Subject CT8

CMP Upgrade 2015/16

CMP Upgrade

This CMP Upgrade lists all significant changes to the Core Reading and the ActEd material since last year so that you can manually amend your 2015 study material to make it suitable for study for the 2016 exams. It includes replacement pages and additional pages where appropriate.

Alternatively, you can buy a full replacement set of up-to-date Course Notes at a significantly reduced price if you have previously bought the full price Course Notes in this subject. Please see our 2016 *Student Brochure* for more details.

This CMP Upgrade contains:

- all changes to the Syllabus objectives and Core Reading
- changes to the ActEd Course Notes, Series X Assignments and Question and Answer Bank that will make them suitable for study for the 2016 exams.

1 Changes to the Syllabus Objectives and Core Reading

1.1 Syllabus Objectives

No changes have been made to the syllabus objectives.

1.2 Core Reading

Chapter 2

The Core Reading in Sections 1.3 (The budget constraint) and 1.4 (How consumers choose) has been deleted. (This appears on pages 6 to 8 in the 2015 Course Notes.) The corresponding points in the summary box on page 39 have also been deleted from the Course Notes.

Chapter 3

The Core Reading in Section 2 (Relationship between dominance concepts and utility theory) has been deleted. (This appears on pages 8 to 10 in the 2015 Course Notes.)

Chapter 5

Page 11

The second sentence has been corrected to read as follows:

With values of –1 and 1, it is possible to obtain risk-free portfolios with zero standard deviation of return, as shown in Question 5.5.

Pages 16-17

The Core Reading in the sub-section on matrix notation (beginning with "**These $N+2$ equations** ..." on page 16 and ending with "... **quadratic in E_P**" on page 17) has been deleted.

Chapter 7

A short section has been added at the end of this chapter outlining the uses of CAPM and the Arbitrage Pricing Theory:

Replacement pages 19-33, which replace pages 19-30 in the 2015 version of the Course Notes, are therefore included at the end of this document.

Chapter 9

The Core Reading in this chapter has been edited and amended extensively. The revised chapter of the Course Notes is therefore included in the replacement pages at the end of this document.

Chapter 14

Some of the Core Reading on pages 8-14 has been changed, with consequent changes to the Course Notes. Replacement pages 7-14 are therefore included at the end of this document.

Chapter 16

Some of the Core Reading on pages 10-14 has been changed, with consequent changes to the Course Notes. In particular, the existing Question 16.8 has been deleted and three new questions added (Questions 16.8 to 16.10 in the 2016 version of the chapter.) The revised chapter of the Course Notes is therefore included in the replacement pages at the end of this document.

2 *Changes to the ActEd Course Notes*

Chapter 2, pages 37-38

The solution to the exam-style question at the end of Chapter 2 has been rewritten with the expected returns and variances in decimal form rather than in % and %% units respectively. This ensures there is no inconsistency between the units of the two terms in the utility function, which would affect the value of α.

The revised solution is included in the replacement pages at the end of this document.

Chapter 6, page 10

The equations in the middle of the page relating to the estimation of beta and alpha have been corrected. Replacement pages 9 and 10 are included at the end of this document.

Chapter 6, page 28

The solution to Question 6.11 has been corrected and now reads as follows:

Recall that the *expected* return is expected return on security i is given by:

$$E_i = \alpha_i + \beta_i E_M$$

So, provided α_i is non-zero, which is usually the case, if β_i doubles, the expected return will not also double.

ie $\quad E_i' = \alpha_i + 2\beta_i E_M \neq 2\left(\alpha_i + \beta_i E_M\right) = 2E_i$

Chapter 16, page 39

There is a typo in the third line of the solution to Question 16.9. The formula for D_t is missing an S_0 and should say $D_t = S_0\, e^{X_t}$.

Chapter 17, page 28

The existing Question 17.17, which is still relevant, has been replaced by the following question.

Question

Summarise the characteristics of the Vasicek, Cox-Ingersoll-Ross (CIR) and Hull & White models.

Solution

The following table summarises the characteristics of the Vasicek, Cox-Ingersoll-Ross (CIR) and Hull & White models.

	Vasicek	CIR	Hull & White
Arbitrage-free	Yes	Yes	Yes
Positive interest rates	No	Yes	No
Mean-reverting interest rates	Yes	Yes	Yes
Easy to price bonds and derivatives	Yes	Yes[(1)]	Yes
Realistic dynamics	No[(2)]	No[(2)]	No[(2)]
Adequate fit to historical data	No	No	Yes
Easy to calibrate accurately to current data	No	No	Yes
Can price a wide range of derivatives	No[(3)]	No[(3)]	No[(3)]

Notes:

(1) Although the CIR model is harder to use that the other two models, it is more tractable than models with two or more factors.

(2) All three models produce perfectly correlated changes in bond prices, which is inconsistent with the empirical evidence, and fail to model periods of high and low interest rates and high and low volatility.

(3) All three models can be used to price short-term, straightforward derivatives (eg European call options), but not complex derivatives.

3 Changes to the Q&A Bank

Solution 5.7 (ii)

A typo has been corrected in the notation used in the final line at the bottom of page 14, which now says $V_1(2) = 65.568$.

Likewise, a typo has been corrected in the notation used in the third line at the top of page 15, which now says $V_0 = 51.922$.

Solution 5.9 (iv)

Typos have been corrected in the first two formulae at the top of page 19. Replacement pages 19 and 20 are included at the end of this document.

4 Changes to the X assignments

No changes have been made to the X Assignments.

5 Other tuition services

In addition to this CMP Upgrade you might find the following helpful with your study.

5.1 Study material

We offer the following study material in Subject CT8:

- Online Classroom
- Flashcards
- Sound Revision
- Revision Notes
- ASET (ActEd Solutions with Exam Technique) and Mini-ASET
- Mock Exam
- Additional Mock Pack.

For further details on ActEd's study materials, please refer to the 2016 *Student Brochure*, which is available from the ActEd website at **www.ActEd.co.uk**.

5.2 Tutorials

We offer the following tutorials in Subject CT8:

- a set of Regular Tutorials (lasting two or three full days)
- a Block Tutorial (lasting two or three full days)
- a Revision Tutorial (lasting one full day)
- Live Online Tutorials (lasting three full days)
- Live Online Revision Tutorials (lasting half a day or a whole day).

For further details on ActEd's tutorials, please refer to our latest *Tuition Bulletin*, which is available from the ActEd website at **www.ActEd.co.uk**.

5.3 Marking

You can have your attempts at any of our assignments or mock exams marked by ActEd. When marking your scripts, we aim to provide specific advice to improve your chances of success in the exam and to return your scripts as quickly as possible.

For further details on ActEd's marking services, please refer to the 2016 *Student Brochure*, which is available from the ActEd website at **www.ActEd.co.uk**.

5.4 Feedback on the study material

ActEd is always pleased to get feedback from students about any aspect of our study programmes. Please let us know if you have any specific comments (*eg* about certain sections of the notes or particular questions) or general suggestions about how we can improve the study material. We will incorporate as many of your suggestions as we can when we update the course material each year.

If you have any comments on this course please send them by email to **CT8@bpp.com** or by fax to **01235 550085**.

Solution

This is Subject 109, April 2003, Question 7.

(i) ***Maximising the investor's expected utility***

Assuming that all of the investor's money is invested, and hence the portfolio weights sum to one, the expected return and variance of a portfolio consisting of a proportion x_A of wealth held in Asset A, and a proportion $1-x_A$ of wealth held in Asset B are:

$$\begin{aligned} E_P &= x_A E_A + (1-x_A)E_B \\ &= 0.06x_A + 0.08(1-x_A) \\ &= 0.08 - 0.02x_A \end{aligned}$$

and:

$$\begin{aligned} V_P &= x_A^2 V_A + x_B^2 V_B + 2x_A x_B \sigma_A \sigma_B \rho_{AB} \\ &= 0.0004x_A^2 + 0.0025(1-x_A)^2 + 0.0010x_A(1-x_A) \\ &= 0.0019x_A^2 - 0.0040x_A + 0.0025 \end{aligned}$$

Therefore the investor's expected utility is:

$$\begin{aligned} E_\alpha(U) &= E(r_p) - \alpha Var(r_p) \\ &= 0.08 - 0.02x_A - \alpha\left(0.0019x_A^2 - 0.0040x_A + 0.0025\right) \end{aligned}$$

We can maximise this function of x_A by differentiating and setting to zero:

$$\frac{dE}{dx_A} = -0.02 - \alpha\left(0.0038x_A - 0.0040\right) = 0$$

$$\Leftrightarrow \qquad x_A = \frac{20\alpha - 100}{19\alpha}$$

or:

$$x_A = \frac{20}{19} - \frac{100}{19\alpha}$$

NB The second-order derivative is:

$$\frac{d^2E}{dx_A^2} = -0.0038\alpha < 0$$

which confirms that we have a maximum.

(ii) ***Show that the investor selects an increasing proportion of Asset A***

Differentiating the formula for the optimal value of x_A in terms of α gives:

$$\frac{dx_A}{d\alpha} = \frac{100}{19\alpha^2} > 0$$

This confirms that as α increases, so x_A, the proportion of wealth held in Asset A, increases too.

2 The single-index model

2.1 Definition

The single-index model as described below is a special case of the multifactor model that includes only one factor, normally the return on the investment market as a whole. It is based upon the fact that most security prices tend to move up or down with movements in the market as a whole. It therefore interprets such market movements as the major influence upon individual security price movements, which are consequently correlated only via their dependence upon the market.

The single-index model is sometimes also called the *market model*. Note that other single-index or one-factor models are possible, in which the single index is a variable other than the market.

The single-index model expresses the return on a security as:

$$R_i = \alpha_i + \beta_i R_M + \varepsilon_i$$

where:

- R_i **is the return on security** i
- α_i, β_i **are constants**
- R_M **is the return on the market**
- ε_i **is a random variable representing the component of** R_i **not related to the market.**

Question 6.5

How can we interpret α_i and β_i?

Under the model ε_i **is uncorrelated with** R_M **and** ε_i **is independent of** ε_j **for all** $i \neq j$**.** It is also normal to set α_i such that $E(\varepsilon_i) = 0$ for $i = 1, \ldots, N$.

So:

- $E[\varepsilon_i] = 0$
- $Cov[\varepsilon_i, \varepsilon_j] = 0$ for all $i \neq j$
- $Cov[\varepsilon_i, R_M] = 0$ for all i.

2.2 *Results of the single-index model*

For any particular security, α and β can be estimated by time series regression analysis.

In order to estimate α and β for security i, the historical returns produced over say $t = 1, \ldots, N$ monthly intervals for both security i, R_{it}, and the market, R_{Mt} say, are required. We can then use regression analysis – as discussed in Subject CT3 – based upon the equation:

$$R_i = \alpha_i + \beta_i R_M + \varepsilon_i$$

Regressing R_i on R_M leads to the following estimates of α_i and β_i:

$$\hat{\beta}_i = \frac{\sigma_{iM}}{\sigma_M^2} = \frac{\sum_{t=1}^{N}[(R_{it} - \overline{R}_i)(R_{Mt} - \overline{R}_M)]}{\sum_{t=1}^{N}(R_{Mt} - \overline{R}_M)^2}$$

and $\quad \hat{\alpha}_i = \overline{R}_i - \hat{\beta}_i \overline{R}_M$

where $\overline{R}_i = \dfrac{1}{N}\sum_{t=1}^{N} R_{it}$ and $\overline{R}_M = \dfrac{1}{N}\sum_{t=1}^{N} R_{Mt}$.

The results of the regression analysis can also be used to estimate $V_{\varepsilon i}$. In each case, there is the usual problem that future values may differ from estimates of past values. An alternative approach is simply to use subjective estimates in the model – though even these are likely to be informed by estimates based on historical data.

The expected return and variance of return on security *i* and the covariance of the returns on securities *i* and *j* are given by:

$$E_i = \alpha_i + \beta_i E_M \tag{6.1}$$

$$V_i = \beta_i^2 V_M + V_{\varepsilon i} \tag{6.2}$$

and $$C_{ij} = \beta_i \beta_j V_M \tag{6.3}$$

where $V_{\varepsilon i}$ is the variance of ε_i .

In order to apply APT, we need to define a suitable multi-index model and then we need to come up with the correct factor forecasts. In practice, any factor model that is good at explaining the return of a diversified portfolio should suffice as an APT model. The exact specification of the factor model may not be important. What is important is that the model contains sufficient factors to capture movement in the important dimensions.

The hard part is the factor forecasts: finding the amount of expected excess return to associate with each factor. The simplest approach is to calculate a history of factor returns and take their average. This implicitly assumes an element of stationarity in the market, *ie* the factors had the same influences in the past as they do today.

A *structural* approach postulates some relationship between specific variables. The variables can be macroeconomic, fundamental or market related. Practitioners tend to prefer structural models since it allows them to connect the factors with specific variables and therefore link their investment experience and intuition to the model.

***Statistical* models are based on statistical analysis of historical data. Academics tend to prefer the pure statistical approach, since they can avoid putting their prejudgements into the model.**

Any test of APT predictions must incorporate a factor model of security returns and is, in effect, a joint test of the equilibrium theory *and* the appropriateness of the selected factor model.

APT implies that, when all pervasive factors are taken into account, the remaining portion of return on a typical security should be expected to equal the risk-free interest rate. Testing this implication is possible in principle, but difficult in practice. Here, "pervasive" means "relevant" or "influential".

A more promising test concerns the prediction that security expected returns will be related only to sensitivities to pervasive factors. In particular, there should be no relationship between expected returns and securities' non-factor risks.

4 Uses of CAPM and the APT

As its name suggests, the capital asset pricing model (CAPM) can be used to price assets, both financial securities and other assets such as capital projects. If the beta of an asset can be estimated from past data, then the security market line can be used to estimate the prospective return that the asset should offer given its systematic risk. This return can then be used to discount projected future cash flows and so price the asset. For a new asset with no past history, the beta from a similar asset could be used.

The arbitrage pricing theory (APT) can be used for both passive and active portfolio management.

Question 7.16

What is the difference between passive and active portfolio management?

For example, it may be desired to track a share index but without holding all of the constituent shares in the index. In this instance, APT could be used to find a smaller sample of shares with the same sensitivities to the same factors as the index. This sample should then broadly replicate the future performance of the index.

More generally, by using APT to estimate the exposure of a portfolio to different risk factors, the extent of that exposure can be managed as required.

Finally, as with CAPM, APT can be used to estimate the expected return on a financial security given its exposure to the various risk factors modelled. This return can then be used to discount projected future cash flows and so price the security and determine if it appears to be under-valued or over-valued.

More detail on both CAPM and APT can be found in the book by Elton & Gruber listed in the Study Guide.

Question 7.17

How might you estimate the beta of a quoted share in practice?

5 Exam-style question

We finish this chapter with an exam-style question based on the capital asset pricing model (CAPM).

Question

This question is taken from Subject CT8 April 2005 Question 3.

An investor has the choice of the following assets that earn rates of return as follows in each of the four possible states of the world:

State	*Probability*	*Asset 1*	*Asset 2*	*Asset 3*
1	0.2	5%	5%	6%
2	0.3	5%	12%	5%
3	0.1	5%	3%	4%
4	0.4	5%	1%	7%
Market capitalisation		10,000	17,546	82,454

Determine the market price of risk assuming CAPM holds.

Define all terms used. [6]

Solution

The market price of risk is:

$$\frac{E_M - r}{\sigma_M}$$

where:

- r is the risk-free interest rate
- E_M is the expected return on the market portfolio consisting of all risky assets
- σ_M is the standard deviation of the return on the market portfolio.

Since Asset 1 always gives the same return of 5%, it is risk-free. So the risk-free interest rate is $r = 5\%$.

Assets 2 and 3, with the capitalisations shown, constitute the market portfolio of risky assets.

The total capitalisation of the market is $17{,}546 + 82{,}454 = 100{,}000$. The table below shows the possible returns on this market portfolio:

State	*Probability*	*Return*	*Return (%)*
1	0.2	$5\% \times 17{,}546 + 6\% \times 82{,}454 = 5{,}824.54$	5.82454%
2	0.3	$12\% \times 17{,}546 + 5\% \times 82{,}454 = 6{,}228.22$	6.22822%
3	0.1	$3\% \times 17{,}546 + 4\% \times 82{,}454 = 3{,}824.54$	3.82454%
4	0.4	$1\% \times 17{,}546 + 7\% \times 82{,}454 = 5{,}947.24$	5.94724%

So the expected return is:

$$E_M = 0.2 \times 5.82454\% + \cdots + 0.4 \times 5.94724 = 5.794724\%$$

The variance of the returns is:

$$\begin{aligned}\sigma_M^2 &= 0.2 \times (5.82454\%)^2 + \cdots + 0.4 \times (5.94724\%)^2 - (5.794724\%)^2 \\ &= 0.454020\%\% = (0.673810\%)^2\end{aligned}$$

So the market price of risk is:

$$\frac{E_M - r}{\sigma_M} = \frac{5.794724\% - 5\%}{0.673810\%} = 1.179$$

This shows the extra expected return (over and above the risk-free rate) per unit of extra risk taken (as measured by the standard deviation) by investing in risky assets.

This page has been left blank so that you can keep the chapter summaries together for revision purposes.

Chapter 7 Summary

Assumptions of the CAPM (including MPT assumptions)

- Investors make their decisions purely on the basis of expected return and variance. So all expected returns, variances and covariances of assets are known.
- Investors are non-satiated and risk-averse.
- There are no taxes or transaction costs.
- Assets may be held in any amounts.
- All investors have the *same* fixed one-step time horizon.
- All investors make the *same* assumptions about the expected returns, variances and covariances of assets.
- All investors measure returns consistently (*eg* in the same currency or in the same real/nominal terms).
- The market is perfect.
- All investors may lend or borrow any amounts of a risk-free asset at the same risk-free rate r.

The extra assumptions of CAPM from MPT move away from thinking about individual investors to assumptions about the entire economy. CAPM is an *equilibrium model*.

Results of the CAPM

- All investors have the same efficient frontier of risky assets.
- The efficient frontier collapses to a straight line in $E-\sigma$ space in the presence of the risk-free asset.
- All investors hold a combination of the risk free asset and the same portfolio of risky assets *M*.
- *M* is the market portfolio – it consists of all assets held in proportion to their market capitalisation.

