CT8 – P C – 15

Combined Materials Pack

ActEd Study Materials: 2015 Examinations

Subject CT8

Contents

Study Guide for the 2015 exams
Course Notes
Question and Answer Bank
Series X Assignments*

***Note:** The Series X Assignment Solutions should also be supplied with this pack unless you chose not to receive them with your study material.

If you think that any pages are missing from this pack, please contact ActEd's admin team by email at **ActEd@bpp.com**.

How to use the Combined Materials Pack

Guidance on how and when to use the Combined Materials Pack is set out in the *Study Guide for the 2015 exams*.

Important: Copyright Agreement

ISBN 978-1-4727-1799-3

9 781472 717993

2015 Study Guide

Subject CT8

Introduction

This Study Guide contains all the information that you will need before starting to study Subject CT8 for the 2015 exams. **Please read this Study Guide carefully before reading the Course Notes,** even if you have studied for some actuarial exams before.

When studying for the UK actuarial exams, you will need:

- a copy of the **Formulae and Tables for Examinations of the Faculty of Actuaries and the Institute of Actuaries, 2nd Edition (2002)** – these are often referred to as simply the "Yellow Tables"

- a "permitted" **scientific calculator** – you will find the list of permitted calculators on the profession's website. Please check the list carefully, since it is reviewed each year.

These are both available from the Institute and Faculty of Actuaries' eShop. Please visit **www.actuaries.org.uk**.

Contents

1 The Subject CT8 course structure

There are four parts to the Subject CT8 course. The parts cover related topics and have broadly equal lengths. The parts are broken down into chapters.

The following table shows how the parts, the chapters and the syllabus items relate to each other. The end columns show how the chapters relate to the days of the regular tutorials. This table should help you plan your progress across the study session.

Part	Chapter	Title	No. of pages	Syllabus objectives	Half day	2 full days	3 full days
	0	Introduction to financial economics	9	none			
	1	The efficient markets hypothesis	22	(vi)			
	2	Utility theory	51	(i) 1-7			
1	3	Stochastic dominance and behavioural finance	31	(i) 8-9	1		
	4	Measures of investment risk	27	(ii)			1
	5	Portfolio theory	44	(iii)		1	
	6	Models of asset returns	28	(iv)			
	7	Asset pricing models	30	(v)			
	8	Brownian motion and martingales	29	(viii) 1			
2	9	Stochastic calculus and Ito processes	41	(viii) 2-5	2		
	10	Stochastic models of security prices	40	(vii)			
	11	Introduction to the valuation of derivative securities	42	(ix) 1-3			2
	12	The Greeks	18	(ix) 12			
3	13	The binomial model	63	(ix) 4-7 & 11	3		
	14	The Black-Scholes option pricing formula	41	(ix) 8-10		2	
	15	The 5-step method in discrete time	40	(ix) 8			3
4	16	The 5-step method in continuous time	42	(ix) 8 & 11	4		
	17	The term structure of interest rates	42	(x)			
	18	Credit risk	34	(xi)			

2 ActEd study support

Successful students tend to undertake three main study activities:

1. *Learning* – initial study and understanding of subject material

2. *Revision* – learning subject material and preparing to tackle exam-style questions

3. *Rehearsal* – answering exam-style questions, culminating in answering questions at exam speed without notes.

Different approaches suit different people. For example, you may like to learn material gradually over the months running up to the exams or you may do your revision in a shorter period just before the exams. Also, these three activities will almost certainly overlap.

ActEd offers a flexible range of products to suit you and let you control your own learning and exam preparation. The following table shows the products that ActEd produces. Note that not all products are available for all subjects.

LEARNING	LEARNING & REVISION	REVISION	REVISION & REHEARSAL	REHEARSAL
Course Notes	Q&A Bank	Flashcards	Revision Notes	Mock Exam A
	X Assignments	Sound Revision	ASET	Additional Mock Pack (AMP)
	Combined Materials Pack (CMP)	MyTest	Revision Tutorials	Mock / AMP Marking
	X Assignment Marking			
	Tutorials			
	Online Classroom			

The products and services available for Subject CT8 are described below.

"Learning" products

Course Notes

The Course Notes will help you develop the basic knowledge and understanding of principles needed to pass the exam. They incorporate the complete Core Reading and include full explanation of all the syllabus objectives, with worked examples and short questions to test your understanding.

Each chapter includes the relevant syllabus objectives, a chapter summary and, where appropriate, a page of important formulae or definitions.

"Learning & revision" products

Question and Answer Bank

The Question and Answer Bank provides a comprehensive bank of questions (including some past exam questions) with full solutions and comments.

The Question and Answer Bank is divided into five parts. The first four parts include a range of short and long questions to test your understanding of the corresponding part of the Course Notes. Part five consists of 100 marks of exam-style questions.

X Assignments

The four Series X Assignments (X1 to X4) cover the material in Parts 1 to 4 respectively. Assignments X1 and X2 are 80-mark tests and should take you two and a half hours to complete. Assignments X3 and X4 are 100-mark tests and should take you three hours to complete. The actual Subject CT8 examination will have a total of 100 marks.

Combined Materials Pack (CMP)

The Combined Materials Pack (CMP) comprises the Course Notes, the Question and Answer Bank and the Series X Assignments.

The CMP is available in **eBook** format for viewing on a range of electronic devices. eBooks can be ordered separately or as an addition to paper products. Visit **www.ActEd.co.uk** for full details about the eBooks that are available, compatibility with different devices, software requirements and printing restrictions.

CMP Upgrade

The purpose of the CMP Upgrade is to enable you to amend last year's study material to make it suitable for study for this year. In most cases, it lists all significant changes to the Core Reading and ActEd material so that you can manually amend your notes. The upgrade includes replacement pages and additional pages where appropriate.

However, if a large proportion of the material has changed significantly, making it inappropriate to include *all* changes, the upgrade will only *outline* what has changed. In this case, we recommend that you purchase a replacement CMP (printed copy or eBook) or Course Notes at a significantly reduced price.

The CMP Upgrade can be downloaded free of charge from our website at **www.ActEd.co.uk**. Alternatively, if the upgrade contains a large number of pages, you may prefer to purchase a hard copy from us at a minimal price to cover production and handling costs.

A separate upgrade for eBooks is not produced but a significant discount is available for retakers wishing to re-purchase the latest eBook.

X Assignment Marking

We are happy to mark your attempts at the X assignments. Marking is not included with the Assignments or the CMP and you need to order it separately. You can submit your scripts by email, fax or post. Your script may be marked electronically, in which case you will be able to download your marked script via a secure link on the internet. Otherwise your marked script will be returned to you in the post.

Don't underestimate the benefits of doing and submitting assignments:

- Question practice during this phase of your study gives an early focus on the end goal of answering exam-style questions.

- You're incentivised to keep up with your study plan and get a regular, realistic assessment of progress.

- Objective, personalised feedback from a high quality marker will highlight areas on which to work and help with exam technique.

In a recent study, it was found that students who attempt more than half the assignments have significantly higher pass rates.

Series Marking

Series Marking applies to a specified subject, session and student. If you purchase Series Marking, you will **not** be able to defer the marking to a future exam sitting or transfer it to a different subject or student.

We typically send out full solutions with the Series X Assignments. However, if you order Series Marking at the same time as you order the Series X Assignments, you can choose whether or not to receive a copy of the solutions in advance. If you choose not to receive them with the study material, you will be able to download the solutions via a secure link on the internet when your marked script is returned (or following the final deadline date if you do not submit a script).

If you are having your attempts at the assignments marked by ActEd, you should submit your scripts regularly throughout the session, in accordance with the schedule of recommended dates set out in information provided with the assignments. This will help you to pace your study throughout the session and leave an adequate amount of time for revision and question practice.

The recommended submission dates are realistic targets for the majority of students. Your scripts will be returned more quickly if you submit them well before the final deadline dates.

Any script submitted *after* the relevant final deadline date will not be marked. It is your responsibility to ensure that scripts are received by ActEd in good time.

Marking Vouchers

Marking Vouchers give the holder the right to submit a script for marking at any time, irrespective of the individual assignment deadlines, study session, subject or person.

