are rational. Hence if one rational investor invests at some point on the CML, then so should all of them. They will all obtain the same rate of return per risk, i.e., the slope of the CML.

If all investors invest on the CML, then they should hold the same risky portfolio, i.e., the relative ratios of the risky assets they hold will be the same for all investors.

For this reason the risky part of the portfolio is called the **market portfolio** or **optimum portfolio** and the slope of the maximal straight line is called the **market price of risk**.

Note that the particular point that each investor chooses to be on the CML will depend on their own *utility function*, i.e., how much risk they are prepared to accept. (see full notes for details of the *utility function*).

- If the investor chooses to invest only in risky assets, then the proportion of their wealth invested in the riskless assest will be $\lambda_0 = 0$.
- If they wish to increase their return (and risk) they will invest more in the risky asset, i.e., increase λ_p, by borrowing from the bank $\lambda_0 < 0$ $(\lambda_0 + \lambda_p = 1)$.
- If they wish to reduce their risk (and return) then they may place some of their wealth in the bank $(\lambda_0 > 0)$.

Recall that the risk-reward structure of a combined portfolio Π is given by

$$\bar{R}_\Pi = E[\lambda_0 R_0 + \lambda_p R_p] = \lambda_0 \bar{R}_0 + \lambda_p \bar{R}_p$$

where $\bar{R}_0$ is the risk-free rate and $\bar{R}_p$ is the expected return on the optimal risky portfolio, and

$$\lambda_0 + \lambda_p = 1.$$

Now the CML will be given by

$$\bar{R}_\Pi - \bar{R}_0 = \frac{(\bar{R}_p - \bar{R}_0)}{\sigma_p} \sigma_\Pi$$

Similarly (since the riskless asset has no variance)

$$\sigma_\Pi = \mid \lambda_p \mid \sigma_p,$$

(the plus or minus sign depending on whether λ_p is greater than or less than zero; the latter possibly applying as usual when short selling is allowed). This makes the relationship between the risk and the reward particularly simple, for we have

$$\bar{R}_\Pi = \left(1 - \frac{\sigma_\Pi}{\sigma_p}\right) \bar{R}_0 + \frac{\sigma_\Pi}{\sigma_p} \bar{R}_p \qquad (\lambda_p > 0)$$

$$\bar{R}_\Pi = \left(1 + \frac{\sigma_\Pi}{\sigma_p}\right)\bar{R}_0 - \frac{\sigma_\Pi}{\sigma_p}\bar{R}_p \qquad (\lambda_p < 0).$$

The consequent risk-reward diagram is shown in figure 5.10.

The line of maximal slope is called the **capital market line** and its slope is termed the **market price of risk** (see figure 5.11). The market price of risk measures how much the portfolio profits by above the risk free rate for accepting a certain level of risk. It is tangential to the efficient frontier for risky securities.

The combination of a riskless asset with a risky portfolio is of crucial importance in the **Capital Asset Pricing Model** which are covered in the full notes.

5.6 Example of calculation of optimal portfolio and CML: 2 risky + 1 riskless asset

Three assets are available to an investor with the following characteristics

Asset	$\bar{R}_i$	σ_i
A	10%	20%
B	5%	0%
C	6%	10%

- All the assets are uncorrelated. Clearly B is a riskless asset.
- The return on a portfolio consisting of A and B only will be

$$\bar{R}_{\text{AB}} = \lambda\bar{R}_A + (1-\lambda)\bar{R}_B = 0.1\lambda + 0.05(1-\lambda),$$

with standard deviation

$$\sigma_{\text{AB}} = \lambda\sigma_A = 0.2\lambda \equiv \sigma.$$

Thus the portfolio opportunity set and efficient frontier for investments in A and B only is a straight line given by eliminating λ between the above two equations to obtain:

$$\bar{R}_{AB} = \frac{\sigma_{AB}}{\sigma_A}\bar{R}_A + \left(1 - \frac{\sigma_{AB}}{\sigma_A}\right)\bar{R}_B = 0.1 \times \frac{\sigma_{AB}}{0.2} + 0.05(1 - \frac{\sigma_{AB}}{0.2}) = 0.25\sigma_{AB} + 0.05.$$

- The return on a portfolio consisting of A and C only is given by

$$\bar{R}_{\text{AC}} = 0.1\mu + 0.06(1-\mu),$$

with standard deviation

$$\sigma_{\mathrm{AC}} = \sqrt{0.2^2\mu^2 + 0.1^2(1-\mu)^2}.$$

The efficient frontier is no longer a straight line.

- The efficient frontier of a portfolio consisting of A, B and C will be a straight line that passes through $(\bar{R}_{\mathrm{ABC}}, \sigma_{\mathrm{ABC}}) = (0.05, 0)$ and is tangential to the portfolio opportunity space of A and C. This straight line has a slope that maximises

$$\theta = \frac{\bar{R}_{\mathrm{AC}} - 0.05}{\sigma_{\mathrm{AC}}},$$

(see diagram above).

Hence we must maximise

$$\theta(\mu) = \frac{0.1\mu + 0.06(1-\mu) - 0.05}{\sqrt{0.2^2\mu^2 + 0.1^2(1-\mu)^2}},$$

with respect to μ. The value of μ that achieves this is the optimal portfolio.

$$\begin{aligned}
\frac{d\theta}{d\mu} &= \frac{(\sqrt{0.2^2\mu^2 + 0.1^2(1-\mu)^2})(0.04) - (0.04\mu + 0.01)\frac{(0.2^2\mu - 0.1^2(1-\mu))}{\sqrt{0.2^2\mu^2+0.1^2(1-\mu)^2}}}{0.2^2\mu^2 + 0.1^2(1-\mu)^2} = 0 \\
\mu &= \cdots \\
\Rightarrow \mu &= \frac{5}{9}
\end{aligned}$$

Hence the optimal portfolio is composed of a 5/9 proportion of A and a 4/9 proportion of C.

The corresponding value of θ is $\sqrt{29}/20 = 0.269258\cdots$

- The efficient frontier consisting of A, B and C is thus the straight line

$$\bar{R} = 0.269258\sigma + 0.05.$$

DIAGRAM!!!

5.7 Example of calculation of optimal portfolio and CML: 3 risky + 1 riskless asset

Suppose that we now have three risky assets S_1, S_2 and S_3 and a riskless asset S_0. The associated returns and sigmas are (all in percent)

$$\bar{R}_1 = 14, \quad \bar{R}_2 = 8, \quad \bar{R}_3 = 20, \quad \bar{R}_0 = 5, \quad \sigma_1 = 6, \quad \sigma_2 = 3, \quad \sigma_3 = 15$$

whilst the correlations are

$$\rho_{12} = \rho_{21} = 0.5, \quad \rho_{13} = \rho_{31} = 0.2, \quad \rho_{23} = \rho_{32} = 0.4.$$

Problem: show that the capital market line is given by

$$\bar{R} = \frac{1}{20} + \frac{29\sqrt{2}}{\sqrt{609}}\sigma \sim 0.05 + 1.66\sigma,$$

the optimal portfolio is

$$X_1 = 14/18, \quad X_2 = 1/18, \quad X_3 = 3/18$$

and the market price of risk and the portfolio risk are about 1.66 and 0.058 respectively.

Solution: OK, well first we have to work out the σ_{ij}. As usual we have $\sigma_{ii} = \sigma_i^2$ and $\sigma_{ij} = \sigma_i \sigma_j \rho_{ij}$ and working these out (and remembering that a percentage is a 1/100th!) we get

$$(\sigma)_{ij} = \frac{1}{10^4}\begin{pmatrix} 36 & 9 & 18 \\ 9 & 9 & 18 \\ 18 & 18 & 225 \end{pmatrix}.$$

The portfolio sigma is given by

$$\sigma_P = \sqrt{\sum_{i,j} X_i X_j \sigma_{ij}} = \sqrt{\vec{X}^T \sigma \vec{X}}$$

where $\vec{X} = (X_1, X_2, X_3)^T$ and $\sigma = (\sigma_{ij})$. Thus

$$\sigma_P = \frac{1}{100}\sqrt{36X_1^2 + 18X_1X_2 + 36X_1X_3 + 9X_2^2 + 36X_2X_3 + 225X_3^2}.$$

We also have

$$R_P - R_0 = \sum_{i=1}^{3} X_i(R_i - R_0) = \frac{1}{100}(9X_1 + 3X_2 + 15X_3)$$

and so

$$\theta = \frac{R_P - R_0}{\sigma_P} = \frac{9X_1 + 3X_2 + 15X_3}{\sqrt{36X_1^2 + 18X_1X_2 + 36X_1X_3 + 9X_2^2 + 36X_2X_3 + 225X_3^2}};$$

our job is to maximise θ subject to $X_1 + X_2 + X_3 = 1$.

OK well first we differentiate with respect to each X_i and set the resultant equal to zero to look for a turning point. Consider what we'll get when we do this: writing the top of θ as A say and the bottom as $\sqrt{B}$, then we find (with $' = \partial/\partial X_1$)

$$\theta' = \frac{\sqrt{B}A' - AB^{-1/2}B'/2}{B} = \frac{1}{\sqrt{B}}\left(A' - \frac{A}{2B}B'\right) = 0.$$

Things that are differentiated will be different depending on the differentiation variable, but because of the homogeneous nature of the functions involved, we can regard A/B as being equal to a "scale factor" (call it λ). This is a trick, but it avoids the need to use Lagrange multipliers and the answer we get is correct in the end.

For the above to be zero we need $A' - AB'/2B = 0$, and thus $A' = \lambda B'/2$. In this case this gives

$$\lambda(36X_1 + 9X_2 + 18X_3) = 9.$$

We now scale the X's by setting $Z_1 = \lambda X_1$, $Z_2 = \lambda X_2$, $Z_3 = \lambda X_3$. Doing the corresponding thing for each of the other variables gives us two more equations and we finally end up with the system

$$36Z_1 + 9Z_2 + 18Z_3 = 9$$

$$9Z_1 + 9Z_2 + 18Z_3 = 3$$

$$18Z_1 + 18Z_2 + 225Z_3 = 15$$

and so now we have three linear equations in the Z's. Solving these in the normal way (e.g., Gaussian elimination) gives

$$Z_1 = \frac{2}{9}, \quad Z_2 = \frac{1}{63}, \quad Z_3 = \frac{1}{21}$$

and now we use the fact that $X_1 + X_2 + X_3 = 1$. This tells us that

$$Z_1 + Z_2 + Z_3 = \lambda$$

and thus

$$\lambda = \frac{2}{9} + \frac{1}{63} + \frac{1}{21} = \frac{2}{7}.$$

Now we can get the X's from the Z's to find that

$$X_1 = \frac{14}{18}, \quad X_2 = \frac{1}{18}, \quad X_3 = \frac{3}{18}.$$

This is the optimum portfolio weighting, and the expected return is

$$X_1\bar{R}_1 + X_2\bar{R}_2 + X_3\bar{R}_3 = 11/75 \sim 0.14666666.$$

The portfolio risk σ_P is also calculated just by "putting the values in" and gives a value of $\sqrt{1218}/600$, which is about 0.05816. The market price of risk is just the value of θ with all the solutions inserted, which gives $29\sqrt{2}/\sqrt{609} \sim 1.66$, whilst the capital market line is given by

$$\bar{R} = \bar{R}_0 + \theta_{max}\sigma = \frac{1}{20} + \frac{29\sqrt{2}}{\sqrt{609}}\sigma \sim 0.05 + 1.66\sigma.$$

This type of calculation could be generalised in an obvious way to n-risky + 1 riskless assets. (you would then probably solve the problem numerically).

5.8 Portfolio management when the expected return is specified

So far we have seen that, when a risky portfolio is combined with a riskless asset, an optimum portfolio can be identified. However, there may be other definitions of optimality than simply "the most return for the least risk". There are many approaches to portfolio selection topics of this sort, and for simplicity we will only consider a fairly elementary example of what can be achieved.

Suppose that we allow short sales and have a portfolio Π composed of three risky assets but no riskless asset. Suppose also that, as usual, we are given the expected returns $\bar{R}_1$, $\bar{R}_2$ and $\bar{R}_3$ of the three risky assets and also the covariance matrix

$$C = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}$$

which, as usual, we assume to be positive definite. The portfolio proportions of Π are, as usual, given by λ_1, λ_2 and λ_3, where $\lambda_1 + \lambda_2 + \lambda_3 = 1$. As in the previous section, it is relatively easy to write down the reward and risk for the portfolio: we have

$$\bar{R}_\Pi = \lambda_1 \bar{R}_1 + \lambda_2 \bar{R}_2 + \lambda_3 \bar{R}_3$$

and

$$\sigma_\Pi^2 = (\lambda_1, \lambda_2, \lambda_3) \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}.$$

Now, however, suppose that we *INSIST that the portfolio enjoys a certain level R of expected return.* This might happen, for example, if we are a large pension fund that has promised a given return. We now have a new problem: minimise σ_Π^2 subject to the two constraints

$$\lambda_1 + \lambda_2 + \lambda_3 = 1 \tag{5.2}$$

and

$$\lambda_1 \bar{R}_1 + \lambda_2 \bar{R}_2 + \lambda_3 \bar{R}_3 = R. \tag{5.3}$$

The easiest way to to solve this problem is to use two Lagrange multipliers α and β. We set

$$L = \sigma_\Pi^2 + \alpha(\lambda_1 + \lambda_2 + \lambda_3 - 1) + \beta(\lambda_1 \bar{R}_1 + \lambda_2 \bar{R}_2 + \lambda_3 \bar{R}_3 - R)$$

This looks rather odd at first. Why introduce more variables? However, we now look for the minimum of L, by solving the (simultaneous linear) equations in λ_1, λ_2, λ_3, α and β given by

$$\frac{\partial L}{\partial \lambda_1} = \frac{\partial L}{\partial \lambda_2} = \frac{\partial L}{\partial \lambda_3} = \frac{\partial L}{\partial \alpha} = \frac{\partial L}{\partial \beta} = 0.$$

Performing the partial derivatives with respect to α and β, we obtain

$$\frac{\partial L}{\partial \alpha} = \lambda_1 + \lambda_2 + \lambda_3 - 1 = 0$$

and

$$\frac{\partial L}{\partial \beta} = \lambda_1 \bar{R}_1 + \lambda_2 \bar{R}_2 + \lambda_3 \bar{R}_3 - R = 0$$

Clearly if these are to vanish we satisfy both the constraint $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and the required return condition $R = \lambda_1 \bar{R}_1 + \lambda_2 \bar{R}_2 + \lambda_3 \bar{R}_3$. The vanishing of the other partial derivatives of L with respect to $\lambda_i,\ i = 1, 2, 3$, in fact minimise the portfolio risk function σ_Π^2, but now subject to the simultaneous satisfaction of the $\lambda-$sum constraint and the required return condition. Just what the resulting values of α and β are is not ultimately important.

Note that the specification of R may set limits on whether or not the required return may be accomplished without short selling.

5.8.1 An example

We indicate how the theory of the previous section applies by using an example. Suppose that the expected returns for the three risky assets are $\bar{R}_1 = 2$, $\bar{R}_2 = 3$ and $\bar{R}_3 = 5$, and the covariance matrix is

$$C = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 5 \end{pmatrix}.$$

Then the portfolio return is

$$\bar{R}_\Pi = 2\lambda_1 + 3\lambda_2 + 5\lambda_3$$

and by doing the sum (5.8) we find that the portfolio risk is

$$\sigma_\Pi^2 = \lambda_1^2 + 2\lambda_1\lambda_2 + 2\lambda_2^2 + 2\lambda_2\lambda_3 + 5\lambda_3^2$$

The Lagrangian L is

$$L = \lambda_1^2 + 2\lambda_1\lambda_2 + 2\lambda_2^2 + 2\lambda_2\lambda_3 + 5\lambda_3^2 + \alpha(2\lambda_1 + 3\lambda_2 + 5\lambda_3 - R) + \beta(\lambda_1 + \lambda_2 + \lambda_3 - 1)$$

and so the first three minimisation equations are

$$\frac{\partial L}{\partial \lambda_1} = 2\lambda_1 + 2\lambda_2 + 2\alpha + \beta = 0$$

$$\frac{\partial L}{\partial \lambda_2} = 2\lambda_1 + 4\lambda_2 + 2\lambda_3 + 3\alpha + \beta = 0$$

$$\frac{\partial L}{\partial \lambda_3} = 2\lambda_2 + 10\lambda_3 + 5\alpha + \beta = 0.$$

These equations can be solved in the normal way to reveal that

$$\lambda_1 = -\alpha - \frac{5}{8}\beta, \quad \lambda_2 = \frac{1}{8}\beta, \quad \lambda_3 = -\frac{1}{2}\alpha - \frac{1}{8}\beta.$$

The other two minimisation equations now give

$$\frac{\partial L}{\partial \alpha} = 2\lambda_1 + 3\lambda_2 + 5\lambda_3 - R = -\frac{9}{2}\alpha - \frac{3}{2}\beta - R = 0$$

and

$$\frac{\partial L}{\partial \beta} = \lambda_1 + \lambda_2 + \lambda_3 - 1 = -\frac{3}{2}\alpha - \frac{5}{8}\beta - 1 = 0$$

so that

$$\alpha = \frac{8}{3} - \frac{10}{9}R, \qquad \beta = -8 + \frac{8}{3}R.$$

We now know that

$$\lambda_1 = \frac{7}{3} - \frac{5}{9}R, \quad \lambda_2 = \frac{1}{3}R - 1, \quad \lambda_3 = -\frac{1}{3} + \frac{2}{9}R.$$

Putting these values into the risk, we find that

$$\sigma_\Pi^2 = 4 - \frac{8}{3}R + \frac{5}{9}R^2 = \frac{4}{5} + \frac{5}{9}(R - \frac{12}{5})^2.$$

The minimum risk is thus attained when $R = \frac{12}{5}$ (in which case $\lambda_1 = 1$, $\lambda_2 = -1/5$ and $\lambda_3 = 1/5$ so short selling is required). We also note that if there is to be no short selling (many pension funds do not allow this) then we need all of the λ_i to be greater than zero, which we quickly see requires that

$$3 \leqslant R \leqslant \frac{21}{5}.$$

Chapter 6

Random Walks, Brownian Motion and Stochastic models of security prices

6.1 Introduction

In this chapter we touch on the idea of representing asset prices using a random variable approach. This will form the basis of pricing models of derivatives.

Most equations you have met so far in your courses have been *deterministic*.

Definition (Deterministic system): A deterministic system is one in which there is no random (uncertain) contribution.

An example of a deterministic system is the interest rate model in the first chapter. Given the force of interest, $\delta(t)$ and the initial investment M_0 at time $t = 0$ it is possible to determine unambiguously the amount of money in the bank at any time t in the future by solving the differentail equation.

$$\frac{dM}{dt} = \delta(t), \quad M(0) = M_0$$

For example if $\delta(t)$ is equal to a **constant** δ then $M(t) = e^{\delta t} M_0$. Much of classical applied mathematics is based on deterministic systems. Of course such systems can often be very complicated, and if they are chaotic it is not *practically* possible to determine their state at any time in the future. However in principle, if you run the system from the same initial value with a (theoretical) perfect computer (which has infinite precision), you will always end up with the same answer.

The factors affecting the movement of stock prices on the markets are are so numerous and complex that a deterministic model of their movements is out of the question. A stochastic approach must be taken.

Definition (stochastic system): A *stochastic* system contains a random input in its evolution and hence has an uncertain future.

As an introduction to such systems, consider the following:

Suppose at time t, an asset has a price X. At the next instant of time $t + dt$, the price changes from X to $X + dX$. Let us suppose that there are two influences on this change, a predictable and an unpredictable one. Then we can set

- Predictable, risk free contribution in time $dt = \mu dt$.
- Unpredictable, stochastic contribution in time $dt = \sigma dB_t$

where, μ is related to the **drift** of an asset, σ is related to the **volatility** of an asset. B_t is a random number sampled from a certain normal distribution (see below). Combining these together we get

$$dX = \mu dt + \sigma dB_t,$$

or

- change in X in time dt = deterministic drift + random volatility

To determine dX, after each time step dt random numbers dB_t must be drawn from a suitable distribution. The value of dX appears to wander and carries out a random walk. We cannot predict X at any future time unless we carry out the whole time evolution explicitly for each dt.

The above equation is called a **stochastic differential equation**. The clockwork calculus that solved the deterministic equation for risk-free investment does not work here, since we don't know how to integrate over a set of random numbers: this would be effectively a sum of random numbers, which is itself random!

Moreover, as we have seen in the previous chapter, successive repetitions of the evolution of X which all start at $t = 0$ with the same given $X(0) = X_0$, (say), μ and σ will (with probability 1) generate different values of $X(t)$ in each repetition, since dX is an independent random variable (which does not depend deterministically on time). If $\sigma = 0$ the equation collapses to a deterministic equation of the risk-free investment type.

Note: Only in the deterministic case are we able to make definite statements about the price of assets in the future. In fact most of continuous time finance is based on that premise (e.g., no insider dealing!). What we do try to make statements about is the likelihood of possible outcomes.

This particular type of stochastic evolution is a form *of Brownian motion.* Originally Brownian motion was used to describe the complicated motion of particles undergoing collisions with many neighbouring particles in a gas or liquid. Here the particle is the asset price. This is another example of why "boring" blue-sky research should be funded for the future!

We will meet other types of stochastic models for price evolution later. To introduce the concepts, before discussing the continuous version of asset price dynamics, we consider a discrete model.

6.2 Random walks

In what follows we use the notation $N(\mu, \sigma^2)$ to denote a normal distribution of random numbers with mean μ and variance σ^2 (see handout).

We start by considering a random walk, $B_n(t)$ in discrete time t:

For n a positive integer, let us define the *binomial process* $B_n(t)$ to have

- $B_n(0) = 0$,
- spacing between time steps to be $1/n$,
- up and down jumps to be equal and of size $1/\sqrt{n}$
- probabilities P_{up}, P_{down} of up or down jumps everywhere equal to 1/2.

Thus, if X_i, $i = 1, 2, 3, \ldots$, is a sequence of independent binomial random variables taking the values $+1$ or -1 with equal probability, then the value of B_n at the i^{th} time step of length $1/n$ is

$$B_n\left(\frac{i}{n}\right) = B_n\left(\frac{i-1}{n}\right) + \frac{X_i}{\sqrt{n}}.$$

The time steps can be decreased towards zero (and therefore towards a continuous process) by increasing n. This increases the sampling rate. The corresponding evolution of B_n is shown in figure 6.1 for various increasing values of n.

There appears to be some order in the general pattern of the B_n as n increases: even though the individual values of B_n are different for each n, qualitatively the overall shape of the graph appears similar. This can be rationalised as follows.

Consider the distribution of B_n at time t=1. By induction, the are $n+1$ possible values it can take. These will range from $-\sqrt{n}$ to $+\sqrt{n}$.

Theorem: The mean of the possible values at $t = 1$ will be zero and the variance will be 1.

Proof: Clearly from the definition above,

$$B_n(1) = \sum_{i=0}^{n} \frac{X_i}{\sqrt{n}}$$

where the X_i are independently distributed random variables. Each of the

$$\bar{X}_i = \sum_{j=1}^{2} X_i^{(j)} P_j = +1 \times \tfrac{1}{2} - 1 \times \tfrac{1}{2} = 0,$$

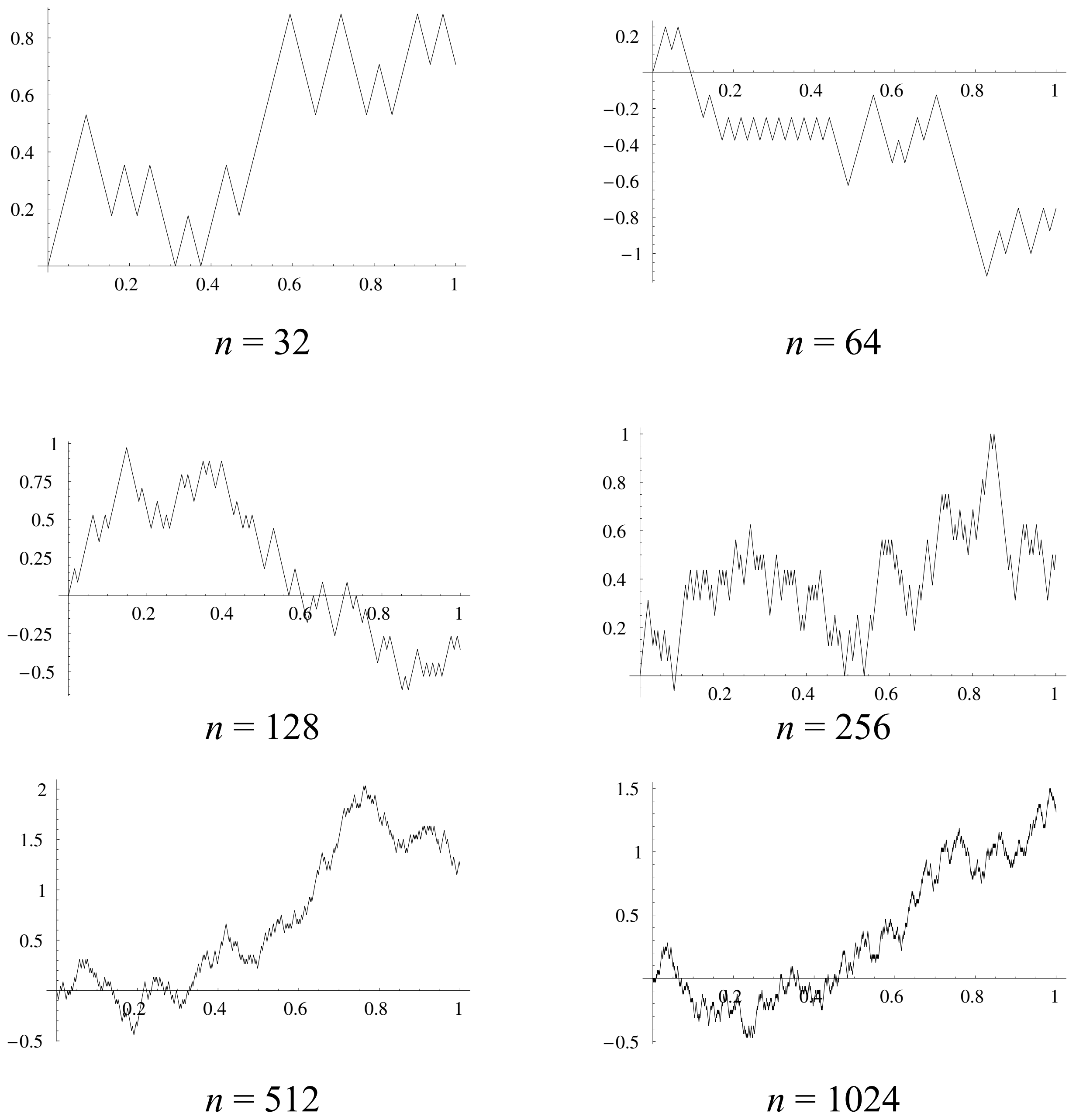

Figure 6.1: Binomial random walks, for increasing numbers of steps n (decreasing discrete time intervals).

and a variance of

$$\sigma_i^2 = \sum_{j=1}^{2} \left(X_i^{(j)} - \bar{X}_i \right)^2 P_j = \left[(1-0)^2 \times \tfrac{1}{2} + (-1-0)^2 \times \tfrac{1}{2} \right] = 1.$$

Thus by the theorem on the mean and variance of sums of n independent random variables (on handout) the mean of the sum of the n independent X_i is zero and the variance of that sum is $n \times 1 = n$. Taking into account the $1/\sqrt{n}$ in the definition of $B_n(1)$ we see that the mean of the distribution of $B_n(1)$ is still 0, but the variance is now 1. QED.

What happens as n becomes very large, i.e., the time-steps ($= 1/n$) become small? By the central limit theorem, these binomial distributions tend towards the unit normal distribution $N(0,1)$ as n becomes large. Thus the distribution of $B_n(1)$ becomes $N(0,1)$ as n becomes large.

Given this distribution for $B_n(1)$, how do we find the distribution for $B_n(t)$? There are t units of 1 in time t. Thus we must take t times the random variables in the sum above and extend the summation to nt. the value of $B_n(t)$ can then be written as

$$B_n(t) = \left(\frac{\sum_{i=1}^{nt} X_i}{\sqrt{n}} \right) = \sqrt{t} \left(\frac{\sum_{i=1}^{nt} X_i}{\sqrt{nt}} \right) = B_n(1)\sqrt{t}$$

Now, the distribution of the ratio inside the second brackets tends, by the central limit theorem, again to the normal $N(0,1)$ as n become large. (Just change variable to $m = nt$ and let m become large, the t is then irrelevant inside the bracket and we are back to the calculation of $B_n(1)$. This was the reason for extracting the $\sqrt{t}$ in the second bracket. Thus we effectively have in the above equation the standard transformation $B_n(t) = B_n(1)\sqrt{t}$, which from the handout on normal distributions converts the $B_n(1)$ from an $N(0,1)$ random variable to a $N(0,t)$ one.

In words this means that the values of an unbiased, up/down random walk will diffuse out at a rate of $\sqrt{t}$ with time, the mean value still being zero.

The conclusion is that in the "limit" of infinitesimal time-steps the discrete tree binomial can be modelled by the appropriate normal distribution. Specifically we have that:

- Each random walk $B_n(t)$has the property that its future movements away from a particular position are independent of where that position is (and by implication its history up to that point).
- Any future displacement $B_n(s+t) - B_n(s)$ is binomially distributed with zero mean and variance t.
- As the sampling rate increases the distribution tends to $N(0,t)$ by the central limit theorem.

6.3 Brownian motion

Mathematically these concepts can be brought together in the continuous time definition of Brownian motion:

Definition (Standard Brownian motion or Wiener Process): The stochastic process $B = (B_t : t \geq 0)$ is a Brownian (or Wiener) process if and only if

- B_t has **independent increments:** the increment $B_t - B_s$ is independent of the history of what the process did up to time s, whenever $s < t$;
- B_t has **stationary increments:** the distribution $B_t - B_s$ depends only on $t - s$;
- B_t has **Gaussian increments:** the distribution $B_t - B_s$ is $N(0, t-s)$;
- B_t is continuous (NB it is not smooth!);
- $B_0 = 0$,

The term "Brownian motion" can also refer to a process W_t which satisfies the first, second and fourth conditions above, but which has a distribution for $W_t - W_s = N(\mu(t-s), (t-s)\sigma^2)$, where μ is termed the **drift coefficient** and σ the **diffusion coefficient**.

In fact either the third or fourth condition can be dropped as it can be shown to be a consequence of the others.