The *separation theorem* suggests that the investor's choice of portfolio of risky assets is independent of their utility function.

Capital market line

$$E_p = r + \frac{\sigma_P}{\sigma_M}(E_M - r)$$

Market price of risk

$$MPR = \frac{E_M - r}{\sigma_M}$$

Security market line

$$E_p = r + \beta_P(E_M - r)$$

where:

$$\beta_P = \frac{Cov\left(R_P, R_M\right)}{Var\left(R_M\right)}$$

The main limitations of the basic CAPM are that most of the assumptions are unrealistic and that empirical studies do not provide strong support for the model. However there is some evidence to suggest a linear relationship between expected return and systematic risk.

Arbitrage pricing theory (APT)

$$E[R_i] = \lambda_0 + \lambda_1 b_{i,1} + \lambda_2 b_{i,2} + \ldots + \lambda_L b_{i,L}$$

APT moves away from thinking about individual investors in the multifactor model to thinking about the entire economy. In this sense, it is an equilibrium model, although the general result is derived from the principle of *no arbitrage*.

The principal strength of the APT approach is that it is based on the no-arbitrage conditions.

APT is extremely general, which is both a strength and a weakness:

- Although it allows us to describe equilibrium in terms of any multi-index model, it gives us no evidence as to what might be an appropriate multi-index model.
- APT tells us nothing about the size or the signs of the λ's.

Chapter 7 Solutions

Solution 7.1

The assumptions of mean-variance portfolio theory are that:

- all expected returns, variances and covariances of pairs of assets are known
- investors make their decisions purely on the basis of expected return and variance
- investors are non-satiated
- investors are risk-averse
- there is a fixed single-step time period
- there are no taxes or transaction costs
- assets may be held in any amounts, (with short-selling, infinitely divisible holdings, no maximum investment limits).

Solution 7.2

If investors:

- have the same estimates of the expected returns, standard deviations and covariances of securities over the one-period horizon and
- are able to perform correctly all the requisite calculations

then they will all arrive at the same opportunity set and hence the same efficient frontier of risky securities.

Solution 7.3

Firstly, we can use the formula for the expected return of the portfolio to express x_A in terms of E_P:

$$E_P = x_A E_A + x_B E_B = E_P = x_A E_A + (1 - x_A) E_B$$

$$\Leftrightarrow \qquad x_A = \frac{E_P - E_B}{E_A - E_B}$$

Also, recall from Chapter 5 that the variance of the portfolio return is:

$$V_P = x_A{}^2 V_A + x_B{}^2 V_B + 2x_A x_B C_{AB}$$

Because Portfolio B is risk-free, $V_B = 0$ and $C_{AB} = 0$. So the above equation simplifies to:

$$V_P = x_A{}^2 \sigma_A^2 = \left(\frac{E_P - E_B}{E_A - E_B} \right)^2 \sigma_A^2$$

$$\Leftrightarrow \qquad \sigma_P = \left(\frac{E_P - E_B}{E_A - E_B} \right) \sigma_A = \frac{\sigma_A}{E_A - E_B} E_P - \frac{\sigma_A E_B}{E_A - E_B}$$

This is a straight line in (E_P, σ_P) space.

Solution 7.4

The last question shows that the new efficient frontier must be a straight line. We also know that it must intercept the *E*-axis at *r* because the risk-free asset is efficient. (You cannot find a portfolio with the same zero variance and a higher return.)

If this straight line isn't at a tangent then it either passes above the old efficient frontier or it passes below it. It cannot pass above since there is no portfolio that exists here! If it passes below then it is not an efficient frontier because you can find portfolios of risky assets that have a higher expected return for the same variance. Therefore the only possibility is that it must be at a tangent.

Solution 7.5

We know that the points $(0,r)$ and (σ_M, E_M) are on the straight line and so its equation is:

$$\frac{E_P - r}{\sigma_P - 0} = \frac{E_M - r}{\sigma_M - 0}$$

We rearrange this to get the equation given for the capital market line.

Solution 7.6

(i) ***Beta of Security A***

This is given by:

$$\beta_A = \frac{Cov(R_A, R_M)}{V_M} = \frac{\rho_{AM}\,\sigma_A\,\sigma_M}{\sigma_M^2}$$

$$= \frac{0.75 \times 0.04 \times 0.05}{0.05^2} = 0.6$$

(ii) ***Expected return of Security A***

This is given by:

$$E_A = r + \beta_A(E_M - r)$$

$$= 0.05 + 0.6(0.10 - 0.05)$$

$$= 0.08$$

ie 8%.

Solution 7.7

If this were not the case, then it would be possible to make an instantaneous, risk-free profit – *ie* an arbitrage profit. For example, suppose that the beta relationship held, but that the expected return from Portfolio 3 was less than E_3. Then it would be possible to make arbitrage profits by selling Portfolio 3 and using the proceeds to buy equal amounts of Portfolios 1 and 2. Starting with a zero initial sum, you could end up with a positive net expected return and hence a risk-free profit.

In practice, we would expect arbitrageurs to notice the price anomaly and act exactly as above. Thus, the price of Portfolio 3 would be driven down (and hence its expected return up) and the prices of Portfolios 1 and 2 up (and hence their expected returns down) until the arbitrage possibility was eliminated – with the security market line relationship again holding.

Recall that the capital asset pricing model is an *equilibrium* model. Therefore short-term deviations from the predicted expected returns may be possible when the market is out of equilibrium.

Solution 7.8

The beta factor of any portfolio *i* is defined as $Cov[R_i, R_M]/V_M$. Hence, for the market portfolio:

$$\beta_M = Cov[R_M, R_M]/V_M = V_M/V_M = 1$$

This must be the case since the return on the market is perfectly correlated with itself (*ie* the correlation coefficient equals one).

Conversely, the risk-free asset has, by definition, neither systematic nor specific risk and so its beta must be zero.

Now, the excess expected return on the market portfolio over and above the risk-free rate is $E_M - r$, whilst the excess systematic risk is $\beta_M = 1$. Hence, $E_M - r$ must also be the gradient of the security market line. As all portfolios lie on the security market line, for any portfolio with a beta β_P the excess expected return over and above the risk-free rate will equal $\beta_P \times (E_M - r)$.

Consequently, the *total* expected return on the same portfolio must be equal to:

$$E_P = \beta_P \times (E_M - r) + r$$

ie $a_0 = r$ and $a_1 = E_M - r$

Solution 7.9

You *may* recall that Subject CT1 defines arbitrage as follows:

Arbitrage is generally described as a risk-free trading profit. More accurately, an arbitrage opportunity exists if either:

- an investor can make a deal that would give her or him an immediate profit, with no risk of future loss
- an investor can make a deal that has zero initial cost, no risk of future loss, and a non-zero probability of a future profit.

Solution 7.10

Consider a portfolio with weight x in Asset A and $1-x$ in Asset B. This will have equal sensitivity to factors I_1 and I_2 if:

$$2x + (1-x) = x + 2(1-x)$$

$$\Leftrightarrow \quad x = ½$$

ie if it is equally weighted between Assets A and B.

Hence:

$$E[R_C] = 4.5 + 1.5I_1 + 1.5I_2$$

ie the price of C is determined by reference to the price (or expected return) of a portfolio of 50% in A and 50% in B.

The purpose of this question is to illustrate how, the prices of assets under APT, are determined by equivalent portfolios, in terms of the exposure to the underlying factors (that generate the investment returns).

Solution 7.11

Systematic risk is that element of the unpredictability of investment returns that cannot be eliminated by diversification.

Solution 7.12

It is valid to represent the portfolios in this way precisely because the only features of a portfolio that *are* of concern to investors are expected return E and the systematic risks relating to each of factors I_1 and I_2. Investors' portfolio choices are therefore a function of these three variables and no others.

Solution 7.13

Just as any two points in any two-dimensional space define a straight line, so any three points in three-dimensional space define a unique plane – unless they all lie on the same straight line. The general equation for a plane in $E - b_1 - b_2$ space has been rearranged to make expected return a function of the sensitivities to the two factors that determine systematic risk.

Solution 7.14

λ_k is the additional expected return or *risk premium* stemming from a unit increase in $b_{i,k}$, the sensitivity of security *i*'s investment return to index *k*.

Solution 7.15

(i) The fact that $b_{i,2}$ is greater than $b_{i,1}$ means that the expected return on the security is more sensitive to the influence of Factor 2 than it is to the influence of Factor 1. Equivalently, we can interpret this as indicating that the security has more exposure to Factor 2 than to Factor 1.

(ii) The expected return is given by:

$$\begin{aligned} E_i &= \lambda_0 + \lambda_1 b_{i,1} + \lambda_2 b_{i,2} \\ &= 3 + 0.9 \times 0.5 + 0.8 \times 1.5 \\ &= 4.65\% \end{aligned}$$

Solution 7.16

Passive portfolio management involves holding a portfolio of assets in order to meet a specified investment objective and only changing that portfolio reactively in response to a change in the objective. For example, this might be in response to a change in the constituents of the index being tracked or the liabilities being matched.

Active portfolio management involves actively identifying and trading mispriced assets in order to make excess risk-adjusted returns, *eg* returns in excess of a particular index.

Solution 7.17

To estimate the beta of the share you could obtain the returns on the share in question (R_i) over each of the last sixty months say, together with the monthly returns over the same period on an appropriate market-wide index (R_M). For example, in the UK you might use the FTSE All-Share Index.

The beta of the share could then be estimated as:

$$\beta_i = \frac{Cov\left(R_i, R_M\right)}{Var\left(R_M\right)}$$

Chapter 9

Stochastic calculus and Ito processes

Syllabus objectives

(viii) Define and apply the main concepts of Brownian motion (or Wiener processes).

2. Demonstrate a basic understanding of stochastic differential equations, the Ito integral, diffusion and mean-reverting processes.

3. State Ito's formula and be able to apply it to simple problems.

4. Write down the stochastic differential equation for geometric Brownian motion and show how to find its solution.

5. Write down the stochastic differential equation for the Ornstein-Uhlenbeck process and show how to find its solution.

0 Introduction

This chapter is concerned with stochastic calculus, in which continuous-time stochastic processes are described using stochastic differential equations. As is the case in a non-stochastic setting (*eg* in mechanics), these equations can sometimes be solved to give formulae for the functions involved. The solutions to these equations often involve *Ito integrals*, which we will look at in some detail. The other key result from stochastic calculus is Ito's Lemma. This is used to determine the stochastic differential equation for a stochastic process whose values are a function of another stochastic process.

Diffusions are a generalisation of Brownian motion in which the constraint that the increments are independent is dropped. However, a slightly weaker condition, known as the "Markov property", is retained. Such processes can be thought of as Brownian motion where the drift and diffusion coefficients are variable.

The main example given is the *Ornstein-Uhlenbeck process*. It is mean-reverting, that is, when the process moves away from its long-run average value, there is a component that tends to pull it back towards the mean. For this reason the process can be used to model interest rates, which are usually considered to be mean-reverting.

As another example we will discuss *geometric Brownian motion*. This can be used to model share prices. Here the log of the share price is assumed to follow Brownian motion, so the model is sometimes known as the *lognormal model*. The advantages of this model over Brownian motion itself are discussed.

1 Stochastic calculus

1.1 Introduction

Newton originally developed calculus to provide the necessary mathematics to handle his laws describing the motion of bodies. His second law, for example, can be written as $F = ma$, where F is the force applied to a body, m its mass, and a is the resulting acceleration. The acceleration is the time derivative of velocity, which is in turn the time derivative of position. We therefore arrive at the differential equation:

$$F = m\ddot{x} = m\frac{d^2x}{dt^2}$$

This is the model of the motion of the body.

This is all very well so long as the path followed by a body or particle is sufficiently smooth to differentiate, as is generally the case in Newtonian mechanics. However, there are situations when the paths followed are not sufficiently smooth. For example, you may have studied impulses *eg* when two snooker balls collide. In these cases the velocity of the objects involved can change suddenly and $\frac{dx}{dt}$ is not a differentiable function.

As mentioned at the start of the previous chapter, the original Brownian motion referred to the movement of pollen grains suspended in a liquid. Each pollen grain is very light, and therefore jumps around as it is bombarded by the millions of molecules that make up the liquid, giving the appearance of a very random motion. A stochastic model of this behaviour is therefore appropriate. A deterministic model in terms of the underlying collisions wouldn't be very practical.

The sample paths of this motion are not sufficiently smooth however. As we have seen, they are differentiable nowhere. Therefore, a description of the motion as a differential equation in the usual sense is doomed to failure. In order to get around this, a new *stochastic calculus* has to be developed. This turns out to be possible and allows the formulation of stochastic differential equations (SDEs).

The stochastic differential equations that we deal with will be continuous-time versions of the equations used to define time series, *ie* stochastic processes operating in discrete time. For example, you may recall that a zero-mean random walk X_t can be defined by an equation of the form:

$$X_t = X_{t-1} + \sigma Z_t \quad \text{or} \quad X_t - X_{t-1} = \sigma Z_t$$

where Z_t is a standard normal random variable. The Z_t's in this equation are called *white noise*.

This is a stochastic difference equation: a "difference" equation, since it involves the difference $X_t - X_{t-1}$, and "stochastic" because the white noise terms are random. It can be "solved" to give:

$$X_t = X_0 + \sigma \sum_{s=1}^{t} Z_s$$

In continuous time, the analogue of a zero-mean random walk is a zero-mean Brownian motion, say W_t. The change in this Brownian motion over a very short time period (in fact, an infinitesimal time period) will be denoted by $dW_t = W_{t+dt} - W_t$. Since Brownian motion increments are independent, we can think of dW_t as a continuous-time white noise. In fact, we have:

$$\operatorname{cov}\left(dW_s, dW_t\right) = \begin{cases} 0 & s \neq t \\ \sigma^2 dt & s = t \end{cases}$$

For a standard Brownian motion this would be:

$$\operatorname{cov}\left(dB_s, dB_t\right) = \begin{cases} 0 & s \neq t \\ dt & s = t \end{cases}$$

We therefore have the stochastic differential equation:

$$dW_s = \sigma dB_s$$

This can be solved by integrating both sides between 0 and *t* to give:

$$W_t - W_0 = \sigma \int_0^t dB_s$$

$$\Leftrightarrow \qquad W_t = W_0 + \sigma \int_0^t dB_s$$

Compare this to the discrete-time case $X_t = X_0 + \sigma \sum_{s=1}^{t} Z_s$.

The analogy is that the dB_s process is considered as a continuous-time white noise process and, because we're working in continuous time, we need to integrate, rather than sum the terms. The existence, meaning and properties of such integrals are discussed in this section, together with some more interesting examples.

1.2 The Ito Integral

When attempting to develop a calculus for Brownian motion and other diffusions, one has to face the fact that their sample paths are nowhere differentiable (see Section 1.5 of Chapter 8).

A direct approach to stochastic integrals like $\int_0^t Y_s dB_s$ is therefore doomed to failure.

An integral like this in which we are integrating with respect to Brownian motion is called an *Ito integral*. It is the fact that we are integrating with respect to the Brownian motion that is the problem. Integration of random variables with respect to a deterministic variable *x* can be dealt with in the standard way.

A quick review of some basic integration results and notation will be helpful.

An integral such as $\int_a^b dx = b - a$ should be very familiar. The integral sign can be interpreted simply as a summation. The summands are the *dx* expressions. These represent small changes in the value of x. Therefore, the integral just says that summing up all the small changes in *x* between *a* and *b* gives the total change $b - a$.

Similarly, the integral $\int_a^b df(x) = f(b) - f(a)$ just gives the total change in the function f as x varies between a and b. This notation may be less familiar, but this is what it means.

The integral $\int_a^b g(x)\,df(x)$ can be evaluated directly if $f(x)$ is a differentiable function since then we can use $\int_a^b g(x)\,df(x) = \int_a^b g(x)\frac{df}{dx}dx$. However, if we want to integrate $\int_0^t Y_s dB_s$ where Y_t is a (possibly random) function of t, and B_t is a standard Brownian motion, then we cannot apply the above method, as B_t is not differentiable.

However, such Ito integrals can be given a meaning for a suitable class of F_s-measurable random integrands Y_s. This involves a method of successive approximation by step functions.

We will illustrate this approach assuming that the integrand is not random, but just a deterministic function $f(s)$.

Ito integrals for deterministic functions

Firstly, when we integrate the constant function $f(s) \equiv 1$ we expect:

$$\int_0^t dB_s = [B_s]_0^t = B_t - B_0 = B_t$$

This is basically what integration means. We add up the (infinitesimal) increments dB_s to get the overall increment $B_t - B_0$. It is worth noting that the increments of Brownian motion are just normal random variables, and furthermore, the increments over disjoint time periods are independent. In "adding" up the increments dB_s we are effectively summing independent normal random variables. Moreover, the increment dB_s should have a $N(0, ds)$ distribution. Again this is consistent with the value of the integral (B_t), which has a $N(0,t)$ distribution.

Also, any constant multiple of the integrand should just multiply the integral. For example:

$$\int_0^t 2dB_s = 2\left[B_s\right]_0^t = 2\left(B_t - B_0\right) = 2B_t$$

Finally, the integrals should add up in the usual way over disjoint time periods. For example, if we have the function:

$$f(s) = \begin{cases} 1 & 1 \le s < 2 \\ 2 & 2 \le s < 3 \\ 0 & \text{otherwise} \end{cases}$$

Then: $\displaystyle\int_1^3 f(s)\,dB_s = \int_1^2 1dB_s + \int_2^3 2dB_s = \left[B_s\right]_1^2 + \left[2B_s\right]_2^3 = 2B_3 - B_2 - B_1$

More generally:

$$\int_1^t f(s)\,dB_s = \begin{cases} \displaystyle\int_1^t 1dB_s = B_t - B_1 & 1 \le t < 2 \\ \displaystyle\int_1^2 1dB_s + \int_2^t 2dB_s = 2B_t - B_2 - B_1 & 2 \le t < 3 \end{cases}$$

Example

What distribution does $\int_{1.5}^{2.5} f(s)\,dB_s$ have?