Marking Vouchers can be used for any assignment. They are valid for four years from the date of purchase and can be refunded at any time up to the expiry date.

Although you may submit your script with a Marking Voucher at any time, you will need to adhere to the explicit Marking Voucher deadline dates to ensure that your script is returned before the date of the exam. The deadline dates are provided with the assignments.

If you live outside the UK you must ensure that your last script reaches the ActEd office earlier than this to allow the extra time needed to return your marked script.

Tutorials

ActEd tutorials are specifically designed to develop the knowledge that you will acquire from the course material into the higher-level understanding that is needed to pass the exam.

ActEd runs a range of different tutorials including face-to-face tutorials at various locations and Live Online tutorials. Full details are set out in ActEd's Tuition Bulletin, which is available from the ActEd website at **www.ActEd.co.uk**.

Regular and Block Tutorials

In preparation for these tutorials, we expect you to have read the relevant part(s) of the Course Notes before attending the tutorial so that the group can spend time on exam questions and discussion to develop understanding rather than basic bookwork.

You can choose ***one*** of the following types of tutorial:

- **Regular Tutorials** (two or three days) spread over the session.

- **A Block Tutorial** (two or three consecutive days) held two to eight weeks before the exam.

Online Classroom

The Online Classroom acts as either a valuable add-on or a great alternative to a face-to-face or Live Online tutorial.

At the heart of the Online Classroom in each subject is a comprehensive, easily-searched collection of over 100 tutorial units. These are a mix of:

- teaching units, helping you to really get to grips with the course material, and

- guided questions, enabling you to learn the most efficient ways to answer questions and avoid common exam pitfalls.

The best way to discover the Online Classroom is to see it in action. You can watch a sample of the Online Classroom tutorial units on the ActEd website at **www.ActEd.co.uk**.

"Revision" products

For most subjects, there is ***a lot of material*** to revise. Finding a way to fit revision into your routine as painlessly as possible has got to be a good strategy! Flashcards and Sound Revision are inexpensive options that can provide a massive boost. They can also provide a variation in activities during a study day, and so help you to maintain concentration and effectiveness.

Flashcards

Flashcards are a set of A6-sized cards that cover the key points of the subject that most students want to commit to memory. Each flashcard has questions on one side and the answers on the reverse. We recommend that you use the cards actively and test yourself as you go.

Flashcards are available in **eBook** format for viewing on a range of electronic devices. eBooks can be ordered separately or as an addition to paper products. Visit **www.ActEd.co.uk** for full details about the eBooks that are available, compatibility with different devices, software requirements and printing restrictions.

Sound Revision

It is reported that only 30% of information that is read is retained but this rises to 50% if the information is also heard.

Sound Revision is a set of three audio CDs, designed to help you remember the most important aspects of the Core Reading.

The CDs cover the majority of the course, split into a number of manageable topics based on the chapters in the Course Notes. Each section lasts no longer than a few minutes so it's perfect for the train, tube, or car journey on the way to work, or where taking folders and course notes is not practical.

Choice of revision product

Different students will have preferences for different revision products.

So, what might influence your choice between these study aids? The following questions and comments might help you to choose the revision products that are most suitable for you:

Flashcards

- Do you have a regular train or bus journey?

  Flashcards are ideal for regular bursts of revision on the move.

- Do you want to fit more study into your routine?

  Flashcards are a good option for "dead time", eg using flashcards on your phone or sticking them on the wall in your study.

- Do you find yourself cramming for exams (even if that's not your original plan!)?

  Flashcards are an extremely efficient way to do your pre-exam memorising.

Sound Revision

- Do you have some regular time where carrying other materials isn't practical, *eg* commuting, at the gym, walking the dog?

  Sound Revision is an ideal "hands-free" revision tool.

- Do you have a preference for auditory learning, *eg* do you remember conversations more easily than emails?

  Sound Revision will suit your preferred style and be especially effective for you.

Both

- Are you retaking a subject?

  Using a different selection of revision products than on a previous attempt can be a major aid to keeping your revision fresh and effective.

"Revision & rehearsal" products

Revision Notes

ActEd's Revision Notes have been designed with input from students to help you revise efficiently. They are suitable for first-time sitters who have worked through the ActEd Course Notes or for retakers (who should find them much more useful and challenging than simply reading through the course again).

The Revision Notes are a set of six A5 booklets – perfect for revising on the train or tube to work. Each booklet covers one main theme or a set of related topics from the course and includes:

- Core Reading with a set of integrated short questions to develop your bookwork knowledge
- relevant past exam questions with concise solutions from the last ten years
- detailed analysis of key past exam questions (selected for their difficulty), and
- other useful revision aids.

ActEd Solutions with Exam Technique (ASET)

The ActEd Solutions with Exam Technique (ASET) contains ActEd's solutions to the previous four years' exam papers, *ie* eight papers, plus comment and explanation. In particular it will highlight how questions might have been analysed and interpreted so as to produce a good solution with a wide range of relevant points. This will be valuable in approaching questions in subsequent examinations.

A "Mini-ASET" will also be available in the summer session covering the April Exam only.

Revision Tutorials

Revision Tutorials are intensive one-day face-to-face or Live Online tutorials in the final run-up to the exam.

They give you the opportunity to practise interpreting and answering past exam questions and to raise any outstanding queries with an ActEd tutor. These courses are most suitable if you have previously attended Regular Tutorials or a Block Tutorial in the same subject.

Details of how to apply for ActEd's tutorials are set out in our *Tuition Bulletin*, which is available from the ActEd website at **www.ActEd.co.uk**.

"Rehearsal" products

Mock Exam A

Mock Exam A is a 100-mark mock exam paper and is a realistic test of your exam preparation. It is based on Mock Exam A from last year but it has been updated to reflect any changes to the Syllabus and Core Reading.

Additional Mock Pack (AMP)

The Additional Mock Pack (AMP) consists of two further 100-mark mock exam papers – Mock Exam B and Mock Exam C. This is ideal if you are retaking and have already sat Mock Exam A, or if you just want some extra question practice.

Mock / AMP Marking

We are happy to mark your attempts at Mock Exam A or the mock exams included within the AMP. The same general principles apply as for the X Assignment Marking. In particular:

- Mock Exam Marking is available for Mock Exam A and it applies to a specified subject, session and student

- Marking Vouchers can be used for Mock Exam A or the mock exams contained within the AMP; please note that attempts at the AMP can **only** be marked using Marking Vouchers.

Recall that:

- marking is not included with the products themselves and you need to order it separately
- you can submit your scripts by email, fax or post
- your script may be marked and returned to you electronically, or marked and returned by post.

Queries and feedback

From time to time you may come across something in the study material that is unclear to you. The easiest way to solve such problems is often through discussion with friends, colleagues and peers – they will probably have had similar experiences whilst studying. If there's no-one at work to talk to then use ActEd's discussion forum at **www.ActEd.co.uk/forums** (or use the link from our home page at **www.ActEd.co.uk**).

Our online forum is dedicated to actuarial students so that you can get help from fellow students on any aspect of your studies from technical issues to study advice. You could also use it to get ideas for revision or for further reading around the subject that you are studying. ActEd tutors will visit the site from time to time to ensure that you are not being led astray and we also post other frequently asked questions from students on the forum as they arise.

If you are still stuck, then you can send queries by email to **CT8@bpp.com** (but we recommend that you try the forum first). We will endeavour to contact you as soon as possible after receiving your query but you should be aware that it may take some time to reply to queries, particularly when tutors are away from the office running tutorials. At the busiest teaching times of year, it may take us more than a week to get back to you.

If you have many queries on the course material, you should raise them at a tutorial or book a personal tuition session with an ActEd tutor. Information about personal tuition is set out in our current brochure. Please email **ActEd@bpp.com** for more details.

If you find an error in the course, please check the corrections page of our website (**www.ActEd.co.uk/Html/paper_corrections.htm**) to see if the correction has already been dealt with. Otherwise please send details via email to **CT8@bpp.com** or send a fax to **01235 550085**.