The relationship between standard Brownian motion and Brownian motion is the same as that between $N(0,1)$ and $N(\mu t, \sigma^2 t)$. The two are related by the transformation

$$W_t = W_0 + \sigma B_t + \mu t \sim W_0 + \mu t + \sigma\sqrt{t}N(0,1) \sim W_0 + N(\mu t + \sigma^2 t)$$

Many properties of standard Brownian motion can be demonstrated using the decomposition:

$$B_s = B_t + (B_s - B_t), \qquad s > t,$$

a decomposition in which the first term is known at time t and the second is independent of everything up until time t.

Properties of Brownian/Wiener processes include:

- Cov(B_s, B_t) =min(s,t), since for $s > t$

  $$\text{Cov}(B_s, B_t) = \text{Cov}(B_t + (B_s - B_t), B_t) = \underbrace{\text{Cov}(B_t, B_t)}_{=\text{var}(B_t)=t} + \text{Cov}(B_s - B_t, B_t) = t + 0$$

  since $\text{Cov}(B_s, B_t) = t$ and $\text{Cov}(B_s - B_t, B_t) = 0$.

- B_t is a Markov process, due to the independent increment property.
- B_t is a **martingale**:

  $$E_{\mathbb{Q}}\left[B_s|\mathcal{F}_t\right] = E_{\mathbb{Q}}\left[B_t + (B_s - B_t)|\mathcal{F}_t\right] = E_{\mathbb{Q}}\left[B_t|\mathcal{F}_t\right] + E_{\mathbb{Q}}\left[(B_s - B_t)|\mathcal{F}_t\right] = B_t + 0 = B_t$$

  by definition. We will deal more with martingales later.[1]
- Brownian motion (in $d \leq 2$) will eventually hit every real value, however large or negative. The same is not true for $d > 2$.[2] Once Brownian motion hits a value, it will hit it again infinitely often in the future.
- Standard Brownian motion has the **scaling property** (see above). If

  $$B_1(t) = \frac{1}{c} B_{ct},$$

  $B_1(t)$ is also a standard Brownian motion. This means the sampling rate of Brownian motion is not so important: it "looks"the same on any scale. In fact it is a fractal.
- If $B_2(t)$ is defined by

  $$B_2(t) = tB_{1/t},$$

  then $B_2(t)$ is also a standard Brownian motion. (This is the **time inversion** property of Brownian motion.)
- Although B is continuous everywhere, it is (with probability 1) nowhere differentiable, i.e., the values of W are zig-zagged.
- By implication, in the limit as time becomes continuous, $dB_t \equiv B_{t+dt} - B_t$ is distributed as $N(0, dt)$.

6.4 Representing asset prices

These properties have important consequences for the use of Brownian motion as a model for stock prices.

First suppose that the stock was deterministic with share price S_t at time t. Then in an infinitesimal time step dt the stock price jumps dS_t would obey the equation

$$\frac{dS_t}{S_t} = \mu dt \qquad \Longleftrightarrow \qquad S_t = S_0 e^{\mu t}.$$

[1]**Definition** (Martingales): A process S is a martingale with respect to a measure $\mathbb{P}$ and a filtration $\mathcal{F}_i$ if

$$E_{\mathbb{P}}\left(S_j|\mathcal{F}_i\right) = S_i \qquad \forall\ i \leq j.$$

In words it means the following. S is a martingale, with respect to a set of probabilities $\mathbb{P}$, if the expected value of S at time j in the future, conditional on its history up until any earlier earlier time i, is merely the value of the process at time i. In other words, the expected value in the future is the value now. It means that there is effectively no overall bias or drift up or down in future estimates.

[2]Hence mines floating randomly on the surface of the sea will eventually hit a ship. However mines drifting randomly below the surface of the sea may never hit a submarine.

Clearly stock prices are not deterministic. A model for stock prices could then be **geometric Brownian motion**

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

where $dB \sim N(0, dt)$ and μ is the drift of the asset and σ is the volatility of the asset. What then is the distribution of the random variable S_t? It will be shown below that S_t is drawn from a **lognormal distribution**, or equivalently, $\ln S_t$ obeys a normal distribution:

$$\ln S_t - \ln S_u \sim N\left([\mu - \sigma^2/2][t-u], \sigma^2[t-u]\right), \qquad t \geq u.$$

Alternatively by standard properties of normal distributions this can be written as

$$\ln S_t - \ln S_u \sim \left(\mu - \frac{\sigma^2}{2}\right)(t-u) + \sigma\sqrt{t-u}\ X$$

where

$$X \sim N(0,1).$$

From above is possible to write

$$S_t = S_u e^{\left(\mu - \frac{\sigma^2}{2}\right)(t-u) + \sigma\sqrt{t-u}\ X}.$$

Note that this is a random variable since X is a random variable drawn from a normal distribution. Thus this is not a deterministic prediction of a security price in the future. We must give up any hope of using this formula to predict S_t. It can be viewed as an algorithm for simulating the security price. Each run of the simulation will generate a different possible set of asset prices (see diagram below).

At each time step a random X is drawn from the appropriate distribution and the price of the security is updated. This step is iterated for as long as required and a sequence of random prices S is formed, that is one simulation of the price movement.

If time is wound back and the whole process repeated, the random nature of X at each step means that a different price sequence is obtained.

None of these price histories will necessarily be the actual observed price over future time, (there is a vanishing probability that any will be) but an ensemble of all the simulated price histories will have the expected future mean value and variance as the actual security price (assuming that the underlying present values of the parameters μ and σ are correct).

Therefore we can only hope to obtain statistical/probabilistic information about the future movement of security prices. Hence the more relevant questions to ask is what is the probability distribution of S_t, what is its mean, and what is its variance? We answer that in the next section.

Before we do that we need to fit the model to the experimental data for a particular stock, i.e., estimate values for the parameters μ and σ. This can be done in several ways. One way is a direct approach using the logs of the returns.

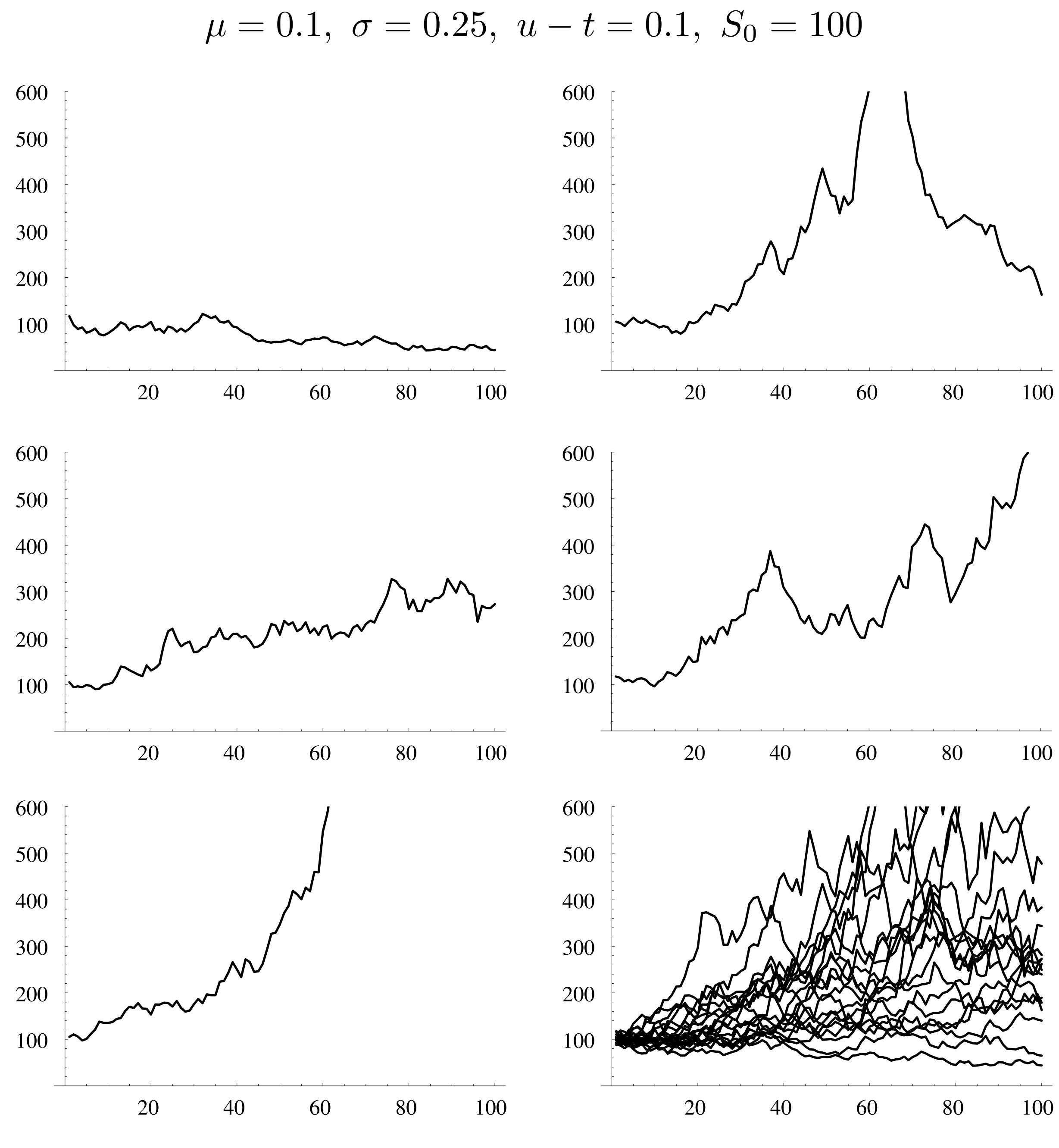

Figure 6.2: Comparison different realisations of the simulation $S_u = S_t \exp(X_{u-t})$ where $X_{u-t} \sim N[\mu(u-t), \sigma^2(u-t)]$ and $\mu = 0.1$, $\sigma = 0.25$, $u - t = 0.1$ and $S_0 = 100$. The bottom right figure shows the superposition of 20 such realisations using the same paramaters.

For example, given a time series of $n+1$ stock prices $S_1, S_2, \cdots S_{n+1}$ at equally placed time intervals dt, then the approximate values of μ and σ^2 can be estimated from

$$R_j = \frac{S_{j+1} - S_j}{S_j} = \left(\frac{S_{j+1}}{S_j} - 1 \right) \approx \ln \left(1 + \left[\frac{S_{j+1}}{S_j} - 1 \right] \right) = \ln \left(\frac{S_{j+1}}{S_j} \right)$$

Hence we have

$$\mu = \frac{1}{ndt} \sum_{j=1}^{n} R_j \approx \frac{1}{ndt} \sum_{j=1}^{n} \ln \left(\frac{S_{j+1}}{S_j} \right),$$

where we have assumed that the jumps are small and have used the Maclaurin expansion for $\ln(1+x)$, $|x| < 1$. A similar calculation gives

$$\sigma^2 \approx \frac{1}{(n-1)dt} \sum_{j=1}^{n} \left\{ \ln \left(\frac{S_{j+1}}{S_j} \right) - \mu \right\}^2 .$$

Of course the question of how large, or small n should be is of crucial importance: too large and you risk including completely irrelevant information, to small and you may neglect important long term trends. This procedure is also based on discrete time sampling. If the interval between sampling times is large, the security prices may fluctuate wildly in-between the sample points. Hence the values given by this approach may not be a true reflection of the underlying behaviour.

Indirect methods are also possible as we may see later on.

On a short timescale, the motion of an arbitrary stock appears to follow a random walk. On a longer timescale there may be some discernible drift, but as the saying goes **past performance is no guarantee of future success**. In other words, just because a stock as been, on average rising over the past six months, it does not mean to say that it will continue to do so. We **cannot** predict stock prices.

Whether or not markets are efficient, the experience of active fund managers shows that it is surprisingly difficult to consistently outperform the market. To reflect this it is often considered good practice to construct models that are consistent with markets being efficient.

6.4.1 Expectation value and variance of lognormal prices

In this section we prove the following results:

$$\begin{aligned} E[S_t|S_u] &= S_u e^{\mu(t-u)}, \qquad t > u \\ \mathrm{Var}[S_t|S_u] &= S_u^2 e^{2\mu(t-u)} \left(e^{\sigma^2(t-u)} - 1 \right) \qquad t > u. \end{aligned}$$

These can be proved as follows.

Consider a random variable $Z \sim N(\alpha, \beta^2)$. We can decompose this as

$$Z \sim \alpha + \beta X, \qquad X \sim N(0,1).$$

What is the mean and variance of the function $e^{-(\alpha+\beta X)}$, given that X is drawn from a standard normal distribution?

First recall that if $X \sim N(0,1)$ then X is distributed according to the pdf

$$p(X)dX = \frac{1}{\sqrt{2\pi}} e^{-X^2/2} dX.$$

We know that the expectation of any (integrable) function $f(X)$ is then

$$E[f(X)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(X) e^{-X^2/2} dX.$$

We also know that

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-Y^2/2} dY = 1.$$

Hence we have the following:

$$\begin{aligned}
E[e^{-(\alpha+\beta X)}] &= e^{-\alpha} E[e^{-\beta X)}] \\
&= \frac{e^{-\alpha}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\beta X} e^{-X^2/2} dX \\
&= \frac{e^{-\alpha}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\beta X - X^2/2} dX \\
&= \frac{e^{-\alpha}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(X+\beta)^2/2} e^{\beta^2/2} dX \qquad \text{(completing the square)} \\
&= \frac{e^{-\alpha+\beta^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-Y^2/2} dY \qquad \text{(changing variables to } Y = X + \beta) \\
\Rightarrow E[e^{-(\alpha+\beta X)}] &= e^{-\alpha+\beta^2/2},
\end{aligned}$$

Note that it's **minus** α, but **plus** $\beta^2/2$, regardless of the sign of β. This is a very important result and you should know its derivation.

The variance of $e^{-(\alpha+\beta X)}$ can also be derived as follows:

$$\begin{aligned}
\mathrm{Var}[e^{-(\alpha+\beta X)}] &= E[\left(e^{-(\alpha+\beta X)} - E[e^{-(\alpha+\beta X)}]\right)^2] \\
&= E[e^{-2(\alpha+\beta X)} + E[e^{-(\alpha+\beta X)}]^2 - 2e^{-2(\alpha+\beta X)} E[e^{-(\alpha+\beta X)}]] \\
&= E[e^{-2(\alpha+\beta X)}] - E[e^{-(\alpha+\beta X)}]^2 \\
&= E[e^{-2(\alpha+\beta X)}] - e^{-2(\alpha+\beta^2/2)} \\
&= e^{-2\alpha+(2\beta)^2/2)} - e^{-2(\alpha+\beta^2/2)} \\
\Rightarrow \mathrm{Var}[e^{-(\alpha+\beta X)}] &= e^{-2\alpha+\beta^2} \left(e^{\beta^2} - 1\right).
\end{aligned}$$

This is another very important result and you should also know its derivation.

Now, suppose that the share price S_t does indeed obey the following lognormal distribution

$$\ln S_t - \ln S_u \sim \left(\mu - \sigma^2/2\right)(t-u) + \sigma\sqrt{t-u}\; X$$

where

$$X \sim N(0,1).$$

Then the share price S_t behaves as

$$S_t = S_u e^{\left(\mu-\sigma^2/2\right)(t-u)+\sigma\sqrt{t-u}\; X}.$$

Then by comparison with the above form $e^{-(\alpha+\beta X)}$, we have

$$S_t = S_u e^{-(\alpha+\beta X)},$$

where (mindful of the minus signs),

$$\alpha = -\left(\mu - \sigma^2/2\right)(t-u), \qquad \beta = -\sigma\sqrt{t-u}.$$

Hence using the above formulae for expectation (and using the fact that S_u is a given constant), we have

$$\begin{aligned}
E[S_t|S_u] &= E[S_u e^{-(\alpha+\beta X)}] \\
&= S_u E[e^{-(\alpha+\beta X)}] \\
&= S_u e^{(\mu-\sigma^2/2)(t-u)+\sigma^2(t-u)/2} \\
\Rightarrow E[S_t|S_u] &= S_u e^{\mu(t-u)}.
\end{aligned}$$

Using the above formula for variance, we similarly have

$$\begin{aligned}
\mathrm{Var}[S_t] &= \mathrm{Var}[S_u e^{-(\alpha+\beta X)}] \\
&= S_u^2 \mathrm{Var}[e^{-(\alpha+\beta X)}] \\
&= S_u^2 e^{-2\alpha+\beta^2}\left(e^{\beta^2} - 1\right) \\
&= S_u^2 e^{2(\mu-\sigma^2/2)(t-u)+\sigma^2(t-u)}\left(e^{\sigma^2(t-u)} - 1\right) \\
\Rightarrow \mathrm{Var}[S_t] &= S_u^2 e^{2\mu(t-u)}\left(e^{\sigma^2(t-u)} - 1\right).
\end{aligned}$$

Hence the peak of the probability distribution function "spreads out" around the mean as time $t-u$ into the future increases. This makes logical sense: the further we try to go into the future, the more chance the asset value has to take up a wider range of possible values.

The figures below show a comparison of the pdfs for the normal and lognormal distributions. The lognormal is a skewed bell-shaped distribution.

6.4.2 Examples of use of lognormal distribution

Probability of a fall in share price

Suppose a (non banking) share is thought to obey the lognormal distribution with an annual drift $\mu = 0.05$, (5%)and an annual volatility $\sigma = 0.2$ (20%). What is the probability that the share price will be less than the value now in six month's time?

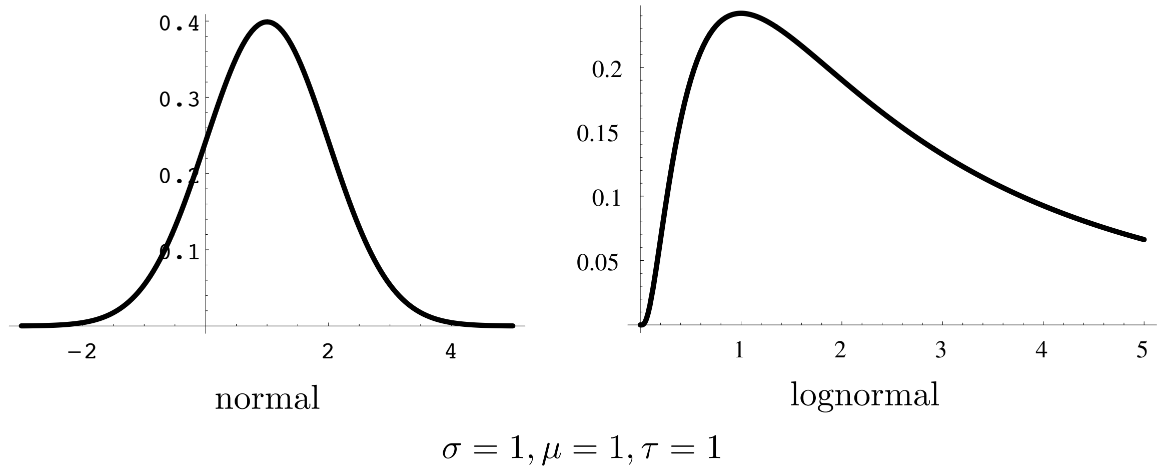

Figure 6.3: Comparison of normal and lognormal distributions for identical values of σ, μ, τ.

We use the lognormal distribution

$$\ln S_t \sim \ln S_u + (\mu - \sigma^2/2)(t-u) + \sigma\sqrt{t-u}\, X, \qquad X \sim N(0,1),$$

and set $u = 0$ (now) to simplify the expression to

$$\ln S_t - \ln S_0 \sim (\mu - \sigma^2/2)(t) + \sigma\sqrt{t}X.$$

We use the monotonic properties of ln to observe that

$$S_t \leq S_0 \qquad \Longleftrightarrow \qquad \ln S_t \leq \ln S_0.$$

Hence the probability we want is:

$$P(S_t \leq S_u) = P(\ln S_t \leq \ln S_0) = P(\ln S_t - \ln S_0 \leq 0).$$

Substituting for X we can write the probabilities in terms of the (easier to handle) normal $N(0,1)$ distribution that X is drawn from:

$$\begin{aligned} P\left(\ln S_t - \ln S_0 \leq 0\right) &= P\left((\mu - \sigma^2/2)(t-0) + \sigma\sqrt{t-0}\, X \leq 0\right) \\ &= P\left(X \leq \frac{(\sigma^2/2 - \mu)\sqrt{t}}{\sigma}\right) \\ &= \Phi\left(\frac{(\sigma^2/2 - \mu)\sqrt{t}}{\sigma}\right), \end{aligned}$$

where Φ is the cdf of the normal $N(0,1)$ distribution. Inserting the value of $\mu = 0.05$, $\sigma = 0.2$ and $t = 1/2$ (6 months) we have

$$P(S_{1/2} \leq S_0) = \Phi(-0.106\ldots) = 0.457765\ldots.$$

Note that the probability of a fall is not 1/2: this is because there is an expected drift of $\mu = 0.05$ (and the lognormal is not a symmetric distribution). Note that we cannot say for sure whether the share price will rise or fall it may still do both. All we could say is that if we could re-run history an infinite amount of times then in $0.457765\ldots$ of the times the asset price would be the same or less at $t = 1/2$.

Calculation of expected share price

The expected value of the asset price in six months time is given from the formula derived in the previous section:

$$S_t = S_0 e^{\mu t} \qquad \Rightarrow \qquad S_{1/2} = S_0 e^{0.05\times 1/2} = 1.02532\ldots S_0.$$

Note that this is not the actual value that we expect the share to be at in 6 months time! Equally past performance is no guarantee of future success and if we use a moving average for μ based on a window of asset prices, μ will obviously change with time and the situation is a lot more complicated.

Note that the variance of the asset price about the expected value will increases exponentially with time t into the future:

$$\mathrm{var}\,(S_t|S_0) = S_0^2 e^{2\mu t}\left(e^{\sigma^2 t} - 1\right)$$

This is sensible: it seems likely that asset price will be closer to the predicted expected value tomorrow than it would be six months into the future, where the asset price has had more time to explore all its possible values.

Calculation of probability of percentage rise

Suppose that we now ask for the probability that the asset price will rise by at least $x\%$ in time t. We require:

$$P\left(S_t > \left(1 + \frac{x}{100}\right) S_0\right).$$

Now consider the following:

$$\ln S_t - \ln S_0 = \ln\left[\left(1 + \frac{x}{100}\right) S_0\right] - \ln S_0 = \ln\left(1 + \frac{x}{100}\right) \equiv R, \text{ say}$$

Hence, by the monotonicity of ln and using

$$\ln S_t - \ln S_0 \sim (\mu - \sigma^2/2)t + \sigma\sqrt{t}X,$$

we can recast the problem in terms the more easily handled normally distributed standard $X \sim N(0,1)$ as:

$$\begin{aligned}
P\left(S_t > \left(1 + \frac{x}{100}\right) S_0\right) &= P(\ln S_t - \ln S_0 > R) \\
&= P\left((\mu - \sigma^2/2)(t) + \sigma\sqrt{t}X > R\right) \\
&= P\left(X > \frac{R - (\mu - \sigma^2/2)t}{\sigma\sqrt{t}}\right) \\
&= 1 - P\left(X \leq \frac{R - (\mu - \sigma^2/2)t}{\sigma\sqrt{t}}\right) \\
&= 1 - \Phi\left(\frac{R - (\mu - \sigma^2/2)t}{\sigma\sqrt{t}}\right)
\end{aligned}$$

For example with $\mu = 0.05$, $\sigma = 0.2$, $t = 1/2$ and $x = 4$, so that $R = 1.04$ we have $X < 0.171266\ldots$. The probability of a 4% rise in 6 months would then be $1 - \Phi(0.171266\ldots) = 0.432007\ldots$.

6.4.3 Discussion of validity of lognormal distribution

Empirical results for testing the log random walk model are mixed. As the model incorporates independent returns over disjoint intervals, it is impossible to use past history to deduce that prices are cheap or dear at any time. This implies weak form market efficiency, and is consistent with empirical observations that technical analysis does not lead to excess performance.

However, more detailed analysis reveals several weaknesses in the log random walk model. The most obvious is that estimates of σ vary widely according to what time period is considered, and how frequently the samples are taken.

Whilst the non-normal distribution clearly provides an improved description of the returns observed, in particular more extreme events are observed than is the case with the log- normal model. The improved fit to empirical data comes at the cost of losing the tractability of working with normal (and log-normal) distributions.

The plot below shows the observed changes in the FTSE All Share index against those which would be expected if the returns were lognormally distributed. The substantial difference demonstrates that the actual returns have many more extreme events, both on the upside and downside, than is consistent with the lognormal model. This is often called the problem of "fat tails" (in the probability distribution). Super- imposed is a simulation where the continuous return comes from a quintic polynomial distribution whose parameters have been chosen to give the best fit to the data.

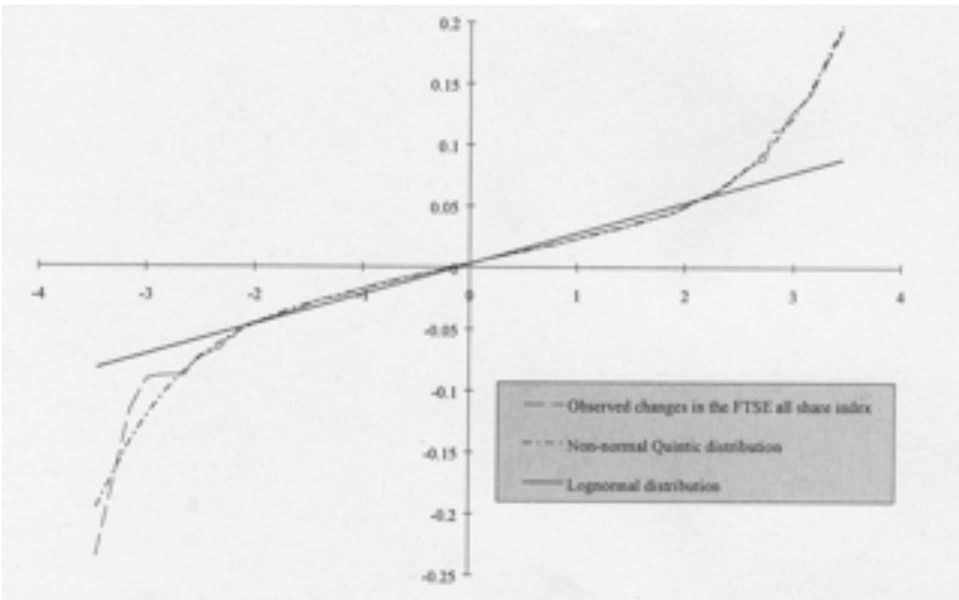

Figure 6.4: Comparison of actual share prices vs theoretical lognormal distribution.

We can also take some evidence from option prices. The Black-Scholes formula (which we discuss later) describes option prices in terms of anticipated values of volatility over the term of the option. Given observed option prices in the market, it is possible to work backwards to the implied volatility, that is, the value of σ which is consistent with observed option prices (we return to this topic later). Examination of historic option prices suggests that volatility expectations fluctuate markedly over time. Hence the probability distribution will also fluctuate over time.

One way of modelling this behaviour is to take volatility as a process in its own right. This can explain why we have periods of high volatility and periods of low volatility. One class of models with this feature is known as **ARCH** - standing for **autoregressive conditional heteroscedasticity.**

A more contentious area relates to whether the drift parameter μ is constant over time. There are good theoretical reasons to suppose that μ should vary over time. It is reasonable to suppose that investors will require a risk premium on equities relative to bonds. As a result, if interest rates are high, we might expect the equity drift, μ, to be high as well.

One unsettled empirical question is whether markets are **mean reverting**, or not. A mean reverting market is one where rises are more likely following a market fall, and falls are more likely following a rise. There appears to be some evidence for this, but the evidence rests heavily on the aftermath of a small number of dramatic crashes. Furthermore, there also appears to be some evidence of momentum effects, which imply that a rise one day is more likely to be followed by another rise the next day.

A further strand of empirical research questions the use of the normality assumptions in market returns. In particular, market crashes appear more often than one would expect from a normal distribution. While the random walk produces continuous price paths, jumps or discontinuities seem to be an important feature of real markets. Furthermore, days with no change, or very small change, also happen more often than the normal distribution suggests.

Fractals are also used to model share prices. Random walks have a fractal dimension of 1/2. Empirical investigations of market returns often reveal a fractal dimension around 1.4.

It is important to appreciate that many of the empirical deviations from the random walk do not imply market inefficiency. For example, periods of high and low volatility could easily arise if new information sometimes arrived in large measure and sometimes in small. Market jumps are consistent with the arrival of information in packets rather than continuously. Even mean reversion can be consistent with efficient markets. After a crash, many investors may have lost a significant proportion of their total wealth; it is not irrational for them to be more averse to the risk of losing what remains. As a result, the prospective equity risk premium could be expected to rise.

Many orthodox statistical tests are based around assumptions of normal distributions. Once we reject normality, we also have to re-test various hypotheses. In particular, the evidence for time-varying mean and volatility is greatly weakened. These apparent effects would be expected to arise as artefacts of a fractal process.

We can distinguish two ways of looking at the properties of time series models. For a given quantity (for example, the force of inflation) we can imagine a two-dimensional table in which each row is one simulation and each column corresponds to a future projection date.

All the simulations start from the same starting position, that is determined by reference to market conditions on the run date.

A **cross-sectional property** fixes a time horizon and looks at the distribution over all the simulations. For example, we might consider the distribution of inflation next year. Implicitly, this is a distribution conditional on the past information which is built into the initial conditions, and is, of course, common to all simulations. If those initial conditions change, then the implied cross-sectional distribution will also change. As a result, cross-sectional properties are difficult to validate from past data, since each year of past history typically started from a different set of conditions. However, the prices of derivatives today should reflect market views of a cross-sectional distribution. Cross-sectional information can therefore sometimes be deduced from the market prices of options and other derivatives.

A **longitudinal property** picks one simulation and looks at a statistic sampled repeatedly from that simulation over a long period of time. For example, we might consider one simulation and fit a distribution to the sampled rates of inflation projected for the next 1000 years. For some models, this longitudinal distribution will converge to some limiting distribution as the time horizon lengthens; furthermore this limiting distribution is common to all simulations. Such convergence results are sometimes called ergodic theorems; the resulting distribution is an **ergodic** distribution. Unlike cross-sectional properties, longitudinal properties do not reflect market conditions at a particular date but, rather, an average over all likely future economic conditions.

Most statistical properties computed from historical data are effectively cross-sectional properties.

In a pure random walk environment, asset returns are independent across years and also (as for any model) across simulations. As a result, cross-sectional and longitudinal quantities coincide. For example, it is common to see historic volatility used as an input to option pricing models. The historic volatility estimates a longitudinal standard deviation, while option pricing requires cross-sectional inputs. To equate the two is valid in a random walk setting, but not for more general models which we now consider.