Solution

$$\int_{1.5}^{2.5} f(s)\,dB_s = \int_{1.5}^{2} 1dB_s + \int_{2}^{2.5} 2dB_s = \left[B_s\right]_{1.5}^{2} + \left[2B_s\right]_{2}^{2.5} = \left(B_2 - B_{1.5}\right) + 2\left(B_{2.5} - B_2\right)$$

Now we know that the two terms on the RHS are independent normal random variables with distributions $N(0,0.5)$ and $N(0,2)$ respectively. (Remember that constants square when taking the variance.) It follows that the original integral has distribution:

$$\int_{1.5}^{2.5} f(s)\,dB_s \sim N(0,2.5)$$

In summary, for these simple cases, we can think of the integration as summing independent normal variables.

It should be obvious from the above that we can integrate any function $f(s)$ that is piecewise constant by splitting the integral up into the constant pieces, and then adding up the answers (assuming this turns out to be finite).

Even without introducing random integrands, the problem of how to integrate more general functions, such as $\int_a^b f(s)\,dB_s$, remains.

This general integral can be thought of as the continuous-time limit of a summation. Consider discretising the interval $[a,b]$ using discrete times $\{s_0, s_1, \ldots, s_n\}$ where $s_0 = a$ and $s_n = b$. The infinitesimal increments can then be replaced by the finite increments $\Delta B_s = B_s - B_{s-1}$. Assuming n is large, $f(s)$ can be approximated by $f(s_{i-1})$ where $s_{i-1} \le s < s_i$. The integral above can then be defined as the limit of a summation:

$$\int_a^b f(s)\,dB_s = \lim_{n\to\infty} \sum_{i=1}^{n} f(s_{i-1})\,\Delta B_{s_i}$$

What is the distribution of this integral? Again, the approximation helps. The distribution of each summand is known:

$$f(s_{i-1})\Delta B_{s_i} \sim N\left(0, f^2(s_{i-1})(s_i - s_{i-1})\right)$$

And since the summands are independent we get:

$$\int_a^b f(s)\,dB_s \sim \lim_{n\to\infty}\sum_{i=1}^{n} N\left(0, f^2(s_{i-1})(s_i - s_{i-1})\right) \sim \lim_{n\to\infty} N\left(0, \sum_{i=1}^{n} f^2(s_{i-1})(s_i - s_{i-1})\right)$$

where we just assume this limit makes sense. In fact, as n gets large, the finite increment $s_i - s_{i-1}$ becomes the infinitesimal increment ds, and the summation becomes an integral. Therefore:

$$\int_a^b f(s)\,dB_s \sim N\left(0, \int_a^b f^2(s)\,ds\right)$$

Thus, the integral of any deterministic function $f(s)$ with respect to the Brownian motion is normally distributed with zero mean, and variance given by an ordinary integral.

Once you get used to the notation, you needn't revert to the summation notation – just interpret the integral directly as a sum. For example, since $dB_s \sim N(0, ds)$, we must have $f(s)\,dB_s \sim N\left(0, f^2(s)\,ds\right)$ since we're just multiplying a normal random variable by a constant. (We are thinking of s as being fixed when we do this.) Furthermore, these random variables, for different values of s, are independent and, since independent normal random variables have an additive property, we arrive at:

$$\int_a^b f(s)\,dB_s \sim N\left(0, \int_a^b f^2(s)\,ds\right)$$

One final property of the integral is that it is a martingale when considered as a process with respect to t, *ie* if we define the process $X_t = \int_a^t f(s)\,dB_s$. Intuitively, since the process has zero-mean increments it should continue "straight ahead" on average.

Mathematically, if $u < t$:

$$E\left[X_t \mid F_u\right] = E\left[\int_a^u f(s)\,dB_s + \int_u^t f(s)\,dB_s \,\middle|\, F_u\right] = X_u + E\left[\int_u^t f(s)\,dB_s \,\middle|\, F_u\right] = X_u$$

The last equality follows because we know that $\int_u^t f(s)\,dB_s$ is normal with mean zero.

Note that a similar approach can be used if we have an Ito integral involving a random integrand Y_s.

It can be shown that an Ito integral has the following properties:

(i) $\left\{\int_0^t Y_s\,dB_s : t \geq 0\right\}$ **is a martingale**

(ii) $E\left[\int_0^t Y_s\,dB_s\right] = 0$

(iii) $E\left[\left(\int_0^t Y_s\,dB_s\right)^2\right] = E\left[\int_0^t Y_s^2\,ds\right]$

ie $E\left[\left(\int_0^t Y_s\,dB_s\right)^2\right] = E\left[\int_0^t Y_s^2\,ds\right] = \int_0^t E\left[Y_s^2\right]ds$

The last equation follows here because the expectation of a sum (*ie* an integral here) is the sum of the expectations.

(iv) The sample paths of $\int_0^t Y_s\,dB_s$ **are continuous.**

We saw earlier that if the integrand is deterministic, then the Ito integral has a normal distribution, with a mean of zero and a variance of $\int_0^t Y_s^2\,ds$.

However, in the stochastic case the integral does not have to be normally distributed, since the stochastic process introduces another source of randomness.

Note that the method of successive approximation by step functions is not one that would be used in practice. It is a theoretical device for proving that such an integral exists. The practical applications of the theory use Ito's Lemma (Result 9.1 on page 16 below) to identify the integrated process.

The Ito integral is well defined. Some of its properties are in sharp contrast to those of usual integrals. For instance, $\int_0^t B_s dB_s$ cannot equal $\tfrac{1}{2}B_t^2$ as the latter is not a martingale, and would be in contravention of property (i) above. In fact, we shall see below that $\int_0^t B_s dB_s = \frac{1}{2}(B_t^2 - t)$, which we know is a martingale.

Recall that we showed earlier that $B_t^2 - t$ was a martingale. Multiplying by a constant, such as ½, preserves the martingale property. So $\frac{1}{2}(B_t^2 - t)$ is also a martingale.

Even though no proper stochastic differential calculus can exist (as opposed to stochastic *integral* calculus) **because of Result 8.1 from Chapter 8, it is common practice to write equations like:**

$$X_t = X_0 + \int_0^t Y_s dB_s + \int_0^t A_s ds \tag{9.1}$$

(where A_s is a deterministic process) in differential notation, namely:

$$dX_t = Y_t dB_t + A_t dt \tag{9.2}$$

We emphasise that (9.2) is purely a shorthand form of (9.1).

If you are going to be able to express stochastic integral equations in differential form then it is important that you can do the same in the usual non-stochastic case. So we'll do a bit of revision of "ordinary" calculus.

Example

Rewrite the equation $x(t) = x(0) + \int_0^t y(s)\,df(s) + \int_0^t a(s)\,ds$ in differential form, explaining the method.

Solution

We have: $$x(t) = x(0) + \int_0^t y(s)\,df(s) + \int_0^t a(s)\,ds$$

and therefore also: $$x(t+dt) = x(0) + \int_0^{t+dt} y(s)\,df(s) + \int_0^{t+dt} a(s)\,ds$$

Subtracting we get: $$x(t+dt) - x(t) = \int_t^{t+dt} y(s)\,df(s) + \int_t^{t+dt} a(s)\,ds$$

ie $$dx(t) = y(t)\,df(t) + a(t)\,dt$$

The derivation of (9.2) by analogy should be apparent.

Question 9.1

Express the general Brownian motion $X_t = \sigma B_t + \mu t$ in the form of the integral equation (9.1). What does this become in differential form?

We have shown that a stochastic calculus exists and satisfies certain properties. However, with standard calculus we have rules that allow us to integrate and differentiate, *eg* the product rule, the quotient rule, and the chain rule.

The key result of stochastic calculus is Ito's lemma. This is the stochastic calculus version of the chain (function-of-a-function) rule and the only rule that we will need.

To be consistent with what is to follow, we will first derive the chain rule for standard calculus. Suppose we have a function-of-a-function $f(b_t)$ and we want to find $\frac{d}{dt} f(b_t)$. We first write down Taylor's theorem to second-order:

$$\delta f(b_t) = f'(b_t)\,\delta b_t + \tfrac{1}{2} f''(b_t)(\delta b_t)^2 + \ldots$$

You may find it helpful to refer to the formulae in Section 1.2 of the *Tables* if you are unfamiliar with Taylor series.

Now dividing by δt and letting $\delta t \to 0$ gives:

$$\frac{df(b_t)}{dt} = f'(b_t)\frac{db_t}{dt} + \lim_{\delta t \to 0} ½ f''(b_t)\frac{(\delta b_t)^2}{\delta t}$$

Since:

$$\lim_{\delta t \to 0} \frac{(\delta b_t)^2}{\delta t} = \lim_{\delta t \to 0} \frac{\delta b_t}{\delta t} \times \delta b_t = \frac{db_t}{dt}\left(\lim_{\delta t \to 0} \delta b_t\right) = 0$$

the second term on the right-hand side must vanish, giving the chain rule:

$$\frac{df(b_t)}{dt} = f'(b_t)\frac{db_t}{dt}$$

or in different notation:

$$df(b_t) = f'(b_t)\,db_t$$

What does this become if we replace the function b_t by the non-differentiable Brownian motion B_t? The analysis starts in much the same way. We can write Taylor's theorem to second-order as:

$$\delta f(B_t) = f'(B_t)\delta B_t + ½ f''(B_t)(\delta B_t)^2 + \cdots$$

Now, in the standard case, taking the limit $\delta t \to 0$ effectively involves replacing δ by d and ignoring second-order and higher-order terms. However, with Brownian motion, it turns out that the second-order term $(dB_t)^2$ cannot be ignored. In fact, it must be changed to dt, *ie* "$(dB_t)^2 = dt$". This is not rigorous, but is a useful rule of thumb. What we end up with is therefore:

$$df(B_t) = f'(B_t)\,dB_t + ½ f''(B_t)\,dt$$

This is Ito's lemma for functions of Brownian motion, *ie* it tells us how to differentiate functions of Brownian motion. Note, however, that this statement must be interpreted in terms of integrals, since Brownian motion is not differentiable.

Example

Find the stochastic differential equation for B_t^2.

Solution

Applying the above formula we have:

$$d\left(B_t^2\right) = 2B_t dB_t + \tfrac{1}{2}2dt = 2B_t dB_t + dt$$

What does this actually mean? As we keep saying, this can only be interpreted sensibly in terms of integrals. If we integrate both sides from 0 to *s*, say, we get:

$$\int_0^s d\left(B_t^2\right) = \int_0^s 2B_t dB_t + \int_0^s dt$$

The left-hand side and second term on the right-hand side can be evaluated:

$$\left[B_t^2\right]_0^s = B_s^2 = \int_0^s 2B_t dB_t + s$$

Finally, rearranging this equation tells us that:

$$\int_0^s B_t dB_t = \tfrac{1}{2}\left(B_s^2 - s\right)$$

This last example shows how Ito's lemma can be used to evaluate Ito integrals.

The above version of Ito's lemma only dealt with functions of Brownian motion. We now generalise this to consider all (time-homogenous) diffusion processes. Diffusion processes X_t are processes that satisfy stochastic differential equations of the form:

$$dX_t = \mu\left(X_t\right)dt + \sigma\left(X_t\right)dB_t$$

We may refer to a process satisfying such an SDE as an *Ito process*, since for our purposes this is just an alternative description of a diffusion process.

Let's compare these processes with a general Brownian motion $W_t = \mu t + \sigma B_t$. The Brownian motion satisfies the differential equation:

$$dW_t = \mu dt + \sigma dB_t$$

(Note this follows straight from the equation $W_t = \mu t + \sigma B_t$. We don't need to apply Ito's lemma here.) Comparing the two stochastic differential equations shows that a general diffusion process is simply a Brownian motion whose drift and diffusion can change with X_t.

When dealing with a function of a diffusion process $f\left(X_t\right)$ we can proceed as with the functions of Brownian motion. Taylor's theorem to second-order is:

$$\delta f\left(X_t\right) = f'\left(X_t\right)\delta X_t + \tfrac{1}{2} f''\left(X_t\right)\left(\delta X_t\right)^2 + \cdots$$

which we can write in the limit as:

$$df\left(X_t\right) = f'\left(X_t\right)dX_t + \tfrac{1}{2} f''\left(X_t\right)\left(dX_t\right)^2$$

higher-order terms being 0 in the limit. (We haven't proved this, but it is true.)

We would next want to substitute in for dX_t in terms of dt and dB_t. Rather than do so now, we will introduce one final complication. This is that the form of Ito's lemma that we will consider is for functions not just of a diffusion process, but also functions that explicitly depend on time – in other words, functions of the form $f\left(t, X_t\right)$.

Before proceeding to the result itself, we will quickly remind ourselves of the classical chain rule in two variables. If we have $f\left(x(t), y(t)\right)$, then:

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

In particular, if we have $f\left(t, x(t)\right)$, then this becomes:

$$\frac{df}{dt} = \frac{\partial f}{\partial t}\frac{dt}{dt} + \frac{\partial f}{\partial x}\frac{dx}{dt} \qquad \text{or equivalently} \qquad df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx$$

Result 9.1 (Ito's Lemma)

Let $\{X_t, t \geq 0\}$ be of the form (9.2) and let $f: R \times R \to R$ be twice partially differentiable with respect to x and once with respect to t. Then $f(t, X_t)$ is also of the form of (9.2) with:

$$df(t, X_t) = \frac{\partial f}{\partial x} Y_t dB_t + \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} A_t + \tfrac{1}{2} \frac{\partial^2 f}{\partial x^2} Y_t^2 \right] dt \qquad (9.3)$$

On grounds of notational compactness we have used the notation $\frac{\partial f}{\partial t}$ to mean $\frac{\partial f}{\partial t}(t, X_t)$, and $\frac{\partial f}{\partial x}$ to mean $\frac{\partial f}{\partial x}(t, X_t)$ *etc.* In some textbooks you will see the alternative form $\frac{\partial f}{\partial X_t}$ to represent $\frac{\partial f}{\partial x}(t, X_t)$. However, this is slightly too casual and can lead to confusion. The correct way to apply Ito's Lemma is thus to derive the partial derivatives of the deterministic function $f(t, x)$ and then evaluate these at the random point (t, X_t).

Note how the right-hand side of the above formula is constructed.

- **use Taylor's formula in two variables to second-order terms to write:**

$$df(t, X_t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \tfrac{1}{2} \frac{\partial^2 f}{\partial x^2} (dX_t)^2$$

- **substitute $dX_t = Y_t dB_t + A_t dt$ in the above**

- **simplify second-order terms using the following "multiplication table":**

	dt	dB_t
dt	0	0
dB_t	0	dt

One can check that this procedure reproduces (9.3).

One interpretation of Ito's formula is that any (sufficiently well-behaved) function of a diffusion process, is another diffusion process, with drift and diffusion given by:

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} A_t + \tfrac{1}{2} \frac{\partial^2 f}{\partial x^2} Y_t^2 \quad \text{and} \quad \frac{\partial f}{\partial x} Y_t$$

The following example should be familiar from earlier in this section, though the presentation is slightly different.

Example

Suppose that $X_t = B_t$, so that $dX_t = dB_t$ and that *f* is given by $f(t, X_t) = B_t^2 - t$. Note that $dX_t = dB_t$ can be rewritten as $dX_t = 1\,dB_t + 0\,dt$ in order to ascertain that $Y_t = 1$ and $A_t = 0$. Then, using Ito's Lemma:

$$d(B_t^2 - t) = 2B_t dB_t + [-1 + 2B_t \times 0 + ½ \times 2]dt$$

This can also be rearranged to be written in the form:

$$d(B_t^2) = 2B_t dB_t + dt$$

Note that it can be ambiguous as to how to choose the diffusion X_t. For example, suppose you want to know the drift and diffusion coefficient of the process defined by:

$$S_t = e^{B_t + t}$$

In other words, we want to find dS_t. We could take $X_t = B_t + t$ so that $dX_t = dB_t + dt$ and $(dX_t)^2 = dt$. In this case $S_t = f(X_t)$ where $f(x) = e^x$, and therefore $f'(x) = f''(x) = e^x$. There is no explicit *t* dependence in this case.

It follows that:

$$\begin{aligned} dS_t = df(X_t) &= f'(X_t)dX_t + \tfrac{1}{2} f''(X_t)(dX_t)^2 \\ &= e^{X_t}(dB_t + dt) + \tfrac{1}{2} e^{X_t} dt \\ &= S_t\left(dB_t + \tfrac{3}{2} dt\right) \end{aligned}$$

On the other hand, we could take $X_t = B_t$. Then $dX_t = dB_t$ and therefore $(dX_t)^2 = dt$.

Question 9.2

What would the function $f(x,t)$ be this time? Show that you get the same answer for dS_t.

Example

A process X_t satisfies the stochastic differential equation:

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

Deduce the stochastic differential equation for the process X_t^3.

Solution

By Ito's lemma:

$$d(X_t^3) = 3X_t^2\sigma(X_t)\,dB_t + \left[3X_t^2\mu(X_t) + 3X_t\sigma^2(X_t)\right]dt$$

Question 9.3

Verify the result in the last example by applying Taylor's theorem and working with the multiplication table given in Ito's lemma.

1.3 Stochastic differential equations

Stochastic differential equations can be used to define a continuous-time stochastic process. These are called diffusion models or Ito process models. In this section we will look at two examples used in financial economics: geometric Brownian motion (which we met previously in Chapter 8, Section 1.5) and the Ornstein-Uhlenbeck process.

Geometric Brownian motion (revisited)

Consider the stochastic differential equation:

$$dS_t = \alpha S_t \, dt + \sigma S_t \, dB_t \tag{9.4}$$

This is known as a *stochastic differential equation (SDE)* since it is a differential equation where one or more of the elements is a stochastic process (in this case B_t is the stochastic process).

It is shorthand for the integral equation:

$$S_t = S_0 + \alpha \int_0^t S_s ds + \sigma \int_0^t S_s dB_s$$

In these equations, S_t is a geometric Brownian motion. This is the standard model for the price of a risky asset like a stock.

We can use Ito's lemma to solve this equation to find an explicit formula for S_t.