Each year ActEd tutors work hard to improve the quality of the study material and to ensure that the courses are as clear as possible and free from errors. We are always happy to receive feedback from students, particularly details concerning any errors, contradictions or unclear statements in the courses. If you have any comments on this course please email them to **CT8@bpp.com** or fax them to **01235 550085**.

The ActEd tutors also work with the profession to suggest developments and improvements to the Syllabus and Core Reading. If you have any comments or concerns about the Syllabus or Core Reading, these can be passed on via ActEd. Alternatively, you can send them directly to the Institute and Faculty of Actuaries' Examination Team by email to **education.services@actuaries.org.uk**.

3 How to study to pass the exams

The CT Subject exams

The Core Reading and exam papers for these subjects tend to be very technical. The exams themselves have many calculation and manipulation questions. The emphasis in the exam will therefore be on *understanding* the mathematical techniques and applying them to various, frequently unfamiliar, situations. It is important to have a feel for what the numerical answer should be by having a deep understanding of the material and by doing reasonableness checks.

Subjects CT2 and CT7 are more "wordy" than the other subjects, including an "essay-style" question or two in Subject CT7.

As a high level of mathematics is required in the courses it is important that your mathematical skills are extremely good. If you are a little rusty you may wish to consider buying the Foundation ActEd Course (FAC) available from ActEd. This covers all of the mathematical techniques that are required for the CT Subjects, some of which are beyond A-Level (or Higher) standard. It is a reference document to which you can refer when you need help on a particular topic.

You will have sat many exams before and will have mastered the exam and revision techniques that suit you. However it is important to note that due to the high volume of work involved in the CT Subjects it is not possible to leave all your revision to the last minute. Students who prepare well in advance have a better chance of passing their exams on the first sitting.

Unprepared students find that they are under time pressure in the exam. Therefore it is important to find ways of maximising your score in the shortest possible time. Part of your preparation should be to practise a large number of exam-style questions under timed exam conditions as soon as possible. This will:

- help you to develop the necessary understanding of the techniques required
- highlight the key topics, which crop up regularly in many different contexts and questions
- help you to practise the specific skills that you will need to pass the exam.

There are many sources of exam-style questions. You can use past exam papers, the Question and Answer Bank (which includes many past exam questions), assignments, mock exams, the Revision Notes and ASET.

Overall study plan

We suggest that you develop a realistic study plan, building in time for relaxation and allowing some time for contingencies. Be aware of busy times at work, when you may not be able to take as much study leave as you would like. Once you have set your plan, be determined to stick to it. You don't have to be too prescriptive at this stage about what precisely you do on each study day. The main thing is to be clear that you will cover all the important activities in an appropriate manner and leave plenty of time for revision and question practice.

Aim to manage your study so as to allow plenty of time for the concepts you meet in this course to "bed down" in your mind. Most successful students will probably aim to complete the course at least a month before the exam, thereby leaving a sufficient amount of time for revision. By finishing the course as quickly as possible, you will have a much clearer view of the big picture. It will also allow you to structure your revision so that you can concentrate on the important and difficult areas of the course.

A sample CT subject study plan is available on our website at:

www.ActEd.co.uk/Html/help_and_advice_study_plans.htm

It includes details of useful dates, including assignment deadlines and tutorial finalisation dates.

Study sessions

Only do activities that will increase your chance of passing. Try to avoid including activities for the sake of it and don't spend time reviewing material that you already understand. You will only improve your chances of passing the exam by getting on top of the material that you currently find difficult.

Ideally, each study session should have a specific purpose and be based on a specific task, *eg* "*Finish reading Chapter 3 and attempt Questions 1.4, 1.7 and 1.12 from the Question and Answer Bank*", as opposed to a specific amount of time, *eg "Three hours studying the material in Chapter 3*".

Try to study somewhere quiet and free from distractions (*eg* a library or a desk at home dedicated to study). Find out when you operate at your peak, and endeavour to study at those times of the day. This might be between 8*am* and 10*am* or could be in the evening. Take short breaks during your study to remain focused – it's definitely time for a short break if you find that your brain is tired and that your concentration has started to drift from the information in front of you.

Order of study

We suggest that you work through each of the chapters in turn. To get the maximum benefit from each chapter you should proceed in the following order:

1. Read the Syllabus Objectives. These are set out in the box on page 1 of each chapter.

2. Read the Chapter Summary at the end of each chapter. This will give you a useful overview of the material that you are about to study and help you to appreciate the context of the ideas that you meet.

3. Study the Course Notes in detail, annotating them and possibly making your own notes. Try the self-assessment questions as you come to them. Our suggested solutions are at the end of each chapter. As you study, pay particular attention to the listing of the Syllabus Objectives and to the Core Reading.

4. Read the Chapter Summary again carefully. If there are any ideas that you can't remember covering in the Course Notes, read the relevant section of the notes again to refresh your memory.

It's a fact that people are more likely to remember something if they review it several times. So, do look over the chapters you have studied so far from time to time. It is useful to re-read the Chapter Summaries or to try the self-assessment questions again a few days after reading the chapter itself.

You may like to attempt some questions from the Question and Answer Bank when you have completed a part of the course. It's a good idea to annotate the questions with details of when you attempted each one. This makes it easier to ensure that you try all of the questions as part of your revision without repeating any that you got right first time.

Once you've read the relevant part of the notes and tried a selection of questions from the Question and Answer Bank (and attended a tutorial, if appropriate) you should attempt the corresponding assignment. If you submit your assignment for marking, spend some time looking through it carefully when it is returned. It can seem a bit depressing to analyse the errors you made, but you will increase your chances of passing the exam by learning from your mistakes. The markers will try their best to provide practical comments to help you to improve.

To be really prepared for the exam, you should not only know and understand the Core Reading but also be aware of what the examiners will expect. Your revision programme should include plenty of question practice so that you are aware of the typical style, content and marking structure of exam questions. You should attempt as many questions as you can from the Question and Answer Bank and past exam papers.

Active study

Here are some techniques that may help you to study actively.

1. Don't believe everything you read! Good students tend to question everything that they read. They will ask "why, how, what for, when?" when confronted with a new concept, and they will apply their own judgement. This contrasts with those who unquestioningly believe what they are told, learn it thoroughly, and reproduce it (unquestioningly?) in response to exam questions.

2. Another useful technique as you read the Course Notes is to think of possible questions that the examiners could ask. This will help you to understand the examiners' point of view and should mean that there are fewer nasty surprises in the exam room! Use the Syllabus to help you make up questions.

3. Annotate your notes with your own ideas and questions. This will make you study more actively and will help when you come to review and revise the material. Do not simply copy out the notes without thinking about the issues.

4. Attempt the questions in the notes as you work through the course. Write down your answer before you refer to the solution.

5. Attempt other questions and assignments on a similar basis, *ie* write down your answer before looking at the solution provided. Attempting the assignments under exam conditions has some particular benefits:

   - It forces you to think and act in a way that is similar to how you will behave in the exam.
   - When you have your assignments marked it is *much* more useful if the marker's comments can show you how to improve your performance under exam conditions than your performance when you have access to the notes and are under no time pressure.
   - The knowledge that you are going to do an assignment under exam conditions and then submit it (however good or bad) for marking can act as a powerful incentive to make you study each part as well as possible.
   - It is also quicker than trying to write perfect answers.

6. Sit a mock exam four to six weeks before the real exam to identify your weaknesses and work to improve them. You could use a mock exam written by ActEd or a past exam paper.

4 Frequently asked questions

What knowledge of earlier subjects should I have?

A: The Course Notes are written on the assumption that students have studied Subjects CT1 to CT7. However, it is possible to study any of these earlier subjects in parallel with CT8. Most students find CT8 quite a tough course compared to some of the other CT subjects and so a good grasp of the material in the earlier subjects, especially Subjects CT1, CT3 and CT6 is essential. Some of the material in CT2, CT4 and CT7 is also useful.

Q: What level of mathematics is required?

A: Some of the maths required for this subject is quite advanced – up to degree standard. The techniques covered in Subjects CT1, CT3 and CT6 will be treated as assumed knowledge and the theory will build on these. You will find the course (and the exam!) much easier if you feel comfortable with the mathematical techniques used in these earlier subjects and you can apply them confidently.