6.5 Auto-regressive models

6.5.1 Mean reversion

A random walk process can be expected to grow arbitrarily large with time. If share prices follow a random walk, with positive drift, then those share prices would be expected to tend to infinity for large time horizons.

However, there are many quantities which should not behave like this. For example, we do not expect interest rates to jump off to infinity, or to collapse to zero. Instead, we would expect some mean reverting force to pull interest rates back to some normal range. In the same way, while dividend yields can change substantially over time, we would expect them, over the long run, to form some stationary distribution, and not run off to infinity. Similar considerations apply to the annual rate of growth in prices or in dividends. In each case, these quantities are not independent from one year to the next; times of high interest rates or high inflation tend to bunch together, i.e. the models are auto-regressive.

One method of modelling this is to consider a vector of **mean reverting processes**. These processes might include (log) yields, or the instantaneous growth rate of income streams. The reason for the log transformation is to prevent negative yields.

The appropriate parameters for mean reverting processes would normally be calibrated using time series analysis of historic data. For convenience, the model quantities may be arranged into a pyramid structure, in which the elements can be calibrated one at a time, avoiding some of the complexities of full multivariate analysis.

In this framework, capital values are computed as the income level on an asset divided by its yield. This structure, with some modifications, forms the basis of the Wilkie model of the UK and elsewhere, the Finnish actuarial model, Thompson's South African model, Tower's Perrin's CAP:Link model of the US and other economies, the Falcon asset model, and various other models world-wide.

6.5.2 The Wilkie Model

The Wilkie model aims to explain the (linked) behaviour of

- $I(t)$, the force of inflation during year t;
- $Y(t)$, the equity dividend yield at the end of year t;
- $\ln K(t)$, the log of the force of dividend growth during year t;
- $\ln R(t)$, the log of the force real yield on perpetual index linked bonds at the end of year t.

It does so by sum of matrices (an "ARMA model" in technical jargon) whose entries include moving averages of the actual inflation, actual dividends, long term interest rates etc. a string of normally distributed random variables and a host of rather strange looking numerical constants derived from historic UK data.

It is incredibly complicated and the masochistic among you may find full details of the Wilkie model in his paper

"More on a Stochastic Asset Model for Actuarial Use", BAJ Volume 1, pages 777 - 964.

6.6 The main alternative models

The models described so far are all statistically based models. The model structure is derived from past time series, together with some intuition regarding what model formulae look reasonable.

However, these statistical models can produce some odd results. It can be useful to impose additional economic constraints on model behaviour. For example, one simulation could possibly produce two assets as follows:

Asset	Prospective Income	Expected income growth p.a.	Market value
A	\$60	2%	\$2,000
B	\$60	3%	\$1,500

If the income stream from asset B is expected to grow faster than the income stream from asset A, we would think asset B should trade at a higher price, that is, a lower initial yield. Otherwise, investors would all favour asset B, and the price of B would rise relative to A until equilibrium held. The end result might be as follows:

Asset	Prospective Income	Expected income growth p.a.	Market value
A	\$60	2%	\$1,500
B	\$60	3%	\$2,000

In an efficient market, we might expect differential growth rates to be equal to the yield differential, perhaps plus or minus some premium for risk when comparing asset classes with different risk characteristics.

We can calibrate mean reverting models such that these efficient market constraints hold approximately. However, the imposition of a particular algebraic form means that we cannot model a fully efficient market, except in trivial special cases (for example, when yields are fixed and not stochastic). To model an efficient market, models for log yields would need to be slightly non-linear. However, the required non-linearity is modest, and it is difficult to test empirically whether the non-linearity is present, or not. In particular, the fact that the Wilkie model, fitted to a particular data set, is a model of an inefficient market, does not rule out the possibility that an efficient market model could just as easily be fitted to the same data.

The decision to adopt an efficient or inefficient market model has important implications for model output. For example, statistical models often produce yields which fluctuate more than would be the case if markets were efficient. Such models steer investors into tactical switches, from low yielding assets to high yielding ones - a strategy which, on the basis of past data, would have been profitable in several countries including the UK. Recommended asset allocations will thus fluctuate substantially between different run dates, according to market conditions at the time. In practice, this problem is usually mitigated by re-calibrating the model to give a predetermined answer each time it is run, while ignoring the impact of future re-calibrations on the dynamics of the economy. For example, in the context of UK pension schemes, a statistical model might be calibrated to suggest 75% equities, while a comparable scheme in the US or the Netherlands would be 60% or 40% respectively. These figures are derived in order to reproduce the average scheme asset allocation in each country. This calibration technique is a crude example of an equilibrium approach.

An alternative to purely statistical models is to give more weight to economic theory. If we are convinced by the theoretical or empirical arguments for market efficiency or purchasing power parity, we may wish to use these theories to guide the construction of stochastic investment models. For example, market efficiency considerations would force us to identify two factors in dividend growth - one component which strongly mean reverts in the short term, and a longer term component which mean reverts more slowly. Statistical analysis of historic dividend growth enables us to be reasonably precise about the short term effect, but this effect alone is insufficient to explain observed yield volatility. The long term effect drives longer term expectations, and hence yields, but is swamped by the short term effect if dividend statistics are considered in isolation. In other words, we can use the volatility of observed yields, together with notions of market efficiency, to infer the long term behaviour of income indices. In this way, we can achieve a model which is statistically adequate, but which can also be rationalised in an efficient market framework.

Any stochastic investment model, combined with a suitable business model, can provide simulated distributions of a wide range of future cash flows and capital items. These distributions can be computed for a range of different strategies. Some strategies will turn out to be high risk / high return and others low risk / low return. Depending on the precise definitions of risk and return, different strategies will appear optimal. This causes a problem in using the models to support decisions. Although asset projection is well developed, interpreting the output remains more of an art than a science.

The advantage of using more economic theory is that it gives us a more concrete way of interpreting model output. For example, if we model a market which is broadly governed by rational pricing rules, we can apply those same pricing rules to simulated output from a model. This gives us a market-based way of comparing strategies, and deciding which strategy is most valuable. The difficulty with this approach is that the model's optimal strategy may not be the strategy that managers wish to follow. In this context, a more flexible judgmental approach may better meet the clients needs.

6.7 Estimating parameters for asset pricing models

The estimation of parameters is one of the most time-consuming aspects of stochastic asset modelling.

The simplest case is the purely statistical model, where parameters are calibrated entirely to past time series. Provided the data is available, and reasonably accurate, the calibration can be a straightforward and mechanical process (see above).

Of course, there may not always be as much data as we would like, and the statistical error in estimating parameters may be substantial. Furthermore, there is a difficulty in interpreting data which appears to invalidate the model being fitted. For example, what should be done when fitting a gaussian model in the presence of large outliers in the data? Perhaps the obvious course of action is to reject the hypothesis of normality, and to continue building the model under some alternative hypothesis. After all, in many applications, the major financial risks lie in the outliers, so it seems foolish to ignore them. In practice, a more common approach to outliers is to exclude them from the statistical analysis, and focus attention instead on the remaining residuals which appear more normal. The model standard deviation may be subjectively nudged upwards after the fitting process, in order to give some recognition to the outliers which have been excluded.

It is common practice in actuarial modelling to use the same data set to specify the model structure, to fit the parameters, and to validate the model choice. A large number of possible model structures are tested, and testing stops when a model which is found which passes a suitable array of tests. Unfortunately, in this framework, we may not be justified in accepting a model simply because it passes the tests. Many of these tests (for example, tests of stationarity) have notoriously low power, and therefore may not reject incorrect models. Indeed, even if the ”true” model was not in the class of models being fitted, we would still end up with an apparently acceptable fit, because the rules say we keep generalising until we find one. This process of generalisation tends to lead to models which wrap themselves around the data, resulting in an understatement of future risk, and optimism regarding the accuracy of out-of-sample forecasts. For example, Huber recently compared the out-of-sample forecasts of the Wilkie model to a naive ”same as last time” forecast over a 10 year period. The naive forecasts proved more accurate.

In the context of economic models, the calibration becomes more complex. The objective of such models is to simplify reality by imposing certain stylised facts about how markets would behave in an ideal world. This theory may impose constraints, for example on the relative volatilities of bonds and currencies. Observed data may not fit these constraints perfectly. In these cases, it is important to prioritise what features of the economy are most important to calibrate accurately for a particular application.

Notwithstanding these comments in the previous chapter, we shall henceforth use the log-normal distribution as the main model of asset price movement.

Chapter 7

Brownian motion and stochastic calculus

7.1 Introduction

In the previous chapter we touched on the idea of representing asset prices using a random variable approach. This will form the basis of pricing models of derivatives. However before we attempt this we introduce here some more essential mathematical and statistical concepts. Some of you will have met similar techniques in the stochastic calculus course, others will not have.

7.2 Stochastic Calculus

7.2.1 Non-differentiable differentials

Consider the Brownian process X, given by

$$X_t = \sigma B_t + \mu t.$$

where μ and σ are (for the time being only) constant drift and volatility of the process.

We can zoom in on the local behaviour of X at a time t, by taking small and smaller time-steps. However, from above we have been told that Brownian motion is fractal, i.e., it appears qualitatively (and qualitatively) the same at whatever magnification we choose. Unlike a deterministic smooth Newtonian curve, where we gradually see the curve becoming linear under magnification, the stochastic process always has zig-zags.

A stochastic differential therefore has both a deterministic and a random increment. If the time-step we take is dt the above process has an infinitesimal change of

$$dX_t = \sigma_t dB_t + \mu_t dt.$$

Notation: In the limit of continuous time we tend to write dB_t as the infinitesimal, but random jump, $B(t+dt) - B(t)$ as $dt \to 0$. In line with the definitions above, it is therefore distributed as $N(0, dt)$.

As above, we interpret the stochastic process as a rule that allows us to predict the statistical distribution of the dX_t, rather than to predict the individual values.

Hence we have:

$$E[dX_t] = E[\sigma_t dB_t + \mu_t dt] = \sigma_t E[dB_t] + E[\mu_t dt] = \sigma_t \times 0 + \mu dt = \mu dt$$

and so since $dB_t \sim N(0, dt)$, from the property of adding random (and deterministic) vairables (see probability handout) we have:

$$dX_t \sim N(\mu dt, \sigma^2 dt).$$

In other words dX_t is distributed normally with mean μdt and variance $\sigma^2 dt$.

We now turn our attention to the "solution" of such SDEs. Since the equation is stochastic, we cannot find finds a unique solution, in the sense that we would look for in deterministic differential equations. By "solution" we mean the prediction of the process that X_t follows, and hence its statistical distribution.

7.2.2 Integral form of a stochastic process

Definition (stochastic differential equation, SDE): When the volatility σ and drift μ depend on the random component B only through X_t, such as $\sigma_t = \sigma(X_t, t)$, $\mu_t = \mu(X_t, t)$, where $\sigma(x, t)$, $\mu(x, t)$ are deterministic functions, the equation

$$dX_t = \sigma(X_t, t)\, dB_t + \mu(X_t, t)\, dt$$

is called a **stochastic differential equation (SDE)**.

This can be integrated formally to give a **stochastic process**:

$$X_t = X_0 + \int_0^t \sigma_s dB_s + \int_0^t \mu_s ds.$$

where B_s are independent samples from a random distribution. (Note that in the first integrand we write B_s rather than B_t since B_s plays the role of a dummy variable, just as we have written ds in the second: the final answer is, in fact, a function of t due to the integration limits.)

The parameters σ and μ are random processes which depend on the history of X such that

$$\int_0^t \left(\sigma_s^2 + |\mu_s|\right) ds < \infty$$

for all times t (with probability 1).

Just what we mean by an integral over random numbers will be explained below. Note that a stochastic integral is a much more complicated beast than a deterministic one. It is the limit of the sum over a set of continuous random variables. (Note that this is different from an integral over a pdf, which is a deterministic integral!).

In general it is far easier to write down a SDE for a particular process than to solve it explicitly. As with deterministic differential equations, solutions may not exist.

Example: Let σ, μ be constants and the SDE for X be

$$dX_t = \sigma dB_t + \mu dt$$

Given an initial condition $X_0 = 0$, the solution is simply

$$X_t = \sigma B_t + \mu t.$$

Here we have assumed, as in deterministic calculus, that σdB_t is the total derivative of σB_t. However if we now consider the SDE

$$dX_t = X_t \left(\sigma dB_t + \mu dt \right),$$

how do we solve this?

7.2.3 How not to integrate stochastic integrals

Health Warning: The following section may upset some readers firmly rooted in the concepts of deterministic calculus you learned at school. Read this section at your own peril. If you are worried that you might be of a stochastically nervous disposition, skip to the next section. **You have been warned!**

Ordinary deterministic calculus has the chain rule, product rule, integration by parts, integration by substitution etc. How do we evaluate integrals with respect to random numbers?

Consider the example of a function f that depends on some underlying Brownian motion $f(B_t) = B_t^2$.

Under the rules of ordinary calculus we have that

$$f(B_t) = B_t^2 \quad \Rightarrow \quad df_t = d\left(B_t^2\right) = 2B_t dB_t.$$

Is this still the result when B_t is a Brownian motion?

The short, but non-intuitive answer is **NO!**

To see this we check the integration procedure: a deterministic calculus approach suggests we should have (with $B_0 = 0$)

$$d\left(B_t^2\right) = 2B_t dt \qquad \Rightarrow \int_0^t d\left(B_s^2\right) = 2\int_0^t B_s dB_s \qquad \Rightarrow B_t^2 = 2\int_0^t B_s dB_s.$$

But now examine what we mean by the integral over the Brownian B. First divide up the time interval into discrete time-steps

$$\left\{0, \frac{t}{n}, \frac{2t}{n}, \frac{3t}{n}, \ldots, \frac{(n-1)\,t}{n}, t\right\}$$

for some (large) integer n. We can then approximate the integral with a sum by follows:

$$2\int_0^t B_s dB_s = \lim_{n\to\infty}\left\{2\sum_{k=0}^{n-1} \underbrace{B\left(\frac{kt}{n}\right)}_{\sim N(0,kt/n)} \underbrace{\left[B\left(\frac{(k+1)t}{n}\right) - B\left(\frac{kt}{n}\right)\right]}_{\sim N(0,t/n)}\right\}.$$

A quick look at the terms in the sum shows that since the means of B and the increment in B in the square brackets are both zero, then the sum is in fact proportional to the covariance[1]

$$\operatorname{Cov}\left(B\left(\frac{(k+1)t}{n}\right) - B\left(\frac{kt}{n}\right), B\left(\frac{kt}{n}\right)\right)$$

But due to the **independent increments** property of Brownian motion this covariance is zero. Hence the expected value, or mean, of the sum is zero. Taking the limit as $n \to \infty$, the integral also then has a zero mean.[2]

Now, if this integral is,as suggested, equal to B_t^2, then B_t^2 should have zero mean as well.

Unfortunately, since B_t is distributed as the normal distribution $N(0,t)$, with mean 0 and variance $\sigma_t^2 = t$, we have

$$t = \operatorname{Var}[B_t] = E\left[(B_t - E[B_t])^2\right] = E\left[(B_t - 0)^2\right] = E\left[(B_t)^2\right]$$

Hence, for $t > 0$, the mean value of B_t^2 is t and so is most certainly not zero.

Thus the LHS of the proposed identity

$$B_t^2 = 2\int_0^t B_s dB_s$$

[1]This can be explained as follows. If there are two random variables X_k and Y_k for which both $E[X_k] = E[Y_k] = 0$, then $\sum_k X_k Y_k = \sum_k (X_k - 0)(Y_k - 0) = \operatorname{cov}(X_k, Y_k)$. We see that this applies to the above result for Brownian motion since $E[B(kt/n)] = E[B((k+1)t/n)] = 0$ (hence $E[B((k+1)t/n) - B(kt/n)] = 0$).

[2]Remember we are integrating functions of random variables and so the integral will be a function of a random variable, and hence a random variable! Consequently we have to discuss the statistical properties of the integral.

has a mean of t, but the RHS has a mean of 0. The LHS and RHS of a stochastic identity cannot have different means and still be an identity. Hence we find that

$$B_t^2 \neq 2\int_0^t B_s dB_s,$$

and the assumptions that we made above about how to integrate, drawn from deterministic calculus, cannot be correct.

7.2.4 What went wrong?

The resolution of this problem comes when we examine the assumptions we make when we perform an ordinary deterministic integration. The Taylor expansion of a deterministic, smooth function $f(x)$ has the elemental infinitesimal form:

$$df(x) = f'(x)dx + \frac{1}{2!}f''(x)(dx)^2 + \frac{1}{3!}f''(x)(dx)^3 + \cdots.$$

This gives us the infinitesimal change df in the function $f(x)$ when we change the deterministic variable x by dx.

We can also use it to integrate df. We assume that since dx is infinitesimally small then $(dx)^2$ and higher order terms are effectively so small so as to be negligible with respect to dx. Hence, in the limit as $dx \to 0$, the integral becomes

$$f - f(0) \equiv \int_{f_0}^{f} df = \int_0^x \frac{df}{dx}dx + \text{negligible terms} \to \int_0^x \frac{df}{dx}dx.$$

If we put $f(x) = x^2$ we get the formula we used in the section above.

However, with Brownian motion, the higher order terms $(dB_t)^2$ **are not negligible**. To see this consider the following. Use the same discrete time steps as above but now use the approximation to the integral as:

$$\int_0^t (dB_t)^2 = \lim_{n\to\infty}\sum_{k=1}^{n}\left[B\left(\frac{kt}{n}\right) - B\left(\frac{(k-1)t}{n}\right)\right]^2.$$

This is just like reading the integral as it stands and just substituting

$$dB_t \approx B\left(\frac{tk}{n}\right) - B\left(\frac{t(k-1)}{n}\right).$$

Now define,

$$Z_{n,k} = \frac{B(tk/n) - B(t(k-1)/n)}{\sqrt{t/n}}.$$

For fixed n, the $Z_{n,1}, Z_{n,2}, Z_{n,3}, \ldots$, is a set of independent normally distributed random variables selected from $N(0,1)$. This is because the increments $B(tk/n) - B(t(k-1)/n)$

are $N(0, t/n)$ from the third property of Brownian motion above, and from section 7.3 above the scaling by $\sqrt{t/n}$ alters this to $N(0,1)$. Thus we have

$$\int_0^t (dB_t)^2 = \lim_{n\to\infty} t \sum_{k=1}^{n} \frac{Z_{n,k}^2}{n}.$$

As n tends to infinity, and time becomes continuous, from the law of large numbers, the RHS sum converges to have the following mean,

$$E\left[\sum_{k=1}^{n} \frac{Z_{n,k}^2}{n}\right] \to \sum_{k=1}^{n} E\left[\frac{Z_{n,k}^2}{n}\right] = \sum_{k=1}^{n} \frac{E\left[Z_{n,k}^2\right]}{n} = \sum_{k=1}^{n} \frac{1}{n} = \frac{n}{n} = 1.$$

Hence the RHS has a mean value of $t \times 1 = t$. This agrees with the predicted value of the LHS, so we have the following statistical equality

$$E\left[\int_0^t (dB_t)^2\right] = t.$$

or in differential form,

$$E\left[(dB_t)^2\right] = dt.$$

In other words, we **cannot** neglect terms of order $(dB_t)^2$ in the Taylor expansion of df above. That we intuitively did so was because of the notation: they were negligible because it was apparently a square of a small number. The stochastic nature of the differential means that such terms are not negligible if the result of the integration is supposed to have the correct mean.

Note that the above identity, in words, means that in the limit as $dt \to 0$ the sum (= integral) of the squares of random numbers drawn from $N(0, dt)$ converges with probability 1 to have a mean of t. Thus the output from a stochastic integral is really a set of functions involving means of probability distributions.

For ease of simplification in calculations, we often just assume expectation values and just use the shortcut replacement

$$(dB_t)^2 = dt.$$

Notation: One sometimes sees $\int \sqrt{dt}$ in stochastic situations, which looks stupid from the deterministic point-of-view. From above we simply interpret this as $\int dB_t$.

A simple way to rationalise the fact that the size of the increment in the Brownian motion dB_t in time dt is about $\sqrt{dt}$ is to return to the binomial definition of the motion: There the increments (dB_t) were of size $1/\sqrt{n}$ and the time partitions dt were of size $1/n$.

It turns out that higher order terms $(dB_t)^3$ **are negligible** with respect to the leading orders, e.g., $E\left[(dB_t)^3\right]$ is the size of $(dt)^{3/2}$, which is negligible w.r.t. dt as $dt \to 0$.

In conclusion: as defined in the context above, stochastic integrals are limits of sums of random variables as the time step vanishes. They are thus ways of using random numbers to generate other random numbers which are often distributed with more complicated probability distributions.

We thus arrive at the most important differential formula in stochastic calculus.

7.3 Itô's Formulae

Itô's great contribution was to write down the analogous stochastic version of Taylor's theorem, as found everywhere in deterministic calculus, at A-level and in first year undergraduate courses.

7.3.1 Itô's Formula in 1D

Theorem (Ito's lemma): Let X be a stochastic process, which satisfies the differential form

$$dX_t = Y_t dB_t + Z_t dt$$

where $Y_t = Y(X_t, t)$, $Z_t = Z(X_t, t)$ and B_t is a standard Brownian motion. Let f be a deterministic, twice differentiable function. Then $f(X_t)$ is also a stochastic process and has the differential form

$$df = Y_t f'(X_t) dB_t + \left(Z_t f'(X_t) + \frac{1}{2} Y_t^2 f''(X_t) \right) dt,$$

where the first and second derivatives f', f'' are the usual deterministic ones.

Proof: By the arguments above we cannot immediately neglect terms in $(dX_t)^2$. Thus the Taylor expansion of f has the form

$$df(X_t) = f'(X_t)\, dX_t + \frac{1}{2} f''(X_t)\, (dX_t)^2 + O\left((dX_t)^3\right)$$

where X_t is a stochastic variable. Then we substitute in for $dX_t = Y_t dB_t + Z_t dt$, to get

$$\begin{aligned} df(X_t) &= f'(X_t) dX_t + \frac{1}{2} f''(X_t)\, (dX_t)^2 + O\left((dX_t)^3\right) \\ &= f'(X_t)\, (Y_t dB_t + Z_t dt) + \frac{1}{2} f''(X_t)\, (Y_t dB_t + Z_t dt)^2 + O\left((dX_t)^3\right) \\ &= f'(X_t)\, (Y_t dB_t + Z_t dt) + \frac{1}{2} f''(X_t) \left(Y_t^2 (dB_t)^2 + 2 Y_t Z_t dB_t dt + Z_t^2 (dt)^2\right) + O\left((dX_t)^3\right) \end{aligned}$$

We now use the shortcut "multiplication table"

	dt	dB_t
dt	$(dt)^2$	$(dt)^{3/2}$
dB_t	$(dt)^{3/2}$	dt

to give an estimate of the size (in the statistical sense) of the products of the infinitesimals involved.

From above we recall that terms in dB_t are of order $\sqrt{dt}$. Hence terms in $dB_t dt$ are of order $(dt)^{3/2}$. Thus in the limit as $dt \to 0$, both these terms, and those in $(dt)^2$ and higher order terms in $(dX_t)^3$ etc. are negligible wrt those in dt.

However the terms in $(dB_t)^2$ are of order dt. We replace the former with the latter, since the expected values of these two quantities will be the same.

Hence with this replacement and ignoring the higher order contributions we can collect up like terms to obtain

$$df(X_t) = Y_t f'(X_t) dB_t + \left(Z_t f'(X_t) + \frac{1}{2} Y_t^2 f''(X_t) \right) dt.$$

as required.

7.3.2 Example: calculation of stochastic integral

Itô's lemma gives us an important short cut to calculating stochastic integrals. Consider our old friend:

$$2 \int_0^t B_t dB_t.$$

Let $f(X_t) = X_t^2$ where we set $X_t = B_t$. Then trivially we have

$$dX_t = dB_t = 1 \times dB_t + 0 \times dt,$$

so we identify $Y_t = 1, Z_t = 0$ in Itô. Consequently applying Itô's formula above, we have

$$\begin{aligned}
df(X_t) &= Y_t f'(X_t) dB_t + \left(Z_t f'(X_t) + \frac{1}{2} Y_t^2 f''(X_t) \right) dt \\
\Rightarrow df(B_t) &= Y_t f'(B_t) dB_t + \left(Z_t f'(B_t) + \frac{1}{2} Y_t^2 f''(B_t) \right) dt \\
\Rightarrow d\left(B_t^2\right) &= 1 \times 2B_t dB_t + \left(0 \times 2B_t + \frac{1}{2} 1 \times 2 \right) dt \\
\Rightarrow d\left(B_t^2\right) &= 2B_t dB_t + dt
\end{aligned}$$

Hence we can integrate both sides to get

$$\begin{aligned}
\int_0^t d\left(B_t^2\right) &= 2\int_0^t B_s dB_s + \int_0^t ds \\
\Rightarrow B_t^2 - B_0 &= 2\int_0^t B_s dB_s + t \\
\Rightarrow B_t^2 - 0 &= 2\int_0^t B_s dB_s + t \\
\Rightarrow \int_0^t B_s dB_s &= \frac{1}{2}\left(B_t^2 - t\right).
\end{aligned}$$

Which is a very different result from the usual deterministic integral, but both LHS and RHS converge statistically to have the same expected value.

7.3.3 Simple Examples of the use of Itô's lemma

More commonly in the course we shall use Itô to evaluate derivatives of functions of random variables.

Example 1:

Suppose that

$$dX_t = \sigma dB_t + \mu dt,$$

so that $Y_t = \sigma$ and $Z_t = \mu$.

Then if $f(X) = X^2$ we find that

$$df = 2\sigma X_t dB_t + \left(2\mu X_t + \sigma^2\right) dt$$

Note that the deterministic version of this derivative would be

$$df = 2X_t dX_t$$

This can be reconciled with the Itô result by setting $\sigma = 0$ in the $dX_t = \sigma dB_t + \mu dt$, whereupon X_t become a deterministic variable so that $dX_t = \mu dt$. Hence we obtain

$$df = 2\mu X_t dt,$$

which is effectively the chain rule result of differentiating $f = \{X(\mu t)\}^2$ with respect to t.

Example 2:

Suppose now that the underlying stochastic variation in X_t obeys the SDE

$$dX_t = X_t\left(\sigma dB_t + \mu dt\right),$$

so that $Y_t = \sigma X_t$ and $Z_t = \mu X_t$. This is termed **geometric Brownian motion**.

Then if $f(X) = X^2$ from we find that Itô's lemma gives

$$\begin{aligned} df &= 2\sigma X_t^2 dB_t + \left(2\mu X_t^2 + \sigma^2 X_t^2\right) dt \\ &= X_t^2(2\sigma dB_t + (2\mu + \sigma^2)dt). \end{aligned}$$

Example 3:

Suppose again that

$$dX_t = X_t\left(\sigma dB_t + \mu dt\right),$$

so that $Y_t = \sigma X_t$ and $X_t = \mu$. Then if $f(X) = \log X$ we find from Itô's lemma

$$df = \sigma dB_t + (\mu - \frac{1}{2}\sigma^2)dt.$$

Another use of Itô's lemma is to generate SDEs from stochastic processes.

Example 4:

Calculate the SDE for the exponential Brownian process

$$f = \exp(\sigma B_t + \mu t),$$

where σ and μ are constants.

- **Solution:** Set $f\left(X_t\right) = \exp\left(X_t\right)$ and then with $X_t = \sigma B_t + \mu t$ we have

  $$dX_t = \sigma dB_t + \mu dt.$$

  Comparing this with Itô's lemma

  $$dX_t = Y_t dB_t + Z_t dt,$$

  we have $Y_t = \sigma$, $Z_t = \mu$. Hence the full lemma then gives us

  $$\begin{aligned} df &= \sigma f'(X_t)dB_t + \left(\mu f'(X_t) + \frac{1}{2}\sigma^2 f''(X_t)\right) dt \\ &= \sigma e^{X_t} dB_t + \left(\mu e^{X_t} + \frac{1}{2}\sigma^2 e^{X_t}\right) dt \\ &= f\left[\sigma dB_t + \left(\mu + \frac{1}{2}\sigma^2\right) dt\right] \\ \Rightarrow \frac{df}{f} &= \sigma dB_t + \left(\mu + \frac{1}{2}\sigma^2\right) dt. \end{aligned}$$

One can sometimes use Ito's lemma the other way round to deduce the process from the SDE. This is analogous to solving an ode by integrating, and as such in general is quite difficult to achieve analytically. Sometimes it is possible to spot the solution by guesswork and then use Itô's lemma to prove that it is a solution. Such a solution to an SDE is called a *diffusion*. (The motivation for this will appear later.)

Example 5:

Consider now the similar geometric Brownian motion SDE we used earlier for asset price movements:

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

Now we have

$$\ln S_t - \ln S_0 \sim N\left(\left[\mu - \sigma^2/2\right]t, \sigma^2 t\right).$$

Hence by comparison of the SDEs for S_t and f:

$$\frac{df}{f} = \sigma dB_t + \left(\mu + \frac{1}{2}\sigma^2\right)dt$$

we can claim that

$$\ln f_t - \ln f_0 \sim N\left(\mu t, \sigma^2 t\right).$$

Example 6:

Note that we must **not** assume that

$$\frac{df}{f} = d(\ln f),$$

as (a sloppy interpretation of) deterministic calculus might tempt us to do.

- To see this we use Itô to derive the SDE for $d(\ln f)$ and observe that it is **not** the same as that for df/f:

$$\begin{aligned} d(\ln f_t) &= \frac{d(\ln f_t)}{df_t}df_t + \frac{1}{2}\frac{d^2\left(\ln(f_t)\right)}{df_t^2}(df_t)^2 \\ &= \frac{df_t}{f_t} - \frac{1}{2f_t^2}(df_t)^2 \end{aligned}$$

Now we substitute the SDE for df/f from example 4, use the relationship $dB_t^2 = dt$ and neglect terms of order higher than dt to obtain:

$$\begin{aligned} d(\ln f_t) &= \frac{1}{f_t}f_t\left(\left[\mu + \frac{\sigma^2}{2}\right]dt + \sigma dB_t\right) - \frac{1}{2f_t^2}f_t^2\left(\left[\mu + \frac{\sigma^2}{2}\right]dt + \sigma dB_t\right)^2 \\ &= \left(\left[\mu + \frac{\sigma^2}{2}\right]dt + \sigma dB_t\right) - \frac{1}{2}\left(\left[\mu + \frac{\sigma^2}{2}\right]dt + \sigma dB_t\right)^2 \\ &= \left(\left[\mu + \frac{\sigma^2}{2} - \frac{\sigma^2}{2}\right]dt + \sigma dB_t\right) \\ \Rightarrow d(\ln f_t) &= \mu dt + \sigma dB_t. \end{aligned}$$

Hence we deduce that $d(\ln f_t) \neq df_t/f_t$.