Divide Equation (9.4) by S_t to separate the variables:

$$\frac{1}{S_t} dS_t = \alpha dt + \sigma dB_t$$

If we were dealing with an ordinary differential equation, integration would lead to the expression $\log(S_t / S_0) = \alpha t + \sigma B_t$ and thus to $S_t = S_0 \exp(\alpha t + \sigma B_t)$.

So perhaps the solution to the *stochastic* equation is also based on $\log S_t$. If we consider the stochastic differential of $\log S_t$, we'll find that it does indeed work out.

To solve the problem within stochastic calculus, we use Ito's Lemma to calculate $d \log S_t$.

In preparation for this, from (9.4), let: $Y_t = \sigma S_t$ and $A_t = \alpha S_t$.

Applying Ito's Lemma to the function $f(t,S_t) = \log S_t$ gives:

$$d\log S_t = \frac{1}{S_t}\sigma S_t dB_t + \left[0 + \frac{1}{S_t}\alpha S_t + \frac{1}{2}\left(-\frac{1}{S_t^2}\right)\sigma^2 S_t^2\right]dt$$

$$= \left(\alpha - \frac{1}{2}\sigma^2\right)dt + \sigma dB_t$$

Written as an integral equation, this reads:

$$\log S_t = \log S_0 + \left(\alpha - \tfrac{1}{2}\sigma^2\right)t + \sigma B_t$$

or, finally:

$$S_t = S_0 \exp\left[\left(\alpha - \frac{1}{2}\sigma^2\right)t + \sigma B_t\right]$$

We see that the process S_t satisfying Equation (9.4) is a geometric Brownian motion with parameter $\mu = \alpha - \frac{1}{2}\sigma^2$.

Recall that in Chapter 8, Section 1.5 we defined geometric Brownian motion as $S_t = S_0 e^{\mu t + \sigma B_t}$. We can restate the statistical properties quoted previously with this alternative parameterisation.

Since $\log S_t$ is normally distributed, it follows that S_t / S_0 has a lognormal distribution with parameters $\left(\alpha - \frac{1}{2}\sigma^2\right)t$ and $\sigma^2 t$. The properties of the lognormal distribution give us the expectation and variance of S_t / S_0:

$$E\left(\frac{S_t}{S_0}\right) = \exp\left(\left(\alpha - \tfrac{1}{2}\sigma^2\right)t + \tfrac{1}{2}\sigma^2 t\right) = e^{\alpha t}, \quad \text{var}\left(\frac{S_t}{S_0}\right) = e^{2\alpha t}(e^{\sigma^2 t} - 1)$$

1.4 Diffusion and Ito process models

There may be good reasons for believing that a Brownian model is inadequate to represent the data. For example, the data may appear to come from a stationary process, or may exhibit a tendency to revert to a mean value. As long as there is no indication that the process being modelled is discontinuous, a diffusion model or an Ito process model can be applied. (For our purposes a diffusion and an Ito process are simply alternative representations of a single process.)

Recall that a diffusion or Ito process is just a stochastic process (X_t) defined by a stochastic differential equation of the form:

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

where $\mu(x)$ and $\sigma(x)$ are specified functions.

In this section we generalise Brownian motion by abandoning the assumption of independence of the increments while retaining the Markov property. The process $\{X_t, t \geq 0\}$ with state space $S = R$ (the real numbers) **is said to be a *time-homogeneous diffusion process* if it is a Markov process, with continuous sample paths and we can write its SDE as:**

$$\boldsymbol{dX_t = \mu(X_t)dt + \sigma(X_t)dB_t}$$

where $\mu(X_t)$ is the drift coefficient and $\sigma(X_t)$ is the volatility or diffusion coefficient

One can think of a diffusion as being locally like Brownian motion with drift, but with a variable drift coefficient $\mu(x)$ and diffusion coefficient $\sigma(x)$.

In other words, if we pick a particular time t_0, when the value of the process is x_0, we know that at times close to t_0, the process will have a drift close to $\mu(x_0)$ and a volatility close to $\sigma(x_0)$. So, during a short time period around t_0, the drift and volatility will be roughly constant, *ie* the process is behaving like Brownian motion.

This extra flexibility is very valuable for modelling purposes.

The processes here are said to be *time-homogeneous* because the functions $\mu()$ and $\sigma()$ do not depend explicitly on the time t. There is also a wider family of time-inhomogeneous diffusion processes (which we won't look at here) defined by the equation:

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

where the drift and volatility depend on both the value of the process and the time.

Question 9.4

What were the functions $\mu(x)$ and $\sigma(x)$ for geometric Brownian motion?

The Ornstein-Uhlenbeck process

Consider for instance the *spot rate of interest* R_t. If we model it as $R_t = r_t + X_t$ where r_t is a deterministic central rate and X_t is a random fluctuation, it is natural to demand that X_t displays some mean-reversion tendency.

In other words, as the interest rate diverges from the mean, there is a factor that pulls it back again.

This is achieved by choosing for X_t the diffusion with coefficients:

$$\mu(X_t) = -\gamma X_t$$

$$\sigma(X_t) = \sigma$$

for some $\gamma > 0$.

Note that since γ is positive, the drift acts in the opposite direction to the current value of X_t. This causes X_t to revert towards the mean 0. R_t will then revert towards r_t.

Such a process is known as an *Ornstein-Uhlenbeck process*.

Fitting a diffusion model involves estimating the drift function $\mu(X_t)$ and the diffusion function $\sigma(X_t)$. Estimating arbitrary drift and diffusion functions is virtually impossible unless a very large quantity of data is to hand.

Example

For example, with the interest rate model described above we would have to first estimate r_t, which represents the "trend" value of interest rates at time t. We could then calculate historical values of $R_t - r_t$, which equals X_t. We would then need to look at all the occasions when X_t was approximately –0.02 say (corresponding to the times when interest rates were 2% below their trend value) and estimate what the typical drift was. We would then have to repeat this procedure for all possible deviations.

It is much more usual to specify a parametric form of the mean or the variance and to estimate the parameters.

We now use Ito's Lemma to revisit the Ornstein-Uhlenbeck process. In view of the drift and diffusion coefficients $\mu(X_t) = -\gamma X_t$, $\sigma(X_t) = \sigma$ we may define the Ornstein-Uhlenbeck process as the solution to the equation:

$$dX_t = -\gamma X_t dt + \sigma dB_t \qquad (9.5)$$

where γ and σ are positive parameters.

One way to solve this equation is to use the method of "variable" parameters, as the Core Reading does below. We will also see that you can solve it by using an integrating factor.

The solution of this equation begins by using standard differential equation methods.

Since $ce^{-\gamma t}$ is the general solution of (the deterministic equation) **$dX_t = -\gamma X_t dt$, we look for a solution of (9.5) in the form:**

$$X_t = U_t e^{-\gamma t}$$

In the non-stochastic case the solution is the parameter c multiplied by $e^{-\gamma t}$. The approach here is to "guess" that the solution to the stochastic version might be the same but with some function in place of the parameter c, *ie* we have "varied the parameter".

Hence, we have:

$$dU_t = d\left(e^{\gamma t} X_t\right)$$

We use Ito's Lemma to determine $d\left(e^{\gamma t} X_t\right)$. In preparation for this, from (9.5), let: $Y_t = \sigma$ and $A_t = -\gamma X_t$.

Applying Ito's Lemma to the function $f(t, X_t) = e^{\gamma t} X_t$ gives:

$$\begin{aligned} d\left(e^{\gamma t} X_t\right) &= e^{\gamma t}\,\sigma dB_t + \left[\gamma e^{\gamma t} X_t + e^{\gamma t}(-\gamma X_t) + \frac{1}{2} \times 0 \times \sigma^2\right] dt \\ &= \sigma e^{\gamma t} dB_t. \end{aligned}$$

Written as an integral equation, this reads:

$$e^{\gamma t}X_t = e^{\gamma 0}X_0 + \int_0^t \sigma e^{\gamma s}\,dB_s$$

ie $$X_t = X_0 e^{-\gamma t} + \sigma\int_0^t e^{\gamma(s-t)}\,dB_s \qquad (9.6)$$

The stochastic differential equation $dX_t = -\gamma X_t dt + \sigma dB_t$ can also be solved using the integrating factor $e^{\gamma t}$.

Question 9.5

Show that you get the same solution if you solve the SDE using an integrating factor.

The properties of X_t are easily extracted from (9.6) using:

Result 9.2

Let $f:[0,\infty)\to R$ be a deterministic function, then:

(i) $M_t = e^{\int_0^t f(s)dB_s - \frac{1}{2}\int_0^t f^2(s)ds}$ **is a martingale.**

(ii) $\int_0^t f(s)\,dB_s$ **has a normal distribution with zero mean and variance** $\int_0^t f^2(s)\,ds$.

Part (i) is a simple generalisation of the fact that $e^{\lambda B_t - \frac{1}{2}\lambda^2 t}$ is a martingale. Part (ii) follows immediately from (i) because, since martingales have a constant mean:

$$E\left[e^{\lambda\int_0^t f(s)dB_s - \frac{1}{2}\lambda^2\int_0^t f^2(s)ds}\right] = 1$$

since $E[M_t] = M_0 = 1$

So: $$E\left[e^{\lambda \int_0^t f(s)dB_s}\right] = e^{\frac{1}{2}\lambda^2 \int_0^t f^2(s)ds}$$

which is the moment generating function of the $N\left(0, \int_0^t f^2(s)ds\right)$ **distribution.**

Note that alternatively we can show (ii) as we did before. Part (i) can then be shown in exactly the same way that you prove that $e^{\lambda B_t - \frac{1}{2}\lambda^2 t}$ is a martingale, only this time we need to invoke part (ii).

As a result, the probability distribution of X_t **is** $N\left(X_0\, e^{-\gamma t}, \frac{\sigma^2}{2\gamma}(1-e^{-2\gamma t})\right)$**, and the long-term distribution is** $N\left(0, \frac{\sigma^2}{2\gamma}\right)$**.**

To obtain the long-term distribution, we just let $t \to \infty$.

Question 9.6

Show that the distribution of X_t is $N\left(X_0\, e^{-\gamma t}, \frac{\sigma^2}{2\gamma}(1-e^{-2\gamma t})\right)$.

It is instructive to compare these properties to those of an AR(1) process of the form:

$$X_n = \alpha X_{n-1} + e_n$$

where e_n **is a white noise with zero mean and variance** σ_e^2**. This process has mean and variance:**

$$E[X_n] = \alpha^n X_0 \ , \qquad \mathrm{var}[X_n] = \sigma_e{}^2 \frac{(1-\alpha^{2n})}{(1-\alpha^2)}$$

These values coincide with those of the Ornstein-Uhlenbeck process if we put:

$$\alpha = e^{-\gamma} \ , \qquad \frac{\sigma_e^2}{1-\alpha^2} = \frac{\sigma^2}{2\gamma}$$

In fact, the correspondence goes deeper: the Ornstein-Uhlenbeck process is the continuous equivalent of an AR(1) process in the same way as Brownian motion is the continuous limit of a random walk.

2 Exam-style question

Question

A derivatives trader is modelling the volatility of an equity index using the following discrete-time model:

Model 1: $\sigma_t = 0.12 + 0.4\sigma_{t-1} + 0.05\varepsilon_t$, $t = 1, 2, 3, \ldots$

where σ_t is the volatility at time t years and $\varepsilon_1, \varepsilon_2, \ldots$ are a sequence of independent and identically-distributed random variables from a standard normal distribution. The initial volatility σ_0 equals 0.15.

(i) Determine the long-term distribution of σ_t. [3]

The trader is developing a related continuous-time model for use in derivative pricing. The model is defined by the following stochastic differential equation (SDE):

Model 2: $d\sigma_t = -\alpha(\sigma_t - \mu)dt + \beta dW_t$

where σ_t is the volatility at time t years, W_t is standard Brownian motion and the parameters α, β and μ all take positive values.

(ii) (a) Show that for this model:

$$\sigma_t = \sigma_0 e^{-\alpha t} + \mu(1 - e^{-\alpha t}) + \int_0^t \beta e^{-\alpha(t-s)} dW_s$$

(b) Hence determine the numerical value of μ and a relationship between the parameters α and β if it is required that σ_t has the same long-term mean and variance under each model.

(c) State another consistency property between the models that could be used to determine precise numerical values for α and β. [7]

The derivative pricing formula used by the trader involves the squared volatility $V_t = \sigma_t^2$, which represents the variance of the returns on the index.

(iii) Determine the SDE for V_t in terms of the parameters α, β and μ. [2]

[Total 12]

Solution

(i) ***Long-term probability distribution***

Since σ_0 has a fixed value, and the ε_t's are normally distributed, then σ_t is a linear combination of independently distributed normal distributions. Hence σ_t will also have a normal distribution.

The long-term mean can be found by taking expectations:

$$E[\sigma_t] = 0.12 + 0.4E[\sigma_{t-1}] + 0$$

In the long-term, $E[\sigma_t] = E[\sigma_{t-1}]$, hence $E[\sigma_t] = 0.2$.

The long-term variance can be found by taking variances:

$$\text{var}(\sigma_t) = (0.4)^2 \text{ var}(\sigma_{t-1}) + (0.05)^2 \times 1$$

In the long-term, $\text{var}(\sigma_t) = \text{var}(\sigma_{t-1})$, hence $\text{var}(\sigma_t) = \dfrac{(0.05)^2}{1-(0.4)^2} = 0.002976$.

So the long-term distribution is $N(0.2, 0.002976)$, *ie* normal with mean 20% and standard deviation 5.46% ($= \sqrt{0.002976}$).

An alternative method of finding the mean and variance of σ_t is to use repeated substitution to obtain the equation:

$$\begin{aligned}\sigma_t &= 0.05\left[\varepsilon_t + 0.4\varepsilon_{t-1} + 0.4^2\varepsilon_{t-2} + \cdots + 0.4^t\varepsilon_0\right] \\ &\quad +0.12\left[1 + 0.4 + 0.4^2 + \cdots + 0.4^t\right] \\ &\quad +0.4^t\sigma_0\end{aligned}$$

The mean and variance of σ_t can be obtained by taking expectations of both sides in each case and summing the resulting geometric progression, which can be assumed to be infinite as t gets very large.

(ii)(a) ***Solving the SDE***

This is an example of an Ornstein-Uhlenbeck SDE.

Rearranging to separate the terms involving σ_t gives:

$$d\sigma_t + \alpha\sigma_t dt = \alpha\mu dt + \beta dW_t$$

Now multiply through by the integrating factor $e^{\alpha t}$:

$$e^{\alpha t} d\sigma_t + \alpha e^{\alpha t} \sigma_t dt = \alpha e^{\alpha t} \mu dt + \beta e^{\alpha t} dW_t$$

ie $$d(e^{\alpha t} \sigma_t) = \alpha e^{\alpha t} \mu dt + \beta e^{\alpha t} dW_t$$

Renaming the variable (*ie* swapping the letter "s" for the letter "t") and integrating between 0 and t, we get:

$$\left[e^{\alpha s} \sigma_s \right]_0^t = \int_0^t \alpha e^{\alpha s} \mu ds + \int_0^t \beta e^{\alpha s} dW_s$$

So: $$e^{\alpha t} \sigma_t - \sigma_0 = \mu(e^{\alpha t} - 1) + \int_0^t \beta e^{\alpha s} dW_s$$

Rearranging and dividing by the integrating factor gives:

$$\sigma_t = \sigma_0 e^{-\alpha t} + \mu(1 - e^{-\alpha t}) + \int_0^t \beta e^{-\alpha(t-s)} dW_s$$

(ii)(b) ***Parameter values***

The mean of the Ito integral is zero.

This is a general property of Ito integrals. It follows because $E[dW_s] = 0$.

Since $e^{-\alpha t} \to 0$ as $t \to \infty$, we can see from the formula derived in part (ii)(a) that, according to Model 2, the long-term mean of σ_t is μ. So for consistency with Model 1, we need $\mu = 0.2$.

The variance of the Ito integral is:

$$\begin{aligned}
\text{var}\left(\int_0^t \beta e^{-\alpha(t-s)} dW_s\right) &= \int_0^t \text{var}\left(\beta e^{-\alpha(t-s)} dW_s\right) \\
&= \int_0^t \left(\beta e^{-\alpha(t-s)}\right)^2 \text{var}\left(dW_s\right) \\
&= \int_0^t \beta^2 e^{-2\alpha(t-s)} ds \\
&= \left[\frac{\beta^2}{2\alpha} e^{-2\alpha(t-s)}\right]_0^t \\
&= \frac{\beta^2}{2\alpha}\left(1 - e^{-2\alpha t}\right)
\end{aligned}$$

In the long term, as $t \to \infty$, this becomes $\dfrac{\beta^2}{2\alpha}$. This is also the long-term variance of σ_t, since the other terms in the formula for σ_t are deterministic. So, for consistency with Model 1, we need:

$$\frac{\beta^2}{2\alpha} = 0.002976 \quad ie \quad \frac{\beta^2}{\alpha} = 0.005952$$

This is equivalent to:

$$\beta = 0.07715\sqrt{\alpha} \quad \text{or} \quad \alpha = 168\beta^2$$

(ii)(c) ***Another property***

We could choose the parameter values so that, in the long term, the correlation between the volatility in consecutive years is the same under both models.

Note that, since Model 2 has 3 parameters, equating the mean and variance is not sufficient to pin down the parameter values, which is why we need a third condition as well.

(iii) ***SDE for V_t***

Let $f(x) = x^2$, so that:

$$f'(x) = 2x \quad \text{and} \quad f''(x) = 2$$

Then, using the Taylor series formula, the infinitesimal increments of V_t are:

$$\begin{aligned} dV_t = d(\sigma_t^2) = df(\sigma_t) &= 2\sigma_t d\sigma_t + \tfrac{1}{2} \times 2(d\sigma_t)^2 \\ &= 2\sigma_t \left[-\alpha(\sigma_t - \mu)dt + \beta dW_t\right] + \tfrac{1}{2} \times 2\beta^2 dt \\ &= \left\{-2\alpha\sigma_t(\sigma_t - \mu) + \beta^2\right\} dt + 2\beta\sigma_t dW_t \\ &= \left\{-2\alpha V_t + 2\alpha\mu\sqrt{V_t} + \beta^2\right\} dt + 2\beta\sqrt{V_t} dW_t \end{aligned}$$

Chapter 9 Summary

Ito integrals

Ito integrals of the form $\int_0^t f(s)\, dB_s$, where $f(s)$ is deterministic:

- cannot be integrated directly
- can often be simplified using Ito's Lemma
- have a normal distribution, with mean zero and variance $\int_0^t f^2(s)\, ds$.