If you feel that you need to brush up on your mathematical skills before starting the course, you may find it useful to recap on Subjects CT1, CT3 and CT6. You may also find it helpful to study the Foundation ActEd Course (FAC) or read an appropriate textbook. The full syllabus for FAC, a sample of the Course Notes and an Initial Assessment to test your mathematical skills can be found on our website at **www.acted.co.uk**.

Q: Which Subject 103 and Subject 109 questions are relevant to Subject CT8?

A: Subject 109 was the predecessor to CT8, whilst the material that appears in Chapters 8 and 9 of CT8 previously appeared in Subject 103. A list of the questions from these courses that are relevant to the current CT8 syllabus is given in Chapter 0 of the Course Notes.

Note also that the material in Chapters 2 and 3 of the CT8 Course on *expected utility* and *stochastic dominance* was included in Subject CT7 from 2005 to 2009. In addition, the material on *behavioural finance* in Chapter 3 was added to CT8 for the 2015 exams. This previously appeared, and still appears, in Subject ST5. Chapter 0 therefore also includes a list of the CT7 and ST5 exam questions that are relevant to the current CT8 syllabus.

For a comprehensive guide to exactly which questions are relevant you can purchase the Subject CT8 Revision Notes. These carefully select from the past papers (including Subjects 103, 109, CT7 and ST5) all the parts of questions that are relevant for Subject CT8. In addition we provide comprehensive solutions together with advice on exam technique and maximising your chances of exam success.

Q: What calculators am I allowed to use in the exam?

A: Please refer to **www.actuaries.org.uk** for the latest advice.

Q: What should I do if I discover an error in the course?

A: If you find an error in the course, please check our website at:

www.acted.co.uk/Html/paper_corrections.htm

to see if the correction has already been dealt with. Otherwise please send details via email to **CT8@bpp.com** or send a fax to **01235 550085**.

5 Core Reading and the Syllabus

Core Reading

The Syllabus for Subject CT8, and the Core Reading that supplements it, has been written by the Institute and Faculty of Actuaries to state the requirements of the examiners. The relevant individual Syllabus Objectives are included at the start of each course chapter and a complete copy of the Syllabus is included in Section 6 of this Study Guide. We recommend that you use the Syllabus as an important part of your study. The purpose of Core Reading is to give the examiners, tutors and students a clear, shared understanding of the depth and breadth of treatment required by the Syllabus. In examinations students are expected to demonstrate their understanding of the concepts in Core Reading. Examiners have the Core Reading available when setting papers.

Core Reading deals with each syllabus objective. Core Reading covers what is needed to pass the exam but the tuition material that has been written by ActEd enhances it by giving examples and further explanation of key points. The Subject CT8 Course Notes include the Core Reading in full, integrated throughout the course. Here is an excerpt from some ActEd Course Notes to show you how to identify Core Reading and the ActEd material. **Core Reading is shown in this bold font.**

Note that in the example given above, the index *will* fall if the actual share price goes below the theoretical ex-rights share price. Again, this is consistent with what would happen to an underlying portfolio.

This is ActEd text

After allowing for chain-linking, **the formula for the investment index becomes:**

This is Core Reading

$$I(t) = \frac{\sum_i N_{i,t} P_{i,t}}{B(t)}$$

where $N_{i,t}$ is the number of shares issued for the *i*th constituent at time *t*;

$B(t)$ is the base value, or divisor, at time *t*.

Core Reading accreditation

The Institute and Faculty of Actuaries would like to thank the numerous people who have helped in the development of this material and in the previous versions of Core Reading.

The following source has been consulted in developing the Core Reading:

Interest Rate Models An Introduction, Cairns Andrew J. G., Princeton University Press, 2004. ISBN: 0691118949

Changes to the Syllabus and Core Reading

The Syllabus and Core Reading are updated as at 31 May each year. The exams in April and September 2015 will be based on the Syllabus and Core Reading as at 31 May 2014.

We recommend that you always use the up-to-date Core Reading to prepare for the exams.

The Institute and Faculty of Actuaries' Copyright

Past exam papers

You can download some past exam papers and Examiners' Reports from the profession's website at **www.actuaries.org.uk**.

Further reading

The exam will be based on the relevant Syllabus and Core Reading and the ActEd course material will be the main source of tuition for students.

However, some students may find it useful to obtain a different viewpoint on a particular topic covered in Subject CT8. The following list of further reading for Subject CT8 has been prepared by the Institute and Faculty of Actuaries. This list is not exhaustive and other useful material may be available.

Baxter, Martin & Andrew Rennie, *Financial calculus; An introduction to derivative pricing*, Cambridge University Press, 1996. (244 pages) ISBN: 978-0521552899.

Cairns, Andrew J. G, *Interest rate models: An introduction*, Princeton University Press, 2004. ISBN: 978-0691118949

Panjer, Harry H *et al* (ed), *Financial economics: with applications to investments, insurance and pensions*, The Actuarial Foundation, 2001. (669 pages) ISBN: 978-0938959489.

Elton, Edwin J, Martin J Gruber, Stephen J Brown and William N. Goetzmann, *Modern portfolio theory and investment analysis* (8th edition), John Wiley, 2010 (727 pages) ISBN: 978-0470505847.

Hull, John C, *Options, futures and other derivatives* (8th edition), Pearson Education, 2011 (847 pages) ISBN: 978-0273759072.

6 Syllabus

The full Syllabus for Subject CT8 is given here. To the right of each objective are the chapter numbers in which the objective is covered in the ActEd course.

Aim

The aim of the Financial Economics subject is to develop the necessary skills to construct asset-liability models and to value financial derivatives. These skills are also required to communicate with other financial professionals and to critically evaluate modern financial theories.

Links to other subjects

Concepts introduced in:

- Subject CT1 – Financial Mathematics
- Subject CT4 – Models
- Subject CT7 – Economics

are used in this subject.

Topics introduced in this subject are further developed in:

- Subject CA1 – Actuarial Risk Management
- Subject ST6 – Finance and Investment Specialist Technical B.
- Subject ST9 – Enterprise Risk Management.

Other Specialist Technical subjects provide an application for some of the hedging and derivative pricing techniques as well as asset-liability modelling.

Objectives

On completion of the course the trainee actuary will be able to:

(i) Describe and discuss the application of utility theory to economic and financial problems.

1. Explain the meaning of the term "utility function". (Chapters 2 and 3)

2. Explain the axioms underlying utility theory and the expected utility theorem.

3. Explain how the following economic characteristics of consumers and investors can be expressed mathematically in a utility function:
   - non-satiation
   - risk aversion, risk neutrality and risk seeking
   - declining or increasing absolute and relative risk aversion

4. Discuss the economic properties of commonly used utility functions.

5. Discuss how a utility function may depend on current wealth and discuss state dependent utility functions.

6. Explain the concept of utility maximisation and hence explain the traditional theory of consumer choice.

7. Perform calculations using commonly used utility functions to compare investment opportunities

8. State conditions for absolute dominance and for first and second-order dominance and discuss their relationship with utility theory.

9. Discuss the key findings in behavioural finance.

(ii) Discuss the advantages and disadvantages of different measures of investment risk. (Chapter 4)

1. Define the following measures of investment risk:
   - variance of return
   - downside semi-variance of return
   - shortfall probabilities
   - Value at Risk (VaR) / Tail VaR

2. Describe how the risk measures listed in (ii)1 above are related to the form of an investor's utility function.

3. Perform calculations using the risk measures listed above to compare investment opportunities.

4. Explain how the distribution of returns and the thickness of tails will influence the assessment of risk

(iii) Describe and discuss the assumptions of mean-variance portfolio theory and its principal results. (Chapter 5)

1. Describe and discuss the assumptions of mean-variance portfolio theory.

2. Discuss the conditions under which application of mean-variance portfolio theory leads to the selection of an optimum portfolio.

3. Calculate the expected return and risk of a portfolio of many risky assets, given the expected return, variance and covariance of returns of the individual assets, using mean-variance portfolio theory.