- Note that we can also use the SDE for $d(\ln f_t)$ to derive the statistical distribution (again) of $\ln f_t$, this time directly by integrating each side, assuming μ and σ are constants:

$$\begin{aligned}
d(\ln f_t) &= \mu dt + \sigma dB_t \\
\int_0^t d(\ln f_s) &= \int_0^t \mu ds + \int_0^t \sigma dB_s \\
\int_0^t d(\ln f_s) &= \mu \int_0^t ds + \sigma \int_0^t dB_s \\
[\ln f_s]_0^t &= \mu [s]_0^t + \sigma [B_t - B_0] \\
\ln f_t - \ln f_0 &= \mu t + \sigma (B_t - B_0) \\
\Rightarrow \ln f_t - \ln f_0 &= \mu t + \sigma B_t \\
\Rightarrow \ln f_t - \ln f_0 &\sim N(\mu t, \sigma^2 t)
\end{aligned}$$

as above, since $B_0 = 0$ by definition and using the property of addition of random (and deterministic) variables.[3]

This is an example of the solution of a stochastic differential equation: the solution is the probability distribution of the integral of the LHS (here $\ln f_t$).

Example 7:

Solve the SDE

$$df = \sigma f dB_t$$

- **Solution:** By comparing with the above result, we see that after division by f, the stochastic term on the RHS is the same. In fact the SDE for f derived in example 4 above can be reduced to the present one if we choose there $\mu = -\sigma^2/2$, so eliminating the mean drift term. Consequently by comparison with the equation for the exponential Brownian process above we conjecture that a solution (up to an overall constant prefactor) is

$$f = \exp\left(\sigma B_t - \frac{1}{2}\sigma^2 t\right).$$

or,

$$f(X_t) = \exp(X_t), \qquad X_t = \sigma B_t - \frac{1}{2}\sigma^2 t.$$

By direct inspection of X_t we have

$$E[X_t] = E\left[\sigma B_t - \frac{1}{2}\sigma^2 t\right] = \sigma E[B_t] - E\left[\frac{1}{2}\sigma^2 t\right] = 0 - \frac{1}{2}\sigma^2 t$$

and since $B_t \sim N(0,1)$ we have by the addition of random (and deterministic variables) that

$$X_t \sim N\left(-\sigma^2 t/2, \sigma^2 t\right).$$

[3]If $B_t \sim N(0,t), \Rightarrow \sigma B_t \sim N(0, \sigma^2 t), \Rightarrow \mu t + \sigma B_t \sim N(\mu t, \sigma^2 t)$ for deterministic σ, μ, t.

The alternative (shortcut) to this result is to compare the SDE here with that for f in example 4 with $\mu = -\sigma^2/2$ and $f_0 = 1$,

$$X_t = \ln f_t - \ln f_0 = \ln f_t \sim N\left(-\sigma^2 t/2, \sigma^2 t\right).$$

7.4 Generalisations of Itô's Lemma

7.4.1 The Product Rule

The deterministic version of the differentiation of a product $f_t g_t$ can be written as $d\left(f_t g_t\right) = f_t dg_t + g_t df_t$.

In the stochastic world, there are two subcases:

- If $X_t,\ Y_t$ follow the same Brownian motion B, i.e.,

$$\begin{aligned} dX_t &= \sigma_t dB_t + \mu_t dt \\ dY_t &= \rho_t dB_t + \nu_t dt \end{aligned}$$

  then

$$d\left(X_t Y_t\right) = X_t dY_t + Y_t dX_t + \sigma_t \rho_t dt$$

- If $X_t,\ Y_t$ follow the different independent Brownian motions, e.g.,

$$\begin{aligned} dX_t &= \sigma_t dB_t + \mu_t dt \\ dY_t &= \rho_t d\tilde{B}_t + \nu_t dt \end{aligned}$$

  then

$$d\left(X_t Y_t\right) = X_t dY_t + Y_t dX_t,$$

  as in the usual deterministic case.

The derivation of these results can be demonstrated in several ways (of varying degrees of rigour).

7.4.2 Inclusion of deterministic dependence in Itô's lemma

For option pricing the most important generalisation of Itô's lemma is to the multivariate case, when one of the variables is deterministic.

Theorem (Ito's lemma plus deterministic dependence): Let X_t be a stochastic process, which satisfies the differential form

$$dX_t = Y_t dB_t + Z_t dt$$

where $Y_t = Y(X_t, t), Z_t = Z(X_t, t)$ and B_t is a standard Brownian motion. Let f be a deterministic, twice differentiable function, which now also depends explicitly on time. Then $f(X_t, t)$ is also a stochastic process and has the differential form

$$df = \left(Y_t \frac{\partial f}{\partial X}\right) dB_t + \left(\frac{\partial f}{\partial t} + Z_t \frac{\partial f}{\partial X} + \frac{1}{2} Y_t^2 \frac{\partial^2 f}{\partial X^2}\right) dt,$$

where the first and second partial derivatives $\partial f / \partial X$, $\partial f / \partial t$, $\partial^2 f / \partial X^{\,2}$ are the usual deterministic ones.

Proof: As above, but starting with the multidimensional Taylor expansion:

$$df(X_t, t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial X} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial X^2} (dX_t)^2 . \tag{7.1}$$

Since $dX_t = O\left(\sqrt{dt}\right)$, these are the leading order terms, with the term containing the first t derivative being of the same size as the term with the second X_t derivative. The cross terms involving $dt dX_t$ are $O\left((dt)^{3/2}\right)$ and so are of higher order and can be neglected. The steps of the analysis then proceed through as above.

This result is vitally important for continuous time option pricing and the derivation of the Black-Scholes model!

Example:

Now we have a deterministic dependence in Itô we illustrate a more general approach to solving SDEs by the following example.

Consider the stochastic function $f_t(t, B_t)$ which satisfies the SDE:

$$df_t = f_t^3 dt - f_t^2 dB_t, \qquad f_0(0, B_0 = 0) = 1.$$

Note that we are told that $f_t(X_t, t) = f_t(B_t, t)$. Hence, by comparison with $dX_t = Y_t dB_t + Z_t dt$, we have $dX_t = dB_t$, i.e., $Y_t = 1, Z_t = 0$. This also implies that $X_t = B_t$. Thus, from

Itô's formula f_t must also satisfy, with $dB_t^2 = dt$,

$$\begin{aligned}
df_t &= \frac{\partial f_t}{\partial t}dt + \frac{\partial f_t}{\partial B_t}dB_t + \frac{1}{2}\frac{\partial^2 f_t}{\partial B_t^2}dB_t^2 \\
&= \frac{\partial f_t}{\partial t}dt + \frac{\partial f_t}{\partial B_t}dB_t + \frac{1}{2}\frac{\partial^2 f_t}{\partial B_t^2}dt \\
&= \left(\frac{\partial f_t}{\partial t} + \frac{1}{2}\frac{\partial^2 f_t}{\partial B_t^2}\right)dt + \frac{\partial f_t}{\partial B_t}dB_t
\end{aligned}$$

If we compare the terms in dt and dB_t between the SDE and Itô we find that f_t must satisfy two partial differential equations:

$$\begin{aligned}
dt: &\qquad \frac{\partial f_t}{\partial t} + \frac{1}{2}\frac{\partial^2 f_t}{\partial B_t^2} = f_t^3 \\
dB_t: &\qquad \frac{\partial f_t}{\partial B_t} = -f_t^2
\end{aligned}$$

Partial differential equations are usually quite difficult to solve. However in this case we can solve the system as follows.

Take the second (dB_t) equation and differential again with respect to B_t (as if it were a deterministic variable):

$$\frac{\partial f_t}{\partial B_t} = -f_t^2 \qquad \Rightarrow \qquad \frac{\partial^2 f_t}{\partial B_t^2} = -2f_t\frac{\partial f_t}{\partial B_t} \qquad \Rightarrow \qquad \frac{\partial^2 f_t}{\partial B_t^2} = -2f_t\left(-f_t^2\right) = 2f_t^3$$

Substitute this into the first (dt) equation:

$$\frac{\partial f_t}{\partial t}+\frac{1}{2}\frac{\partial^2 f_t}{\partial B_t^2} = f_t^3 \;\Rightarrow\; \frac{\partial f_t}{\partial t}+\frac{1}{2}\left(-2f_t\frac{\partial f_t}{\partial B_t}\right) = f_t^3 \;\Rightarrow\; \frac{\partial f_t}{\partial t}+\frac{1}{2}\left(2f_t^3\right) = f_t^3 \;\Rightarrow\; \frac{\partial f_t}{\partial t} = 0.$$

The conclusion is that f_t does not explicitly depend on t and is purely a function of B_t. Hence returning to the dB_t equation, actually we no longer have a partial differential equation, but an ordinary one:

$$\frac{\partial f_t}{\partial B_t} = -f_t^2 \qquad \Rightarrow \qquad \frac{df_t}{dB_t} = -f_t^2$$

This can be solved formally as a variables separable equation (MATH1052!) as follows (again treating the variables as if they were deterministic):

$$\int \frac{df_t}{f_t^2} = -\int dB_t \qquad \Rightarrow \qquad -\frac{1}{f_t} = -B_t + C \qquad \Rightarrow \qquad f_t = \frac{1}{B_t - C}$$

Using the fact that $B_0 = 0$ and the given initial data $f_0 = 1$ we have

$$1 = \frac{1}{0 - C} \qquad \Rightarrow \qquad C = -1.$$

Inserting this value in the solution we arrive at the solution:

$$f_t = \frac{1}{B_t + 1}.$$

7.4.3 The Ornstein-Uhlenbeck process

The following is an example of a more sophisticated calculation using Itô on a process that is of major importance in stochastic finance.

Example:

Consider the **Ornstein-Uhlenbeck** process

$$dX_t = -\gamma X_t dt + \sigma dB_t.$$

where γ and σ are positive constants. (Note that there is no X_t in front of the stochastic σdB_t bit and so this is **not** geometric Brownian motion.)

The solution begins by observing that the general solution of a deterministic equation $dX_t = -\gamma X_t dt$ is $X_t = ce^{-\gamma t}$. Hence we hope that the stochastic nature of the equation doesn't change this much and we look for a solution of the form to the SDE of the form:

$$X_t = U_t e^{-\gamma t}.$$

Using Itô we have:

$$\begin{aligned} dU_t &= d\left(e^{\gamma t} X_t\right) = \gamma e^{\gamma t} X_t dt + e^{\gamma t} dX_t, \\ &= \gamma e^{\gamma t} X_t dt + e^{\gamma t}\left(-\gamma X_t dt + \sigma dB_t\right), \\ &= \sigma e^{\gamma t} dB_t \end{aligned}$$

Hence we can say that

$$U_t = U_0 + \sigma \int_0^t e^{\gamma s} dB_s$$

and so

$$X_t = e^{-\gamma t} U_t = X_0 e^{-\gamma t} + \sigma \int_0^t e^{\gamma(s-t)} dB_s.$$

This result will be useful later on when we discuss interest rate models.

7.5 Itô and stochastic integrals

Suppose that we wish to compute the statistical properties of the stochastic integral

$$\int_0^t g(s) dB_s$$

for some deterministic function $g(s)$, where s is the dummy (deterministic) time variable where B_t is standard Brownian motion.

According to the definition of a stochastic integral, if we discretise time as

$$\left\{0, \frac{t}{n}, \frac{2t}{n}, \ldots \frac{it}{n}, \ldots, t\right\}$$

we can write this as

$$\int_0^t g(s)dB_s = \lim_{n\to\infty}\left\{\sum_{i=0}^{n} g\left(\frac{it}{n}\right)\left[B\left(\frac{(i+1)t}{n}\right) - B\left(\frac{it}{n}\right)\right]\right\}.$$

Now by the properties of Brownian motion, the quantity in square brackets is distributed as

$$B\left(\frac{(i+1)t}{n}\right) - B\left(\frac{it}{n}\right) \sim N(0, t/n).$$

By the property of addition of random variables (see probability handout), the sum will also be distributed normally with mean zero. Hence, in the limit the integral

$$\int_0^t g(s)dB_s$$

is distributed normally, with mean zero:

$$E\left[\int_0^t g(s)dB_s\right] = 0.$$

To find the variance (and hence identify the precise normal distribution) we need to find

$$\begin{aligned}
\mathrm{Var}\left[\int_0^t g(s)dB_s\right] &= E\left[\left(\int_0^t g(s)dB_s - E\left[\int_0^t g(s)dB_s\right]\right)^2\right] \\
&= E\left[\left(\int_0^t g(s)dB_s\right)^2\right] - \underbrace{E\left[\int_0^t g(s)dB_s\right]^2}_{=0^2} \\
&= E\left[\left(\int_0^t g(s)dB_s\right)^2\right].
\end{aligned}$$

To simplify this we use Itô's isometry for a stochastic or deterministic function $Z(t)$:

$$E\left[\left(\int_a^b Z(t)dB_t\right)^2\right] = E\left[\int_a^b Z^2(t)dt\right],$$

the proof of which can be found by using the formal definition of a stochastic integral (above), the independence of Brownian increments and the fact that $E[dB_t^2] = dt$. The proof is not required. For a deterministic function the beauty of the Itô's isometry is that integral on the left hand side is (being now an ordinary deterministic integral over a deterministic variable) extremely easy to evaluate!

$$E\left[\left(\int_0^t g(s)dB_s\right)^2\right] = E\left[\int_0^t g^2(s)ds\right],$$

Hence for deterministic g, we have

$$\text{var}\left[\int_0^t g(s)dB_s\right] = E\left[\int_0^t g^2(s)ds\right] = \int_0^t g^2(s)ds.$$

Therefore we have that

$$\int_0^t g(s)dB_s \sim N\left(0, \int_0^t g^2(s)ds\right).$$

7.5.1 Distribution of the integrated Ornstein-Uhlenbeck process

We solved the Ornstein-Uhlenbeck SDE to generate the process

$$X_t = X_0 e^{-\gamma t} + \sigma \int_0^t e^{\gamma(s-t)} dB_s.$$

We know from above that the stochastic integral on the RHS will be normally distributed with mean zero. Since the other components of the RHS are deterministic, the LHS will thus be distributed normally with a mean and variance to be found.

The mean is given by:

$$\begin{aligned}
E[X_t] &= E\left[X_0 e^{-\gamma t}\right] + \sigma E\left[\int_0^t e^{\gamma(s-t)} dB_s\right] \\
&= X_0 e^{-\gamma t} + \sigma \times 0 \\
E[X_t] &= X_0 e^{-\gamma t}.
\end{aligned}$$

The variance is given by:

$$\begin{aligned}
E\left[(X_t - E[X_t])^2\right] &= E\left[\left(X_0 e^{-\gamma t} + \sigma \int_0^t e^{\gamma(s-t)} dB_s - E[X_t]\right)^2\right] \\
&= \sigma^2 E\left[\left(\int_0^t e^{\gamma(s-t)} dB_s\right)^2\right] \\
&= \sigma^2 e^{-2\gamma t} E\left[\left(\int_0^t e^{\gamma s} dB_s\right)^2\right] \\
&= \sigma^2 e^{-2\gamma t} E\left[\int_0^t e^{2\gamma s} ds\right] \qquad \text{(by Ito isometry)} \\
&= \sigma^2 e^{-2\gamma t} \int_0^t e^{2\gamma s} ds \\
&= \sigma^2 e^{-2\gamma t} \left[\frac{e^{2\gamma s}}{2\gamma}\right]_0^t \\
&= \frac{\sigma^2 \left(1 - e^{-2\gamma t}\right)}{2\gamma}.
\end{aligned}$$

Hence we can deduce that the integrated Ornstein-Uhlenbeck process is normally distributed as

$$X_t \sim N\left(X_0 e^{-\gamma t}, \frac{\sigma^2\left(1 - e^{-2\gamma t}\right)}{2\gamma}\right).$$

Chapter 8

Introduction to the valuation of derivative securities

"Derivatives are financial weapons of mass destruction"[1]

8.1 Introduction

In one way or another, most of this unit is concerned with the **valuation** of things. In the first part of the unit we saw how to value assets and portfolios of assets. The valuation of financial instruments such as this is relatively straightforward. We come now however to the question of how to value **derivative** products. In general, a derivative product depends for its value on the value of some other asset or assets. Because of the huge increase in interest in these products in recent years, much effort has been expended on finding their "fair value".

> *"The philosopher Thales, believing six months in advance that the spring weather would be exceptionally favourable to olive growth, negotiated the right to rent oil presses for the next harvest at a bargain rate from press owners who wanted to hedge their bets against a poor harvest. When spring brought a lush olive harvest, Thales rented the machines to others at a much higher rate, thereby making a profit to support his philosophical enquiries"*

(Translation of Aristotle "Politics", 5th Century BC: Thales had purchased a call option!)

8.2 Basic Options: Calls and Puts

The simplest sort of derivative product is a "vanilla" (i.e. an "unexciting flavoured") option. **Calls** and **puts** are vanilla options, and the process of selling such options is termed "writing"

[1] Warren Buffett (by common consent the most successful share trader ever): Berkshire Hathaway Annual Report 2002

the option. The **underlying asset** (referred to henceforth as just "the underlying") is most often a share or the an index (for example, shares in BP or the value of the FTSE 100).

Every option has an **expiry date** (usually denoted by T) and a **strike price** (often abbreviated simply to "**strike**" and usually denoted by K). Note: in the financial press the Expiry is usually simply given as a month (e.g., "January Calls"). By tradition the *actual day* of expiry is the 3rd Friday of each month.

- A **European Call** on an underlying gives the holder the right, but not the obligation, to buy the underlying (to "**exercise**") at the price K at the time T.
- A **European Put** on an underlying gives the holder the right, but not the obligation, to sell the underlying (to "**exercise**") at the price K at the time T.

Notes:

- Since they are easier to analyze, we have decided to consider first so-called **European Options**. **American** options run along similar lines, but may be exercised at ANY time before expiry. (Important point: do not misunderstand–the terms "American" and "European" simply indicate a particular type of financial deal. The names are a historical accident and do not imply anything about where the options are bought and sold.)
- A Call or a Put option is a valuable thing to possess; we should therefore expect to have to pay for this privilege.
- The writer has an OBLIGATION to settle up at expiry; the holder has no obligation to exercise though.
- Normally no shares are actually bought or sold (though they can be). The option acts like a straight bet between the writer and the holder. Initially, the holder pays the writer for the option. If the holder does not exercise at expiry then no further money changes hands and the holder is out of pocket; if the holder does exercise then the writer has to pay the holder what he or she owes him.
- The real issue here is how the writer can ELIMINATE RISK, ("hedge"); for technically the writer is exposed to possible infinite losses.
- There is a close connection between options and "spread bets" that are becoming increasingly popular with the bookies. With spread bets it is also possible to hedge (to a certain extent) and even sometimes to make a risk free profit when nearby bookmaker's shops make errors. Your course coordinator is investigating!
- BIG MONEY is spent on options. The often-quoted figure is that the total worldwide spend on derivative securities per year was \$10,000 *billion*, and that was way back in 1992! By anybody's standards, 10^{13} dollars is a lot of money. More recently the takeover kings of the 1980s and the dotcom boomers of the 1990s have been replaced

by a new elite of city workers who specialise in derivatives. In the last quarter of 2003, UK banks had derivative product liabilities of £1,000 billion (one terapound!). This is roughly the same as the value of the whole of the UK's output for a year.

- It seems that derivative's share of the total trade is increasing. Between 2000 and 2003 the value of UK shares traded remained fairly constant at about £7.5 billion a day, but the number of FTSE 100 index futures traded on the Euronext LIFFE exchange in London rose from 49,605 per day to 79,110 per day: about £3.2 billion. If you add in options, interest rate swaps, and everything else then the total is closed to the £7.5 billion for all share transactions. On average, every day 60,000 tons of robusta coffee is notionally traded in London!

8.3 Simple example of the use of options

First, a simple example: Figure 8.1 is a (not very good) copy of the options page of the Times for 17th November 1998. (The date on which this paragraph was written.) Let's pretend that we are going to risk a little money and see what happens. Pick two companies, say ASDA and Zeneca. For each company prices for Calls, and the Puts, are listed for Jan, Apr and July. The current price is shown under the Company name, and the strike price where it says "series". So how shall we gamble? Well here's a "Company assessment".

ASDA: The share price has risen steadily over the last few years, but dropped by about one quarter in mid 1998 (along with almost everything else). The broker's estimates have recently downgraded the share from a "BUY" to a "HOLD" and 4 directors have sold shares in the last 5 months. PREDICTION: The shares will go down. To make a profit therefore we have to buy PUTS. The current price is 158 and January PUTS at a strike of 160 are 12.5p, so let's buy 4,000 of them (cost £500).

ZENECA: Two brokers have recently rated the share "UNDP" (underpriced) and while the profit estimates for 1998 are nothing special, there's a big increase forecast for 1999. In addition, on 29th October approval was given for "Nolvadex" (a drug supposed to reduce the incidence of breast cancer - and so far the only one of its type). The potential upside of this is huge, so let's take a bullish view. PREDICTION: The shares will rise. The current share price is 2297.5 and Calls for January 1999 are on sale at 199.5p (call it 200p) with a strike of 2200. So we'll buy 500 of them (cost £1000).

JUDGEMENT DAY: On expiry in January 1999 the share prices were ASDA: 161.25, ZENECA: 2617. So what happens now? Well the ASDA share price has not fallen as we'd hoped. There's no point selling something for 160 when it's worth more, so we do not exercise the PUTS. We have to throw them away, losing £500. ZENECA did well, though. The difference between the current share price and the strike is 2617p-2200p = 417p, and 500 Calls are therefore worth $500 \times 4.17 =$ £2085. Why? This is because we exercise the

LIFFE OPTIONS

	Series	Calls Jan	Calls Apr	Calls Jul	Puts Jan	Puts Apr	Puts Jul
ASDA	140	23½	29½	33½	4	8½	10
(*158)	160	12	19	23½	12½	17½	19
Abbey Nat	1200	100½	146	189	66	109	135
(*1222½)	1250	76½	122	165	91	134½	160½
Allid Dom	500	41	62	72½	27½	41½	52
(*508)	550	20	39	51½	57	70	80
BAA	650	50	77	90½	26½	44½	58½
(*658)	700	28	52½	66½	53½	70	84
BAT Ind	500	64½	—	—	25½	—	—
(*632½)	650	39	—	—	47½	—	—
Barclays	1200	127	177	214½	72½	126	150½
(*1240½)	1250	102½	153½	192	96	151½	175½
Bass	750	55½	83½	102	51	58	80½
(*766)	800	34½	61½	80½	81	96	108½
Boots	950	75	108½	138½	43	62	67½
(*971½)	1000	51½	83½	114	66	84½	112
Br Airways	360	48	66	78	22½	36	46½
(*384)	390	32½	52	54	35	50½	62
BP	900	75½	108	131	42½	57½	85½
(*922½)	950	52	83	107	66	92½	111
Br Steel	100	8½	13½	15½	5	10	16½
(*102)	110	5	10	12	14½	16	23
C&W	650	83	116	136½	39	65	79½
(*688)	700	58½	93	114	64	91½	105½
CGU Plc	900	90	128	160½	59½	105	124
(*920)	950	66	104	139	85½	131½	150½
GlaxoW	1800	156	223	277½	82	136	172
(*1855)	1900	106½	174½	229½	131	186	222½
HSBC	1450	143½	192½	234	96	149½	181½
(*1477)	1500	119½	167	211	121	175	207½
Halifax	800	76	100	116½	43½	70	78½
(*821½)	850	51½	75½	92½	59½	97½	104
ICI	550	79½	105½	119	34½	57½	75
(*590½)	600	54	82	97	59	94½	101½
Land Sec	800	66½	91	104	20	36½	50½
(*844½)	850	47	65	78	44	59½	72½
M&S	420	47	60	71½	13½	22½	31
(*448)	460	25	43½	52½	30½	47½	50½
Natl Pwr	500	61½	79	96	15	25½	34
(*549½)	550	32½	52	59½	35½	47½	57½
Nat West	900	107½	130	159	49	75	95
(*947)	950	82	106	135½	73	100½	121
Reed Intl	460	37½	47	62½	30	51	51½
(*463)	500	21	31½	46	53½	76½	74½
Reuters	600	55	82	103½	49	74	84½
(*605)	650	34½	60½	82	76	102½	112½
Rvl&Sun	460	61	81½	95½	24½	49	55½
(*489½)	500	41	62	77½	44	70	76½
Safeway	280	20	30½	36½	17½	26½	33
(*283)	300	12	22	25	29	37½	46
Sainsbury	500	43	60	69½	23	34½	42½
(*516½)	550	21½	38	47	51½	61½	69½
Shell	330	39½	52	61	9	16½	25
(*359½)	360	21½	34½	46½	21	29½	39½
SmKl Bch	700	59½	103	121	32	53	55
(*731)	750	43½	79	97½	55½	78½	90½
Std Char	600	80	108	137½	49	81½	99
(*627½)	650	57½	95½	110	76½	108½	126½
Thames W	1168	75	99	—	47	56	[illegible]
(*1184)	1223	48	72½	—	73	94½	—
Vodafone	800	99	141	170½	47½	76	96
(*848½)	850	74	118½	148½	71½	101½	122½
Woolwich	360	26½	39	47½	15½	26	33
(*367½)	390	13	25½	34	33	42½	49
Zeneca	2200	199½	292	350	101	168	204½
(*2297½)	2300	148	237½	302	149	217½	254½

	Series	Calls Nov	Calls Feb	Calls May	Puts Nov	Puts Feb	Puts May
BTR	100	7	16½	20	½	6	12
(*106½)	110	1½	11½	15	4½	13½	17½
Blue Circ	280	18½	38½	47½	1	17½	28
(*298)	300	4½	28	38	7	27½	38
Br Aero	450	21	—	—	2	—	—
(*468½)	475	6	—	—	12	—	—
Br Telcm	800	29½	82	111	2½	50½	67½
(*827½)	850	3½	56	87	26½	76½	93
Cadbury	850	43½	99½	124	1½	43	65
(*892½)	900	8	74	99	16½	66½	89½
Carlton Cms	420	13	46	58½	3½	31	42
(*429½)	460	1	28½	41	31	52½	64½
Diageo	650	16	71½	93	9	55½	77
(*657½)	700	1	49½	71½	43½	83½	105
Energy Gp	800	40½	—	—	—	—	—
(*840)	850	½	—	—	10½	—	—
Gallaher	420	29½	53	63½	½	19	29½
(*449)	460	2½	31½	43½	13½	37½	49½
GEC	460	24	56½	75	1	29	41½
(*482½)	500	2	36½	56½	19½	49½	62
Hanson	420	16	44	57	1½	24½	37
(*434½)	460	½	24½	38	26	44½	55
Imp Tob	650	18½	45½	66½	2	35	43½
(*666½)	700	1	24½	46	33½	64½	71½
Kingfshr	525	16	59	—	5	40	—
(*534)	550	4	47½	65½	18	53	71
Lasmo	160	13½	25	31½	½	9½	13½
(*173)	180	1½	15	22	8½	17½	23
Ladbroke	220	5	24½	34	6½	21½	31½
(*220)	240	½	17	26	21	33½	43½
LucasVar	200	4½	19	24½	3	14	18½
(*201½)	220	—	11	16	18½	26	30
P&O	600	20½	63½	80	3½	37½	57½
(*617)	650	1	40½	57	34	64½	85½
Prudentl	800	28	94½	123½	5½	60½	84½
(*824)	850	3½	71	100	31½	86	110½
Rio Tinto	650	41½	81½	99	1	30½	53½
(*690½)	700	6½	56	73½	15	53½	76½
R-Royce	220	9	26½	34	1½	14½	21½
(*228)	240	½	17½	24½	13	25	32
Tesco	167	10½	20	—	½	9	—
(*177)	183	1	12½	—	6	17	—
Utd Bisc	238	11	24½	—	1	11½	—
(*248)	261	1	14½	—	12½	23½	—

	Series	Calls Jan	Calls May	Calls Aug	Puts Jan	Puts May	Puts Aug
All&Leic	850	62½	117	139½	29	58	71
(*893½)	900	55	97½	115	51	82	95½

	Series	Calls Dec	Calls Mar	Calls Jun	Puts Dec	Puts Mar	Puts Jun
BG	390	28½	48½	63	10½	20½	35
(*407½)	420	13½	34	49	24	35½	50
BSkyB	460	41½	51	72	9½	30½	38
(*492½)	500	17½	39½	51	26	47½	55½
Br Bio	40	4	8½	11½	4	7½	10
(*40½)	45	2	6½	9½	6	10½	13
Centrica	110	9	14	19	3	6½	8½
(*116)	120	4	9	14	6	11½	13½
Dixons	600	44½	72	92	14½	40½	50½
(*624)	650	20½	47½	68	39½	66½	76
EMI	330	38	53	69	15½	31	40½
(*349½)	360	23½	39½	56	30	46	56½
Granada	800	70	111½	131½	29	66½	73
(*841)	850	43	88	108	51	93	98½
GUS	500	45	68	84	15½	35½	41
(*630)	650	19½	44	59	39½	60½	66
Lloyds TSB	750	67½	102½	128	28½	66½	85
(*779)	800	35½	78½	105	53	92	111
Lonrho	300	25½	37½	42½	5½	17½	23
(*319)	330	9	22½	26½	20½	34	39
Norwich	420	28	54½	65½	15	33½	45
(*430½)	460	11	36½	48	38½	55½	67
Orange	600	43½	81	105	35½	67½	81
(*602½)	650	24	60	83	66	95	108½
Railtrck	1700	84	167	224½	65	128½	168½
(*1710½)	1800	43	122	176	125	185½	221
Scot Pwr	600	49½	74	88½	9½	23	34
(*645½)	650	17½	46½	61	30½	45	57
Tarmac	100	10½	15½	17½	3½	7½	11½
(*106½)	110	5	11	13	9	12½	16½
Tomkins	260	23½	34½	42½	8	18	23½
(*274½)	280	12½	25	33	17½	28½	34
Unilever	600	52½	80½	96½	13½	30½	42
(*638)	650	24½	54	72½	32½	54	56

November 16 Total: 31347 Calls: 18714 Puts: 12633

FTSE INDEX (*5519)

Strike	Calls Nov	Calls Dec	Calls Jan	Calls Feb	Calls Jun	Puts Nov	Puts Dec	Puts Jan	Puts Feb	Puts Jun
5350	188½	319	402½	479	—	19	125	190½	251½	—
5400	147	282½	370½	448½	635	28½	139	208½	270½	421½
5450	111½	252	341	418	—	43½	158½	226	289½	—
5500	80	223	311	387½	574	62½	179½	247	309	455
5550	53½	194	281	357	—	86½	204	266½	328½	—
5600	34½	165	251½	326½	512½	118	226½	287	348½	494½
5650	20½	143	223½	296	—	155½	254	310	369½	—
5700	11½	122	201½	273½	456	196	284	337½	396½	532½

November 16 Total: 5578 Calls: 4147 Puts: 1431

*Underlying security price

Figure 8.1: The options page of the "Times" on 17th November 1998

calls and buy the shares at 2200p each. However this is below the current market price and so we can sell them for 2617p, making a profit (excluding dealing charges) of 417p per option/share. The total balance sheet after our little gamble is thus:

- SPENT: £500 + £1000 = £1500
- RETURN: £0 + £2085 = £2085,

a healthy profit of £585, which is close to a thirty nine per cent return on our investment. Of course, dealing charges and the bid/ask spread have been ignored, but they will not reduce the profit here substantially.