Ito's lemma

Ito's Lemma can be used to differentiate a function *f* of a stochastic process X_t .

If $dX_t = \mu dt + \sigma dB_t$, then:

- $$df(X_t) = \left[\mu \frac{\partial f}{\partial X_t} + \tfrac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial X_t^2} \right] dt + \sigma \frac{\partial f}{\partial X_t} dB_t$$

  This form is used when the new process depends only on the *values* of the original process.

- $$df(X_t, t) = \left[\mu \frac{\partial f}{\partial X_t} + \tfrac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial X_t^2} + \frac{\partial f}{\partial t} \right] dt + \sigma \frac{\partial f}{\partial X_t} dB_t$$

  This form is used when there is *explicit time-dependence*, *ie* the new process depends on the value of the original process *and* the time.

Alternatively, *Taylor's formula* to the second-order can be used to write:

$$df\left(t, X_t\right) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \tfrac{1}{2}\frac{\partial^2 f}{\partial x^2}\left(dX_t\right)^2$$

into which $dX_t = \mu dt + \sigma dB_t$ can be substituted. The second-order terms can then be simplified using the multiplication table:

	dt	dB_t
dt	0	0
dB_t	0	dt

Ornstein-Uhlenbeck process

Mean-reverting financial quantities (such as interest rates) can be modelled using the *Ornstein-Uhlenbeck process*.

The process is defined by the SDE:

$$dX_t = -\gamma X_t dt + \sigma dB_t$$

where γ is a positive parameter.

It can be shown that the formula for the process itself is:

$$X_t = e^{-\gamma t}U_t = X_0 e^{-\gamma t} + \sigma\int_0^t e^{-\gamma(t-s)}dB_s$$

Also, the probability distribution of X_t is $N\left(X_0\, e^{-\gamma t}, \dfrac{\sigma^2}{2\gamma}(1-e^{-2\gamma t})\right)$, and the long-term distribution is $N\left(0, \dfrac{\sigma^2}{2\gamma}\right)$.

Chapter 9 Solutions

Solution 9.1

The general Brownian motion $X_t = \sigma B_t + \mu t$ can be written:

$$X_t = \int_0^t \sigma \, dB_s + \int_0^t \mu \, ds$$

or in differential form:

$$dX_t = \sigma dB_t + \mu dt$$

The latter is also suggested by "differentiating" the equation $X_t = \sigma B_t + \mu t$ with respect to *t*. However, since B_t is not differentiable, the differential form is really only shorthand notation for the integral form.

Solution 9.2

If we take $X_t = B_t$ then $f(x,t) = e^{x+t}$. It follows that:

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial x^2} = f(x,t) = e^{x+t}$$

So we get:

$$\begin{aligned} dS_t = df(t, B_t) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dB_t + \tfrac{1}{2} \frac{\partial^2 f}{\partial x^2} (dB_t)^2 \\ &= S_t \left(dB_t + \tfrac{3}{2} dt \right) \end{aligned}$$

which is the same answer as before.

Solution 9.3

In general, Taylor's formula to second-order is given by:

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + O\left(h^3\right)$$

With the given notation, $df(x) = f(x+h) - f(x)$ and $h = dx$ this becomes:

$$df(x) = f'(x)\,dx + ½ f''(x)(dx)^2 + O\left((dx)^3\right)$$

For the specific case of $f(x) = x^3$ this gives:

$$dX_t^3 = 3X_t^2 dX_t + 3X_t (dX_t)^2$$

ignoring terms above order 2. But $dX_t = \sigma(t)\,dB_t + \mu(t)\,dt$ so that:

$$(dX_t)^2 = \sigma^2(t)\,dt$$

by using the multiplication table.

This gives:

$$\begin{aligned} dX_t^3 &= 3X_t^2\left(\sigma(t)\,dB_t + \mu(t)\,dt\right) + 3X_t\sigma^2(t)\,dt \\ &= 3X_t^2\sigma(t)\,dB_t + \left(3X_t\sigma^2(t) + 3X_t^2\mu(t)\right)dt \end{aligned}$$

as Ito's formula gave us.

Solution 9.4

The stochastic differential equation defining geometric Brownian motion is:

$$dS_t = \alpha S_t dt + \sigma S_t dB_t$$

So: $\mu(x) = \alpha x$ and $\sigma(x) = \sigma x$

Solution 9.5

We want to solve the equation:

$$dX_t = -\gamma X_t dt + \sigma dB_t$$

or $$dX_t + \gamma X_t dt = \sigma dB_t$$

Multiplying through by the integrating factor $e^{\gamma t}$ and then changing the dummy variable to s gives:

$$e^{\gamma s} dX_s + \gamma e^{\gamma s} X_s ds = \sigma e^{\gamma s} dB_s$$

The left-hand side is now the differential of a product. So we have:

$$d\left(e^{\gamma s} X_s\right) = \sigma e^{\gamma s} dB_s$$

Now we can integrate between 0 and *t* to get:

$$e^{\gamma t} X_t - e^{\gamma 0} X_0 = \sigma \int_0^t e^{\gamma s} dB_s$$

Finally, we can rearrange this to get the desired form:

$$X_t = X_0 e^{-\gamma t} + \sigma \int_0^t e^{-\gamma(t-s)} dB_s$$

Solution 9.6

From Equation (9.6) we have:

$$X_t = X_0 e^{-\gamma t} + \sigma \int_0^t e^{\gamma(s-t)} dB_s$$

Now, consider the first term. Because it is deterministic we have:

$$E\left[X_0 e^{-\gamma t}\right] = X_0 e^{-\gamma t} \text{ and } V\left[X_0 e^{-\gamma t}\right] = 0$$

Result 9.2(ii) states that $\int_0^t f(s)\, dB_s$ has a normal distribution with zero mean and variance $\int_0^t f^2(s)\, ds$. Applying this result, we get:

$$E[X_t] = X_0 e^{-\gamma t} + 0$$

and $$V[X_t] = \sigma^2 \int_0^t e^{2\gamma(s-t)} ds = \frac{\sigma^2}{2\gamma}\left[e^{2\gamma(s-t)}\right]_0^t = \frac{\sigma^2}{2\gamma}\left(1 - e^{-2\gamma t}\right)$$

2 The Black-Scholes model

2.1 Introduction

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.

2.2 The underlying SDE

Suppose that we have a European call option on a non-dividend-paying share S_t which is governed by the stochastic differential equation (SDE):

$$dS_t = S_t(\mu dt + \sigma dZ_t)$$

where Z_t is a standard Brownian motion.

The share price process is therefore being modelled as a geometric Brownian motion or lognormal model, as discussed previously. The constants μ and σ are referred to as the *drift* and *volatility* parameters respectively.

Investors are allowed to invest positive or negative amounts in this share. Investors can also have holdings in a risk-free cash bond with price B_t at time t.

This is governed by the ordinary differential equation:

$$dB_t = rB_t dt$$

where r is the (assumed-to-be) constant risk-free rate of interest. Hence:

$$S_t = S_0 \exp\left[(\mu - \frac{1}{2}\sigma^2)t + \sigma Z_t\right]$$

$$B_t = B_0 \exp(rt)$$

> ***Question 14.3***
>
> Let X_t be a diffusion process:
>
> $$dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dZ_t$$
>
> where Z_t is a standard Brownian motion. Find the stochastic differential equation for $f(t, X_t)$ using Ito's Lemma and check your answer using a Taylor's series expansion for two variables.

To check this solution, we define the function $g(t, Z_t) = S_0 \exp[(\mu - \frac{1}{2}\sigma^2)t + \sigma Z_t]$ so that $S_t = g(t, Z_t)$. Note that this is a function of t and Z_t.

Recall that the Ito process for standard Brownian motion is:

$$dZ_t = 1\,dZ_t + 0\,dt$$

Now apply Ito's Lemma to $g(t, Z_t)$, which we write as g for notational ease, that is:

$$\begin{aligned} dg &= \sigma g dZ_t + \left[\left(\mu - \frac{1}{2}\sigma^2\right)g + 0 + \frac{1}{2}\sigma^2 g\right]dt \\ &= \mu g dt + \sigma g dZ_t \end{aligned}$$

Replacing $g = g(t, Z_t)$ with S_t, we have the original SDE and the check is complete:

$$dS_t = \mu S_t dt + \sigma S_t dZ_t$$

Since Z_t is normally distributed, S_t is log-normally distributed with all of the usual properties of that distribution.

From the formula given for S_t we can deduce that:

$$\log S_t = \log S_0 + \left(\mu - ½\sigma^2\right)t + \sigma Z_t \sim N\left[\log S_0 + \left(\mu - ½\sigma^2\right)t, \sigma^2 t\right]$$

So S_t has a lognormal distribution with parameters $\log S_0 + \left(\mu - ½\sigma^2\right)t$ and $\sigma^2 t$. **This price process is sometimes called *a log-normal process*, *geometric Brownian motion* or *exponential Brownian motion*.**

2.3 The Black-Scholes formula

Let $f(t,s)$ be the price at time t of a call option given:

- **the current share price is $S_t = s$**
- **the time of maturity is $T > t$**
- **the exercise price is K.**

Proposition 14.1 (The Black-Scholes formula)

For such a call option:

$$f(t,S_t) = S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$$

where:

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

and:

$$d_2 = d_1 - \sigma\sqrt{T-t}$$

and $\Phi(z)$ is the cumulative distribution function of the standard normal distribution.

This formula is the Garman-Kohlhagen formula found on page 47 of the Tables. Here the dividend rate q is equal to zero.

For a put option we also have $f(t,s) = Ke^{-r(T-t)}\Phi(-d_2) - S_t\Phi(-d_1)$ where d_1 and d_2 are as defined above.

Question 14.4

Starting from the formula for the price of a call option given in Proposition 14.1, use put-call parity to derive the formula just given for the price of a put option.

We will give two proofs of this result for the call option**, one here using the partial differential equation (PDE) approach and the other in Chapter 16 using the martingale approach.**

2.4 The PDE approach

Here we use Ito's Lemma to derive an expression for the price of the derivative as a function, f, of the underlying share price process S_t. Here S_t again refers to the share price excluding any dividends received. This method involves the construction of a risk-free portfolio, which in an arbitrage-free world must yield a return equal to the risk-free rate of return.

We first use Ito's Lemma to write a stochastic differential equation (SDE) for the change in the derivative price as a function of the change in the share price. Here $df(t, S_t)$ means the change in the value of the derivative over a very small time period.

An expression for $df(t,S_t)$

Given the Ito process:

$$dS_t = \mu S_t dt + \sigma S_t dZ_t$$

which is the SDE for geometric Brownian motion, with drift and volatility functions for S_t of μS_t and σS_t respectively,

then applying Ito's Lemma to the function $f(t,S_t)$, we have:

$$df(t,S_t) = \frac{\partial f}{\partial s}\sigma S_t dZ_t + \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial s}\mu S_t + \frac{1}{2}\frac{\partial^2 f}{\partial s^2}\sigma^2 S_t^2\right]dt$$

On grounds of notational compactness we have used the notation $\frac{\partial f}{\partial t}$ to mean $\frac{\partial f}{\partial t}(t,S_t)$, and $\frac{\partial f}{\partial s}$ to mean $\frac{\partial f}{\partial s}(t,S_t)$ *etc.* In some textbooks you will see the alternative form $\frac{\partial f}{\partial S_t}$ to represent $\frac{\partial f}{\partial s}(t,S_t)$. However, this is slightly too casual and can lead to confusion. The correct way to apply Ito's Lemma is thus to derive the partial derivatives of the deterministic function $f(t,s)$ and then evaluate these at the random point (t,S_t).

The risk-free portfolio

Suppose that at any time t, $0 \le t < T$, we hold the following portfolio:

- **minus one derivative**
- **plus $\frac{\partial f}{\partial s}(t,S_t)$ shares.**

Let $V(t, S_t)$ be the value of this portfolio, that is:

$$V(t, S_t) = -f(t, S_t) + \frac{\partial f}{\partial s} S_t$$

The pure investment gain over the period $(t, t+dt]$ is the change in the value of the minus one derivative plus the change in the value of the holding of $\partial f / \partial s$ units of the share, *ie*:

$$dV(t, S_t) = -df(t, S_t) + \frac{\partial f}{\partial s} dS_t$$

We can get to the right-hand side of this equation by noting that $\frac{\partial f}{\partial s}$, which represents the number of shares held in the portfolio over the time interval $[t, t+dt)$, is constant. Now, we can show that:

$$-df(t, S_t) + \frac{\partial f}{\partial s} dS_t = -\left\{\frac{\partial f}{\partial s}\sigma S_t dZ_t + \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial s}\mu S_t + \frac{1}{2}\frac{\partial^2 f}{\partial s^2}\sigma^2 S_t^2\right]dt\right\}$$
$$+\frac{\partial f}{\partial s}\left[\mu S_t dt + \sigma S_t dZ_t\right]$$

Question 14.5

Derive the right-hand side of the previous equation.

After cancelling some terms on the right-hand side of the equation we are left with:

$$-df(t, S_t) + \frac{\partial f}{\partial s} dS_t = \left(-\frac{\partial f}{\partial t} - \frac{1}{2}\frac{\partial^2 f}{\partial s^2}\sigma^2 S_t^2\right)dt$$

We are assuming here that there is no net investment into or out of the portfolio.

Note that this portfolio strategy is not self-financing: that is, the pure investment gain derived above is not equal to the instantaneous change in the value of the portfolio, $dV(t, S_t)$.

Note also that the right-hand side of the above expression, which represents the pure investment gain over the interval $[t, t+dt)$, involves the share price, which is random. However, the value of S_t is known at time t. Likewise, if we know the relevant functions, then the values of the derivatives $\dfrac{\partial f}{\partial t}$ and $\dfrac{\partial^2 f}{\partial s^2}$ are also known at time t. So, the expression on the right-hand side involves no terms whose values are unknown at time t.

Now note that the expression for $-df(t,S_t)+\frac{\partial f}{\partial s}dS_t$ involves dt but not dZ_t so that the instantaneous investment gain over the short interval t to $t+dt$ is risk-free.

Given that the market is assumed to be arbitrage-free, this rate of interest must be the same as the risk-free rate of interest on the cash bond. (If this was not true then arbitrage opportunities would arise by going long in cash and short in the portfolio (or vice versa) with zero cost initially and a sure, risk-free profit an instant later.)

Therefore we must have, for all t and $S_t > 0$, the alternative expression:

$$-df(t,S_t)+\frac{\partial f}{\partial s}dS_t = rV(t,S_t)dt$$

Question 14.6

Derive the right-hand side of this equation.

Recall that $V(t,S_t) = -f(t,S_t)+\dfrac{\partial f}{\partial s}dS_t$ is the value of the portfolio. We now have two different expressions for $-df(t,S_t)+\dfrac{\partial f}{\partial s}dS_t$. If we equate these, we get:

$$\left(-\frac{\partial f}{\partial t}-\frac{1}{2}\frac{\partial^2 f}{\partial s^2}\sigma^2 S_t^2\right)dt = r\left(-f+\frac{\partial f}{\partial s}S_t\right)dt$$

$$\Rightarrow \qquad \frac{\partial f}{\partial t}+rS_t\frac{\partial f}{\partial s}+\frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 f}{\partial s^2} = rf$$

This is known as the Black-Scholes PDE. So, we have a non-stochastic partial differential equation (PDE) that can be solved to determine the value of the derivative.

The value of the derivative is found by specifying appropriate boundary conditions and solving the PDE.

The boundary conditions are:

$$f(T,S_T)=\max\{S_T-K,0\} \quad \text{for a call option}$$

$$f(T,S_T)=\max\{K-S_T,0\} \quad \text{for a put option}$$

We have now constructed a partial differential equation that the value of any fairly priced derivative based on the underlying share must satisfy. This means that if a proposed model for the fair price of a derivative does not satisfy the PDE it is not an accurate model.

Finally, we can try out the solutions given in the proposition for the value of a call and a put option. We find that they satisfy the relevant boundary conditions and the PDE.

As an example we could check that the Black-Scholes formula for the fair price of a European call option:

- satisfies the Black-Scholes PDE and
- satisfies the boundary condition $f(T,S_T)=\max\{S_T-K,0\}$

You do not need to be able to do this because the details are beyond Subject CT8. You will meet them if you study Subject ST6.

Question 14.7

Two clever scientists called Giggs and Beckham propose a formula for the fair price of a European put option based upon the assumptions underlying the Black-Scholes model. What are the two things we need to check to verify that it is an accurate formula?

Question 14.8

Why do you think the Black-Scholes PDE contains only three of the six Greeks?

Question 14.9

A forward contract is arranged where an investor agrees to buy a share at time T for an amount K. It is proposed that the fair price for this contract at time t is:

$$f(S_t, t) = S_t - Ke^{-r(T-t)}$$

Show that this:

(i) satisfies the boundary condition

(ii) satisfies the Black-Scholes PDE.

Intuitive interpretation of the PDE

We now return to the intuitive interpretation of the Greeks, this time seeing it within the context of the Black-Scholes PDE.

From the Black-Scholes PDE we have:

$$\frac{\partial f}{\partial t} + rS_t \frac{\partial f}{\partial s} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 f}{\partial s^2} = rf$$

or $$\Theta + rS_t\Delta + \frac{1}{2}\sigma^2 S_t^2\Gamma = rf$$

Recall from Chapter 12 that, a portfolio for which the weighted sum of the deltas of the individual assets is equal to zero is sometimes described as *delta-neutral*. Also recall that if Γ is small, then Δ will change only slowly over time and so the adjustments needed to keep a portfolio delta-neutral will be minimal.

So, if the delta and gamma of a portfolio are both zero then Θ is the risk-free rate of growth of the portfolio.

Chapter 16

The 5-step method in continuous time

Syllabus objectives

(ix) Demonstrate a knowledge and understanding of the properties of option prices, valuation methods and hedging techniques.