4. Explain the benefits of diversification using mean-variance portfolio theory.

(iv) Describe and discuss the properties of single and multifactor models of asset returns. (Chapter 6)

1. Describe the three types of multifactor models of asset returns:
   - macroeconomic models
   - fundamental factor models
   - statistical factor models

2. Discuss the single index model of asset returns.

3. Discuss the concepts of diversifiable and non-diversifiable risk.

4. Discuss the construction of the different types of multifactor models.

5. Perform calculations using both single and multifactor models.

(v) Describe asset pricing models, discussing the principal results and assumptions and limitations of such models. (Chapter 7)

1. Describe the assumptions and the principal results of the Sharpe-Lintner-Mossin Capital Asset Pricing Model (CAPM).

2. Discuss the limitations of the basic CAPM and some of the attempts that have been made to develop the theory to overcome these limitations.

3. Discuss the assumptions, principal results and limitations of the Ross Arbitrage Pricing Theory model (APT).

4. Perform calculations using the CAPM.

(vi) Discuss the various forms of the Efficient Markets Hypothesis and discuss the evidence for and against the hypothesis. (Chapter 1)

1. Discuss the three forms of the Efficient Markets Hypothesis and their consequences for investment management.

2. Describe briefly the evidence for or against each form of the Efficient Markets Hypothesis.

(vii) Demonstrate a knowledge and understanding of stochastic models of the behaviour of security prices. (Chapter 10)

1. Discuss the continuous time lognormal model of security prices and the empirical evidence for or against the model.

2. Discuss the structure of autoregressive models of security prices and other economic variables, such as the Wilkie model, and describe the economic justification for such models.

3. Discuss the main alternatives to the models covered in (vii)1. and (vii)2. above and describe their strengths and weaknesses.

4. Perform simple calculations involving the models described above.

5. Discuss the main problems involved in estimating parameters for asset pricing models:
   - data availability
   - data errors
   - outliers
   - stationarity of underlying time series
   - the role of economic judgement.

(viii) Define and apply the main concepts of Brownian motion (or Wiener processes). (Chapters 8 and 9)

1. Explain the definition and basic properties of standard Brownian motion (or Wiener process).

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.

(ix) Demonstrate a knowledge and understanding of the properties of option prices, valuation methods and hedging techniques. (Chapters 11 to 16)

1. State what is meant by arbitrage and a complete market.

2. Outline the factors that affect option prices.

3. Derive specific results for options that are not model-dependent:
   - Show how to value a forward contract
   - Develop upper and lower bounds for European and American call and put options
   - Explain what is meant by put-call parity

4. Show how to use binomial trees and lattices in valuing options and solve simple examples.

5. Derive the risk-neutral pricing measure for a binomial lattice and describe the risk-neutral pricing approach to the pricing of equity options.

6. Explain the difference between the real-world measure and the risk-neutral measure. Explain why the risk-neutral pricing approach is seen as a computational tool (rather than a realistic representation of price dynamics in the real world).

7. State the alternative names for the risk-neutral and state-price deflator approaches to pricing.

8. Demonstrate an understanding of the Black-Scholes derivative-pricing model:
   - Explain what is meant by a complete market
   - Explain what is meant by risk-neutral pricing and the equivalent martingale measure.
   - Derive the Black-Scholes partial differential equation both in its basic and Garman-Kohlhagen forms.
   - Demonstrate how to price and hedge a simple derivative contract using the martingale approach.

9. Show how to use the Black-Scholes model in valuing options and solve simple examples.

10. Discuss the validity of the assumptions underlying the Black-Scholes model.

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.

12. Demonstrate an awareness of the commonly used terminology for the first, and where appropriate second, partial derivatives (the Greeks) of an option price.

(x) Demonstrate a knowledge and understanding of models of the term structure of interest rates. (Chapter 17)

1. Describe the desirable characteristics of a model for the term structure of interest rates.

2. Describe, as a computational tool, the risk-neutral approach to the pricing of zero-coupon bonds and interest-rate derivatives for a general one-factor diffusion model for the risk-free rate of interest.

3. Describe, as a computational tool, the approach using state-price deflators to the pricing of zero-coupon bonds and interest-rate derivatives for a general one-factor diffusion model for the risk-free rate of interest.

4. Demonstrate an awareness of the Vasicek, Cox-Ingersoll-Ross and Hull & White models for the term structure of interest rates.

5. Discuss the limitations of these one-factor models and show an awareness of how these issues can be addressed.

(xi) Demonstrate a knowledge and understanding of simple models for credit risk. (Chapter 18)

1. Define the terms credit event and recovery rate.

2. Describe the different approaches to modelling credit risk: structural models, reduced form models, intensity-based models.

3. Demonstrate a knowledge and understanding of the Merton model.

4. Demonstrate a knowledge and understanding of a two-state model for credit ratings with a constant transition intensity.

5. Describe how the two-state model can be generalised to the Jarrow-Lando-Turnbull model for credit ratings.

6. Describe how the two-state model can be generalised to incorporate a stochastic transition intensity.

CT8 Index

Chapter 0

Introduction to financial economics

Syllabus objectives

There are no syllabus objectives covered by this chapter.

0 Introduction

This chapter gives:

- some useful background info on the CT8 Course and the relevant past exam questions (Sections 1 and 2)
- a brief introduction to financial economics and also explains some possible different approaches in achieving its goals (Sections 3 to 7).

You may find it useful to re-read this chapter again when you have finished the whole course because it helps to explain how the different elements of the course fit together.

1 Syllabus Objectives, Core Reading and Course Notes

The CT8 Course Notes are divided into 18 chapters, plus this introductory chapter, each containing the relevant Syllabus Objectives and also the Core Reading, which appears in **bold font**.

The *Syllabus* is intended to set out exactly what you need to know, understand and be able to do come the exam.

The purpose of *Core Reading* is to give examiners, tutors and students a clear, shared understanding of the depth and breadth of treatment required by the Syllabus. In the exam, you will be expected to demonstrate your knowledge and understanding of the concepts in Core Reading.

The *Course Notes* then aim to give further explanation of the key points within the Core Reading and include examples and self-assessment questions to help clarify the main points and principles involved. Each chapter also includes one or more exam-style questions towards the end.

Remember though that it is the Core Reading on which you are being tested and so it is the Core Reading on which you should focus when revising and preparing for the exam.

2 Past exam questions

The Subject CT8 Course came into being for the 2005 exams. It largely replaced Subject 109, which was introduced in 2000, although some of the material in the current course previously appeared in Subject 103. The relevant past exam questions prior to 2005 are therefore to be found in these two courses.

We think that the Subject 109 questions from 2000 to 2004 are probably all relevant to the Subject CT8 syllabus *except* for the following:

- April 2000: Q2
- April 2001: Q8
- April 2002: Q4, Q5, Q8
- September 2002: Q4(ii)-(iii)
- April 2003: Q2, Q4(i), Q9(ii)-(iv)
- April 2004: Q10(iii)-(iv), Q11(vi)
- September 2004: Q3(i), Q8(iii)-(iv).

In addition, the following Subject 103 questions from 2000 to 2004 *may* also be relevant to the current Subject CT8 syllabus:

- April 2000: Q2, Q5
- September 2000: Q4
- April 2001: Q7
- September 2001: Q2, Q9
- April 2002: Q2
- September 2002: Q4, Q8
- April 2003: Q3, Q5
- April 2004: Q1
- September 2004: Q4, Q6, Q10.

Note also that the material in Chapters 2 and 3 on expected utility and stochastic dominance was included in Subject CT7 from 2005 to 2009. (Prior to 2005, this material was included in Subject 109.) The following CT7 exam questions are therefore relevant to the current CT8 syllabus:

- April 2005: Q16, Q27, Q28
- September 2005: Q3, Q8, Q12
- April 2006: Q8, Q18
- September 2006: Q6, Q29
- April 2007: none
- September 2007: Q4, Q6, Q10, Q31
- April 2008: none
- September 2008: Q5, Q6, Q27, Q38
- April 2009: Q11
- September 2009: Q11, Q13, Q31.