8.4 Why buy options?

Why do holders of options want to hold them? And why do writers of options want to write them? There are actually quite a few reasons and it is important to understand them if one is to understand the way that the options market works.

As far as the holder is concerned:

- The holder may have taken a view on the market and be buying for reasons of speculation. If the holder believes that the share value will increase/decrease, he or she buys calls/puts; the inherent high gearing means that prospective profits and losses may be large even for relatively modest share movements.
- Suppose that an investor holds certain shares as long-term investments. To insure against a temporary fall in share price the investor might hedge to a certain extent by buying some puts. An alternative to doing this would be to sell the shares now and buy them back at a later date; but if the investor's view is wrong then (a) they will cost more to rebuy and (b) transaction costs on two deals will have to be paid.

On the other hand, as far as the writer is concerned:

- The writer expects to make a risk-free profit. This may be achieved buy selling options at a price slightly *above* the "true value". The risk may then be hedged away until the option expires.

There are actually many more complicated reasons why holders would wish to hold options and writers would wish to write them; most are outside the scope of this unit however. The reader is referred to the unit book list for further information.

8.5 Payoffs for vanilla options

Before attempting to value the simplest form of option, let us consider briefly what is at stake when an option is bought and how the payoff depends on the option structure and the price of the underlying asset.

Starting with a European call, we note that at expiry when $t = T$ the holder of the call has to decide whether or not to exercise the call. If the share price S_T is below the strike K then obviously this would be foolish, for it makes no sense buying something for more than it is worth. Thus the payoff will be 0 whenever $S_T \leqslant K$. If $S_T > K$ at expiry however, then evidently the option should be exercised; the payoff for a European call (and thus the value of the option at time $t = T$) is thus

$$c_T = \max(S_T - K, 0).$$

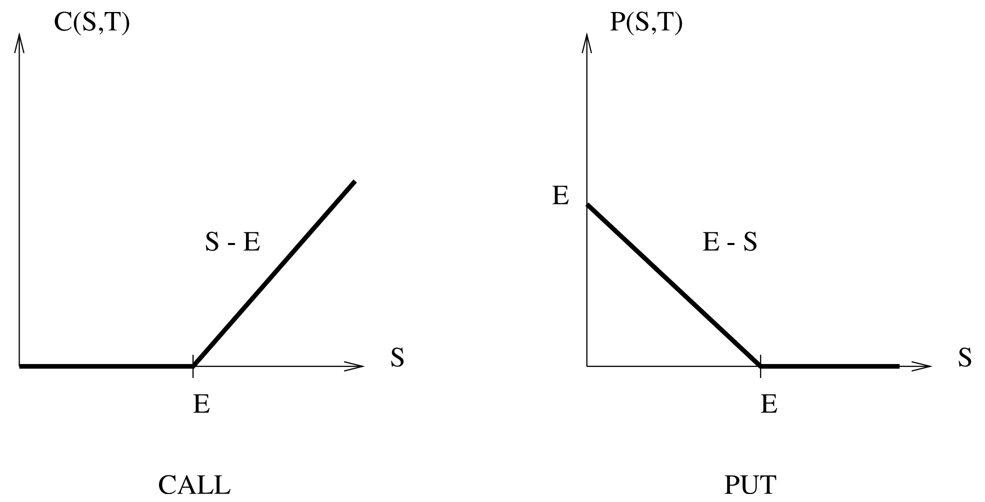

Figure 8.2: Payoff diagrams (from the holder's point of view) for Call and Put options

For a European put, the situation is similar, but reversed. Now it only makes sense to exercise the option if $S_T < K_T$ at expiry. The value of the option is thus

$$p_T = \max(K - S_T, 0).$$

The payoff diagrams for both calls and puts are shown in figure (8.2).

In figure (8.3) we present the same diagrams, but this time from the writer's point of view. Note that for a call, the writer's loss is potentially unbounded. Taking a "naked" position (i.e. making no attempt to hedge) is thus very risky for an option writer. Note also that in both figure (8.2) and figure (8.3) we have shown only the payoff and not the profit. To draw similar pictures for the profit the initial investment (cost of the option) must be taken into account.

Note that options may be combined together to form more complicated payoffs with somewhat exotic names. See the problem sheets.

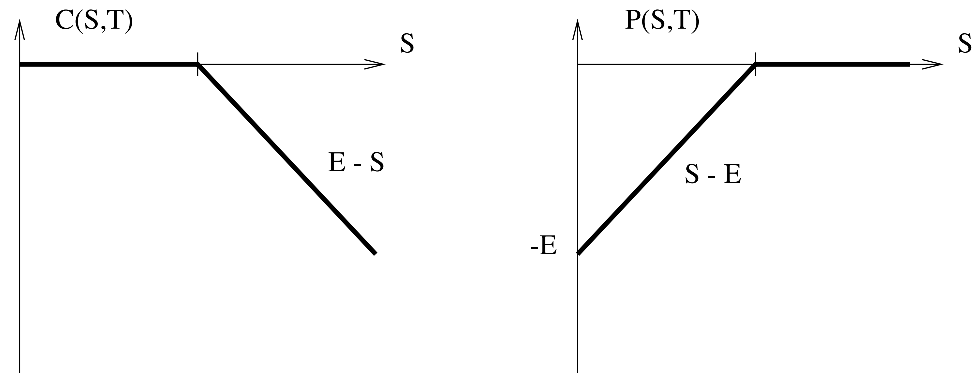

Figure 8.3: Payoff diagrams (from the writer's point of view) for Call and Put options

8.6 Gearing and a trivial option valuation problem

Soon we will start to set up the mechanism for valuing an option, but to obtain an idea of the sort of results that we will be seeking, consider for a moment a greatly simplified and unrealistic case. Suppose we were to buy one call option on a Company (whose current share price is 100p) with a given expiry date and a strike of 100p. Suppose also that now somehow we *know* that, at expiry, the share price of the Company will either be 80p or 120p, each with probability 0.5. How much is this option worth?

This is simple to answer. There is a half chance that the price will be 120p at expiry, in which case we can buy for 100p something that is worth 120p. Our profit here is thus 20p. On the other hand, if the coin comes down the other way, then the profit will be zero as the option will expire worthless (there is no point buying something only worth 80p for 100p). The expected profit P is thus

$$P = \frac{1}{2}(20) + \frac{1}{2}(0) = 10\text{p}.$$

Ignoring interest rates and other complications, it therefore seems sensible that the option should be valued at 10p. Of course, real option valuation is a lot harder than this as we never know the future distribution of share prices; the ideas involved are essentially the same however.

The other point well illustrated by this example concerns the *gearing* involved. Let's consider the percentage profit involved in buying the share and contrast it with that involved in buying the option. If we buy the share then there is a half chance of making 20 percent on the deal and a half chance of losing 20 percent. Buying the option is *far more risky*, for we either make a profit of 100 percent (we spent 10p on the option, the payoff was 20p and so the profit was 10p) or we lose everything. The high gearing reflects the fact that option prices respond in a much more violent way to changes in the underlying asset price; for speculation purposes, options are not for the faint-hearted!

.

8.7 Summary of Notation

It is convenient to collect together here the notation we shall adopt for options henceforth.

- t is the current time;
- S_t is the underlying asset price at time t;
- K is the strike or exercise price;
- T is the option expiry date;
- σ is the volatility (standard deviation) of the underlying asset price;
- c_t is the price at time t of a European call option on an underlying asset;
- p_t is the price at time t of a European put option on an underlying asset;
- C_t is the price at time t of an American call option on an underlying asset;
- P_t is the price at time t of an American put option on an underlying asset;
- r is the risk-free continuously-compounding rate of interest (assumed constant);

8.8 Factors affecting option prices

A number of mathematical models are used to value options. One of the more widely used is the **Black-Scholes** model. We shall study this later. It uses five parameters to value an option on a non-dividend-paying share. For the moment we just list the qualitative reasons why these are important. The five parameters are:

- **Underlying share price** S_t:

  Consider a Eurocall. The payoff at expiry is $\max(S_T - K, 0)$. During the lifetime of the option, $t < T$ clearly as S_t approaches the strike price K the likelihood of a change in the payout at expiry will increase. The larger the value of S_t the greater the chance of a non-zero payout at expiry T.

  Similar arguments apply to a Europut, except in reverse (the lower the value of S_t the greater the chance of a payout).

- **Strike price** K:

  For a call option, the higher the strike price, the smaller the likely payout at expiry. For a put option, the lower the strike, the smaller the ultimate payout.

- **Time to expiry** $T - t$:

  The longer the time to expiry, the greater the chance that the underlying share price can move significantly in favour of the holder of the option before expiry. So the value of an option will increase with term to maturity $(T - t)$. This increase is moderated slightly by the change in the time of the value of money (e^{rt} inflation).

- **Volatility of the underlying share:** σ

  The higher the volatility of the underlying asset, the greater the chance that the underlying share price can move significantly in favour of the holder of the option before expiry. So the value of an option will increase with the volatility of the underlying share.

- **Interest rates:** r

  An increase in the risk-free rate of interest will result in a higher value for a call option because the money saved by purchasing the option rather than the underlying share can be invested at this higher rate of interest, thus increasing the value of the option. For a put option, higher interest means a lower value.

The basic Black-Scholes model can be adapted to allow for a sixth factor determining the value of an option:

- **Income received on the underlying security:**

  In many cases the underlying security might provide a flow of, say, dividend income. Normally such income is not passed onto the holder of an option. Then the higher the level of income received, the lower is the value of a call option, because by buying the option instead of the underlying share the investor foregoes this income. The reverse is true for a put.

8.9 Arbitrage

One of the central concepts in this section of Financial Economics is that of **arbitrage**.

Put in simple terms, an arbitrage opportunity is a situation where we can make a sure profit with no risk. This is sometimes described as a "free lunch". Put more precisely an arbitrage opportunity means that:

- We can start at time 0 with a portfolio which has a net value of zero (implying that we are long in some assets and short in others).
- At some future time T

- the probability of a loss is 0
- the probability that we make a strictly positive profit is greater than 0.

If such an opportunity existed then we could multiply up this portfolio as much as we wanted to make as large a profit as we desired. The problem with this is that all of the active participants in the market would do the same and the market prices of the assets in the portfolio would quickly change to remove the arbitrage opportunity.

The **principle of no arbitrage** states simply that arbitrage opportunities do not exist.

If we assume that there are no arbitrage opportunities in a market, then it follows that any two securities or combinations of securities that give exactly the same payments must have the same price. This is sometimes called the "Law of One Price".

Note that under this definition a call option is not technically an arbitrage, since the buyer has to pay for the option initially to own it. If the option expires worthless, then the buyer has lost (a finite, restricted amount of) money. Similarly, if the buyer exercises the option, the seller is exposed to losses (unless they hedge the risk away).

Arbitrage arguments can be used in a similar way to **proof by contradiction** to justify the value of financial quantities.

The general approach is to start with the proposed result, e.g., as an equality and show that if it became an inequality (both $>$ and $<$), the inequalities would lead to an arbitrage. Then logically, invoking **the principle of no arbitrage**, the inequalities cannot exist, and so the equality must hold.

Alternatively, if you are proving that an inequality holds (say a $\geq$), then you would have to show that the complementary inequality (here a $<$) would lead to an arbitrage.

The construction of arbitrage situations in the proof is often governed by an argument that proceeds along the following lines:

- sell the expensive side (larger side of the inequality), (short-selling if necesary);
- use the proceed to buy the cheaper side (smaller side of the inequality);
- invest the difference at the risk free rate;
- at the end of the time period (expiry) close the position;
- make a profit from (at least) the invested money, so making a guaranteed profit, hence having an arbitratge.

The general idea can be memorised as **sell expensive, buy cheap**

8.10 Bounds for option prices

The following section uses arbitrage arguments to provide bounds on the values of options. These can be checked against the exact values we derive for the relevant options in the subsequent chapters.

8.10.1 Lower bounds on option prices

European call

Suppose you buy one underlying asset and simultaneously write one Eurocall with strike K and expiry T. At time t the portfolio, Π_c, consists of

- long one asset with value S_t, and
- short one Eurocall with value c_t.

At expiry T, portfolio Π_c has a value

$$\begin{aligned} S_T - c_T &= S_T - \max(S_T - K, 0) = \begin{cases} S_T & \text{if } S_T < K, \\ K & \text{if } S_T \geq K. \end{cases} \\ \Rightarrow S_T - c_T &\leq K. \end{aligned}$$

Clearly the Π_c produces a payoff bounded above by the strike price.

But we can say more: for times $t < T$ the value of Π_c cannot exceed the value of K discounted to t: $Ke^{-r(T-t)}$. Why? If it did, then you would instantly:

- short sell as many portfolios Π_c as you like (at zero initial cost) and,
- invest the proceeds ($> Ke^{-r(T-t)}$ for each Π_c) in the bank at the risk free rate.
- At $t = T$ the money in the bank would have grown with interest to be greater than K for each of the short-sold Π_c.
- At $t = T$ each short sold Π_c must be bought back for a cost (from above) of $\leq K$.
- Hence, since the proceeds in the bank are $> K$ for each short-sold Π_c you will have more money in the bank than you need to buy each back.
- Hence you will have made a guaranteed instantaneous risk-free profit at expiry.
- Hence there is a greater than non-zero (certain) chance of making an instantaneous risk free return at expiry, which is arbitrage and therefore not allowed (under our assumptions).

Hence we must have

$$S_t - c_t \leq Ke^{-r(T-t)} \qquad \Rightarrow \qquad c_t \geq S_t - Ke^{-r(T-t)}, \qquad\qquad t \leq T.$$

This gives us a lower bound for c_t.

European put

A similar argument can be used for put options:

Suppose portfolio Π_p contains

- long one asset with value S_t, and
- long one Europut with value p_t.

At time T portfolio Π_p will be worth

$$\begin{aligned} p_T + S_T &= \max(K - S_T, 0) + S_T = \begin{cases} K & \text{if } S_T \leq K, \\ S_T & \text{if } S_T > K \end{cases} \\ \Rightarrow p_T + S_T &\geq K. \end{aligned}$$

Clearly the Π_p produces a payoff bounded below by the strike price. Thus for times $t < T$ the value of Π_p cannot fall below the value of K discounted to t: $Ke^{-r(T-t)}$, i.e.,

$$p_t + S_t \geq Ke^{-r(T-t)}, \qquad t \leq T.$$

Why?

Suppose that at some time $t < T$ instead we had

$$p_t + S_t < Ke^{-r(T-t)},$$

then:

- At time $t < T$, an investor would borrow $Ke^{-r(T-t)}$ at the risk-free rate r and buy $\Pi_p = p_t + S_t$.
- At $t = T$ the value of Π_p would have grown, with certainty to $p_T + S_T > K$.
- The investor would sell the Π_p and receive more than K.
- The debt from the loan would have risen to $Ke^{-r(T-t)} \times e^{r(T-t)} = K$.
- The investor pays this back with the proceeds of the sale of Π_p at T.
- As the proceeds of the sale exceed the debt K, the investor has made a risk-free (here) guaranteed profit at T.
- Hence, there is an arbitrage and so $p_t + S_t$ can never be less than $Ke^{-r(T-t)}$ for $t < T$.

Hence we must have

$$p_t + S_t \geq Ke^{-r(T-t)} \qquad \Rightarrow \qquad p_t \geq Ke^{-r(T-t)} - S_t, \qquad\qquad t \leq T.$$

This gives us an lower bound for p_t.

The lower bound for an American put option can be increased above that derived above for a European put option. Since early exercise is always possible it turns out that we have

$$P_t \geq K - S_t, \qquad t \leq T.$$

8.10.2 Upper bounds on option prices

European call

A call option gives the holder the right to buy the underlying share for a certain price. The payoff at expiry is $\max(S_T - K, 0)$. Hence at expiry T, the option is always less than the value of the share S_T:

$$c_T \leq S_T.$$

Therefore the value of the call option must be less than or equal to the value of the share:

$$c_t \leq S_t.$$

Why?

Suppose that at some time $t < T$ instead we had

$$c_t \geq S_t,$$

then:

- At time $t < T$, an investor would write (sell) one call option (at no initial cost) and receive c_t from the option buyer.
- The investor would then use $c_t > S_t$ to buy a share for S_t, investing the difference $(c_t - S_t)$at a risk-free rate r in the bank.
- At $t = T$, if $S_T > K$, the option buyer would exercise and demand a payoff of $S_T - K$. But $S_T - K < S_T$, so the investor sells the share they bought for S_T, and pays off the option-buyer with a further profit of $S_T - K$.
- At $t = T$, if $S_T \leq K$, the option buyer would not bother to exercise, the call option expires worthless and the investor would be able to sell the share and receive S_T in profit.
- In either scenario, $S_T > K$, $S_T \leq K$, the investor receives a profit, since even if in the latter case, $S_T = 0$, the investor still makes a profit from the money invested in the bank $(c_t - S_t)e^{r(T-t)}$ at the risk-free rate.
- Hence, there is an arbitrage and so c_t can never be greater than or equal to S_t.

European put

For a European put option the maximum value obtainable at expiry is the strike price K. Hence we have

$$p_T \leq K.$$

Furthermore, the value at $t < T$ satisfies:

$$p_t \leq Ke^{-r(T-t)} \qquad t \leq T.$$

Why?

Suppose that at some time $t < T$ instead we had

$$p_t > Ke^{-r(T-t)},$$

then:

- At time $t < T$, an investor would write (sell) one put option (at no initial cost) and receive p_t from the option buyer.
- The investor would then invest $p_t > Ke^{-r(T-t)}$ at a risk-free rate r in the bank.
- At $t = T$ this investment would have grown to $p_t e^{r(T-t)} > K$.
- At $t = T$, if $K > S_T$, the option buyer would exercise and demand a payoff of $K = S_T$. But $K - S_T < K < p_t e^{r(T-t)}$. Hence the investor can pay off the option buyer with the money invested in the bank and still have a profit.
- At $t = T$, if $K \leq S_T$, the option buyer would not bother to exercise, the call option expires worthless and the investor would have a profit from the investment in the bank of $p_t e^{r(T-t)}$.
- In either scenario, $S_T > K$, $S_T \leq K$, the investor receives a profit.
- Hence, there is an arbitrage and so p_t can never be greater than $Ke^{-r(T-t)}$.
- Thus we must have

$$p_t \leq Ke^{-r(T-t)} \qquad t \leq T.$$

For certain types of stochastic model for S_t we find that we are able to write down explicit formulae for the prices of European call and put options.

From this section we can also see that it is possible to give bounds on American call options on a non- dividend-paying stock. On the other hand, the possibility of early exercise of an American put option presents us with much more complexity. As a result of this there is no corresponding explicit formula for the price of an American put option.

8.11 Put-call parity

Consider the argument we used to derive the lower bounds for European call and put options on a non-dividend-paying stock, with the same expiry and strike price K. This used two portfolios:

- **A:** one call plus cash of $Ke^{-r(T-t)}$;
- **B:** one put plus one share.

Both portfolios have a payoff at the time of expiry $t = T$ of the options of $\max(K, S_T)$. Hence we have

$$c_T + K = p_T + S_T.$$

What about $t < T$? European options cannot be exercised before expiry, but they can still be bought and sold with values c_t and p_t. An amount of cash K at $t = T$ has a value $Ke^{-r(T-t)}$ at $t < T$. Using these facts and the equality of A and B at expiry we can conjecture that the following portfolios should also have the same value at any time $t < T$:

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

This relationship is known as **put-call-parity**.

We can prove this result by showing that in the absence of the equality, arbitrage possibilities exist. We shall demonstrate the existence of arbitrage here for one of the inequalities. The demonstration that the second also leads to an arbitrage is left as an exercise for the student.

- Suppose that at some time $t < T$ instead we had

  $$c_t + Ke^{-r(T-t)} > p_t + S_t.$$

  then:

  - At time $t < T$, an investor would
    - write (sell) one call option (at no initial cost) and receive c_t from the option buyer;
    - borrow $Ke^{-r(T-t)}$ from the bank;
    - use this revenue $c_t + Ke^{-r(T-t)}$ to purchase a put option (on same underlying, with same expiry and strike K) and a share at a total cost of $p_t + S_t$;
    - invest the positive difference $c_t + Ke^{-r(T-t)} - (p_t + S_t) > 0$ in their savings account.
  - At $t = T$ the liabilities of the investor are $-c_T$ (value of the call option) and $-K$ (loan now due to bank with interest).
  - At $t = T$ the assets of the investor are p_T (value of the put option) and S_T (value of the share), plus the savings, which, with interest, has grown to $\left[c_t + Ke^{-r(T-t)} - (p_t + S_t)\right] e^{r(T-t)} > 0$.

- At $t = T$, the net position of the investor is thus

$$-c_T - K + p_T = S_T = \begin{cases} -(S_T - K) - K + 0 + S_T = 0, & S_T > K \\ 0 - K + (K - S_T) + S_T = 0, & S_T \leq K \end{cases}$$

  plus the savings $\left[c_t + Ke^{-r(T-t)} - (p_t + S_t)\right] e^{r(T-t)} > 0$.

- Hence in either possible scenario, $S_T > K$ or $S_T \leq K$, the investor does not only cover the payoff from the call option they have written and repay the loan they have taken out, but also ends up with a guaranteed positive, profit.
- Hence, there is an arbitrage and so $c_t + Ke^{-r(T-t)}$ can never be greater than $p_t + S_t$.

- A similar argument can be carried out for the case

$$c_t + Ke^{-r(T-t)} < p_t + S_t$$

  (exercise for student).

- Since both inequalities are excluded by the principal of no arbitrage, and using the result at $t = T$, we may thus conclude that

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

Although the put-call parity formula expresses an important relationship between the values of calls c_t and puts p_t in terms of known quantities, the discounted stroke price $Ke^{-r(T-t)}$ and price of the underlying S_t at the same time, it does not tell us what c_t and p_t are individually. To calculate values for c_t and p_t we require a model.

In all of these sections, the pricing of derivatives is based upon the principle of **no arbitrage**.

Note that we have made very few assumptions in arriving at these results. No model has been assumed for stock prices. All we have assumed is that we will make use of buy-and-hold investment strategies. Any model we propose for pricing derivatives must, therefore, satisfy both put-call parity and the forward-pricing formula. If a model fails one of these simple tests then it is not arbitrage free.

In the next few chapters we will turn to the actual detailed process of valuing the options exactly. This will involve use of technical mathematical processes. Before we do so, we should how an arbitrage argument can be used to value one type of derivative product, namely a **forward contract**.

8.12 Forward contracts and forward pricing

A **forward contract** is a binding agreement between parties to exchange a set of assets (e.g., shares) at an expiry date T in the future for an agreed price K given that the asset price now, at $t = 0$, isS_0. The key difference from options is that **the option to exercise**

or not is unavailable to the holder: the contract must be fulfilled regardless of market conditions or asset price at expiry.

Forward contracts are the most simple form of derivative contract. It is also the most simple to price in the sense that the forward price can be established without reference to a model for the underlying share price.

Suppose that, besides the underlying share, we can invest in a cash account which earns interest at the continuously compounding rate of r per annum. Recall that the forward price K should be set at a level such that the value of the contract at time 0 is zero (that is, no money changes hands at time 0).

- **Proposition:**

  The fair, or economic forward price, is

  $$K = S_0 e^{rT}.$$

  There are (at least) two ways to prove this:

- **Proof 1:**

  - Suppose, first, that the writer has set the forward price at

    $$K = S_0 e^{rT},$$

    and they

    * issue one forward contract;
    * borrow an amount S_0 in cash (subject to interest at rate r);
    * buy one share, with which to fulfil the contract at T.

    The net cost to the writer time $t = 0$ is then zero.

    At time T the writer will have:

    * one share worth S_T on the open market;
    * a cash debt of $S_0 e^{rT}$;
    * a contract to sell the share at the forward price K.

    Therefore the writer hands over the one share to the holder of the forward contract and receives K. At the same time the writer repays the loan: an amount $S_0 e^{rT}$. Since $K = S_0 e^{rT}$ the writer has made a profit of exactly 0. There is no chance of losing money on this transaction, nor is there any chance of making a positive profit. It is a **risk-free trading strategy**, but there has been no arbitrage.

  - Now suppose instead that

    $$K > S_0 e^{rT}.$$

    and the writer

* issues one forward contract;
* borrows an amount S_0 in cash (subject to interest at rate r);
* buys one share.

The net cost at time $t = 0$ to the writer is zero.

At time T they will have:

* one share worth S_T on the open market;
* a cash debt of S_0e^{rT};
* a contract to sell the share at the forward price K.

Therefore the writer hands over the one share to the holder of the forward contract and receive K. At the same time they repay the loan: an amount S_0e^{rT}. Since $K > S_0e^{rT}$ they have made a guaranteed profit having made no outlay at time 0.

This is **arbitrage**: that is, for a net outlay of zero at time 0 they had a probability of 0 of losing money and a strictly positive probability (1) of making a profit greater than zero.

Instead of issuing one contract at this price, the writer could issue lots of them and make a fortune. In practice a flood of sellers would come in immediately, pushing down the forward price to something less than or equal to S_0e^{rT} . In other words the arbitrage possibility could exist briefly but it would disappear very quickly before any substantial arbitrage profits could be made.

– Now suppose that

$$K < S_0e^{rT}.$$

We follow the same principles: at time $t = 0$. The buyer would

* buy one forward contract;
* short sell one share at a price S_0;
* invest an amount S_0 in cash.

The net value to the buyer at time $t = 0$ is zero.

At time T the buyer would

* have cash of S_0e^{rT};
* pay K ($K < S_0e^rT$) for one share after which our net holding of shares is zero.
* have $S_0e^{rT} - K > 0$ in cash

In other words there would be an instantaneous risk free profit for the buyer. Again this is an example of **arbitrage**, meaning that we should not, in practice, find that $K < S_0e^{rT}$.

- **Proof 2:**

  Let K be the forward price. Now compare the setting up of the following portfolios at $t = 0$:

  - **A:** one long forward contract.
  - **B:** borrow Ke^{-rT} cash and buy one share at S_0.

  If we hold both of these portfolios up to time T, then both have a value of $S_T - K$ at T. By the principal of no arbitrage these portfolios must have the same value at all times before T (otherwise someone could hold a combination of both portfolios that would lead to a guaranteed profit at a time $t < T$). In particular, at time $t = 0$, portfolio **B** has value $S_0 - Ke^{-rT}$. which must equal the value of the forward contract. This can only be zero (the value of the forward contract at $t = 0$) if $K = S_0 e^{rT}$.

Note that if transaction costs are included, for example it costs the buyer an amount F_0 to enter into the contract (for example to pay for the cost of drawing up the contract), then the fair deliver price changes to

$$K = (S_0 - F_0)\, e^{rT}.$$

This is covered on one of the problem sheets. Before you look at the answers, see if you can construct an arbitrage argument for this yourself.

8.13 American call option

The lower bound we derived above for the European call option,

$$c_t \geq S_t - Ke^{-r(T-t)}, \qquad t \leq T,$$

is extremely useful in establishing a key theoretical result associated with the American call option C_t on an asset that is not paying dividends.

- The payoff of the American call option is:

  $$\max(S_t - K, 0), \qquad t \leq T.$$

  Note that this implies that $C_t \geq 0$. If $C_t < 0$, then an investor would buy a call option for a negative price (i.e., be paid to receive it), then immediately exercise the option for either 0 ($S_t \leq K$) or $S_t - K > 0$, ($S_t > K$). Either way this is arbitrage as a guaranteed profit is made, so can't happen. Hence $C_t \geq 0$.

- Anything you do with a European call option on a non-dividend paying underlying asset, you can do with an American call option on the same asset. However, the American call option has more flexibility than the European in that there is an option to exercise early. This flexibility cannot have a negative value, although it may have a zero value. Hence an American call option cannot have less value than a European one.

  Hence

  $$C_t \geq c_t.$$

- Now since
$$c_t \geq S_t - Ke^{-r(T-t)}$$
we have
$$C_t \geq c_t \geq S_t - Ke^{-r(T-t)} > S_t - K.$$

- Thus
$$C_t > S_t - K.$$
Now think what this means. For any time $t \leq T$ the payoff from an American call option is:
$$S_t > K: \qquad \max(S_t - K, 0) = S_t - K < C_t, \tag{8.1}$$
$$S_t \leq K: \qquad \max(S_t - K, 0) = 0 < C_t. \tag{8.2}$$
Hence the payoff of the American call is **less** than the value of holding it, throughout the lifetime of the option: more money can be made by selling the American call option than by holding it! Hence if the holder choses not to sell, then **there is no advantage in the early exercise of the American option**!

- The final result follows from the realisation that if there is no extra value in early exercise of the American call, then you might as well hold a European call. Hence we have
$$C_t = c_t.$$

These results are **not** true if the underlying asset pays a dividend: the time that the American call option has been held affects the amount of dividends that have been paid, and so the overall value of the option. This complicates the issue and so it is no longer obvious that it is not advantageous to exercise early. These results are also not true for an American put.

8.14 What happens in reality?

Of course in practice arbitrage arising from incomplete information such lack of knowledge of changes in the spot price S_0 when the contract was written (market friction), commerical pressures (loss leaders), differences in timing of the valuations, currency fluctuations, interest rate differences (e.g., Icelandic investors borrowing at low rates in Japan to invest at higher rates in the UK), differences in government or market rules, differences in transaction costs, or even plain errors, etc.

That is why trading institutions employ highly paid people called **arbitrageurs**, whose job is to seek out an exploit such "loophole" opportunities.

It is still a fair assumption though that if enough people spot an arbitrage the market will eventually move to eliminate the opportunity. Hence it is an assumption we shall make henceforth.

Chapter 9

Valuation of Options in Continuous time: The Black-Scholes PDE approach

9.1 Introduction

In this chapter we show how to value options using an approach based on continuous time partial differential equations approach. The method is based on the (Nobel prize-winning) work of Black and Scholes, who showed that, under certain assumptions, the value of **any** option satisfies the **same** PDE. The stated payoff from each option generate final, boundary or other conditions that, when taken with appropriate general solutions of the Black-Scholes PDE, determine the values of the specific options.

9.2 Assumptions

In this section we will show how to derive the price of a European call or put option using a model under which share prices evolve in continuous time and are characterised at any point in time by a continuous distribution rather than a discrete distribution.