8. Demonstrate an understanding of the Black-Scholes derivative-pricing model:

- *Derive the Black-Scholes partial differential equation both in its basic and Garman-Kohlhagen forms. (part)*
- *Demonstrate how to price and hedge a simple derivative contract using the martingale approach. (part)*

11. Describe and apply in simple models, including the binomial model and the Black-Scholes model, the approach to pricing using deflators and demonstrate its equivalence to the risk-neutral pricing approach.

0 Introduction

In the last chapter we built up the theory required for an alternative proof, known as the 5-step method, of the derivative pricing formula. We also showed how the structure of this proof works by proving the formula for the binomial model once more. You should remember the main steps we performed.

Question 16.1 (Revision)

What were the five main steps?

In this chapter we will use the 5-step method again to prove the formula $V_t = e^{-r(T-t)} E_Q[X \mid F_t]$. We will then use this to derive the Black-Scholes formula. Later in the chapter we will then extend the theory to incorporate dividends and hence prove the Garman-Kohlhagen formula. Again, you should be familiar with the final results from Chapter 14 and the Garman-Kohlhagen formula, which appears on page 47 of the Tables.

Although some of the theory may appear abstract and purely mathematical, in this chapter we also see how delta-hedging can give the martingale approach a more intuitive appeal. Specifically, we will see that the ϕ component of the replicating portfolio turns out to equal Δ. As a by-product we will deduce the Black-Scholes PDE by an alternative method.

We will also discuss the advantages and disadvantages of the 5-step method as compared to the PDE method from Chapter 14, and look briefly at the state price deflator approach.

1 The 5-step method in continuous time

1.1 Introduction

Having seen the structure of the 5-step method in Chapter 15, we now repeat the steps of the proof in continuous time. To make the theory easier to follow, you should recall, also from Chapter 15, the purpose of the steps in the proof, *ie* recall the worded description of what we are trying to achieve:

- The aim of the 5-step proof is to show that a certain portfolio replicates the derivative at all times.
- If you look at the definition of a replicating strategy you will see that we first need a self-financing portfolio.
- If you look at the definition of a self-financing strategy you will see that we need the holdings in the portfolio to be previsible.
- To show a process is previsible, we have the martingale representation theorem to help us.
- However, the martingale representation theorem requires that we have two martingales.
- The discounted share price process is a Q-martingale but we do need one more.
- The other martingale is constructed from the derivative payoff, discounting this all the way back (past the current time t) to time 0.

Remember that, in continuous time, the share price process is being modelled as a geometric Brownian motion or lognormal model, which we discussed previously.

Question 16.2 (Revision)

Fully describe what it means for a share price to have geometric Brownian motion.

Question 16.3 (Revision)

Derive the Black-Scholes PDE for a dividend-paying share.

1.2 *The martingale approach (the 5-step method)*

When we looked at the binomial model in Chapter 15 we demonstrated how the value of a derivative could be expressed as:

$$V_t = e^{-r(T-t)} E_Q[X \mid F_t]$$

where X is the value of the derivative at maturity and Q is the equivalent martingale measure. This was shown for a discrete-time process with a finite state space. Here we are working in continuous time and with continuous state spaces (that is, S_t can take any value greater than zero).

At this stage we do not introduce dividends. So, we are looking at a non-dividend-paying share in continuous time.

However, the binomial result can be extended in the obvious way to give the following result:

Proposition

Let X be any derivative payment contingent on F_T, payable at some fixed future time T, where F_T is the sigma algebra generated by S_u for $0 \le u \le T$. So F_T is the history of S_t up until time T.

Then the value of this derivative payment at time $t < T$ is:

$$V_t = e^{-r(T-t)} E_Q[X \mid F_t].$$

Proof

We follow the same sequence of steps described in Chapter 15.

Step 1

Establish the unique equivalent measure Q under which the discounted asset price process $D_t = e^{-rt} S_t$ is a martingale.

It can be shown that this measure exists, is unique and that under Q:

$$D_t = D_0 \exp\left(\sigma \tilde{Z}_t - \frac{\sigma^2 t}{2} \right)$$

where $\tilde{Z}$ is a Brownian motion under Q.

This is the same idea as in the discrete case. We want to construct the measure that assigns probabilities to the possible asset price paths in such a way that the discounted asset price is a martingale. In discrete time this was easy because we only had to find a single probability for each branch of the tree. In the continuous case we have a whole continuum of possible paths and the problem is not so straightforward.

It is beyond the syllabus to actually construct the probability measure. We can say something about it, however, based on the result of the following question.

Question 16.4

Let Z_t be a standard Brownian motion. By considering the stochastic differential equation or otherwise, prove that $e^{\sigma Z_t - \frac{1}{2}\sigma^2 t}$ is a martingale.

Since:

$$S_t = S_0 \exp\left((\mu - \tfrac{1}{2}\sigma^2)t + \sigma Z_t\right)$$

it follows that:

$$D_t = e^{-rt} S_t = S_0 \exp\left((\mu - r - \tfrac{1}{2}\sigma^2)t + \sigma Z_t\right)$$

Now in order to be a martingale with respect to Q, we could assign probabilities to $\sigma \tilde{Z}_t = (\mu - r)t + \sigma Z_t$ that make $\tilde{Z}_t$ a standard Brownian motion. We saw in Chapter 15 that the Cameron-Martin-Girsanov theorem can be used to achieve this. So, with respect to the real-world probabilities P, the random process $\tilde{Z}_t$ is a Brownian motion with drift. But, when we assign new probabilities, using the measure Q, we remove this drift.

We then have:

$$D_t = S_0 \exp\left[-\tfrac{1}{2}\sigma^2 t + \sigma \tilde{Z}_t\right]$$

where $\tilde{Z}_t$ is a standard Brownian motion under Q, so that D_t is a martingale with respect to Q by using the previous question.

Hence we can write:

$$S_t = S_0 \exp\left[\left(r - ½\sigma^2\right)t + \sigma \tilde{Z}_t\right]$$

where $\tilde{Z}_t = \left(\dfrac{\mu - r}{\sigma}\right)t + Z_t$

Step 2 (proposition)

Define:

$$\boldsymbol{V_t = e^{-r(T-t)} E_Q[X \mid F_t]}$$

We propose that this is the fair price of the derivative.

Step 3

Let:

$$\boldsymbol{E_t = e^{-rT} E_Q[X \mid F_t] = e^{-rt} V_t}$$

Under Q, E_t is a martingale.

Recall that, as in Chapter 15:

$$\begin{aligned} E_t &= B_n^{-1} E_Q\left[C_n | F_t\right] \\ &= e^{-rn} E_Q\left[C_n | F_t\right] \\ &= e^{-rt} e^{-r(n-t)} E_Q\left[C_n | F_t\right] \\ &= e^{-rt} V_t \end{aligned}$$

where C_n is the claim amount at time *n*, here denoted by *X*.

Q

Question 16.5 (Revision)

Why is E_t a *Q*-martingale?

Step 4

By the martingale representation theorem there exists a previsible process ϕ_t (that is ϕ_t is F_{t^-} measurable) such that:

$$dE_t = \phi_t dD_t$$

As in the discrete case, this application of the martingale representation theorem guarantees that the stock process ϕ_t is previsible.

Step 5

Let:

$$\psi_t = E_t - \phi_t D_t$$

We will again see that this is just the right holding of the cash bond that makes the value of the portfolio held equal to the value of the derivative at that time.

Suppose that at time t we hold the portfolio:

- **ϕ_t units of the underlying asset S_t**
- **ψ_t units of the cash account B_t .**

where $\phi_t dD_t = dE_t$ and $\psi_t = E_t - \phi_t D_t$.

Remembering that $e^{-rt}S_t = D_t$ and $E_t = e^{-rt}V_t$ we can examine the change in the value of this portfolio over the very short time period $[t, t+dt]$.

At time t, the portfolio has value:

$$\phi_t S_t + \psi_t B_t = e^{rt}(\phi_t D_t + \psi_t) = e^{rt} E_t = V_t$$

At time $t + dt$, the portfolio has value:

$$\begin{aligned}
\phi_t S_{t+dt} + \psi_t B_{t+dt} &= e^{r(t+dt)}(\phi_t D_{t+dt} + \psi_t) \\
&= e^{r(t+dt)}(\phi_t D_t + \phi_t dD_t + \psi_t) \\
&= e^{r(t+dt)}(E_t + dE_t) \\
&= e^{r(t+dt)} E_{t+dt} = V_{t+dt}
\end{aligned}$$

Therefore, the change in the value of the portfolio over t up to $t+dt$ is:

$$V_{t+dt} - V_t = dV_t$$

Over the period t up to $t+dt$ the pure investment gain on this portfolio is:

$$\phi_t dS_t + \psi_t dB_t$$

So, because the change in the value of the portfolio is the same as the pure investment gain**, the hedging strategy (ϕ_t, ψ_t) is self-financing.**

We need to check that the portfolio has the correct value at the expiry date.

Furthermore:

$$V_T = E_Q[X \mid F_T] = X\ .$$

Therefore the hedging strategy is replicating, so that $V_t = e^{-r(T-t)}E_Q[X \mid F_t]$ is the fair price at time t for this derivative contract.

As before, V_t, the proposed no-arbitrage value of the derivative at time $t < n$, is equal to:

- the time-t *expectation* of the claim amount paid at time n
- calculated with respect to the *probability measure* Q and
- the *information set* F_t generated by the history of the stock price up to and including time t and
- discounted at the continuously-compounded risk-free rate of return, r.

The formula:

$$V_t = e^{-r(T-t)} E_Q\left[X \mid F_t\right]$$

is a general formula that applies to any derivative on a dividend-paying share. In order to find an expression for any specific derivative we would need to specify the derivative payoff X and then calculate the expectation in the above formula.

1.3 *Delta hedging and the martingale approach*

Recall Chapter 12 where we defined the delta of a derivative as one of the *Greeks*.

Question 16.6 (Revision)

(i) What is the definition of delta?

(ii) In what numerical range would you expect delta to be for:

(a) a call option

(b) a put option?

It is important to mention delta hedging at this stage. In the martingale approach we showed that *there exists* a portfolio strategy (ϕ_t, ψ_t) which would replicate the derivative payoff.

However, **we did not say what ϕ_t actually is or how we work it out.** In fact it turns out to be delta. **This is quite straightforward.**

First we can evaluate directly the price of the derivative $V_t = e^{-r(T-t)} E_Q[X \mid F_t]$ either analytically (as in the Black-Scholes formula) or using numerical techniques.

In general, if S_t represents the price of a *tradeable* asset:

$$\phi_t = \frac{\partial V}{\partial s}(t, S_t) = \Delta$$

As we have already seen**, ϕ_t is usually called the d*elta* Δ of the derivative.** Recall that ϕ_t is the change in the discounted derivative process relative to the change in the discounted share process. If we ignored the discount factors up to the current time then this would be the change in the derivative process relative to the change in the share process, *ie* delta.

The martingale approach tells us that provided:

- **we start at time 0 with V_0 invested in cash and shares**
- **we follow a self-financing portfolio strategy**
- **we continually rebalance the portfolio to hold exactly ϕ_t** (delta) **units of S_t with the rest in cash,**

then we will precisely replicate the derivative payoff, without risk. This is a form of delta-hedging.

1.4 *Example: the Black-Scholes formula for a call option*

The 5-step method has shown us that the fair price at time t for a derivative contract that pays a (random) amount X at time T is $V_t = e^{-r(T-t)} E_Q\left[X|F_t\right]$. We now want to evaluate this expression in the case where the derivative is a European call option on a non-dividend-paying share.

Question 16.7 (Revision)

What is the payoff function for a European call option?

Proposition

The Black-Scholes formula for a call option on a share with no dividends is:

$$V_t = S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$$

where $d_1 = \dfrac{\log\frac{S_t}{K} + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$ **and** $d_2 = d_1 - \sigma\sqrt{T-t}$

Proof

Using risk-neutral valuation, the fair price of the derivative with payoff function X_T is:

$$V_t = e^{-r(T-t)} E_Q\left[X_T | F_t\right]$$

This is the general risk-neutral pricing formula, which we derived using the 5-step method.

The payoff function for a European call option is $\max\left[S_T - K, 0\right]$.

We therefore need to substitute this payoff function into the general risk-neutral pricing formula and work out the resulting expression. To do this, we also need to know the risk-neutral probability of each possible value that this payoff function might take at the expiry date T, conditioned on the current share price S_t.

Question 16.8

Why do we not need to condition on the full past history of the share price process F_t?

So:

$$V_t = e^{-r(T-t)} E_Q\left[\max\left[S_T - K, 0\right] \mid S_t\right]$$

The process S_t is assumed to be a geometric Brownian motion and so has a continuous state space. This means we need to use integration to work out the expected value of the payoff function. Hence:

$$V_t = e^{-r(T-t)} \int_0^{\infty} \max\left[S_T - K, 0\right] f(S_T \mid S_t)\, dS_T$$

where $f(S_T \mid S_t)$ is the (conditional) probability density function for S_T, given S_t, and we are summing over all the possible values of S_T from zero to infinity.

Question 16.9

How are we able to simplify this expression?

This expression can be simplified to give:

$$\boldsymbol{V_t = e^{-r(T-t)} \int_K^{\infty} (S_T - K)\, f(S_T \mid S_t)\, dS_T}$$

which can be written in terms of two integrals as follows:

$$\boldsymbol{V_t = e^{-r(T-t)} \int_K^{\infty} S_T f(S_T \mid S_t)\, dS_T - Ke^{-r(T-t)} \int_K^{\infty} 1 \times f(S_T \mid S_t)\, dS_T}$$

Under the risk-neutral measure *Q*:

$$\boldsymbol{S_T = S_t \exp\left[(r - \tfrac{1}{2}\sigma^2)(T-t) + \sigma^2(\bar{Z}_T - \tilde{Z}_t)\right]}$$

The above expression for S_T in terms of S_t can be derived from the stochastic differential equation for the share price under Q, which is:

$$dS_t = rS_t dt + \sigma S_t d\tilde{Z}_t$$

where r is the risk-free rate and $\tilde{Z}_t$ is a standard Brownian motion process under Q.

Thus:

$$\boldsymbol{\log S_T \mid S_t \sim N\left[\log S_t + (r - \tfrac{1}{2}\sigma^2)(T-t), \sigma^2(T-t)\right]}$$

which tells us the distribution of S_T given the current share price S_t, *ie* it has a lognormal distribution.

So, we can use the formula for the truncated moments of a lognormal distribution on page 18 in the *Tables*.

Recall that in general the share price can take any value from zero to infinity. However, in the above integrals, we are only summing from K to infinity, *ie* we need to evaluate truncated moments of the share price.

When applying the formula from the *Tables*, note that the power of S_T in the first integral is one, whereas that in the second integral is zero – as S_T to the power of zero is equal to one!

Thus:

$$\boldsymbol{V_t = e^{-r(T-t)}\left[e^{\log S_t + (r - \frac{1}{2}\sigma^2)(T-t) + \frac{1}{2}\sigma^2(T-t)}\right]\{\Phi(U_1) - \Phi(L_1)\}}$$
$$\boldsymbol{-Ke^{-r(T-t)}\left[e^0\right]\{\Phi(U_0) - \Phi(L_0)\}}$$

which after some cancelling becomes:

$$\boldsymbol{V_t = S_t\{\Phi(U_1) - \Phi(L_1)\} - Ke^{-r(T-t)}\{\Phi(U_0) - \Phi(L_0)\}}$$

Finally, to finish off, we just need to evaluate the $\Phi(U_1), \Phi(L_1), \Phi(U_0)$ and $\Phi(L_0)$ terms.

Now:

- $\boldsymbol{U_1 \approx \infty}$
- $\boldsymbol{U_0 \approx \infty}$
- $\boldsymbol{L_1 = \dfrac{\log K - \log S_t - (r - ½\sigma^2)(T-t) - \sigma^2(T-t)}{\sigma\sqrt{T-t}} = -d_1}$
- $\boldsymbol{L_0 = \dfrac{\log K - \log S_t - (r - ½\sigma^2)(T-t)}{\sigma\sqrt{T-t}} = -d_2}$

These are obtained using the formulae on page 18 in the *Tables*.

> ***Question 16.10***
>
> Explain why $U_0 \approx \infty$?

Thus:

- $\Phi(U_1) = \Phi(U_0) \approx \Phi(\infty) = 1$

Hence:

$$\begin{aligned}\boldsymbol{V_t} &\boldsymbol{= S_t\{1 - \Phi(-d_1)\} - K e^{-r(T-t)}\{1 - \Phi(-d_2)\}} \\ &\boldsymbol{= S_t\,\Phi(d_1) - K e^{-r(T-t)}\Phi(d_2)}\end{aligned}$$

This is the Black-Scholes formula for the value of a call option on a non-dividend-paying share.

1.5 Example: deriving the Black-Scholes PDE using the martingale approach

Let's now find an actual formula for ϕ_t and look at an alternative method of deriving the Black-Scholes PDE using martingales.

We define:

$$E_t = e^{-rt}V_t \quad \text{and} \quad D_t = e^{-rt}S_t$$

which are both martingales under Q. Then:

- $dD_t = \sigma D_t d\tilde{Z}_t$
- $dS_t = B_t(rD_t dt + dD_t)$ **and**
- $dE_t = -re^{-rt}V_t dt + e^{-rt}dV_t = e^{-rt}(-rV_t dt + dV_t)$.

Question 16.11

Derive these three relationships.