Finally, note also that the material in Chapter 3 on behavioural finance was added to the Course Notes for the 2014 exams. This material was added from Subject ST5 where it still appears. The following ST5 exam questions (from April 2005 to September 2013 inclusive) refer to behavioural finance and so may be of relevance to Subject CT8:

- April 2005: Q6
- September 2005: Q4
- April 2008: Q6 (ii)
- September 2008: Q3
- September 2010: Q8
- October 2011: Q8 (ii) and (iii)
- October 2012: Q3.

3 Economic models

3.1 What is financial economics?

In simple terms, financial economics as presented in this course involves the modelling of financial markets. The choice and accuracy of the model chosen is critically dependent on the *efficiency* of the market – *ie* whether market prices quickly and accurately reflect all available information. This leads us straight to Chapter 1 and one of the most controversial questions in financial economics: are markets efficient? The Core Reading has been reordered in the Course Notes to start with the Efficient Markets Hypothesis (EMH) because it is such an important topic in financial economics and one that has bearing on everything else studied in the course.

Financial economics often develops models based on economic theory that aim to predict security prices and returns. Such models are therefore referred to as *economic models*. The foundation of economic models is the expected utility theory, which is discussed in detail in Chapter 2. This is because all such models start from the basic assumption that investors aim to maximise their expected utility – often as a function of investment risk and return.

One such economic model is the *mean-variance portfolio theory* discussed in Chapter 5. This characterises how a *single* investor makes its investment choice so as to maximise expected utility within a single time period setting.

3.2 Equilibrium models

Rather than looking at just a single investor, *equilibrium* models assume that there are many investors, each aiming to maximise his or her individual expected utility by an appropriate investment choice. The solution to the model then characterises the equilibrium in the investment market as a whole.

For example, the *capital asset pricing model* (Chapter 7) builds on the assumptions of mean-variance portfolio theory and characterises the equilibrium that results in the *investment market* when there is a large number of investors making their investment choices so as to maximise their individual expected utilities. The capital asset pricing model makes predictions about both the investment choices of individual investors and the expected returns yielded by the investments that they choose. As its name suggests, it can also be used to price assets.

4 Statistical models

4.1 General introduction

Conversely, several of the models discussed in the course are not based on any economic theory. These models can be described as *statistical* models. Examples are:

- *multifactor models* (Chapter 6) – which can be used to describe the process generating investment returns

- *arbitrage pricing theory* (Chapter 7) – which can be used to predict the expected returns on different securities. This is often described as an equilibrium model, however, because the assumption of arbitrage-free pricing is consistent with a market in equilibrium.

4.2 No-arbitrage models

A key category of statistical model is that of the no-arbitrage models. The Vasicek term structure model and the Black-Scholes option pricing formula can both be derived using the no-arbitrage approach.

The Vasicek model of the term structure of interest rates (Chapter 17) was originally derived based on the assumption that:

- the short-term interest rate follows a specific stationary stochastic process
- bond markets are arbitrage-free – *ie* they do not permit risk-free profits to be made.

Similarly, the Black-Scholes option pricing formula (Chapters 14 to 16) can be derived using the no-arbitrage approach by assuming that share prices follow a geometric Brownian motion process and that security markets are again arbitrage-free.

If a realistic stochastic process is assumed and investment markets are actually arbitrage-free, then the predictions of the models should be realistic.

5 Economic v statistical models

The key advantage of a statistical model is that the model structure and parameters can be chosen so as to closely fit actual data, whereas an economic model may fit the data less closely. A statistical model may therefore give more realistic predictions of the future, to the extent that the past offers a reliable guide to the future.

Conversely, economic models have the advantage that the results predicted are capable of an economic interpretation. So, whilst a purely statistical model may be able to describe *how* an investment market works, some economic theory is needed to explain *why* it works in that way. In addition, purely statistical models may sometimes produce results that contradict economic principles, such as the assumption of arbitrage-free investment markets. It is therefore up to the modeller or investor to decide which approach is most appropriate in any particular situation.

These ideas are discussed further in Chapter 10.

6 Time series models

The structure and parameters of many investment-related models are determined by reference to past data, using time series analysis as described in Subject CT6. As before, these models can be divided into two broad categories:

1. *statistical models* – in which the model is determined with reference to the historical data, primarily to ensure a good fit to the past data, though there are one or two other pertinent issues as well. A key statistical model is the *lognormal model of security prices* (Chapter 10), which models the process determining share prices.

2. *econometric models* – in which the choices of model structure and/or parameter values are constrained so as to ensure consistency with economic theory.

In practice, many models fall somewhere in between these two extremes. For example, although the Wilkie model of investment returns (mentioned in Chapter 10) was developed so as to ensure a reasonable fit to historical data, some of both the structure and individual parameter values were chosen with reference to economic judgement and intuition, as opposed to strict economic theory.

As before, econometric models have the advantage of an economic interpretation, but they may not fit the past data as closely as purely statistical models.

7 General modelling principles

In practice, a good model is one that:

- provides a reasonable fit to historical data – on the basis that we can then be confident that it provides a reasonable description of the actual process that is to be modelled

- is consistent with economic theory – on the basis that economic theory actually helps us to understand *why* investors and investment markets behave as they do

  is reasonably simple to develop and use – such a model can be described as *parsimonious*.

It is up to the individual modeller to determine the appropriate balance between these, sometimes conflicting, factors as no single model will ever be completely satisfactory. This choice will usually reflect the reason for developing and using the model in the first place.

Chapter 1

The Efficient Markets Hypothesis

Syllabus objectives

(vi) Discuss the various forms of the Efficient Markets Hypothesis and discuss the evidence for and against the hypothesis.

1. Discuss the three forms of the Efficient Markets Hypothesis and their consequences for investment management.

2. Describe briefly the evidence for or against each form of the Efficient Markets Hypothesis.

0 Introduction

In simple terms, an *efficient* security market is one in which the price of every security fully reflects all available information and hence is equal to its "true" investment value. According to the Efficient Markets Hypothesis security markets *are* efficient.

This basic idea has been extended to allow for different forms of market efficiency corresponding to different levels of information. These are:

- *strong form* – market prices reflect all information, whether or not it is publicly available
- *semi-strong form* – market prices reflect all publicly available information
- *weak form* – market prices reflect all of the information contained in historical price data.

The importance of market efficiency derives from the fact that if markets are *inefficient* then investors with better information may be able to obtain higher investment returns. If, however, markets are *efficient*, then it is not possible to identify under- or over-priced securities, which can then be traded to generate excess risk-adjusted returns. Hence, it is not worth trying to do so.

This chapter therefore:

- describes the three different definitions of market efficiency

- discusses the evidence for and against the different forms of the Efficient Markets Hypothesis – which turns out to be largely inconclusive.

Note that past exam questions have often tested the detailed content of the Core Reading from this chapter. A thorough knowledge and understanding of this material is therefore essential.

1 The three forms of the Efficient Markets Hypothesis

1.1 Background

From the 1930s until the early 1960s, there was a widespread folklore about how to make money on the stock market. The dominant theory, going back to Adam Smith in the 1700s, was that markets are essentially fickle, and that prices tend to oscillate around some true or fundamental value.

Starting with the seminal work by Benjamin Graham, traditional investment analysis involved detailed scrutiny of company accounts, to calculate fundamental values, and thus ascertain when a given investment is cheap or dear. The objective would be to buy cheap stocks and sell dear ones. Any excess performance thus obtained would be at the expense of irrational traders, who bought and sold on emotional grounds and without the benefit of detailed analysis.

This detailed analysis is known as *fundamental analysis* and is discussed in more detail in Subject ST5.

By the 1960s, it became clear that these supposedly foolproof methods of investment were not working. Strategies based on detailed analysis did not seem to perform any better than simple buy-and-hold strategies. Attempts to explain this phenomenon gave rise to the *Efficient Markets Hypothesis*, which claims that market prices already incorporate the relevant information. The market price mechanism is such that the trading pattern of a small number of informed analysts can have a large impact on the market price. Lazy (or cost conscious) investors can then take a free ride, in the knowledge that the research of others is keeping the market efficient.

This provides a strong argument in favour of the *passive* investment management style that we discuss below.