We shall make the following assumptions

1. The asset price follows the lognormal random walk given by geometric Brownian motion

   $$dS_t = S_t(\mu dt + \sigma dB_t)$$

   where B_t is a standard Brownian motion.

   From previous work we know that this SDE has a "solution"

   $$S_t = S_0 \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma B_t\right]$$

This can be checked by substitution into Itô's lemma.

2. There is no **arbitrage** and therefore all risk free portfolios earn the same return.

3. The (risk free) interest rate r and the volatility σ are known functions of time for the lifetime of the option and the same for borrowing or lending. Often we will also assume that they are constant.

   Investors can hold a risk-free cash bond with price $\mathcal{B}_t$, at time t. The evolution is governed by the ordinary differential equation:

$$d\mathcal{B}_t = r\mathcal{B}_t dt,$$

   where r is the (assumed-to-be) constant risk-free rate of interest, so that

$$\mathcal{B}_t, = \mathcal{B}_0 \exp rt.$$

4. Unlimited short selling is permitted.

5. There are no taxes or transaction costs. (These are seriously simplifying assumptions and will be revisited later.)

6. The asset may be traded continuously and in infinitesimally small amounts.

7. The asset pays no dividends (this can be relaxed as detailed later on).

This second assumption is crucial in allowing us to develop an equation for the value of a derivative. It has many interpretations and there are many different ways of thinking of it, but probably the simplest is to simply assume

If options were overpriced then nobody would buy them as the money would make more in the bank; if they were underpriced then nobody would use banks and everybody would buy nothing but options.

The key general implication of the underlying assumptions is that the market in the underlying share is complete: that is, all derivative securities have payoffs which can be replicated.

This consequence is at odds with the real world and implies problems with the underlying assumptions.

It is clear that each of these assumptions is unrealistic to some degree, for example:

- Share prices can jump. This invalidates assumption 1. since geometric Brownian motion has continuous sample paths. However, hedging strategies can still be constructed which substantially reduce the level of risk.

- The risk-free rate of interest does vary and in a unpredictable way. However, over the short term of a typical derivative the assumption of a constant risk-free rate of interest is not far from reality. (More specifically the model can be adapted in a simple way to allow for a stochastic risk-free rate, provided this is a predictable process.)

- Unlimited short selling may not be allowed except perhaps at penal rates of interest. These problems can be mitigated by holding mixtures of derivatives which reduce the need for short selling. This is part of a suitable risk management strategy.
- Shares can normally only be dealt in integer multiples of one unit, not continuously and dealings attract transaction costs: invalidating assumptions 4, 5 and 6. Again we are still able to construct suitable hedging strategies which substantially reduce risk.
- Distributions of share returns tend to have fatter tails than suggested by the log-normal model, invalidating assumption 1.

Despite all of the potential flaws in the model assumptions, analyses of market derivative prices indicate that the Black-Scholes model does give a very good approximation to the market at times away from the settlement date.

It is worth stressing here that all models are only approximations to reality. It is always possible to take a model and show that its underlying assumptions do not hold in practice. This does not mean that a model has no use. A model is useful if, for a specified problem, it provides answers which are a good approximation to reality or if it provides insight into underlying processes. In this respect the Black-Scholes is a good model since it gives us prices which are close to what we observe in the market (despite the fact that we can criticise quite easily the individual assumptions) and because it provides insight into the usefulness of dynamic hedging.

9.3 Derivation of Black-Scholes equation

Under these assumptions, we proceed as follows.

Let $V(S_t, t)$ be the price at time t of **any** option given:

- the current share price is S_t,
- the time of maturity is $T > t$.

In what follows, we shall drop the subscript t from S_t to avoid confusion with derivatives.

By Ito's lemma, the random walk followed by V is given by

$$dV = \sigma S V_S dW + \left(\mu S V_S + \frac{1}{2}\sigma^2 S^2 V_{SS} + V_t\right) dt.$$

where we now use the subscript notation to mean, e.g.,

$$V_S = \frac{\partial V}{\partial S_t}, \qquad V_{SS} = \frac{\partial^2 V}{\partial S_t^2}, \qquad \text{etc.}$$

NEED TO CHANGE THIS AROUND SO THAT IT"S FROM THE SELLER'S PERSPECTIVE:

$$\Pi = -V + \delta S$$

THEN EVERYTHING SQUARES UP AND THE WRITE HAS TO HOLD (NOT SHORT) SHARES FOR A CALL.

THE FOLLOWING IS FROM A OPTION HOLDER'S PERSPECTIVE.

Now suppose that we construct a portfolio, Π, which, at time t, consists of

- long one option
- number $-\Delta$ of the underlying asset.

At the moment Δ (pronounced "delta") is just a number. If $\Delta > 0$, then the asset is **short sold** in this portfolio. We shall choose a value for Δ in a moment.

The value of this portfolio at t is

$$\Pi = V + S(-\Delta).$$

and so the jump in the value of the portfolio in one infinitesimal time step dt is

$$d\Pi = dV - \Delta dS = dV - \Delta[\sigma S dW + \mu S dt].$$

Using the Itô result above and substituting for dV in this expression we find that

$$d\Pi = \sigma S(V_S - \Delta)dX + \left(\mu S V_S + \frac{1}{2}\sigma^2 S^2 V_{SS} + V_t - \mu \Delta S\right) dt.$$

First magic step.

The random part of the change $d\Pi$ (at this leading order of Itô approximation) can be completely eliminated by **choosing**

$$\Delta = V_S,$$

(thus defining "the delta") whence

$$d\Pi = \left(V_t + \frac{1}{2}\sigma^2 S^2 V_{SS}\right) dt.$$

With this choice of Δ, at time t (only) the portfolio will evolve in a risk-free way.

Second, and Nobel prize-winning, magic step.

Now if the portfolio evolves instantaneously in a risk-free way, by our arbitrage argument, this change in the value of the portfolio **must** be equal to the return in a time dt of an amount Π invested in a riskless asset. This is evidently $r\Pi dt$ and thus

$$r\Pi dt = \left(V_t + \frac{1}{2}\sigma^2 S^2 V_{SS}\right) dt.$$

But

$$\Pi = V - \Delta S = V - SV_S$$

and thus

$$r(V - SV_S)dt = \left(V_t + \frac{1}{2}\sigma^2 S^2 V_{SS}\right) dt.$$

Dividing by dt and rearranging now gives

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV = 0.$$

This is the (in)famous **Black-Scholes** (BS) equation for the fair value V of an option which first appeared in a 1973 publication and ultimately won Black and Scholes the Nobel prize for Economics. We note a number of important facts about this equation:

- It is very similar to the heat diffusion equation

  $$u_t = \kappa u_{xx}.$$

  where u is the temperature (as a function of position x and time t) of a bar with thermal conductivity κ. This is basically because under the geometric Brownian motion/lognormal assumption, the stock market moves in exactly the same way that heat diffuses, i.e., a random process that slowly spreads out its influence.

- The equation is **linear:** a sum of two solutions of the BS equation is itself a solution.

- The equation is **very** general: subject to the assumptions stated earlier, **any** derivative product whose value depends only upon S and t must satisfy the Black-Scholes equation.

- The growth parameter μ does not appear in Black-Scholes. The value of an option is thus independent of how rapidly an asset grows. It is only the interest rate r and the volatility σ that matter for the value of an option.

- We have yet to specify meaningful boundary conditions for the equation. The particular boundary conditions we impose define the option we want to price. For some boundary conditions (e.g., the Eurocall and Europut) the Black-Scholes equation may be solved in closed form; for most others (e.g., all but one of the American options) a numerical solution is required.

- It turns out that the BS equation is a sort of "backward" heat equation since the sign of the "κ" is negative compared to the standard heat equation. We should therefore normally expect to have to specify two boundary conditions for BS and one "final" (rather than an initial) condition. For example, typically we might expect to have to specify $V(S_1,t) = V_1(t)$ and $V(S_2,t) = V_2(t)$, and $V = V_f(S)$ at $t = T$ say.

9.3.1 Boundary conditions for the Black-Scholes equation

Now we know the basic equation for option pricing we need to consider the types of boundary conditions to apply. Each type of option will have a different boundary condition. Each suitable set of boundary conditions will value a different option.

Let's start off by examining the simplest case, namely the European Call. The value V of this option satisfies the Black-Scholes equation

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV = 0$$

This is a second order linear homogeneous (parabolic) partial differential equation in variables S and t. This type of equation usually requires three boundary conditions (believe me it does). Often this is an initial condition that specifies $V(S,0)$ and two boundary conditions that specify the value of V on the boundaries of the domain e.g., $S = a$ and $S = b$ so that $V(a,t)$ and $V(b,t)$ are known.

The difference with the Black-Scholes system is that we specify not an **initial condition**, but a **final condition**. Why? Well evidently there is one instant of time when we know **for sure** what the value of the option must be, for at expiry when $t = T$ we know that the value of a European call is

$$V(S,T) = \max(S-K,0)$$

where as usual K is the strike price. Note that really we would like to work back from this value at expiry to price the option at an earlier time. Hence a final condition is sensible.

How about boundary conditions? Well there are two: first, it's clear that our dreams of wealth will be dealt a fatal blow if ever the asset price drops to zero, if ever $S = 0$ then the share price can never change again). Thus

$$V(0,t) = 0.$$

On the otherhand, as $S \to \infty$ it is clear that the strike price becomes more and more irrelevant. The second boundary condition is thus

$$V(S,t) \to S \text{ as } S \to \infty.$$

For the Europut, the analogous conditions would be

$$\begin{aligned} V(S,T) &= \max(K-S,0) \\ V(0,t) &= Ke^{-r(T-t)} \\ V(S,t) &\to 0 \text{ as } S \to \infty. \end{aligned}$$

For American options we would have to specify the value of the payoff at the time that the option was exercised, say at $t = t^*$ where $0 < t^* < T$. Since, *a priori* we don't know where t^* is, we don't know where we can apply the final condition. This is called a **free boundary** problem. In general such problems require an extra boundary condition to locate the free boundary, and do not submit themselves to analytic solutions in terms of simple formulae. Solutions can only really be obtained numerically. The same is true for other types of exotic options.

Nevertheless, armed with this knowledge, how would we solve the Black-Scholes equation for the simple European options?

The following section is not examinable in the course, but is included to explain how the Black Scholes equation is solved. You can jump to the solutions in section 13.5 below if you wish.

9.4 Procedure for solving the Black-Scholes equation

9.4.1 Solutions of the Heat Equation

It turns out that in nearly every case the best way to proceed is first to transform the original Black-Scholes problem in some way to the heat diffusion equation, and then to use the existing extensive knowledge about solutions to this equation. Accordingly, before we can proceed, we will have to spend a little time building up our knowledge about different ways of solving the heat diffusion equation.

The one dimensional heat diffusion equation is normally written

$$u_t = \kappa u_{xx}.$$

Here u may be thought of as representing the temperature in a one dimensional rod whose thermal properties are characterised by the parameter κ (which, in reality, is the ratio of the thermal conductivity of the bar to the product of the density and the specific heat of the bar). We shall assume unless otherwise stated that κ is constant.

There are many ways of solving the heat equation. Which method we use in practice tends to depend upon the kind of boundary and initial conditions that are specified. Since it is second order in space and first order in time, it makes sense to specify two boundary conditions and one initial condition. In the cases that we will be interested in, the best way of solving the equation is normally to seek a **similarity solution**. In general there is a massive theory of similarity variables and solutions for general partial differential equations, but the underlying idea is always the same:

We attempt to turn the heat equation, a **partial differential equation** into an **ordinary differential equation** which is posed in terms of a new "similarity" variable η which is some "cunningly-chosen" function of x and t. If we can accomplish this, then the difficulty of the problem is normally reduced, since in general ODEs are easier to solve than PDEs.

So what should we use as a similarity variable for the heat equation? It has long been known that if we define

$$u = U(\eta), \qquad (\eta = \frac{x}{\sqrt{\kappa t}})$$

then

$$u_t = \kappa u_{xx}$$

simplifies to an ODE. We have (using the chain rule)

$$u_t = -\frac{x}{2\sqrt{\kappa}t^{3/2}}U'(\eta)$$

(where $' = d/d\eta$)

$$u_x = -\frac{1}{\sqrt{\kappa t}}U'$$

and

$$u_{xx} = \frac{1}{\kappa t}U''.$$

Thus

$$-\frac{x}{2\sqrt{\kappa}t^{3/2}}U' = \frac{\kappa}{t\kappa}U''$$

and so

$$U'' + \frac{1}{2}\eta U' = 0.$$

Since the variables x and t do not appear explicitly in this last equation, the similarity reduction has succeeded and all that remains is to solve the resulting ODE.

Setting $V = U'$ and then solving the resulting first order equation for V using elementary methods, we find that

$$V(\eta) = Ce^{-\eta^2/4} \qquad (*)$$

and we must now consider what boundary and initial conditions we wish to impose. To start with, let us assume that at $t = 0$ we have $u = 0$, but thereafter we impose $u = 1$ at $x = 0$ and additionally assume that $u \to 0$ as $x \to \infty$. In terms of our new problem formulation, these amount to

$$U(0) = 1, \qquad U(\infty) = 0$$

and we can integrate $(*)$ with respect to η and impose these boundary conditions to get

$$U(\eta) = \frac{1}{\sqrt{\pi}}\int_\eta^\infty e^{-s^2/4}ds.$$

(Note: here we have used the famous result that

$$\int_0^\infty e^{-s^2/4}ds = \sqrt{\pi}).$$

In terms of the original variables, we have therefore shown that a solution to the heat equation is

$$U(x,t) = \frac{1}{\sqrt{\pi}}\int_{x/\sqrt{\kappa t}}^\infty e^{-s^2/4}ds$$

(and this may easily be verified by direct substitution).

Actually, the heat diffusion equation possesses many other similarity solutions; for example, if we try for a solution

$$u = \frac{1}{\sqrt{t}} U(\eta)$$

where η is as defined above, then we find that another solution of the heat equation (with κ set to 1 for convenience, say) is given by

$$u_f(x,t) = \frac{1}{2\sqrt{\pi t}} e^{-x^2/4t}$$

(Again, this may be checked by just substituting it into the equation; alternatively, try for a solution of the heat equation of the form $u = f(\eta)/\sqrt{t}$ where $\eta = x^2/t$. This gives $f + (4+2\eta)f' + 8\eta f'' = 0$, and trying a solution $g(\eta) = \exp(k\eta)$ now gives that k must satisfy $(1+4k)(1+2k\eta) = 0$. Hence $k = -1/4$.)

This "fundamental" solution actually turns out to be very useful indeed, as the constants in it have been chosen so that the integral of u_f with respect to x from $-\infty$ to ∞ is 1.

In fact, the real advantage of this form of the solution is that it may now easily be generalised, for we note that (a) whenever x is non-zero, u_f is zero at time $t = 0$ (the exponential always "wins" over the power), (b) when x *is* zero, u_f is infinite at time $t = 0$, and (c) The area under the curve u_f is 1. Now (a)-(c) are exactly the properties that characterise the "delta function" $\delta(x)$. Thus

$$u_f(x, 0) = \delta(x).$$

Now consider the function

$$u(x,t) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} u_0(s) e^{-(x-s)^2/4t} ds.$$

This clearly satisfies the heat diffusion equation, since it is simply a "superposition" of solutions (and, since the heat diffusion equation linear, we may add solutions together to create other solutions). Further, we note that, because of the δ-function behaviour of the fundamental solution above, we have

$$u(x, 0) = \int_{-\infty}^{\infty} u_0(s)\delta(s-x) ds = u_0(x).$$

What we have shown, therefore, is that the solution to

$$u_t = u_{xx} \qquad (-\infty < x < \infty, \quad t \geqslant 0)$$

with initial data

$$u(x, 0) = u_0(x)$$

is

$$u(x,t) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} u_0(s) e^{-(x-s)^2/4t} ds.$$

The fundamental solution is also known as the **Green function** of the heat equation. It is a general mathematical principle that the convolution of a Green function of a equation with any suitably integral initial data will generate the actual solution of the equation that satisfies the given initial data. The integral solution above is just that type of convolution.

Note that this integral solution can also be interpreted as an expectation of the initial data $u(x, 0) = u_0(x)$ under a normal probability density function

$$p(s) = \frac{1}{2\sqrt{\pi t}} e^{-(x-s)^2/4t}.$$

Of course, there are many other different ways of solving the heat diffusion equation. As far as option valuation is concerned, however, we now have what we require.

9.4.2 Turning Black-Scholes into the heat equation

We're now going to value the European call by using a sequence of simple transformations to reduce it to an example the heat diffusion equation whose solution we already know. Denoting the European call value by $c(S, t)$ we must solve

$$c_t + \frac{\sigma^2 S^2}{2} c_{SS} + rSc_S - rc = 0$$

subject to

$$c(0, t) = 0, \qquad c(S, t) \to S \quad (S \to \infty) \qquad c(S, T) = \max(S - K, 0).$$

To begin we set

$$S = Ke^x, \qquad t = T - \tau/(\frac{1}{2}\sigma^2), \qquad c = Kv(x, \tau)$$

so that the independent variables (t, S) transform to (τ, x). We have

$$\frac{\partial}{\partial t} = -\frac{\sigma^2}{2}\frac{\partial}{\partial \tau}, \qquad \frac{\partial}{\partial S} = \frac{e^{-x}}{K}\frac{\partial}{\partial x}$$

and thus

$$c_t = -\frac{\sigma^2}{2} K v_\tau$$
$$c_S = e^{-x} v_x$$
$$c_{SS} = \frac{e^{-x}}{K}(-e^{-x} v_x + e^{-x} v_{xx}).$$

The governing equation becomes

$$-\frac{\sigma^2}{2} K v_\tau + \frac{\sigma^2}{2} K^2 e^{2x} \frac{e^{-x}}{K}(-e^{-x} v_x + e^{-x} v_{xx}) + rKe^x e^{-x} v_x - rKv = 0$$

and thus

$$v_\tau + v_{xx} + (k-1)v_x - kv = 0$$

where

$$k = r/(\frac{1}{2}\sigma^2).$$

The boundary conditions transform to

$$v(-\infty,\tau) = 0, \qquad v(x,\tau) \to e^x \quad (x \to \infty), \qquad v(x,0) = \max(e^x - 1, 0)$$

and we note that the equation for v depends **only** upon the dimensionless parameters $k = r/\frac{1}{2}\sigma^2$ and the non-dimensional time to expiry $\frac{1}{2}\sigma^2 T$ - the only two genuinely independent parameters in the problem. We now reduce the equation for v to the heat diffusion equation by further setting

$$v = e^{\alpha x + \beta\tau} u(x,\tau)$$

where α and β are to be suitably chosen. We have

$$v_\tau = \beta e^{\alpha x+\beta\tau} u + e^{\alpha x+\beta\tau} u_\tau$$

$$v_{xx} = \alpha^2 e^{\alpha x+\beta\tau} u + 2\alpha e^{\alpha x+\beta\tau} u_x + e^{\alpha x+\beta\tau} u_{xx}$$

and putting it all in and collecting terms gives

$$u_\tau = u_{xx} + (2\alpha + k - 1)u_x + (\alpha^2 + (k-1)\alpha - k - \beta)u.$$

We must therefore take

$$2\alpha + k = 1, \qquad \alpha^2 + (k-1)\alpha - k - \beta$$

and so

$$\alpha = \frac{1-k}{2}, \qquad \beta = \alpha^2 + (k-1)\alpha - k.$$

At last we have reached the heat diffusion equation! The new problem to be solved is

$$u_\tau = u_{xx} \qquad (-\infty < x < \infty, \quad \tau > 0)$$

with

$$u(x,0) = u_0(x) = \max(e^{(k+1)x/2} - e^{(k-1)x/2}, 0).$$

9.4.3 Unravelling the solution

From the penultimate subsection we know how to solve the heat equation (even) with the (strange looking) initial conditions; we simply use the convolution solution. Thus

$$u(x,\tau) = \frac{1}{2\sqrt{\pi\tau}} \int_{-\infty}^{\infty} u_0(s) e^{-(x-s)^2/4\tau} ds$$

where

$$u_0(x) = \max(e^{(k+1)x/2} - e^{(k-1)x/2}, 0).$$

We still have to calculate the integral, however, and "undo" our previous transformations.

To do this, we first set $p = (s - x)/\sqrt{2\tau}$. The integral becomes

$$u(x, \tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u_0(x + p\sqrt{2\tau}) e^{-p^2/2} dp$$

and

$$u_0(x + p\sqrt{2\tau}) = \max[\exp[(k+1)(x + p\sqrt{2\tau})/2] - \exp[(k-1)(x + p\sqrt{2\tau})/2], 0].$$

Now the exponential function is monotonic increasing and $(k+1)/2 > (k-1)/2$ for all k. Thus the quantity $u_0(x + p\sqrt{2\tau})$ is given by

$$\exp[(k+1)(x + p\sqrt{2\tau})/2] - \exp[(k-1)(x + p\sqrt{2\tau})/2]$$

whenever $x + p\sqrt{2\tau} \geqslant 0$, and is otherwise zero. The integrals must therefore be evaluated for $p \geqslant -x/\sqrt{2\tau}$, giving

$$\begin{aligned} u(x, \tau) &= \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{(k+1)(x + p\sqrt{2\tau})/2} e^{-p^2/2} dp \\ &- \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{(k-1)(x + p\sqrt{2\tau})/2} e^{-p^2/2} dp \\ &= I_1 - I_2 \end{aligned}$$

say.

We simplify I_1 by writing

$$\begin{aligned} I_1 &= \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{(k+1)x/2 + (k+1)^2\tau/4} e^{-(p-(k+1)\sqrt{2\tau}/2)^2/2} dp \\ &= \frac{e^{(k+1)x/2 + (k+1)^2\tau/2}}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{-(p-(k+1)\sqrt{2\tau}/2)^2/2} dp. \end{aligned}$$

Now putting $p - (k+1)\sqrt{2\tau}/2 = q$ gives

$$I_1 = \frac{e^{(k+1)x/2 + (k+1)^2\tau/2}}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau} - (k+1)\sqrt{2\tau}/2}^{\infty} e^{-q^2/2} dq$$

and thus, on setting $q = -q$, we find that

$$I_1 = \frac{e^{(k+1)x/2 + (k+1)^2\tau/2}}{\sqrt{2\pi}} \int_{-\infty}^{x/\sqrt{2\tau} + (k+1)\sqrt{2\tau}/2} e^{-q^2/2} dq.$$

At last, this is in a convenient form, for the function

$$\Phi(d_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1} e^{-q^2/2} dq$$

is the cumulative normal distribution function (well known to statisticians, and, more importantly for our purposes, extensively tabulated). Thus

$$I_1 = \exp((k+1)x/2 + (k+1)^2\tau/4)\Phi(d_1)$$

where

$$d_1 = \frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k+1)\sqrt{2\tau}$$

and, since I_2 is just I_1 but with $k+1$ replaced with $k-1$, we have

$$I_2 = \exp((k-1)x/2 + (k-1)^2\tau/4)N(d_2)$$

where

$$d_2 = \frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k-1)\sqrt{2\tau}.$$

Now we have found $u(x,\tau)$; what we really want to know though is $c(S,t)$. We can find this by "undoing" our transformations. We have

$$v = e^{\alpha x + \beta\tau}u(x,\tau)$$

where

$$\alpha = \frac{1-k}{2}, \qquad \beta = -\left(\frac{1+k}{2}\right)^2$$

and

$$c = Kv(x,\tau), \qquad x = \log(S/K), \qquad t = T - \tau/(\frac{\sigma^2}{2}).$$

Thus

$$\begin{aligned} c &= K\exp[(1-k)\log(S/E)/2 + \beta(\frac{1}{2}\sigma^2)(T-t)] \times \\ &\times \left\{\exp[(k+1)x/2 + (k+1)^2\tau/4]\Phi(d_1) - \exp[(k-1)x/2 + (k-1)^2\tau/4]\Phi(d_2)\right\}, \end{aligned}$$

which gives

$$c = K\exp\left[\frac{(1-k)}{2}\log(\frac{S}{K}) - \sigma^2(T-t)\frac{(1+k)^2}{8} + \frac{(1+k)}{2}\log(\frac{S}{K}) + \right.$$
$$\left.\frac{(k+1)^2}{8}\sigma^2(T-t)\right]\Phi(d_1) \quad - \quad K\exp\left[\frac{(1-k)}{2}\log(\frac{S}{K}) - \frac{(1+k)^2}{8}\sigma^2(T-t) + \right.$$
$$\left.\frac{(k-1)}{2}\log(\frac{S}{K}) + \frac{(k-1)^2}{8}\sigma^2(T-t)\right]\Phi(d_2)$$

so that

$$c = K\exp(\log(S/K))\Phi(d_1) - K\exp[-k\frac{\sigma^2}{2}(T-t)]\Phi(d_2).$$

Finally though we have $k = r/(\frac{1}{2}\sigma^2)$, and so

$$c = S\Phi(d_1) - K\exp(-r(T-t))\Phi(d_2)$$

where

$$d_1 = \frac{\log(S/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = \frac{\log(S/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

x	0.0	0.2	0.5	1.0	1.5	2.0	2.5	3.0
$\Phi(x)$	0.500	0.579	0.691	0.841	0.933	0.977	0.994	0.999

Table 9.1: Values of $\Phi(x)$ for various arguments x. Note that values for negative x may also be obtained from these as the function is symmetric.

and our (rather long) journey to the solution for the European Call is over.

Note that the integral involved in the solution cannot be evaluated in closed form, as there is no function which, when differentiated, gives e^{-x^2}. It is thus necessary either to look it up in tables, or to use MAPLE or MATHEMATICA to evaluate it. Table 9.1 gives a few representative values for $\Phi(x)$: others may be obtained using linear interpolation.

9.5 Solutions of the Black-Scholes Equation

9.5.1 Eurocall

The solution for the eurocall (on a non-dividend paying asset) is

$$c(S,t) = S\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$$

where

$$d_1 = \frac{\log(S/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = \frac{\log(S/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

where Φ is the normal cumulative distribution function:

$$\Phi(d_1) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_1} e^{-q^2/2}dq.$$

This can be verified by substitution into the Black-Scholes equation. (Exercise for student!) It is identical to the result obtained by using martingales.

9.5.2 Europut

The Europut option can be solved in a similar way to the Eurocall above. The only thing that changes is the final condition which feeds through into slightly different terms involving the cdf of the normal distribution. The answer is identical to that obtained from the martingale approach:

$$p(S,t) = Ke^{-r(T-t)}\Phi(-d_2) - S\Phi(-d_1).$$

9.6 The delta for calls and puts

Now that we have valued the European vanilla call and put options, we should pause to work out the delta of these options, for this is the crucial quantity as far as hedging is concerned. Remember that the replicating portfolio should hold an amount $-\Delta$ at each stage to ensure that the portfolio value grows as a risk free rate. This is equivalent to ensuring that the valuation is performed under a risk neutral measure $\mathbb{Q}$.

To determine the delta for a Euro-call we start from the expression

$$c(S,t) = S\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$$

where

$$d_1 = \frac{\log(S/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = \frac{\log(S/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

Now from the definition of Δ above have

$$\Delta = \frac{\partial V}{\partial S} = \frac{\partial c(S,t)}{\partial S} = \Phi(d_1) + S\frac{\partial \Phi(d_1)}{\partial S} - Ke^{-r(T-t)}\frac{\partial \Phi(d_2)}{\partial S}.$$

Furthermore

$$\Phi(d_1) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_1} e^{-q^2/2}dq$$

and thus

$$\frac{\partial \Phi(d_1)}{\partial S} = \frac{1}{\sqrt{2\pi}}\frac{\partial d_1}{\partial S}e^{-d_1^2/2} = \frac{1}{\sqrt{2\pi}}\frac{1}{S\sigma\sqrt{T-t}}e^{-d_1^2/2}$$

and also

$$\frac{\partial \Phi(d_2)}{\partial S} = \frac{1}{\sqrt{2\pi}}\frac{1}{S\sigma\sqrt{T-t}}e^{-d_2^2/2}.$$

Thus

$$\begin{aligned}\Delta &= \Phi(d_1) + \frac{1}{\sigma\sqrt{2\pi(T-t)}}e^{-d_1^2/2} - \frac{Ke^{-r(T-t)}}{S\sigma\sqrt{2\pi(T-t)}}e^{-d_2^2/2} \\ &= \Phi(d_1) + \frac{1}{\sigma\sqrt{2\pi(T-t)}}\left[e^{-d_1^2/2} - e^{-\log(S/K)-r(T-t)-d_2^2/2}\right].\end{aligned}$$

Now

$$d_1 - d_2 = \frac{\sigma^2(T-t)}{\sigma\sqrt{T-t}} = \sigma\sqrt{T-t}$$

and thus

$$d_1^2 = d_2^2 + \sigma^2(T-t) + 2d_2\sigma\sqrt{T-t}.$$

So

$$\Delta = \Phi(d_1) + \frac{e^{-d_2^2/2}}{\sigma\sqrt{2\pi(T-t)}}\left[e^{-\frac{\sigma^2}{2}(T-t)-d_2\sigma\sqrt{T-t}} - e^{-\log(S/K)-r(T-t)}\right]$$

$$= \Phi(d_1) + \frac{e^{-d_2^2/2}}{\sigma\sqrt{2\pi(T-t)}}\left[e^{-\frac{\sigma^2}{2}(T-t)-\log(S/K)-r(T-t)+\frac{\sigma^2}{2}(T-t)} - e^{-\log(S/K)-r(T-t)}\right]$$

and thus

$$\Delta = \Phi(d_1),$$

which is a particularly simple result for the delta of a Euro-call.

The delta for a put can be worked out in a similar way, but it is much simpler to use put-call parity;

$$c - p = S - Ke^{-r(T-t)} \qquad \Rightarrow \qquad p = c - S + Ke^{-r(T-t)}$$

Hence for a put

$$\Delta = \frac{\partial p}{\partial S} = \Phi(d_1) - 1.$$

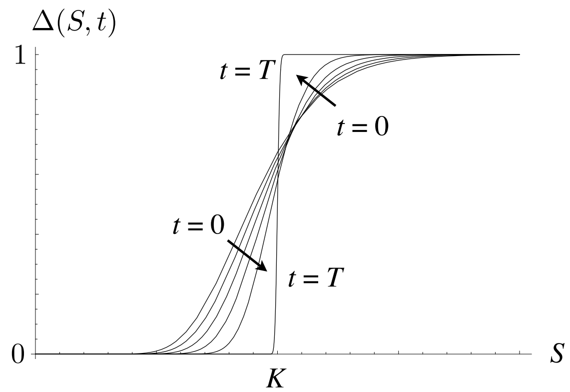

Figure 9.1: The delta of a call option as a function of underlying asset price, plotted for different values of t, $0 < t < T$.