Applying Ito's Lemma to the function $V(t, S_t)$:

$$dV_t = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial s}dS_t + \frac{1}{2}\frac{\partial^2 V}{\partial s^2}(dS_t)^2$$

$$= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial s^2}\right)dt + \frac{\partial V}{\partial s}B_t(rD_t dt + dD_t)$$

$$= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial s^2} + rS_t \frac{\partial V}{\partial s}\right)dt + \frac{\partial V}{\partial s}B_t dD_t$$

Hence:

$$dE_t = e^{-rt}(-rV_t dt + dV_t)$$

$$= e^{-rt}\left[-rV_t dt + \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial s^2} + rS_t \frac{\partial V}{\partial s}\right)dt + \frac{\partial V}{\partial s}B_t dD_t\right]$$

$$= e^{-rt}\left[-rV_t dt + \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial s^2} + rS_t \frac{\partial V}{\partial s}\right)dt\right] + \frac{\partial V}{\partial s}dD_t$$

Now we know that E_t and D_t are both martingales under Q. Therefore by the martingale representation theorem there exists some previsible process ϕ_t such that:

$$dE_t = \phi_t \, dD_t = \sigma \, \phi_t \, D_t d\tilde{Z}_t$$

This can be written as:

$$dE_t = 0dt + \phi_t dD_t$$

We can compare this with the SDE derived above:

$$dE_t = e^{-rt}\left(-rV_t + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial s^2} + rS_t \frac{\partial V}{\partial s}\right)dt + \frac{\partial V}{\partial s} dD_t$$

This means that:

$$\phi_t = \frac{\partial V}{\partial s}$$

and $$-rV_t + \frac{\partial V}{\partial t} + rS_t \frac{\partial V}{\partial s} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial s^2} = 0$$

$$\Leftrightarrow \qquad rV_t = \frac{\partial V}{\partial t} + rS_t \frac{\partial V}{\partial s} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial s^2}$$

(otherwise E_t would not be a martingale under Q).

We recognise the last equation as the Black-Scholes PDE from page 46 in the *Tables*.

It can be shown (but won't be here) **that:**

$$\frac{\partial V}{\partial s} = \Phi(d_1)$$

So this martingale approach has provided an alternative derivation of the Black-Scholes PDE and it has also given us an explicit formula for ϕ_t, in case we wanted to set up a replicating portfolio in real life.

1.6 Advantages of the martingale approach

We have now seen two ways of deriving the Black-Scholes formula for a call option $c_t = S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$. One method involved evaluating the martingale formula $e^{-r(T-t)}E_Q[X \mid F_t]$. The other method involved "solving" the PDE with the correct boundary condition.

The main 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. Of course, we ourselves did not to literally have to guess the solution, we just had to look it up on page 47 of the Tables!

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. So, the point is that, without knowing the formula on page 47 of the Tables there is no easy way we could work it out using the PDE approach, whereas it can be worked out using the martingale approach without knowing it beforehand.

Furthermore the martingale approach also gives us the replicating strategy for the derivative.

Question 16.12

What is the replicating strategy for a European call option?

Question 16.13

You are trying to replicate a 6-month European call option with strike price 500, which you purchased 4 months ago. If $r = 0.05$, $\sigma = 0.2$, and the current share price is 475, what portfolio should you be holding, assuming that no dividends are expected before the expiry date?

Finally, the martingale approach can be applied to any F_T-measurable derivative payment, including path-dependent options (for example, Asian options), whereas the PDE approach, in general, cannot.

An Asian option is one where the payoff depends on the average share price up to expiry.

> ***Question 16.14***
>
> What would you say is the main disadvantage of the martingale approach as compared to the PDE approach?

1.7 Risk-neutral pricing

This (martingale) approach is often referred to as *risk-neutral pricing*. The measure Q is commonly called the *risk-neutral* measure. However, Q is also referred to as the *equivalent martingale* measure because the discounted prices processes $S_t e^{-rt}$ and $V_t e^{-rt}$ are both **martingales under Q,** which is equivalent to the real-world probability measure P in the technical sense we described earlier.

2 The state price deflator approach

Recall that we have:

$$dS_t = S_t[\mu dt + \sigma dZ_t] \text{ under } P$$

and:

$$dS_t = S_t[rdt + \sigma d\tilde{Z}_t] \text{ under } Q$$

where:

$$d\tilde{Z}_t = dZ_t + \gamma dt \text{ and } \gamma = \frac{\mu - r}{\sigma}$$

Corollary to the Cameron-Martin-Girsanov Theorem

There exists a process η_t such that, for any F_T -measurable derivative payoff X at time T , we have:

$$E_Q[X \mid F_t] = E_P\left[\frac{\eta_T}{\eta_t} X \middle| F_t\right]$$

We do not prove this corollary in Subject CT8.

In the present case, where $\tilde{Z}_t$ is a Q -Brownian motion, Z_t is a P -Brownian motion and $d\tilde{Z}_t = dZ_t + \gamma dt$ we have:

$$\eta_t = e^{-\gamma Z_t - \frac{1}{2}\gamma^2 t}$$

Question 16.15

What interesting property does η_t have?

Now if we further define:

$$A_t = e^{-rt}\eta_t$$

the price at time t for the derivative X payable at time T is then:

$$V_t = e^{-r(T-t)}E_Q[X \mid F_t]$$

under the martingale approach and:

$$V_t = e^{-r(T-t)}E_P\left[\frac{\eta_T}{\eta_t}X \middle| F_t\right]$$

$$= \frac{E_P[e^{-rT}\eta_T X \mid F_t]}{e^{-rt}\eta_t}$$

$$= \frac{E_P[A_T X \mid F_t]}{A_t}$$

under the new approach.

The process A_t is called a state price deflator (also *deflator*; *state price density*; *pricing kernel*; or *stochastic discount factor*).

Note that A_t is defined in terms of η_t, which is a function of Z_t. So A_t is a stochastic process linked to the random behaviour of the share price.

Q

Question 16.16

Use Ito's lemma to show that the SDE for A_t is:

$$dA_t = -A_t\left(rdt + \gamma dZ_t\right)$$

A very important point to note is that, for this model, the risk-neutral and the state price deflator approaches give the same price V_t. Theoretically they are the same. They only differ in the way that they present the calculation of a derivative price.

3 The 5-step approach with dividends

3.1 Introduction

We now extend the theory involved in the 5-step method so that it can deal with underlying asset that pays dividends. You may find it useful to review Sections 3.1 and 3.2 from Chapter 14 at this stage. In these sections we discussed how the share price process must be modified to incorporate dividends.

Recall that the solution to the modified SDE was:

$$\tilde{S}_t = S_0 \exp\left[\left(\mu + q - \tfrac{1}{2}\sigma^2\right)t + \sigma Z_t\right]$$

The cash process will stay the same:

$$B_t = e^{rt}$$

3.2 The martingale approach

We have already mentioned that for a continuous-dividend-paying asset S_t the tradable asset is:

$$\tilde{S}_t = \tilde{S}_0 \exp[(\mu + q - \tfrac{1}{2}\sigma^2)t + \sigma Z_t]$$

rather than just S_t.

To price a derivative contingent on this underlying asset we can repeat the steps which allow us to price and replicate the derivative.

Step 1

Find the unique **equivalent martingale measure Q under which:**

$$\tilde{D}_t = e^{-rt}\tilde{S}_t$$

is a martingale.

Again, it is beyond the syllabus to actually construct the probability measure.

ie $$\tilde{D}_t = \tilde{S}_0 \exp[-\tfrac{1}{2}\sigma^2 t + \sigma \tilde{Z}_t]$$

where $\tilde{Z}_t$ is a standard Brownian motion under Q.

Hence we can write:

$$S_t = S_0 \exp[(r - q - \tfrac{1}{2}\sigma^2)t + \sigma \tilde{Z}_t]$$

Step 2 (proposition)

Let:

$$V_t = e^{-r(T-t)} E_Q[X \mid F_t]$$

where X is the derivative payoff at time T. We propose that this is the fair price of the derivative at t.

Note that this is exactly the same formula as before, except that the risk-neutral measure used to calculate the expectation is now the one for a dividend-paying share.

Step 3

Let:

$$E_t = e^{-rT} E_Q[X \mid F_t] = e^{-rt} V_t$$

This is a martingale under Q.

Step 4

By the martingale representation theorem there exists a previsible process $\tilde{\phi}_t$ such that $dE_t = \tilde{\phi}_t d\tilde{D}_t$.

As before without dividends, this application of the martingale representation theorem guarantees that the stock process $\tilde{\phi}_t$ is previsible.

Step 5

Let:

$$\boldsymbol{\psi_t = E_t - \tilde{\phi}_t \tilde{D}_t}$$

At time t we hold the portfolio:

- **$\tilde{\phi}_t$ units of the tradable asset $\tilde{S}_t$**

  (This is equivalent to $\phi_t = e^{qt}\tilde{\phi}_t$ units of S_t)

- **ψ_t units of the cash account.**

where $\tilde{\phi}_t d\tilde{D}_t = dE_t$ and $\psi_t = E_t - \phi_t \tilde{D}_t$.

Remembering that $e^{-rt}\tilde{S}_t = \tilde{D}_t$ and $E_t = e^{-rt}V_t$ we examine the change in the value of this portfolio over the very small time period $[t, t+dt]$.

At time t the value of this portfolio is equal to:

$$\tilde{\phi}_t \tilde{S}_t + \psi_t B_t = e^{rt}\tilde{\phi}_t \tilde{D}_t + e^{rt}\left(E_t - \tilde{\phi}_t \tilde{D}_t\right) = e^{rt}E_t \quad = \boldsymbol{V_t}$$

At time $t+dt$ the value of this portfolio is equal to:

$$\begin{aligned}
\tilde{\phi}_t \tilde{S}_{t+dt} + \psi_t B_{t+dt} &= e^{r(t+dt)}(\tilde{\phi}_t \tilde{D}_{t+dt} + \psi_t) \\
&= e^{r(t+dt)}(\tilde{\phi}_t \tilde{D}_t + \tilde{\phi}_t d\tilde{D}_t + \psi_t) \\
&= e^{r(t+dt)}(E_t + dE_t) \\
&= e^{r(t+dt)}E_{t+dt} = V_{t+dt}
\end{aligned}$$

So the change in the value of the portfolio over the same period is:

$$\begin{aligned} V_{t+dt} - V_t = dV_t &= B_t dE_t + E_t dB_t \\ &= B_t \left[\tilde{\phi}_t d\tilde{D}_t + rE_t dt \right] \\ &= \tilde{\phi}_t d\tilde{S}_t + \psi_t dB_t \end{aligned}$$

The pure investment gain over the period t up to $t + dt$ is:

$$\tilde{\phi}_t d\tilde{S}_t + \psi_t dB_t = B_t \left[\tilde{\phi}_t d\tilde{D}_t + r\left(\tilde{\phi}_t \tilde{D}_t + \psi_t\right) dt \right]$$

So, because the change in the value of the portfolio is the same as the pure investment gain, **the portfolio is self-financing.**

Also $V_T = X$

So, the hedging strategy $(\tilde{\phi}_t, \psi_t)$ is replicating and V_t is the fair price at time t.

Yet again, V_t, the proposed no-arbitrage value of the derivative at time $t < n$, is equal to:

- the time-t *expectation* of the claim amount paid at time n
- calculated with respect to the *probability measure* Q and
- the *information set* F_t generated by the history of the stock price up to and including time t and
- discounted at the continuously-compounded risk-free rate of return r.

3.3 Example: the price of a European call option on a share with dividends

The idea in this section is that we can use the Black-Scholes formulae we have already derived for a call option on a non-dividend-paying share to derive the corresponding (Garman-Kohlhagen) formula when there are dividends.

In the absence of dividends, we know from the derivative pricing formula and the Black-Scholes formula that:

$$c_t = e^{-r(T-t)}E_Q[\max(S_T - K, 0)] = S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$$

where Q is the risk-neutral probability measure for S_t.

In the presence of dividends, we need to work out:

$$V_t = e^{-r(T-t)}E_{\tilde{Q}}[\max(S_T - K, 0)]$$

where $\tilde{Q}$ is the risk-neutral probability measure for $\tilde{S}_t$. Note that the payoff is still based on S_T, not $\tilde{S}_T$. If you exercise the option you just get the basic share without the accumulated dividends.

If we use the relationship $\tilde{S}_T = S_T e^{qT}$ and we define $\tilde{K} = Ke^{qT}$, we can write:

$$\begin{aligned} V_t &= e^{-r(T-t)}E_{\tilde{Q}}[\max(\tilde{S}_T e^{-qT} - \tilde{K}e^{-qT}, 0)] \\ &= e^{-qT} \times e^{-r(T-t)}E_{\tilde{Q}}[\max(\tilde{S}_T - \tilde{K}, 0)] \end{aligned}$$

The part of this expression after the multiplication sign looks exactly like the pricing formula for a call option, except that we have put squiggles on the S_T, K and Q. (Note also that $\tilde{Q}$ is the correct risk-neutral probability measure for $\tilde{S}_t$.)

This means that we can calculate this bit using the Black-Scholes formula, provided that we replace all the S's and K's with $\tilde{S}$'s and $\tilde{K}$'s. This gives us:

$$V_t = e^{-qT} \times [\tilde{S}_t\Phi(d_1) - \tilde{K}e^{-r(T-t)}\Phi(d_2)]$$

where d_1 and d_2 are now calculated based on $\tilde{S}$ and $\tilde{K}$.

This trick works because the only difference between the probability measures Q and $\tilde{Q}$ is that the drift is increased by q. However, this only affects the value of the parameter μ, which doesn't appear in the Black-Scholes formula!

Suppose that:

$$\boldsymbol{X = \max\{S_T - K, 0\} = e^{-qT} \max\{\tilde{S}_T - \tilde{K}, 0\}}$$

where $\tilde{K} = Ke^{qT}$.

By analogy with the non-dividend-paying stock:

$$\begin{aligned} \boldsymbol{V_t} &= \boldsymbol{e^{-r(T-t)} E_Q[X \mid F_t]} \\ &= \boldsymbol{e^{-qT}\left[\tilde{S}_t \Phi(d_1) - \tilde{K}e^{-r(T-t)}\Phi(d_2)\right]} \end{aligned}$$

where:

$$\begin{aligned} \boldsymbol{d_1} &= \boldsymbol{\frac{\log\frac{\tilde{S}_t}{\tilde{K}} + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}} \\ &= \frac{\log\frac{e^{qt}S_t}{e^{qT}K} + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ &= \frac{\log\frac{S_t}{K} - q(T-t) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ &= \boldsymbol{\frac{\log\frac{S_t}{K} + (r - q + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}} \end{aligned}$$

and $\boldsymbol{d_2 = d_1 - \sigma\sqrt{T-t}}$

Also:

$$\boldsymbol{e^{-qT}\tilde{S}_t = S_t e^{-q(T-t)}}$$

and:

$$\boldsymbol{e^{-qT}\tilde{K} = K}$$

$$\Rightarrow \quad \boldsymbol{V_t = S_t e^{-q(T-t)}\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)}$$

So the difference when dividends are present is that we have to "strip out" the dividends with an $e^{-q(T-t)}$ factor, and change r to $r-q$ in the calculation of d_1 and d_2. The formula for a put option works in the same way.

By looking at this formula we can see that it may be optimal to exercise early an American call option on a continuous-dividend-paying stock. This is because the value of the equivalent European call option can be less than the option's intrinsic value. In particular, for any $t<T$, as S_t gets large (relative to *K*), V_t is approximately equal to:

$$S_t e^{-q(T-t)} - Ke^{-r(T-t)} < S_t - K$$

for large enough S_t.

This is because d_1 and d_2 would be large, so that $\Phi(d_1)$ and $\Phi(d_2)$ are approximately equal to 1.

Question 16.17

According to this approximation, how large would S_t need to be?

Question 16.18

Can you spot any other situations where this "reverse" situation could apply?

We can equally derive the price of a European put option on a dividend-paying stock:

ie
$$V_t = Ke^{-r(T-t)}\Phi(-d_2) - S_t e^{-q(T-t)}\Phi(-d_1)$$

where d_1 and d_2 are defined above.

4 Exam-style question

We finish this chapter with an exam-style question on risk-neutral pricing.

Question

You are given that the fair price to pay at time t for a derivative paying X at time T is $V_t = e^{-r(T-t)} E_Q\left[X \middle| F_t\right]$, where Q is the risk-neutral probability measure and F_t is the filtration with respect to the underlying process. The price movements of a non-dividend-paying share are governed by the stochastic differential equation $dS_t = S_t\left(\mu dt + \sigma dB_t\right)$, where B_t is standard Brownian motion under the risk-neutral probability measure.

(i) Solve the above stochastic differential equation. [4]

(ii) Determine the probability distribution of $S_T \mid S_t$. [2]

(iii) Hence show that the fair price to pay at time t for a forward on this share, with forward price K and time to expiry $T-t$, is:

$$V_t = S_t - Ke^{-r(T-t)}$$ [4]

[Total 10]

Solution

(i) ***Solving the SDE***

The given SDE is on page 46 of the *Tables*. Note that, because we are using the risk-neutral measure, $\mu = r$. The process is geometric Brownian motion. To solve it we consider the function $f\left(S_t\right) = \log S_t$.

It is a good idea to know this for the exam.

Applying the Taylor's series formula to the above function, we get:

$$df\left(S_t\right) = d(\log S_t) = \frac{1}{S_t} dS_t + \frac{1}{2}\left(-\frac{1}{S_t^2}\right)(dS_t)^2$$

$$= \frac{1}{S_t}(rS_t\,dt + \sigma S_t\,dB_t) - \frac{1}{2S_t^2}(rS_t\,dt + \sigma S_t\,dB_t)^2$$

$$= r\,dt + \sigma\,dB_t - \tfrac{1}{2}\sigma^2\,dt$$

$$= (r - \tfrac{1}{2}\sigma^2)dt + \sigma dB_t$$

Changing the t's to s's and integrating this equation between limits of $s = 0$ and $s = t$, we get:

$$\left[\log S_s\right]_{s=0}^{s=t} = (r - \tfrac{1}{2}\sigma^2)\int_0^t ds + \sigma\int_0^t dB_s$$

$$\Rightarrow \quad \log S_t - \log S_0 = (r - \tfrac{1}{2}\sigma^2)\,t + \sigma B_t$$

$$\Rightarrow \quad S_t = S_0\,e^{(r - ½\sigma^2)t + \sigma B_t}$$

(ii) ***Probability distribution***

From part (i), we know that, under the risk-neutral probability measure Q,

$$\log S_t - \log S_0 = (r - \tfrac{1}{2}\sigma^2)t + \sigma B_t$$

Since $B_t \sim N(0,t)$ then:

$$\log S_t - \log S_0 \sim N\left((r - ½\sigma^2)t,\ \sigma^2 t\right)$$

Replacing 0 and t with t and $T-t$ we get:

$$\log S_T - \log S_t \sim N\left((r - ½\sigma^2)(T-t),\ \sigma^2(T-t)\right)$$

$$\Rightarrow \quad \log S_T \mid S_t \sim N\left(\log S_t + (r - ½\sigma^2)(T-t),\ \sigma^2(T-t)\right)$$

$$\Rightarrow \quad S_T \mid S_t \sim \log N\left(\log S_t + (r - ½\sigma^2)(T-t),\ \sigma^2(T-t)\right)$$

(iii) ***Fair price for a forward***

We are given that the fair price to pay at time t for a derivative paying X at time T is $V_t = e^{-r(T-t)} E_Q\left[X \middle| F_t\right]$, where Q is the risk-neutral probability measure.