1.2 *The three forms of the Efficient Markets Hypothesis (EMH)*

The academic literature has distinguished different forms of the Efficient Markets Hypothesis, based on a finer dissection of exactly what constitutes *relevant information*. In particular, the following three forms of EMH are commonly distinguished:

1. **Strong form EMH: market prices incorporate all information, both publicly available and also that available only to insiders.**

   If the strong form of the EMH holds, then insider trading is ineffective, in the sense that its application should not enable investors to obtain higher investment returns through information that is known only to themselves and not to the general public.

2. **Semi-strong form EMH: market prices incorporate all publicly available information.**

   If the semi-strong form of the EMH holds, then *fundamental analysis* based upon such information (assuming it is publicly available) should not be able to generate higher investment returns for the investor. Fundamental analysis uses information concerning the issuer of the security (*eg* turnover, profitability, liquidity, level of gearing) and general economic and investment conditions (*eg* real interest rates and inflation) in order to determine the "true" or "fundamental" value of a security and hence whether or not it is cheap or dear.

3. **Weak form EMH: the market price of an investment incorporates all information contained in the price history of that investment.**

   If the weak of the EMH holds, then trading rules based only on historical price data should not be able to generate higher investment returns for the investor. Such trading rules form the basis of *technical analysis*.

Question 1.1

What is the relationship between the three forms of market efficiency?

Strong form efficiency

Stock markets around the world are subject to regulation. Often rules exist to prevent individuals with access to price sensitive information, which is not yet public, using this information for personal gain. Hence senior management involved in take-over talks are often banned from trading in the stock of their company. Such rules would be unnecessary if strong form efficiency held.

Question 1.2

Why would such rules be unnecessary?

Trading on the basis of privileged information that is not publicly available is sometimes known as *insider trading* or *insider dealing*. If insider trading does not occur, then the strong form Efficient Markets Hypothesis cannot hold, as there is then no mechanism by which security prices can incorporate inside information.

Semi-strong form efficiency

The semi-strong form of market efficiency argues that all publicly available information is reflected in the price.

Different stock exchanges have different levels of required disclosure. Hence it would be reasonable to expect different markets to have different levels of efficiency. For example the NYSE (New York Stock Exchange) which requires high level of disclosure, should be more efficient than a market with limited disclosure requirements.

Even if information is publicly available there is a cost involved in obtaining the information quickly and accurately. Any advantage achieved by acting on price relevant information could well be eroded by the cost of obtaining and analysing the information.

In other words, the cost of obtaining additional information could outweigh the additional returns that it might generate. Note that a necessary requirement for efficiency as it has been defined above is that the costs of both acquiring the relevant information and trading on the basis of it should be equal to zero. This must be the case if investors are to trade until security prices do reflect all available information.

Note that just because information is publicly available, it does not mean that everybody has read and understood it, *eg* the contents of the Subject CT8 course. This could be because:

- for many people – who do not wish to be actuaries or investment specialists – the costs of buying a Subject CT8 course outweigh the benefits that it confers
- most of the population is unaware of either the existence or the benefits to be derived from studying Subject CT8, or both!

There is also no commonly accepted definition of what constitutes publicly available information.

This can lead to problems when testing whether a particular market is actually efficient or not.

Access to company accounts is cheap and easy. However, private investors are unable to gain access to senior management of large companies. On the other hand fund managers with large sums at their disposal spend a large amount of time interviewing senior management of organisations which they invest, or intend to invest in. Clearly the fund manager has an advantage in terms of being able to form an opinion on the competence of the management team and the strategy of the company. It is not therefore straightforward to identify when the semi-strong form of market efficiency has been contradicted.

However, in some instances, *eg* the market for residential housing, private investors may be at an informational advantage compared to institutional investors and hence this type of investment is not popular with institutional investors.

Weak form efficiency

The final form of efficiency, weak form, merely claims that the price history of an asset cannot be used to predict out-performance. This form, if true, means that there is nothing to be gained from analysing charts of prices and spotting patterns. Weak form efficiency implies that so called technical analysis (or Chartism) is of no more use than randomly selecting stocks.

Technical analysis is discussed in detail in Subject ST5.

Further refinements of the notion of efficiency

More recently, the notion of efficiency has been further refined, to improve the definition of the returns considered. The most significant refinement has been to consider returns net of various costs – such as the costs of obtaining and using additional information about the company or security.

In this environment, it may be possible to forecast some market movements, but the research costs of making the forecasts (or paying a fund manager to do it for you), plus the transaction costs (brokerage, market impact) of executing the deal must be taken into account. To demonstrate an exploitable opportunity, it must be shown that opportunity is large enough to remain intact even after all these costs are taken into account.

An alternative definition of efficiency sometimes used is therefore that prices reflect all available information up to the point at which the marginal costs and benefits of that information are equal. If these marginal costs and benefits differ between investors, some investors may enjoy an advantage over their peers.

A further consequence is that those investment markets with the most freely available information and the lowest transactions costs are likely to be the most efficient. Thus, government bond markets tend to be more efficient than property markets.

Active versus passive investment management

The question of whether or not markets are efficient has important implications for investment management. Active fund managers attempt to detect exploitable mispricings. Passive fund managers simply aim to diversify across a whole market.

According to the Efficient Markets Hypothesis (EMH), active investment management, with its active trading policy and consequent higher level of management fees, cannot be justified.

> ***Question 1.3***
>
> Why can active management not be justified according to the EMH?

If active investment management cannot be justified, then a more appropriate investment strategy might be to simply match or "track" the market by holding a portfolio whose performance will closely replicate that of the market as a whole. In this way the fund should yield approximately the same level of investment returns as the market, whilst also enjoying the benefits arising from both diversification and lower dealing costs. In practice this is often achieved by matching or tracking an index that is representative of the investment market in question. Such index tracking is a very important example of a *passive* investment management style and is discussed in detail in Subject ST5.

Arithmetic dictates that, in aggregate, active managers must hold the market, so the average actively managed fund is unlikely to diverge far from a passive fund. However, if markets are inefficient, then we would expect those active managers with above average skill to perform better than passive managers do.

The question of market efficiency therefore has a crucial bearing upon the choice of investment management style.

> ***Question 1.4***
>
> Comment on the advantages that could be derived from "insider trading" in a market that is strong form efficient.

2 The evidence for or against each form of the Efficient Markets Hypothesis

2.1 Difficulties with testing the Efficient Markets Hypothesis

Tests of EMH are fraught with difficulty. There is a substantial body of literature proving the existence of mispricings, in contravention of EMH. Unfortunately, there is also a substantial body of literature proving various forms of EMH. Both schools of thought can cite a great deal of empirical evidence and an impressive wealth of statistical tests. It is reasonable to ask, from a philosophical point of view, how it could come about that we have categorical proof of mutually contradictory statements. One possible explanation is that many published tests make implicit, but possibly invalid, assumptions (for example, normality of returns, or stationarity of time series).

Consequently, a test that appears to disprove the Efficient Markets Hypothesis may actually be disproving something else.

We can note that whilst an apparent proof based upon historical data over one period of time might be valid for that particular period, it might not be valid for a subsequent time period, perhaps because the nature of the market or the available information has changed. We can also note that the parties involved in providing proof will have vested interests and may therefore be biased, publishing only those results that support their position.

Some of the differences are purely differences of terminology. For example, do we regard anomalies as disproving EMH, if transaction costs prevent their exploitation?

Thus, although it may in principle be possible to exploit temporary mispricings, it may not be possible in practice after appropriate allowance has been made for both transaction costs and the costs of obtaining information. Whether or not such a finding contradicts the EMH depends upon exactly how we define the EMH.

More subtle is the need to make an appropriate allowance for risk. The EMH is not contradicted by a strategy that produces higher profits than the market portfolio by taking higher risks. The market rewards investors for taking risks so we expect, on average, a high-risk strategy to result in higher returns.

Thus, all else being equal, securities with higher betas *should* yield higher expected returns.

What would contradict the EMH is an investment strategy that provided returns over and above those necessary to compensate an investor for the risk they faced. Unfortunately, there is no universally agreed definition of risk, and no perfectly accurate way of measuring it.

We will consider a number of different measures of investment risk in Chapter 4.

With these caveats in mind we can now consider some empirical work.