9.7 Forwards and futures

Forwards and futures are also derivative products. As such their fair values should obey the Black-Scholes formula. We have already values a forward in the section on arbitrage, but here we value it again as a further example of the use of Black-Scholes.

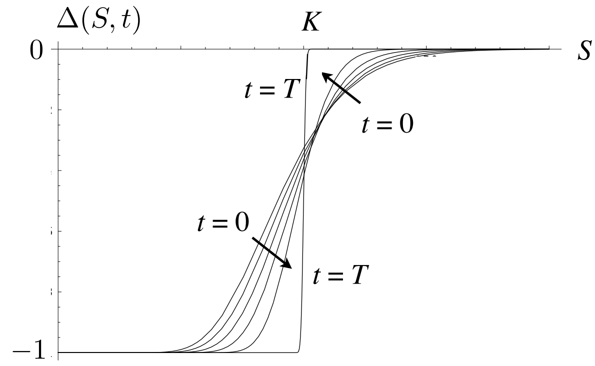

Figure 9.2: The value of a put option as a function of underlying asset price, plotted for different values of t, $0 < t < T$.

The holder of an option, having paid up front for the option, may choose whether or not to exercise at expiry; it is essentially this element of choice that makes the option valuation problem a non-trivial one. A **forward contract** is an agreement between two parties in which one party agrees to buy a specific asset from the other party at a given agreed price (the "forward price") at a specified time (the "delivery date"). No money therefore changes hands until the delivery date, and, unlike an option, there is no "right not to exercise": the deal *must* be done at the delivery date however much either party hates doing it.

The lack of an element of choice actually makes forward contracts much easier to value than options. Suppose that we denote the "fair value" of a forward contract by F and assume that the the contract is agreed at time t_0 when the value of the underlying is $S(t_0)$. Assuming that the interest rate is known, the usual arbitrage argument allows us to determine F.

The party who is short the asset (and so who must deliver at time T) can evidently borrow an amount $S(t_0)$ from the bank, buy the asset now, and use the money F received at exercise to pay back the loan. The fair value of the forward contact is thus simply the cost of the loan, i.e.

$$F = S(t_0)e^{r(T-t_0)} \tag{9.1}$$

The forward contract price (9.1) may be understood in the standard options framework by considering the payoff: at time T the payoff is simply $S - F$ (i.e. what it's worth now minus what you paid for it). Now the Black-Scholes equation

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV = 0$$

with final condition

$$V(S,T) = S - F$$

evidently has the exact solution

$$V(S,t) = S - Fe^{-r(T-t)}$$

(this is easily verified by just putting it into the equation). At time $t = t_0$ the value of the contract is zero (since no money changes hands until expiry) and thus

$$0 = V(S,t_0) = S(t_0) - Fe^{-r(T-t_0)}$$

which once again confirms (9.1)

Finally, how can we modify the theory if the asset pays a constant dividend D_0? We have not considered dividends before, but since we are about to do so, let's introduce the subject by considering dividends for a forward contract. In fact, it is simple to show, using an argument very similar to that above, that in this case,

$$F = S(t_0)e^{(r-D_0)(T-t_0)}.$$

A **future** is very similar to a forward contract, but there are some differences. A forward contract may be set up between any two parties, but futures are usually traded on a specific exchange. The delivery date and the contract size are thus usually subject to the standard rules and customs of the exchange. Also, to protect both parties against non-payment, a "margin requirement" is enforced. Essentially this means that futures contract is evaluated every day and any amount due is immediately paid. In this way the net profit or loss is paid "by installments" over the life of the contract. These technicalities make little difference to valuation, however, and to all intents and purposes the value of a futures contract is still given by (9.1).

9.8 Addition of Options

The Black-Scholes equation is a linear equation. As such, if V_1 and V_2 are solutions of the equation (without the boundary conditions), then so is $\alpha V_1 + \beta V_2$ where α and β are constants. The values of α and β may be determined by applying the boundary conditions. This approach generalises to any number of solutions V_i, $i = 1, 2, \cdots$. The same is true for martingales.

As a result of this linearity property it is often possible to value derivative products with more complicated payoffs by constructing these payoffs from combinations of more basic calls and puts. For example consider a European option with a payoff

$$\text{Payoff at expiry} = \begin{cases} 100 - S_T, & S_T < 100 \\ 0, & 100 \leq S_T \leq 150 \\ S_T - 150, & S_T > 150 \end{cases}$$

- Expiry T: 3 months;
- Current asset price $S = 145$p;
- Risk-free rate $r = 5\%$ per annum;
- Volatility of underlying asset, 20%;
- Underlying asset pays no dividends.
- No early exercise allowed.

This payoff at expiry of this option can be contructed from:

- long one Euro-put, strike price K=100, expiry T= 3 months.
- long one Euro-call, strike price K=150, expiry T= 3 months.

Since the total option must obey Black-Scholes, we can then write down the fair value of this option at any time t as

$$c(S_t, t, K = 150, \sigma, r, T) + p(S_t, t, K = 100, \sigma, r, T)$$

where $\sigma = 0.2, r = 0.05, T = 1/4$.

Exercise for student: Work out the value of this option at $t = 0$ if $S_0 = 145p$.

As a further example, suppose now that the European option had a payoff that was

$$\text{Payoff at expiry} = \begin{cases} 50, & S_T < 100 \\ 150 - S_T, & 100 \leq S_T \leq 150 \\ 0, & S_T > 150 \end{cases}$$

with expiry $T = 3$ months.

Show that this can be valued by

- long one put, strike $K = 150$, expiry $T = 3$ months.
- short one put, strike $K = 100$, expiry $T = 3$ months.

Chapter 10

Further analysis of options in Continuous time:

10.1 Definition of the "Greeks"

10.1.1 The "Greeks"

As well as the delta, there are many other important parameters that may be calculated. These are normally known in the trade as "The Greeks" as they are denoted by..... Greek letters. They measure various different quantities and have great practical importance. Once the fair value of an option has been calculated, they may usually be determined with little difficulty. For example:

- The **Delta**, of any option with value V, as shown above, is defined by

  $$\Delta = \frac{\partial V}{\partial S}$$

  When we consider delta hedging we add up the deltas for the individual assets and derivatives (taking account, of course, of the number of units held of each). If this sum is zero and if the underlying asset prices follow a diffusion then the portfolio is instantaneously risk free.

  If it is intended that the sum of the deltas should remain close to zero (this is what is called delta hedging) then normally it will be necessary to rebalance the portfolio on a regular basis. The extent of this rebalancing depends primarily on **gamma**.

- **The gamma**, Γ, of an option with value V is defined by

  $$\Gamma = \frac{\partial^2 V}{\partial S^2}.$$

  Since we may regard the gamma as the S-derivative of the delta, evidently the gamma represents the sensitivity of the delta to the price of the underlying; its obvious interpretation is as a measure of how often a position has to be rehedged.

If a portfolio has a high value of Γ then it will require more frequent rebalancing or larger trades than one with a low value of Γ. It is recognised that continuous rebalancing of the portfolio is not feasible and that frequent rebalancing increases transaction costs. The need for rebalancing can, therefore, be minimised by keeping Γ close to zero.

- **The "theta"**, Θ, of an option with value V is defined by

$$\Theta = \frac{\partial V}{\partial t}$$

  and measures the rate of change of the option value with time.

  Since time is a variable which advances with certainty (at least in the financial world) it does not make sense to hedge against changes in t in the same way as we do for unexpected changes in the price of the underlying asset.

Delta, gamma and theta all measure the sensitivity of the option to various changes in the **independent** variables; other "Greeks" that are commonly used measure the sensitivity with respect to the **parameters** in the problem. For example:

- **The "vega"** of an option with value V is defined by

$$\text{vega} \quad = \frac{\partial V}{\partial \sigma}$$

  and measures the sensitivity of the option value to volatility σ. Unlike the other names Vega is not a Greek letter and is often denoted by the letter ν. For the underlying asset, S_t, $\nu = 0$.

  The value of a portfolio with a low value of vega will be relatively insensitive to changes in volatility. Put another way: it is less important to have an accurate estimate of σ if vega is low. Since σ is not directly observable, a low value of vega is important as a risk- management tool. Furthermore, it is recognised that σ can vary over time. Since many derivative pricing models assume that σ is constant through time the resulting approximation will be better if ν is small.

- **The "rho"** ρ of an option is defined by

$$\rho = \frac{\partial V}{\partial r}$$

  and measures the sensitivity of the option value to changes in interest rate. The risk-free rate of interest can be determined with a reasonable degree of certainty but it can vary by a small amount over the (usually) short term of a derivative contract. As a result a low value of ρ reduces risk relative to uncertainty in the risk-free

Note that in terms of the Greeks, the Black-Scholes PDE

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV = 0,$$

can be written as

$$\Theta + \frac{1}{2}\sigma^2 S^2 \Gamma + rS\Delta - rV = 0.$$

This is a wonderfully useless observation, since although it has removed any mention of partial derivatives and is therefore pleasing to non-mathematical City-types, you have to know three of the unknown quantities $\Theta, \Gamma, \Delta, V$ to know the fourth. As this is unlikely this notational change is somewhat numerically pointless.

About the only observation you can make from this is that if Δ and Γ of a portfolio are both zero, then Θ is the risk-free rate of growth for the portfolio.

10.1.2 The use of the Greeks in risk management

In the preceeding pages we have discussed the calculation of individual indicators of risk. So how are they used by banks?

We will concentrate here on the three Greeks: delta, gamma and vega.

Suppose that a derivatives department in a bank has positions in N derivatives (for simplicity all written on the same underlying S_t). For derivative i , there are n_I units held (which could be positive or negative) with, per unit price, delta Δ_i, gamma Γ_i, and vega ν_i.

The first line of defence is for banks to consider the delta of the overall portfolio Π

$$\Delta_\Pi = \sum_{i=1}^{N} n_i \Delta_i$$

If the bank's model is absolutely correct and if it is able to rebalance continuously in a way which keeps $\Delta_\Pi = 0$ at all times then the bank will have eliminated all risk.

In practice this does not happen for a variety of reasons, of which we will focus on two.

1. The bank is only able to rebalance at discrete times rather than continuously. As a consequence this introduces hedging errors because the delta of the overall portfolio will gradually drift away from zero in between rebalances.

   If the portfolio gamma is small then this drift and the consequent hedging errors will be very small. However, if the portfolio gamma is large then the drift away from a delta-neutral position could be quite large as will the hedging effors.

   As a result of this banks try to keep the portfolio gamma-neutral as well as delta-neutral. This helps to reduce the impact of rebalancing at discrete times. Thus them aim for

$$\Gamma_\Pi = \sum_{i=1}^{N} n_i \Gamma_i \approx 0.$$

2. The parameter values used by the bank (and even the model itself) may be incorrect. This means, potentially, that the prices for each derivative might not be correct and that the delta-hedging strategy may not be quite right. Thus even if the bank is able to rebalance continuously there may still be hedging errors and bias in prices.

   In equity derivatives the most important parameter which might lead to errors of this type is the volatility σ. The vega ν for an individual derivative tells us how sensitive the price is relative to errors in the value of σ. If the portfolio vega is close to zero then this means the impact of errors in σ will be small. On the other hand if the portfolio vega is large in magnitude then the bank may have significantly over or undervalued its portfolio with resulting risk to the bank.

   As a result of this banks try to keep the portfolio vega-neutral to avoid the risks associated with misspecification of the parameter values. Thus they aim for

$$\nu_{\Pi} = \sum_{i=1}^{N} n_i \nu_i \approx 0.$$

10.2 Dividends

Although we have valued the European Call and Put, we have done so in a rather synthetic case. As anybody who has ever bought shares is aware, life is more complicated than we have made it. The main two weaknesses in our theory so far are that in real life,

- most shares pay dividends and
- it is not usually possible to buy and sell shares for free, and there are transaction costs to be paid.

How can we cope with these in the framework that we have set up?

10.2.1 Including continuous dividends

The inclusion of dividends into the option valuation problem is not difficult provided that certain simplifying assumptions are made concerning the nature of the dividends. Anybody who owns shares will know that dividends are paid at discrete times (normally twice a year) and that their value may vary as a result of a number of factors. In our analysis, we shall initially assume

- That dividends are paid continuously ("drip-fed" into the account of the asset holder). This is an unrealistic assumption, but we will modify it later.
- That the dividend to be paid is known (from the outset) by the holder of the equity and is a constant multiple (call it D_0) of the current asset price.

With these assumptions, things are greatly simplified.

Recall that for reasons of arbitrage, when a dividend is paid, the share price must fall by the amount of the dividend. (Check this for yourself.)

Hence with a continuous dividend, in a time dt the underlying pays $D_0 S_t dt$ in dividend payments. The usual arbitrage argument then shows that the underlying must satisfy the random walk

$$dS = \sigma S dB_t + (\mu - D_0) S dt$$

which reflects the fact that the asset price must fall by the amount of the dividend payment.

It might seem at first that the dividend should have no effect on the price of the option, but a moment's reflection shows that it affects a crucial part of the previous option-pricing argument. For every asset held we receive $D_0 S dt$ in a time dt. If we argue as before, using a portfolio Π comprising an option of value V and an amount $-\Delta$ of the underlying, then the change in the value of our portfolio is now

$$d\Pi = dV - \Delta dS - D_0 S \Delta dt,$$

the last term being the fall in value of the stock component of the portfolio due to the dividend. Note that if $\Delta > 0$, then the stock is short sold. As such the portfolio owner will not be entitled to the dividend and so the value of the portfolio must fall by the corresponding amount.

Proceeding in exactly the same manner as before, but using the new expression for $d\Pi$, we find that the standard Black-Scholes equation for an option (of any type) with value V must be replaced by

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + (r - D_0) S V_S - rV = 0.$$

For the standard option valuation problem, how do the boundary conditions change? Consider the simplest example, namely the European Call. Well the payoff is the same, and if ever S becomes zero, then the option is still worthless. Thus, as in the case with no dividends we still have the conditions

$$V(S,T) = \max(S - E, 0) \quad \text{and} \quad V(0,t) = 0.$$

The other condition **does** change, however: now we need

$$V(S,t) \sim S e^{-D_0(T-t)}$$

as $S \to \infty$, since in this limit the value of the option is equivalent to the asset price, but **without** the dividend income. (If this seems confusing, then remember that if the asset price **now** is S_t then it will not be at expiry, since it will have suffered falls due to dividend payments to be made.)

10.2.2 Valuation of Euro-call under continuous dividends: Garman-Kohlhagen formula

So how do continuous dividends change the European Put and Call prices that we found before? The new equation and boundary condition do not really change things much, and though we could go through the standard transformations much as we did before, there is a quicker way of valuing a Call. Suppose we denote by $c_d(S,t)$ the fair value of a European Call with a constant dividend. Then, from above, c_d satisfies

$$\frac{\partial c_d}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 c_d}{\partial S^2} + (r - D_0) S \frac{\partial c_d}{\partial S} - r c_d = 0$$

with

$$c_d(S,T) = \max(S-K,0), \quad c_d(0,t) = 0,$$
$$c_d(S,t) \sim S e^{-D_0(T-t)} \quad (S \to \infty).$$

The trick is to now put

$$c_d(S,t) = e^{-D_0(T-t)} V(S,t).$$

We have

$$\frac{\partial c_d}{\partial t} = D_0 e^{-D_0(T-t)} V + e^{-D_0(T-t)} \frac{\partial V}{\partial t}$$

and

$$\frac{\partial c_d}{\partial S} = e^{-D_0(T-t)} \frac{\partial V}{\partial S}, \qquad \frac{\partial^2 c_d}{\partial S^2} = e^{-D_0(T-t)} \frac{\partial^2 V}{\partial S^2}.$$

Thence V satisfies

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - D_0) S \frac{\partial V}{\partial S} - (r - D_0) V = 0$$

with

$$V(S,T) = \max(S-K,0), \quad V(0,t) = 0, \quad V(S,t) \sim S \quad (S \to \infty).$$

Now this is EXACTLY the standard (i.e., no dividend) European Call option problem, save for the fact that r has been replaced by $r - D_0$. The solution can therefore be read off as

$$c_d(S,t) = e^{-D_0(T-t)} S \Phi(d_{10}) - e^{-D_0(T-t)} K e^{-(r-D_0)(T-t)} \Phi(d_{20})$$

and thus

$$c_d(S,t) = e^{-D_0(T-t)} S \Phi(d_{10}) - K e^{-r(T-t)} \Phi(d_{20})$$

where

$$d_{10} = \frac{\log(S/E) + (r - D_0 + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$
$$d_{20} = \frac{\log(S/E) + (r - D_0 - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

This solution is called the **Garman-Kohlhagen formula** for a call option on a dividend paying share.

Associated with options on a continuous-dividend-paying asset is one further Greek:

- **The "lambda"**, λ, of an option is
$$\lambda = \frac{\partial V}{\partial D_0}.$$
It represents the sensitivity of the value of the option to the dividend.

10.2.3 Euro-put valuation and Put-call parity formula under continuous dividends

Now that we've valued a European Call with continuous dividends we use the usual strategy for valuing a Put: instead of doing a lot of tedious mathematics, we appeal to arbitrage and use Put-Call parity. The relevant argument with continuous dividends is clear: we consider a portfolio
$$\Pi = p - c + Se^{-D_0(T-t)}.$$
Whatever the value of the underlying at expiry, the pay-off is K, and so the value of the portfolio must be the same as a riskless investment of an amount K, which is $Ke^{-r(T-t)}$. Thus
$$Se^{-D_0(T-t)} + p - c = Ke^{-r(T-t)}.$$
Thence
$$p = Ke^{-r(T-t)} - Se^{-D_0(T-t)} + e^{-D_0(T-t)}S\Phi(d_{10}) - Ke^{-r(T-t)}\Phi(d_{20})$$
which simplifies to
$$p = -Se^{-D_0(T-t)}[1 - \Phi(d_{10})] + Ke^{-r(T-t)}[1 - \Phi(d_{20})].$$
Actually, this may be expressed even more simply. Using standard definitions of integrals we find that for any d
$$1 - \Phi(d) = \Phi(-d)$$
and thus
$$p = -Se^{-D_0(T-t)}\Phi(-d_{10}) + Ke^{-r(T-t)}\Phi(-d_{20}).$$

The following diagrams show the effect of the dividend on the option price. Note that at expiry, $t = T$ the payoff is of the options are still the same as in the case of an asset that does not pay a dividend.

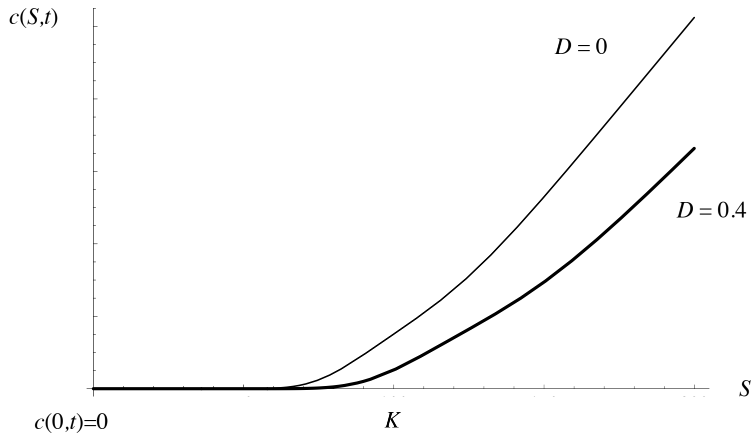

Figure 10.1: Comparison of the European call options at $t = T/2$ for assets with the same strike price K, volatility $\sigma = 0.25$ and with risk free rate $r = 0.05$. The thick line denotes the Euro-call value for an asset that pays a continuous dividend DS where $D = 0.4$. The thin line denotes a Euro-call on an underlying asset that pays no dividend.

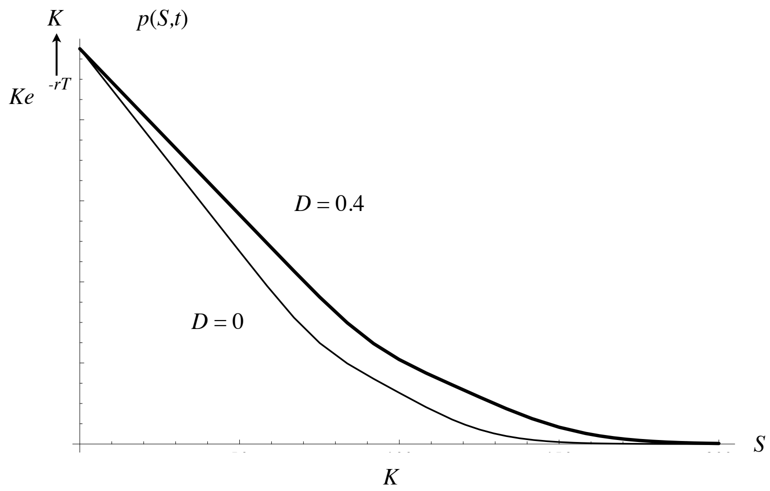

Figure 10.2: Comparison of the European put options at $t = T/2$ for assets with the same strike price K, volatility $\sigma = 0.25$ and with risk free rate $r = 0.05$. The thick line denotes the Euro-put value for an asset that pays a continuous dividend DS where $D = 0.4$. The thin line denotes a Euro-putl on an underlying asset that pays no dividend.

10.3 The “advantages” of the martingale approach

The following section is lifted directly from notes written by the Institute of Actuaries. The footnotes are not.

“The advantage of the martingale approach is that it gives us much more clarity in the process of pricing derivatives. Under the PDE approach we derived a PDE and had to ‘guess’ the solution for a given set of boundary conditions.[1]

Under the martingale approach we have an expectation which can be evaluated explicitly in some cases and in a straightforward numerical way in other cases.[2]

Furthermore the martingale approach also gives us the replicating strategy for the derivative. [3].

Finally, the martingale approach can be applied to any $\mathcal{F}_T$-measurable derivative payment, including path-dependent options (for example, Asian options), whereas the PDE approach, in general, cannot.[4]”

---

[1]Not, of course if you realise that the Black-Scholes equation can be transformed into the heat equation for **any** derivative product. The only thing that changes is the payoff, or equivalently, the initial data in the heat equation If you realise this, then you can use the standard integral representation for the solution of a heat equation involving a given set of initial data. The integral representation derived in the PDE approach is, of course exactly the same as in the Martingale approach. The only difference is the interpretation: for the PDE approach this is a convolution of the initial data with the Green’s function of the heat equation, for the martingale approach it is the expectation integral arising from a normal pdf. The pdf of the martingale approach is the Green’s function of the PDE approach. *Chaqun a son goût*.

[2]The same is true for the PDE case: you just have to convolve the transformed initial data of the heat equation arising from the payoff with the Green’s function/normal distribution pdf. The key test is whether you can evaluate the integral in terms of known functions. In general you will not be able to, whether you start with a PDE or take the martingale route, since ultimately you will have to evaluate the same integral, wherever you started.

[3]So does the PDE approach. In fact you practically have to use the PDE approach to define the ϕ_t of the martingale approach.

[4]This is incorrect. With American or Asian path-dependent options you have to be a little cleverer than in European options, but, as with the Martingale approache for these options, the results can be derived using standard numerical procedures. To see how to do this for the PDE approach is beyond the scope of this course, but it can be found in many texts, e.g., *Derivatives* by Paul Wilmott, published by John Wiley.

Appendix A

Dictionary of Financial Terms

Financial Mathematics and Finance in general is awash with abbreviations and jargon. The purpose of this small "dictionary" is to provide a quick source of reference for terms or abbreviations that may be used during the unit. It also contains many definitions that have nothing to do with the unit content, but apply to general financial matters and are worth knowing anyway.

AIM (Alternative Investment Market) The newest attempt by the Stock Exchange to provide a market for smaller and emerging companies. Companies listed by AIM are often speculative and highly volatile, and their shares are frequently illiquid.

American Option Any option that may be exercised at a time chosen by the holder rather than at a specified expiry date.

Annual General Meeting ("A.G.M.") All shareholders are entitled to attend this annual meeting, which may cover the election and re-election of Directors, approval of the dividend payment, appointment of auditors, Chairman's statement, etc.

Annual Report A report produced every year by all listed Companies containing profit and loss statements, accountants' reports, details of major Company personnel and their salaries and usually a message from the Chief Executive or Managing Director on how he or she expects the next year's trading to go. May contain much other material as well. Each shareholder should receive a copy.

Annualisation The adjustments that have to be made to the calculation of (for example) EPS when the period in question is not equal to 12 months. The methodology may be complicated but amounts essentially to no more than a weighted average.

Annualised Growth Rate A tool used to measure the consistency of a fund and thereby rate its performance. Basically a 5-year AGR is simply the total return over the 5 years divided by five.

Arbitrage The process of getting "something for nothing". For example, if circumstances arose where it was possible to buy a share at one price and instantly sell for a greater price, then this would be arbitrage.

As-You-Like-It Option Another name for a chooser Option.

Asian Option A general name for any sort of option where the price depends on some form of average.

Asset-or-Nothing Option A European Asset-or-Nothing Call option pays the current share price at expiry if this price is above a given strike, and nothing otherwise.

Asset per Share A measure of the asset position of a Company; determined by dividing the total assets by the total number of shares. The only complication arises in deciding exactly what to count as an "asset".

At-the-money The option whose exercise price is closest to the current value of the underlying asset.

Auditor's Report Contained somewhere in a Company's accounts, this usually promises that the financial statements give a true and accurate picture of the Company's financial position. Strictly speaking, the auditors are employed by the shareholders to ensure that the Directors are complying with the Companies Act.

Average Rate Option Closely related to an average strike option, but differs in the structural form of the payoff. Whilst an average strike is the same as a vanilla option but for the fact that the exercise price is replaced by the average price, an average rate option has the same payoff as a vanilla option but with the asset price replaced by the average.

Average Strike Option A path-dependent option whose payoff is the difference between the asset price at expiry and its average over some time period prior to expiry (if this difference is positive) or zero if it is negative.

Barrier Option A general name for a sort of option which either comes into existence or becomes worthless depending on whether or not the underlying asset reaches some prescribed value before it expires.

Bear Market Generally prevailing downward conditions in the market, with share values tending to decrease. A serious Bear Market might lead to a crash. A bear market prevails at present and has done for three years, but there is evidence that things are starting to get better....

Bed & Breakfast The practice (outlawed 4 years ago) of selling shares the night before the end of the tax year and buying them back again the next day in order to use all of one's tax exempt allowance.

Bermudan Option An option that may be exercised early, but only on specified dates.

Beta Factor A figure that gives an indication of how rapidly and consistently a Company's move up and down with the market. The market's beta factor is one: shares with a beta factor in excess of one are more volatile (risky?) than the market.

Beta For a given asset, the correlation between the changes in the asset's return and the changes in the "market portfolio".

Bid/Offer Spread The difference between the buying and selling price.

Binary Option An option with a generalized pay-off. (Originally an option that was effectively a straight bet on whether the underlying asset's price would be above (call) or below (put) a certain price.)

Blue Chip Shares Companies that are large, household names. Businesses with solid track records of growth and large Stock Market valuations. Usually considered to be "safe" shares. May not be quite so safe after 9/11!

Black-Scholes Equation The second order partial differential equation that underlies a large part of the theory of the valuation of options and other derivative products. First published in the open literature by Black & Scholes in 1973.

Bond A contract, paid for up front, which pays a given amount at some fixed date in the future.

Bond Option An option that is identical to a standard option on a share, save for the fact that the underlying asset is a bond.

Bond Pricing The problem of finding the fair value of a bond with a given pay-off and maturity date.

Book Value The notional idea of what a Company would be worth if it sold off all its assets. Greatly complicated by the fact that assets may arise as tangibles (machinery, stock, etc.) and intangibles (patents, copyrights, staff development investment etc.) and it is often not clear what should be included.

Broker's Estimate(s) Brokers regularly publish their estimate for the next year or two of what they predict a given Company's profits will be. The larger the Company, the more broker's estimates can be expected. For example, between 19/3/98 and 28/7/98 there were 16 Broker's estimates for Debenhams PLC. Estimates for pretax profits for 1998 ranged from £132m to £135m and for 1999 from £144m to £154m. Most brokers rated this share a "buy" at 338. (Value on 2/11/98 was 392).

Broker's Forecast Changes The changes in forecast turnover, profit and EPS from one estimate to the next, either by the same or different brokers. If the profit estimates are increasing, then this is normally a sign that things are going better than expected; if they are decreasing then watch out!

Bull Market The opposite of a "bear market". Circumstances where shares tend to rise in value by large amounts. When the "bull is running" investors are happy and actively seek to buy more and more shares. Indices such as the FTSE-100 rise accordingly. Please! come back! Please!

Butterfly Spread A portfolio which is long one call with exercise E_1 and long one put with exercise E_2 where $E_1 < E < E_2$.

Calendar Spread A portfolio of options with different expiry times.

Call Option An option that allows the owner to buy a share for a given price at some given time in the future. Normally bought in hope that the share price will rise, whereupon the owner can buy the asset for less than it is worth.

Cap A loan at the normal floating interest rate but with the added proviso that the interest rate is guaranteed not to go above a particular value (the cap).

Capel-Cure Myers Dividend Survey A famous survey of Companies between 1955 and 1988 which showed conclusively that, on average, share portfolios containing low dividend yielding shares performed much worse than those containing high dividend yielding shares.

Capital Asset Pricing Model (CAPM) A well-respected model that describes the relationship between risk and reward assuming an "efficient market".

Capital Employed A measure of how much money is being used in the operation of a company. Calculated by summing all share capital, reserves, loan stocks, debentures, all borrowings etc.

Capital Expenditure One way that a Company may use its cash flow. A distinction must be made between say replacing old machinery in existing factories and buying new machines for new factories; frequently however it is hard to tell what the capital expenditure has been upon and this may make it difficult to correctly interpret the financial health of a Company.

Caption An option on a cap.

Cash Flow per Share Regarded as a test of future capacity to pay a dividend and essentially measures the volume of cash generated by a Company that is available to pay the dividend with. Basically calculated by dividing the cash flow by the number of shares.

Cash Flow The life blood of a business. Consists essentially of the net cash inflow from all operating activities minus returns on investment and servicing of loans, after adjustments have been made for taxation. Strong cash flow is considered by many to be the single most important pointer to the growth prospects of a Company.

Cash-or-Nothing Option An binary option that amounts to nothing more than a simple bet on the asset price; eg for a call Cash-or-Nothing if the asset value is above the strike the option exerciser receives an agreed payoff, whilst if the asset value is below the strike nothing is paid.

CBOE The Chicago Board Options Exchange - the first organisation to officially trade derivative products (in 1973).

Chairman's Statement A "message from the chairman" which is usually contained in the annual report. It may provide an idea of how things look for the current year, and what might be expected to happen in the future. Such statements are sometimes ludicrously bullish, sometimes sober and reserved.