The random variable payoff of a forward on a non-dividend-paying share is:

$$X = S_T - K$$

Substituting this into the fair price formula, we get:

$$\begin{aligned} V_t = e^{-r(T-t)} E_Q\left[X \middle| F_t\right] &= e^{-r(T-t)} E_Q\left[S_T - K \middle| F_t\right] \\ &= e^{-r(T-t)}\left(E_Q\left[S_T \middle| F_t\right] - K\right) \end{aligned}$$

Now $E_Q\left[S_T \middle| F_t\right]$ is the conditional mean of the random variable S_T, and from part (ii), we know that $S_T \mid S_t \sim \log N\left(\log S_t + (r - ½\sigma^2)(T-t),\ \sigma^2(T-t)\right)$.

Using the formula for the expectation of the lognormal distribution on page 14 of the *Tables*:

$$E_Q\left[S_T \middle| F_t\right] = \mathrm{e}^{\left\{\log S_t + (r - \frac{1}{2}\sigma^2)(T-t) + \frac{1}{2}\sigma^2(T-t)\right\}}$$

$$= S_t e^{r(T-t)}$$

So: $$V_t = e^{-r(T-t)} E_Q\left[X \middle| F_t\right] = e^{-r(T-t)}\left(E_Q\left[S_T \middle| F_t\right] - K\right)$$

$$= S_t - Ke^{-r(T-t)}$$

Chapter 16 Summary

The martingale (5-step) approach (with dividends)

The derivative pricing formula $V_t = e^{-r(T-t)} E_Q[X \mid F_t]$ can be derived using the martingale approach, which consists of five steps:

Step 1

Find the unique equivalent martingale measure Q under which $\tilde{D}_t = e^{-rt} \tilde{S}_t$ is a martingale.

Step 2

Let $V_t = e^{-r(T-t)} E_Q\left[X \mid F_t\right]$ where X is the derivative payoff at time T. This is proposed as the fair price of the derivative at time t.

Step 3

Let $\tilde{E}_t = e^{-rT} E_Q\left[X \mid F_t\right] = e^{-rt} V_t$. This is a martingale under Q.

Step 4

By the martingale representation theorem, there exists a previsible process $\tilde{\phi}_t$ such that $d\tilde{E}_t = \tilde{\phi}_t d\tilde{D}_t$.

Step 5

Let $\psi_t = \tilde{E}_t - \tilde{\phi}_t \tilde{D}_t$ and at time t hold the portfolio consisting of:

- $\tilde{\phi}_t$ units of the tradable $\tilde{S}_t$
- ψ_t units of the cash account.

At time t the value of this portfolio is equal to V_t. Also $V_T = X$. Therefore, the hedging strategy $(\tilde{\phi}_t, \psi_t)$ is replicating and so V_t is the fair price at time t.

Garman-Kohlhagen formula for a European option on a dividend-paying share

Call option

$$f(S_t) = S_t \Phi(d_1) e^{-q(T-t)} - K e^{-r(T-t)} \Phi(d_2)$$

Put option

$$f(S_t) = K e^{-r(T-t)} \Phi(-d_2) - S_t \Phi(-d_1) e^{-q(T-t)}$$

where:

- $$d_1 = \frac{\ln \frac{S_t}{K} + (r - q + ½\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$
- $$d_2 = d_1 - \sigma\sqrt{T-t}$$

The Black-Scholes formulae for a non-dividend-paying share are the same but using $q = 0$. These formulae can be derived by direct evaluation of the expected value using the general option pricing formula found using the 5-step method.

Delta hedging

$$\phi_t = \frac{\partial V}{\partial s}(t, S_t) = \Delta$$

- We start at time 0 with V_0 invested in cash and shares.
- We follow a self-financing portfolio strategy.
- We continually rebalance the portfolio to hold exactly ϕ_t units of S_t with the rest in cash.

By following these steps, we precisely replicate the derivative payoff, without risk.

The martingale approach vs. the PDE approach

- In the PDE approach we have to "guess" the solution, whereas with the martingale approach we do not.
- The martingale approach provides an expectation that can be evaluated explicitly in some cases and in a straightforward numerical way in other cases.
- The martingale approach also gives the replicating strategy for the derivative.
- The martingale approach can be applied to any F_T-measurable derivative payment, whereas the PDE approach cannot always.
- However, the PDE approach is much quicker and easier to construct, and more easily understood.

State price deflator approach

Corollary to the Cameron-Martin-Girsanov Theorem

There exists a process η_t such that, for any F_T-measurable derivative payoff X_T at time T, we have:

$$E_Q[X_T \mid F_t] = E_P\left[\frac{\eta_T}{\eta_t} X_T \middle| F_t\right]$$

The *state price deflator* A_t is defined by:

$$A_t = e^{-rt}\eta_t$$

where $\eta_t = e^{-\gamma Z_t - ½\gamma^2 t}$, which is a martingale under *P*.

Derivative prices can be calculated using the state price deflator formula:

$$V_t = \frac{E_P[A_T X_T \mid F_t]}{A_t}$$

This page has been left blank so that you can keep the chapter summaries together for revision purposes.

Chapter 16 Solutions

Solution 16.1

- We established the equivalent martingale measure Q.
- We proposed a fair price, V_t, for a derivative and its discounted value $E_t = e^{-rt}V_t$.
- We used the martingale representation theorem to construct a hedging strategy (ϕ_t, ψ_t).
- We then showed that this hedging strategy replicates the derivative payoff at time n.
- So V_t was the fair value of the derivative at time t.

Solution 16.2

Geometric Brownian motion means that we are modelling the share price using the continuous-time lognormal model. Alternatively, we can express this in terms of the stochastic differential equation:

$$dS_t = \left[\mu\, dt + \sigma\, dZ_t\right] S_t$$

This can be seen on page 46 of the Tables.

Solution 16.3

We must first perform some preliminaries:

$$dS_t = \left[\mu\, dt + \sigma\, dZ_t\right] S_t$$

$$\begin{aligned}
\Rightarrow \qquad (dS_t)^2 &= \left[\mu^2 (dt)^2 + \sigma^2 (dZ_t)^2 + 2\mu\sigma\, dt\, dZ_t\right] S_t^2 \\
&= \left[\mu^2 (dt)^2 + \sigma^2 (dt) + 2\mu\sigma\, dt\, dZ_t\right] S_t^2 \\
&= \sigma^2 S_t^2 dt
\end{aligned}$$

Using these results and Ito's lemma, we get:

$$df(S_t,t) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial s}dS_t + \frac{1}{2}\frac{\partial^2 f}{\partial s^2}(dS_t)^2$$
$$= \Theta dt + \Delta dS_t + \frac{1}{2}\Gamma\sigma^2 S_t^2 dt$$

Now construct the portfolio:

- minus one derivative and
- plus Δ shares.

The value of this portfolio is:

$$V_t = -f(S_t,t) + \Delta S_t$$

The change in the value of this portfolio over a very short time period *dt* is:

$$dV_t = -df(S_t,t) + \Delta\left(dS_t + qS_t dt\right)$$
$$= -\left\{\Theta dt + \Delta dS_t + \frac{1}{2}\Gamma\sigma^2 S_t^2 dt\right\} + \Delta\left(dS_t + qS_t dt\right)$$
$$= -\left\{\Theta - q\Delta S_t + \frac{1}{2}\Gamma\sigma^2 S_t^2\right\}dt$$

We now notice that the final line does not involve dZ_t and hence it is non-random, depending only on the change in time dt. By the principle of no arbitrage, this portfolio must earn the risk-free rate of interest.

ie $\quad dV_t = rV_t dt$

Putting these last two equations together we get:

$$-\left\{\theta - q\Delta S_t + \frac{1}{2}\Gamma\sigma^2 S_t^2\right\}dt = rV_t dt = r\left(-f(S_t,t) + \Delta S_t\right)dt$$

$$\Leftrightarrow \qquad rf(S_t,t) = \theta + (r-q)\Delta S_t + \frac{1}{2}\Gamma\sigma^2 S_t^2$$

This PDE is on page 46 in the *Tables*.

Solution 16.4

We've already seen that the process $X_t = e^{\sigma Z_t - ½\sigma^2 t}$ satisfies the stochastic differential equation (SDE):

$$dX_t = X_t\left[\left(-½\sigma^2 + ½\sigma^2\right)dt + \sigma dZ_t\right] = \sigma X_t dZ_t$$

It follows that the process X_t has no drift and hence must be a martingale.

Alternatively:

$$E\left[\exp\left(-½\sigma^2 t + \sigma Z_t\right)|F_s\right]$$

$$= E\left[\exp\left(-½\sigma^2 t + \sigma Z_s + \sigma(Z_t - Z_s)\right)|F_s\right]$$

$$= \exp\left(-½\sigma^2 t + \sigma Z_s\right) E\left[\exp\left(\sigma(Z_t - Z_s)\right)|F_s\right]$$

since $\exp\left(-½\sigma^2 t + \sigma Z_s\right)$ is a constant given F_s.

Now, $Z_t - Z_s \sim N(0, t-s)$, so we can use the moment generating function of a normal distribution to note that:

$$E\left[\exp\left(\sigma(Z_t - Z_s)\right)|F_s\right] = M_{N(0,t-s)}(\sigma) = \exp\left(½\sigma^2 (t-s)\right)$$

Together with the above, this gives:

$$E\left[\exp\left(-½\sigma^2 t + \sigma Z_t | F_s\right)\right] = \exp\left(-½\sigma^2 s + \sigma Z_s\right)$$

ie $\quad E[X_t \mid F_s] = X_s$

Solution 16.5

E_t is a martingale with respect to *Q*, since for *s* > 0:

$$E_Q\left[E_{t+s}|F_t\right] = E_Q\left[B_T^{-1} E_Q\left\{C_T | F_{t+s}\right\}|F_t\right] = B_T^{-1} E_Q\left[C_T|F_t\right] = E_t$$

Solution 16.6

(i) ***Definition of delta***

Delta is the rate of change of the value of the derivative with respect to the share price:

$$\Delta = \frac{\partial f}{\partial S_t}$$

(ii)(a) ***Call option***

$$0 \le \Delta \le 1$$

(ii)(b) ***Put option***

$$-1 \le \Delta \le 0$$

Solution 16.7

$$f(S_T, T) = \max\left\{S_T - K, 0\right\}$$

Solution 16.8

The assumption of *independent increments* for Brownian motion means that future values of the share price depend only on the current share price S_t and not the past history of how we arrived at it.

Hence, the share price at the maturity date *T*, S_T, and likewise the derivative payoff at time *T*, depends only on S_t.

Solution 16.9

This expression can be simplified by noting that if $S_T < K$, *ie* the call option is out-of-the-money, then the option payoff is zero. Consequently, we do not need to sum over the range of share prices from zero up to *K* (as adding up lots of zeros just gives zero!).

In addition, if $S_T > K$, *ie* the call option is in-the-money, then the option payoff is just $S_T - K$.

Solution 16.10

Using the formula for U_k with $k=0$ gives:

$$U_0 = \frac{\log(\infty) - \log S_t - (r - ½\sigma^2)(T-t) - \sigma^2 (T-t)}{\sigma\sqrt{T-t}}$$

As $x \to \infty$, $\log(x) \to \infty$. So, the log term is **much** bigger than the other terms in the numerator, which can therefore be ignored. In addition, we can argue that dividing infinity by a finite term, such as $\sigma\sqrt{T-t}$, still gives us infinity.

Hence:

$$U_0 \approx \infty$$

Note that a similar argument can be used for U_1, where the subtraction of $\sigma\sqrt{T-t}$ does not affect the end result.

Solution 16.11

First equation

Since $S_t = S_0 e^{(r-½\sigma^2)t+\sigma\tilde{Z}_t}$, we have:

$$D_t = e^{-rt} S_t = e^{-rt} S_0 e^{(r-½\sigma^2)t+\sigma\tilde{Z}_t} = S_0 e^{-½\sigma^2 t+\sigma\tilde{Z}_t}$$

Let $X_t = -½\sigma^2 t + \sigma\tilde{Z}_t$, so that $dX_t = -½\sigma^2 dt + \sigma d\tilde{Z}_t$ and $D_t = S_0 e^{X_t}$.

Now apply Ito's formula:

$$dD_t = \left\{-½\sigma^2 e^{X_t} + ½\sigma^2 e^{X_t}\right\} dt + \sigma e^{X_t} d\tilde{Z}_t = \sigma D_t d\tilde{Z}_t$$

Second equation

$$D_t = e^{-rt} S_t$$

So $\quad S_t = D_t e^{rt}$

Provided that both of the processes X_t and Y_t are not stochastic, the product rule:

$$d(X_tY_t) = X_t dY_t + Y_t dX_t$$

applies.

We can use this in conjunction with the result from ordinary calculus that $\frac{d}{dt}e^{rt} = re^{rt}$ (or $de^{rt} = re^{rt}dt$) to get:

$$\begin{aligned} dS_t = d(e^{rt}D_t) &= e^{rt}dD_t + D_t de^{rt} \\ &= e^{rt}dD_t + D_t re^{rt}dt \\ &= e^{rt}(dD_t + rD_t dt) \\ &= B_t(rD_t dt + dD_t) \end{aligned}$$

Third equation

$$dE_t = d(e^{-rt}V_t) = e^{-rt}dV_t - V_t re^{-rt}dt = e^{-rt}(-rV_t dt + dV_t)$$

Solution 16.12

Hold:

- $\phi_t = \frac{\partial V}{\partial S} = \Phi(d_1)$ shares and
- $\psi_t = E_t - \phi_t D_t = e^{-rt}(V_t - \phi_t S_t)$ units of the cash bond (*ie* an actual cash amount of $V_t - \phi_t S_t$)

Solution 16.13

Here $T = \frac{6}{12}$, $t = \frac{4}{12}$, $r = 0.05$, $\sigma = 0.2$, $K = 500$ and $S_t = 475$. So you need:

$$\phi_t = \Phi(d_1) = \Phi\left(\frac{\ln(475/500) + (0.05 + 0.2^2/2)\times(6-4)/12}{0.2\sqrt{(6-4)/12}}\right) = 0.314 \text{ shares}$$

and $$V_t - \phi_t S_t = 475\Phi(d_1) - 500e^{-0.05\times 2/12}\Phi(d_1 - 0.2\sqrt{2/12})$$
$$-0.314\times 475 = -142 \text{ cash}$$

Solution 16.14

Because it doesn't require an understanding of stochastic calculus, the PDE approach is quicker and easier to describe, and more easily understood.

Solution 16.15

It's a martingale. See Question 16.4.

Solution 16.16

We have that:

$$A_t = e^{-rt}\eta_t = e^{-\gamma Z_t - (r+½\gamma^2)t}$$

If we let $X_t = -\gamma Z_t - (r+½\gamma^2)t$, so that $dX_t = -\gamma dZ_t - (r+½\gamma^2)dt$, then:

$$A_t = e^{X_t}$$

and we can apply Ito's lemma to get:

$$\begin{aligned} dA_t &= \left[-(r+½\gamma^2)e^{X_t} + ½\gamma^2 e^{X_t}\right]dt - \gamma e^{X_t} dZ_t \\ &= -rA_t dt - \gamma A_t dZ_t \\ &= -A_t(rdt + \gamma dZ_t) \end{aligned}$$

This shows that A_t is just a "randomised" version of the ordinary discount factor e^{-rt}, for which $d(e^{-rt}) = -(e^{-rt})rdt$.

Solution 16.17

If we rearrange this inequality, we have:

$$K\left\{1-e^{-r(T-t)}\right\} < S_t\left\{1-e^{-q(T-t)}\right\}$$

So we would need to have:

$$S_t > \frac{1-e^{-r(T-t)}}{1-e^{-q(T-t)}}K$$

Solution 16.18

One example would be if r is close to zero, but q is high. However, for economic reasons this situation is less likely to occur with the shares of major companies.

The right-hand side of the expression in (iii) can be written as:

$$\sum_{i=1}^{n} x_i^2 C_{ii} + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} x_i x_j C_{ij} \qquad [1]$$

Differentiating this with respect to x_i gives:

$$2x_i C_{ii} + 2\sum_{\substack{j=1 \\ j \neq i}}^{n} x_j C_{ij} = 2\sum_{j=1}^{n} x_j C_{ij} \qquad [1½]$$

Note that here the second term gives a contribution from each summation.

Using the result from (ii) above, this can be written as:

$$2\, C_{iP} \quad \ldots (2) \qquad [½]$$

Hence, equating Equations (1) and (2) and cancelling the 2's gives:

$$\sigma_P \frac{\partial \sigma_P}{\partial x_i} = C_{iP} \qquad [1]$$

Dividing both sides of this equation by σ_P^2:

$$\frac{1}{\sigma_P} \frac{\partial \sigma_P}{\partial x_i} = \frac{C_{iP}}{\sigma_P^2} \qquad [½]$$

The left-hand side now matches the definition of β_{iP} given in the question. So we have:

$$\beta_{iP} = \frac{\partial \sigma_P}{\partial x_i} \frac{1}{\sigma_P} \qquad [½]$$

The definition of β_{iP} given here is:

$$\beta_{iP} = \frac{C_{iP}}{\sigma_P^2} = \frac{\operatorname{cov}(R_i, R_P)}{V_P}$$

So β_{iP} represents the beta of security i relative to portfolio P. The equation we have derived shows us that it is equal to the proportionate change in the standard deviation of the portfolio returns when there is a small change in the portfolio weighting x_i. [1]

[Total 7]