Charting A way of "predicting" how a share price is likely to move by looking for given patterns in the share price graph. Popular in some quarters, but tends to ignore fundamentals, and may be spectacularly wrong on occasions.

Chief Executive Usually simply another name for the Managing Director of a Company: put simply, the boss.

Chooser Option An option that gives its owner the right to buy, for a given amount at a given time, *either* a call or a put with a given exercise price and a give expiry time.

Cliquet Option A complicated and esoteric exotic option.

Collar An undertaking not to charge more than a stated maximum and minimum on a floating-rate loan.

Compound Option An option on an option. Often the underlying asset is a vanilla call or put. May still be valued using the Black-Scholes framework.

Contract Note A formal confirmation of a purchase of shares, bonds, or trust funds.

Convertible A corporate bond or preference share that can be converted into ordinary shares at a set price on set dates.

Corporate Bond Loan stock or "IOUs" issued by public companies to raise capital. The Company usually promises to pay a given sum of money on a given date or dates in the future until the "redemption date" at which point it repays the loan.

Coupon A known cash dividend paid at fixed times during the life of a bond. Very similar to a share dividend.

Crash A rapid and disastrous fall in share prices. The wall Street Crash of 1929 caused the great American depression of the 1930's.

Current Ratio A measure of how effectively current liabilities of a Company are covered by current assets. Obtained by dividing the current assets by the current liabilities. A current ratio below 1 may be a sign of weakness in a Company.

Cyclical Companies in sectors such as Building and Construction, Building Materials, Paper and Packaging, Printing, and Engineering whose share price is particularly sensitive to changes in the economic factors as a whole such as interest rates.

Delta-hedge The process of reducing (or ideally removing) the sensitivity of a portfolio to price changes in the underlying by essentially taking opposite positions on the underlying and on an option on the underlying. The delta hedge is dynamic and requires constant re-hedging as the value of the underlying changes.

Delta-Neutral The name for the position that one takes when delta-hedging an option and an asset.

Delta The Delta of a portfolio Π is defined as $\partial\Pi/\partial S$; it is this delta that is used in delta-hedging.

Derivative A financial instrument that depends for its own value on the value or performance of another (typically a share, bond or interest rate). Examples of derivative products are options, futures and warrants.

Digital Option A simpler kind of exotic option (also known as a "binary") which differs from a vanilla option only in that the payoff at expiry is some function of the asset price.

Director's Dealings All share dealings made by Directors of a Company in their own Company must be notified. Since Directors are by definition in a good position to know the inner workings and prospects of a Company, Directors buying or selling shares is usually considered to be an important pointer to private investors.

Dividend A payment, usually paid once or twice per year, to shareholders to reward them for their loyalty in holding a given share. Equivalent to receiving interest on a share. Usually expressed as a percentage of the nominal price of the share.

Dividend Cover A measure of how secure the dividend is. Calculated by dividing the Company's earnings per share and dividing them by the net dividends per share. A well-covered dividend is normally more likely to be maintained or increased.

Dividend per Share (DPS) Calculated by dividing the total of net declared dividends by the number of shares. Also usually appears as one of the figures forecast by brokers in their estimates.

Dividend Yield The percentage return from the annual income on a share: if one were to buy 1000 shares at £1 each and and the Company paid a (gross) dividend of 13p per share during the year then the dividend yield would be 13%.

Down-and-in Option A barrier option which expires worthless unless a given barrier is reached from above at some time before expiry.

Down-and-out Option A barrier option which expires worthless if a given barrier is reached from above at some time before expiry.

Earnings per Share (EPS) Calculated by dividing the profit after tax (minus minority interests and preference dividends) by the weighted average number of shares in issue. Calculations of EPS have been complicated by the emergence of a number of different "standards" for producing the figure. Companies that are thriving normally show a year-on-year increase in EPS.

Efficient Market Hypothesis The justification for the assertion that asset prices move "randomly": the assumption that (a) the past history of an asset is fully reflected in its current price and (b) markets respond instantly to any new information about an asset.

Enhanced Scrip Dividend An increase in the dividend after a scrip issue.

Equity Another word for share. Also called securities.

European Option A vanilla option that has a set expiry date and cannot be exercised before that date.

Exceptional Items (Also known as "Extraordinary Items") Profits or losses that are not considered to be part of the normals trading of a Company and are therefore normally not included in forecasts or calculations of figures such as EPS and DPS. May include one-off expenses and other unforeseen items.

Ex-dividend The period before a share pays out its dividend. If you buy a share during this period, then you are considered to be too late to receive the dividend, though you still receive it if you sell during this period.

Exercise Strategy The strategy one employs when deciding when to exercise an American option (or any other option where the expiry date is not fixed).

Exotic Option A general name for options other than standard European and American Calls and Puts. Originally referred only to options whose payoff depended upon the asset price history rather than simply the asset price at expiry, but now tends to be applied to any "non-vanilla" options.

Expiry Date The date when the value of an option is finally decided and the option ceases to exist.

Financial Year The twelve month period that a Company chooses to count as its "whole year". This may vary from Company to Company; for example the year end for Highland Distillers (FTSE Mid 250) is 31 August, whilst the year end for Kewill Systems Ltd. (FTSE Smallcap) is 31 March.

Fixed-interest Security Another name for a bond.

Floor A loan at the normal floating interest rate but with the added proviso that the interest rate is guaranteed not to go below a particular value (the floor).

Floortion An option on a floor.

Flotation The act of becoming a Company listed on the Stock Exchange. Recently the areas (amongst others) of Building Societies, Insurance Companies and Football Clubs have all seen a number of flotations (with varying amounts of success).

Foreign Exchange Option Similar to a normal option but the underlying is a foreign currency rather than an equity. For example, a Foreign Exchange Option might consist of the right to buy dollars at an exchange rate of £1 = 1.67 USD at some specified time in the future

Forward Contract An agreement between two parties wherein one party is contracted to buy a given asset from the other at a given price at some given future date.

FRS3 Financial Reporting Standard 3; the latest incarnation of a set of rules and guidelines (produced by the Accounting Standards Board) that dictate exactly how earnings per share should be calculated.

FT All-Share Index An index that attempts to measure the movements in shares as a whole. Does not include Companies smaller than a given size.

FT Mid 250 Index The index for Companies not large enough to qualify for the FTSE 100 index but which are too large for the Smallcap index. In August 1998, FTSE Mid 250 Market Caps ranged from Bryant Gp. (£263 million) to Sema (£3,321 Million)

FTSE 100 The "Premier League" of British Companies, notionally containing the top 100 UK Companies. In August 1998 the top Company in the FTSE 100 was Glaxo Wellcome with a market cap of £66,859 whilst the lowest was LASMO with a market cap of £2,091.

FTSE Fledgling The index for Companies not large enough to qualify for the FTSE Smallcap index but which are nevertheless listed on the full Stock Exchange. In August 1998, Fledgling Market Caps ranged from Versalite (£2 million) to Triad Gp. (£175 Million).

FTSE Smallcap The index for Companies not large enough to qualify for the FT Mid 250 but too large for the Fledgling index. In August 1998, Smallcap Market Caps ranged from Huntingdon (£22 million) to Psion (£459 Million)

Future Similar to a forward contract, but (a) usually handled through an exchange such as LIFFE and (b) settled continuously during the life of the contract in order to protect both parties from possible default.

Gamma The rate of change of the delta of a portfolio Π, defined as $\partial^2\Pi/\partial S^2$. Sometimes use to hedge away smaller order effects in a portfolio.

Gearing The amount of money that a Company has borrowed, expressed as a percentage of its net assets.

Gilt A bond issued by the Government. Known as gilt-edged as they are issued by the most secure borrower of all (theoretically the Government cannot go bust!) gilts are the prime source of borrowing for the PSBR (Public Sector Borrowing Requirement).

Growth Company A Company whose shares everybody would like to buy: the share price goes up and up whilst turnover, profits and EPS increase every year. Often a Company from the FTSE Smallcap or Fledgling indices. Hard to spot, but profitable when you do!

IIMR The Institute of Investment Management and Research; a body which has recommended that analysts adjust the profits that have been worked out using the full version of FRS3. They claim that this is likely to give a fairer assessment of the trading profit made during the year by a Company.

Implied Volatility Since direct measurements of the volatility of individual assets are often hard to perform, an alternative way to estimate the volatility is to assume that "the market knows". Using the current interest rate, we therefore solve the Black-Scholes equations and choose the volatility (the "implied volatility") so that the Black-Scholes price agrees with quoted prices.

In-the-money A call (or put) option whose exercise price is less than (greater than) the current asset price, and so has an intrinsic value.

Inclusion Criteria A set of criteria that a Company must satisfy in order to be included in a given index. Usually includes some measure of a Company's value, longevity, reliability etc. Naturally the criteria for AIM would be less than those for listing on the FTSE Mid 250.

Index Demotion Candidate A Company whose market capitalisation has fallen, making it "too cheap" for its current index (FTSE 100, 250, Smallcap, etc.). Unless things improve, it is likely to be "relegated".

Index Promotion Candidate A Company whose market capitalisation has risen, making it "too expensive" for its current index (FTSE 250, Smallcap, etc.). If the Company is "promoted", (especially to the FTSE 100) then its share price often rises still further as Tracker Funds are forced to buy shares.

Insider Trading The illegal use of privileged information to buy or sell shares before everybody else knows what is happening. Fairly common, but very hard to prove. The Stock Exchange monitors every unusual price rise and fall.

Installment Option A European or American Call or Put option with the additional feature that, as well as paying an initial sum to buy the option, the holder must also pay "installments" during the life of the option. If at any time the holder chooses to stop paying the installments then the option is cancelled.

Intangible Assets Those assets of a Company that are not "viewable" and may only exist in principle or on paper. Includes such things as patents, goodwill, copyright, brand names and newspaper titles (which may nevertheless be of great value).

Interest Cover Ratio Calculated by taking a Company's profits (before interest and taxation) and dividing them by the annual interest charge on any money that the Company may borrow. The figure indicates the Company's ability to continue funding its borrowings through its profits.

Interest Rate In effect, the interest rate governs the price of money. Simply thinking of the current bank rate as the interest rate may be misleading, for the cost of borrowing money also depends crucially upon inflation. Currently both interest rates and inflation are low.

Investment Trust An alternative to unit trusts. Investment trusts employ Managers to invest a fixed share capital in a number of equities. The price of a share in the investment trust responds to the demand for them in the Stock Market. Unlike unit trusts, they can invest in unlisted Companies. For long periods, most investment trust shares trade at a discount to the value of the shares that they hold, which is attractive to many.

ISA The replacement for TESSAs and PEPs that was introduced in April 1999. The ISA has received widespread criticism and is universally regarded as less of a Tex Perk than a TESSA or a PEP.

Junk Bond A bond issued by a company in trouble to raise money. The bond promises spectacularly high returns, but the risk of the company going bust (and the bond therefore becoming worthless) is high. For gamblers only!

Key Dates The dates in a Company's year when certain important events that may influence the share price take place. For example, the interim and final dividend payment dates, the interim results announcement date, and the dates when the preliminary results, annual report and AGM take place.

Ladder Option A complicated esoteric type of exotic option.

LIFFE The London International Financial Futures and Options Exchange, which acts as the Derivative products equivalent to the Stock Market.

Liquidity Related to the number of shares in a given Company that are traded each day. Illiquid stocks may be traded in such small numbers that they are hard to sell. Many of the Companies that are listed on AIM suffer from illiquidity, which adds to the risk of buying them since even if their value rises it may be impossible to sell and thus realise one's profit.

Long Position The opposite of a short position; the state of owning assets (which one might later sell) as opposed to not owning assets that one might wish to sell.

Lookback Option A type of path dependent option whose price depends on some sort of asset price maximum or minimum.

Market Capitalisation ("Market Cap") The nominal total value of Company. Calculated by multiplying the number of share in circulation by the current share price. On 3rd August 1998 the Market Cap of BT (2nd largest Company in Britain) was £56,034 million, whilst the Market Cap of Southampton Leisure (1477th largest) was £18.8 million.

Market Maker The traders who actually determine the price of a share. Usually holding shares themselves, often bought with borrowed money, they make their money from the difference between the price at which they buy and the price for which they sell.

Market Price of Risk The excess return above the risk-free rate of interest on a portfolio that accrues to one when a certain extra level of risk is accepted.

Maturity Date The date when a bond pays out.

MEPS (Maintainable Earnings per Share) A notion used by analysts in calculating estimated future earnings per share. MEPS represent the likely results of a Company given its structure at the time that the calculation is performed.

NMS "Normal Market Size": provides a measure (in thousands) of the average amount of the stock traded in a given Company. NMS bands are allocated by the Stock Exchange according to a formula based on the current and previous year's trading figures. In August 1998 BT (FTSE 100, turnover £15,640m) had an NMS of 150 whilst Polymasc Pharmaceuticals (AIM, turnover £0.53m) had an NMS of 1.

Net Cash Usually calculated by determining the percentage of cash (less borrowings) in relation to the shareholders' funds including intangibles.

Net Gearing The gearing of a Company after certain adjustments have been made.

Niche Market A particularly specialised market, where any Company who become the major manufacturer is likely to effectively corner the market.

Normalised Earnings per Share Similar to EPS (Earnings per share) but adjusted to take into account items such as exceptional tax charges, large-scale restructuring costs and long-term redundancy programmes etc. Normally considered to give a better picture of a Company's position than IIMR Headline EPS.

OFEX market The Off-Exchange market (opened October 1995), often regarded as a nursery for Companies wish eventually to be listed on AIM. Shares of OFEX Companies are frequently illiquid, highly speculative and generally risky.

Olive Option Traditionally supposed to have been the first ever option: the ancient Greeks bought olives for future delivery at a price agreed in the present.

Optimal Exercise Boundary The boundary for an American option above which the option should be held, and below which the option should be exercised.

Optimal Exercise Price The optimal price at which to exercise an American option.

Out-of-the-money A call (or put) option whose exercise price is greater than (less than) the current asset price, and so has no intrinsic value.

Over-trading A normally undesirable situation where earnings per share are expanding rapidly but cash flow is shrinking. Often signifies that large amounts of cash are tied up in debtors who may be unable or unwilling to pay.

P/E Ratio Obtained by dividing the current share price of a Company by EPS. Often cited as a measure of how many years it would take the Company to earn net profits equal to the market value of the Company. A high P/E ratio may often indicate that the market is expecting fast profits growth; a low P/E may mean that lower profits are expected.

Path-Dependent Option Any option whose value depends on some aspect of the history of an asset price, rather than simply upon the asset price at exercise. Examples are barrier options, Asian options and lookback options.

PEG An indicator of growth which is obtained by dividing the P/E ratio by the forecast growth of the Company in percent. A low PEG (under 1, say) is cited by many as the sign of a growth company.

Penny Shares A share whose price is low. (Originally just a few pence, but nowadays most shares priced at under £1 are termed "penny shares"). The upside is that rapid growth and high profits are possible. The downside is that such stocks are often highly volatile and risky.

PEP, Tracker A PEP whose value depends on the current price of a certain index (often the FTSE 100). Similar to buying a "FTSE 100" share. The fund manager "tracks" the index by buying shares from all of its Companies.

PEP A "Personal Equity Plan" (discontinued in April 1999). Currently allows the holder to invest up to £6000 per annum in an approved unit trust or fund and up to £3000 per annum in a single Company. Now replaced by ISAs.

Perpetual Option An option without an expiry date, where the theoretical time horizon is infinite. (For example, a Russian option).

Portfolio Theory The theory of maximising return whilst minimising risk from a portfolio of risky and riskless assets. In theory, for a given group of assets, the theory should determine the optimum amounts of each asset to hold.

Portfolio A collection of shares, options, bonds, etc. held by an individual or an organisation. Pension funds etc. often have a large portfolio in order to spread their possible risk.

Pre-Tax Profits The profits of a Company before tax has been deducted. Held by some to be a truer measure of the success or otherwise of a Company as sometimes complicated tax positions may mask the real performance figures.

Preference Shares Shares with a fixed (usually high) dividend. if the Company becomes insolvent or goes bankrupt, preference shareholders are paid before ordinary shareholders.

Price-to-Book Value Ratio (PBV) Calculated by dividing the current share price of a Company by its book value (ordinary capital + equity reserves) per share. A low PBV may sometimes indicate that a takeover might be around the corner, but the role of intangibles in the calculation is often difficult to interpret.

Price-to-Cash Flow Ratio (PCF) A figure that indicates how much cash flow one is buying per share. Calculated by dividing the Company's market capitalisation by its cash flow. A low PCF is normally very attractive.

Price-to-Research and Development Ratio (PRR) A figure obtained by dividing the Market Capitalisation of a Company by the total expenditure of research and development. Most meaningful for Technology Companies.

Price-to-Sales Ratio (PSR) Calculated by dividing the latest share price by the sales per share. Often cited as a key figure for identifying recovery stocks which may frequently have abnormal PSRs.

Price-to-Tangible Book Value Ratio (PTBV) Obtained by dividing the Company share price by its net asset value per share after intangibles have been deducted. Forms a fairly harsh measure of the Company's net asset value, but its significance depends greatly on the nature of the Company.

Profit Margin The ratio of operating profit to turnover. Generally speaking, a high margin is a sign of a good Company, though of course very high margins may be used by the competition to start a "prices war".

Profits Warning A statement issued by a Company indicating that the profits for the next year or half year are likely to be less than expected. Made with the idea of lessening the damage once the final figures are out, but can still lead to large falls in share price, especially for smaller Companies.

Put Feature A feature of some bonds that allows the owner the right to return the bond to the issuing company for a given amount.

Put Option An option that allows the owner to sell a share for a given price at some given time in the future. Normally bought in hope that the share price will fall, whereupon the owner can sell the asset for more than it is worth.

Put-Call Parity A simple argument that leads to a formula that allows the value of a Call to be expressed in terms of the value of a Put, and vice-versa. The two therefore do not have to be valued separately.

Quanto A complicated financial product where, as well as an option or future structure being present, fees and/or payoffs are in currencies different to that normally used. A quanto is therefore usually a bet on the currency rates as well as on an asset or an index.

Quick Ratio A figure that attempts to estimate what would happen if the company suddenly had to pay off all its existing liabilities. Obtained by dividing the current assets (less stock and work-in-progress) by the current liabilities. A low quick ratio indicates vulnerability.

Recovery Stock A share that has fallen upon hard times but is now likely to recover. Typically to be found when a recession or slump is almost over, often in areas such as building, estate agents, stores etc.

Relative Strength A measure of how the price of a particular share has changed relative to the FTSE All-share index. Sometimes called "price relative" performance.

Rights Issue Used by companies to raise cash. Existing shareholders are offered the chance to buy a certain number of shares (in a 1-for-2 rights issue for example, each 2 shares held entitles you to buy one more), usually at a reduced price. May be used for rescue or expansion purposes.

Risk Neutrality An alternative premise that leads to another way of valuing derivative products.

ROCE Return on Capital Employed. Indicates how much profit the company makes for its capital. Calculated by expressing the before-tax operating profit as a percentage of the year-end capital employed.

Russian Option A perpetual American option that pays out (at any time chosen by the owner) the maximum realised asset price up to that date.

Sales per Share The total sales or operating profit for the last period covered by the annual report divided by the weighted average number of ordinary shares in issue during that period.

Scrip Issue An issue of "free shares" to the shareholders of a Company. There are many reasons why this might happen, but the most obvious is when a Company feels that the price of an individual share is too high. A "2 for 1 scrip", for example, would give each shareholder two extra shares for each share currently held. Of course, the share price may be expected to drop accordingly.

Sector Each listed Company is allotted a "sector" relevant to the sort of business that they are involved in. For example, Airtours are in the "Leisure and Hotels" sector, whilst Coca-Cola come under "Food Producers".

Short Position A position in which one seeks to sell shares that one does not own.

Short Selling The process of selling an asset that one does not own. This may be done through most brokers and is perfectly legal; to actually accomplish the deal a broker will normally "borrow" and then sell somebody else's stock. The most obvious reason for short selling is if one expects the value of an asset to fall. Selling the asset short for today's price and then buying it back again at a future lower price is obviously profitable.

Smile The colloquial name for the shape of the curve produced when implied volatility is plotted against exercise price. Points to shortcomings in the Black-Scholes model which however are not easily identified or fixed.

Spread bet The betting equivalent of a future. Increasingly being used by speculators as no shares are bought (hence no stamp duty is payable) and quick profits can be made. Note: quick losses can be made too!

Stop-Loss When an asset is bought, it is sometimes considered wise to specify a "stop-loss" clause that guarantees that the asset will be sold if ever its price drops below a certain value. This effectively limits the possible loss on an individual asset.

Straddle A particular sort of portfolio designed to take advantage of ones market view. Often made up by being short one asset and long two calls on the asset, or sometimes by being long one call and one put on an asset.

Strike Price The price appearing in an option contract. Also known as the exercise price.

Supershare A binary option which has payoff $1/d$ if the strike E and current asset value S satisfy $E < S < E + d$ at expiry and zero therwise.

Swap An agreement between two parties to exchange the interest rate payments on a given amount (usually called the "principal") for a certain length of time. Normally one party pays the other a fixed rate of interest in return for a variable rate payment.

Swaption An option on a swap.

Take a View A fancy word for making a guess.

Tax Rate The tax rate which a Company pays. Depending on their recent trading record, this may be as much as 70% or, if times are hard, may even be negative (a tax refund).

TESSA Tax Exempt Special Savings Account (discontinued April 1999): a scheme designed by the Conservative Government of the 1980's to "encourage saving". A certain amount of money (up to £9000) could be invested in a Building Society, the interest being tax-free so long as the account was kept open for five years.

Theta A measure of the rate of change of a portfolio's value Π with time. Defined by $\Theta = -\partial\Pi/\partial t$.

Trading Volume The number of shares actually traded in an asset. Often invaluable to know when trying to interpret why a share price has risen or fallen.

Tracker Fund An Investment fund which invests mechanically to mirror the performance of a given index.

Transaction Costs The costs that form the "broker's fee" (as opposed to the bid-offer spread). A typical transaction cost is 1.5-2% for normal purchases and sales of UK equities; for foreign stocks and more complicated financial products it may be much more.

Treasury Stock The "stock" that one nominally buys when buying gilts. For example, "12.5% Treasury Stock 2003-2005" refers to gilts with a coupon of 12.5some time between 2003 and 2005.

Turnover A Company's total sales (excluding VAT) as shown by the last annual report or by the announcement of the preliminary results for the next year.

Two-Colour Rainbow Option A complicated type of option that depends on two underlyings for its value rather than just on a single asset.

Unit Trust An "open-ended" investment fund that can accept any amount of investors' money at either fixed or disparate intervals. The money is invested in a wide range of shares, bonds, options, property etc. The fund is also divided into units that rise and fall in line with the values of the underlying assets.

Up-and-in Option A barrier option which expires worthless unless a given barrier is reached from below at some time before expiry.

Up-and-out Option A barrier option which expires worthless if a given barrier is reached from below at some time before expiry.

Vanilla Option The simplest sort of option; normally a simple call or put, as distinct from other more exotic "flavours".

Vega A measure of the sensitivity of a portfolio value to changes in volatility. Usually defined as $\partial\Pi/\partial\sigma$ where σ is the volatility and Π the value of a given portfolio.

Volatility A measure of the average wildness in the fluctuations of the value of an asset. High volatility shares are very risky, but may allow large profits to be made in a short space of time.

Warrant Derivative securities related to options. Most commonly call options on shares in a Company, where the Company issues new shares to complete the contract if the option is exercised. As far as the Company are concerned, the issue of warrants mimics a deferred rights issue.

Writer The person responsible for issuing an options, futures etc. contract. If the option is exercised, the it is the writer's responsibility to ensure that the contract is honoured.

Yield Curve A measure of future values of interest rates. Almost invariably requires somebody to take a view concerning how interest rates will behave. An important feature in the theory of bond pricing.

Zero Coupon Bond A bond that does not pay coupons.

Appendix B

Basic facts from statistics

Discrete Distributions

For discrete events with probabilities p_j and n possible outcomes taking values x_j

- the probabilities are normalised
$$\sum_{j=1}^{n} p_j = 1,$$
- the mean/expected outcome is
$$\mu = E[x] = \sum_{j=1}^{n} p_j x_j,$$
- has variance
$$\sigma^2 = \sum_{j=1}^{n} p_j \left(x_j - \mu\right)^2 .$$

Continuous Distributions

Any continuous probability distribution $p(x)$ defined for some interval $a \leq x \leq b$

- has a continuous probability density $p(x) \geq 0$ for $a \leq x \leq b$,
- has is normalised
$$\int_a^b p(x) ds = 1,$$

- has a cumulative distribution function $U(x)$ defined by

$$U(x) = \int_a^b p(t)dt$$

such that $U(\alpha)$ is the probability that the random variable $x \leq \alpha$.

- When the integral converges, the mean μ of the distributions is

$$\mu = \int_a^b xp(x)dx.$$

- The variance σ^2 is defined by

$$\sigma^2 = \int_a^b (x-\mu)^2 \, p(x)dx.$$

The variance measures the spread of the random variables about the mean.

Expectation operator

The expectation operator E associated with a continuous probability distributions p is equivalent to the mean integration of the distribution above. The expected value of a function of the continuous random variable $f(x)$ is given by

$$E\left[f(x)\right] = \int_a^b f(x)p(x)dx.$$

General results on probabilities

For a random variables X with expectation $E[X]$ we have

$\mathrm{Var}[x] = E[(x - E[x])^2] = E[x^2 - 2xE[x] + E[x]^2] = E[x^2] - E[x]E[x] + E[x]^2 = E[x^2] - E[x]^2.$

The covariance of two random variables x and y is given by

$$\mathrm{Cov}[x, y] = E[(x - E[x])\,(y - E[y])] = \cdots = E[xy] - E[x]E[y]$$

Covariances add:

$$\lambda_x \mathrm{Cov}[x, z] + \lambda_y \mathrm{Cov}[y, z] = \mathrm{Cov}[(\lambda_x x + \lambda_y y), z]$$

for constants λ_x, λ_y and random variables x, y, z.

Unstandardised normal distribution

The unstandardised normal distribution with mean μ and variance σ^2 has a probability density function

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \qquad -\infty < x < +\infty,$$

and is denoted as $N\left(\mu, \sigma^2\right)$.

Useful integrals associated with normal distributions are:

$$\begin{aligned}
\int_{-\infty}^{+\infty} dx \exp\left(-\beta x^2\right) &= \sqrt{\frac{\pi}{\beta}} \\
\int_{-\infty}^{+\infty} dx\ x \exp\left(-\beta x^2\right) &= 0 \\
\int_{-\infty}^{+\infty} dx\ x^2 \exp\left(-\beta x^2\right) &= \frac{1}{2}\sqrt{\frac{\pi}{\beta^3}}.
\end{aligned}$$

The first integral is useful for checking the normalisation of a pdf. The second is used to calculate the mean. The third calculates the variance.

In general

$$\int_{-\infty}^{+\infty} x^n \exp\left(-\beta x^2\right) = \begin{cases} 0, & n \text{ odd} \\ \frac{\Gamma(n+1/2)}{\beta^{n+1/2}} & n \text{ even} \end{cases}$$

where the gamma function $\Gamma(z)$ is defined recursively by

$$\begin{aligned}
\Gamma(z+1) &= z\Gamma(z), \\
\Gamma(1) &= 1 \\
\Gamma(1/2) &= \sqrt{\pi}
\end{aligned}$$

When z is a positive integer, m say, $\Gamma(m+1) = m!$.

Standardised normal distribution

The standarised normal distribution has mean $\mu = 0$ and variance $\sigma^2 = 1$ and is denoted by $N(0, 1)$.

To convert from $N(\mu, \sigma^2)$ to $N(0, 1)$, a change of variables $x = \sigma X + \mu$ to obtain the new pdf $p(X)$.

$$p(X)dX = \frac{1}{\sqrt{2\pi}} e^{-X^2/2} dX$$

The probability that $X \leq d_1$ is given by the normal cumulative distribution function (cdf) Φ:

$$P(X \leq d_1) = \Phi(d_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1} e^{-X^2/2} dX.$$

Note that Φ can be expressed in terms of the **error function**, erf, a standard statistical special function that can be found in most contemporary mathematical software packages:

$$\Phi(x) = \frac{1}{2}\left\{1 + \mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)\right\}.$$

Mean and variances of sums of random variables

Let $x_1, x_2, \cdots, x_n$ be n **independent** random variables with respective means μ_i and variances σ_i^2 and let $a_1, a_2, \cdots, a_n$ and b be real constants. Let

$$X = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + b.$$

- $E[X] = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n + b$.
- $\mathrm{Var}[X] = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2$.

Corollary: If n independent random variables x_i are all drawn from distributions having the same mean μ and variance σ^2, then the random variable

$$\frac{(x_1 + x_2 + \cdots + x_n)}{n},$$

has mean μ but variance σ^2/n.

Note: If the x_i above are normally distributed, then so is X with the mean $E[X]$ and variances $\mathrm{Var}[X]$ stated.

Appendix C

Further Reading

For the first part of the course I recommend:

- *An Introduction to the Mathematics of Finance: A Deterministic Approach,* S.J. Garrett, 2nd ed., Butterworth-Heinemann, 2013. (available on the course website)
- *An Introduction to the Mathematics of Finance,* J.J. McCutcheon and W.F. Scott, Heinemann, 1986.
- *Theory of Interest,* S.J. Garrett, 3rd ed., Irwin, 2008.
- *Southampton: MATH2040 Lecture Notes* G. Kennedy (available on the course website)

For the second part of the course I recommend

- *Options, futures and other derivatives.* Hull, J. C. 7th ed. Prentice Hall, 2002. 744 pages. ISBN: 9780136015764
- *The Mathematics of Financial Derivatives: A Student Introduction* P Wilmott, S Howison & J Dewynne, Cambridge University Press, 1995.
- *Southampton: MATH3022 Lecture Notes* C. Howls and J. Vickers (available on the course website)

The Institute of Actuaries has a list of further reading at

`http://www.actuaries.org.uk/research-and-resources/pages/ct-exams-reading`

- *Financial calculus. An introduction to derivative pricing.* Baxter, M.; Rennie, A. CUP, 1996. 233 pages. ISBN: 9780521552899

- *Financial economics: with applications to investments, insurance and pensions.* Panjer, H. H. (ed) The Actuarial Foundation, 1998. 669 pages. ISBN: 9780938959489
- *Interest rate models. An introduction.* Cairns, Andrew J. G. Princeton University Press, 2004
- *Modern portfolio theory and investment analysis.* Elton, E. J.; Gruber, M. J.; Brown, S. J. et al. 7th ed. John Wiley, 2007. 728 pages. ISBN: 9780470050828
- *Options, futures and other derivatives.* Hull, J. C. 7th ed. Prentice Hall, 2002. 744 pages. ISBN: 9780136015764