Testing the strong form EMH

This is problematic, as it requires the researcher to have access to information that is not in the public domain.

In order to decide if security prices do reflect all available information we ourselves need to have access to all available information – including information that is not publicly available.

However, studies of directors' share dealings suggest that, even with inside information, it is difficult to out-perform.

Nevertheless, as we mentioned earlier, senior management are often banned from taking advantage of their privileged position.

Testing the weak form EMH

Using price history to try and forecast future prices, often using charts of historical data, is called technical or chartist analysis. Studies have failed to identify a difference between the returns on stocks using technical analysis and those from purely random stock selection after allowing for transaction costs. No credible challenge has emerged to the EMH in its weak form.

Question 1.5

Country X runs a national lottery in which the purchaser of a ticket selects six different numbers from 1 to 50 inclusive. If those same six numbers are then drawn randomly from a hat on live TV, the holder of the ticket wins a share of a large cash sum equal in value to the total ticket sales.

Is the market for lottery tickets weak form efficient?

Testing the semi-strong form EMH

The semi-strong form of the EMH has been where research has concentrated and where the debate is most fierce. We will consider tests of the EMH in two categories, tests of informational efficiency and volatility tests.

2.2 Informational efficiency

The EMH (in its various forms) states that asset prices reflect information. However it does not explicitly tell us how new information affects prices, *eg* the speed and extent to which it does so.

It is also empirically difficult to establish precisely when information arrives – for example, many events are widely rumoured prior to official announcements.

An example here is a merger. Should tests of efficiency be based upon the official announcement date of the merger or the date at which rumours concerning the likelihood of the merger first started – or possibly some date in between?

Many studies show that the market over-reacts to certain events and under-reacts to other events. The over/under-reaction is corrected over a long time period. If this is true then traders could take advantage of the slow correction of the market, and efficiency would not hold.

Over-reaction to events

Some of the effects found by studies can be classified as over-reaction to events, for example:

1. **Past performance – past winners tend to be future losers and vice versa – the market appears to over-react to past performance.**

   Hence it might be possible to make excess profits by selling firms that have performed well recently and buying those that have performed badly. This is sometimes referred to as a *contra-cyclical* investment policy.

2. **Certain accounting ratios appear to have predictive powers, *eg* companies with high earnings to price, cashflow to price and book value to market value – generally poor past performers – tend to have high future returns. Again an example of the market apparently over-reacting to past growth.**

3. **Firms coming to the market; in the US evidence from a number of major financial markets including the UK and the US appears to support the idea that stocks coming to the market by Initial Public Offerings and Seasoned Equity Offerings have poor subsequent long-term performance.**

Under-reaction to events

There are also well-documented examples of under-reaction to events:

1. **Stock prices continue to respond to earnings announcements up to a year after their announcement. This is an example of under-reaction to information which is slowly corrected.**

2. **Abnormal excess returns for both the parent and subsidiary firms following a de-merger. Another example of the market being slow to recognise the benefits of an event.**

3. **Abnormal negative returns following mergers (agreed takeovers leading to the poorest subsequent returns). The market appears to over-estimate the benefits from mergers, the stock price slowly reacts as its optimistic view is proved to be wrong.**

Anomalies

All these effects are often referred to as "anomalies" in the EMH framework.

A fast growing area of research in finance is behavioural finance which investigates whether such anomalies arise due to the irrational behaviour of individual investors. However, this approach is still controversial, with many academics unconvinced that irrational behaviour is an important determinant of aggregate asset pricing. Even if the market is efficient, pure chance is going to throw up some apparent examples of mispricings. We would expect to see as many examples of over-reaction as under-reaction. This is broadly consistent with the literature to date.

Even more important is that the reported effects do not appear to persist over prolonged time periods and so may not represent exploitable opportunities to make excess profits. For example, the small companies effect received attention in the early 1980s. This work showed the out-performance of small companies in the period 1960-80. However, if a strategy based on this evidence had been implemented, the investor would have experienced abnormally low returns throughout the 1980s and early 1990s. During this period no papers appeared claiming that small company returns disproved the EMH.

Other examples of anomalies, for example the ability of accounting ratios to indicate out-performance, are arguably proxies for risk. Once these risks have been taken into account many studies which claim to show evidence of inefficiency turn out to be compatible with the EMH.

Question 1.6

Over the last five years, the shares of Company A have yielded an average investment return twice that of Company B. Does this contradict the Efficient Markets Hypothesis?

2.3 *Volatility tests*

Several observers have commented that stock prices are "excessively volatile". By this they mean that the change in market value of stocks (observed volatility), could not be justified by the news arriving. This was claimed to be evidence of market over-reaction which was not compatible with efficiency.

Excessive volatility therefore arises when security prices are more volatile than the underlying fundamental variables that should be driving them.

Shiller first formulated the claim of "excessive volatility" into a testable proposition in 1981. He considered a discounted cashflow model of equities going back to 1870. By using the actual dividends that were paid and some terminal value for the stock he was able to calculate the *perfect foresight price*, the "correct" equity price if market participants had been able to predict future dividends correctly. The difference between the perfect foresight price and the actual price arises from the forecast errors of future dividends. If market participants were rational we would expect no systematic forecast errors. Also if markets are efficient broad movements in the perfect foresight price should be correlated with moves in the actual price as both react to the same news.

Shiller found strong evidence that the observed level of volatility contradicted the EMH. In other words, that actual security prices were more volatile than perfect foresight prices based upon the present value of future dividends.

However, subsequent studies using different formulations of the problem found that the violation of the EMH only had borderline statistical significance. Numerous criticisms were subsequently made of Shiller's methodology. These criticisms covered:

- **the choice of terminal value for the stock price**
- **the use of a constant discount rate**
- **bias in estimates of the variances because of autocorrelation**
- **possible non-stationarity of the series, *ie* the series may have stochastic trends that invalidate the measurements obtained for the variance of the stock price.**

Although subsequent studies by many authors have attempted to overcome the shortcomings in Shiller's original work there still remains the problem that a model for dividends and distributional assumptions are required. Some equilibrium models now exist which calibrate both to observed price volatility and also observed dividend behaviour. However, the vast literature on volatility tests can at best be described as inconclusive.

2.4 Conclusion

The literature on testing the EMH is vast, and articles can be found to support whatever view you wish to take. It is possible to find research claiming incontrovertible evidence either for or against the EMH.

Question 1.7

An investment market is strong form efficient. Describe what would happen to the price of a company's shares if some positive information about the company becomes known. (Assume that nobody had known about this information in advance).

Question 1.8

Is the following statement *true* or *false*?

"The semi-strong form of the Efficient Markets Hypothesis suggests that no investor will "beat" the market in the long term."

Reference

Shiller, RJ, (1981) "Do stock prices move too much to be justified by subsequent changes in dividends?", *American Economic Review*, June 1981, pp 421–36.

Further reading

A good summary of the ideas underlying the Efficient Markets Hypothesis and the various tests of the hypothesis appear in the book by Elton & Gruber listed in the Study Guide.

3 Exam-style questions

At the end of each chapter, we provide exam-style questions for you to attempt. Some will be actual past exam questions (we will indicate where this is the case); others will be taken from other sources (for example, mock exam papers written by ActEd tutors in past years). You should attempt each exam-style question as far as possible without looking at the solution (which is printed overleaf). This should give you some idea as to whether you have understood the ideas introduced in the chapter itself. We give two exam-style questions here.

Question 1

(i) Describe the three forms of the Efficient Markets Hypothesis (EMH). [2]

(ii) A researcher has analysed the annual returns of equity stocks in a particular country over a 10-year period and has made the following observations:

(a) Annual market returns in consecutive years have a negative correlation of –0.25.

(b) The closing value of the index of the 100 stocks with the highest market capitalisation has been found to be 1% higher on average on Fridays than on Mondays.

(c) Announcements of changes in company's dividend policies typically take three months to become fully reflected in the quoted share price.

(d) The prices of a particular subset of stocks have been consistently observed to fall immediately following a favourable announcement and to rise immediately following an unfavourable announcement.

Discuss these observations in the light of the EMH. [4]

[Total 6]

Solution 1

(i) ***Three forms of Efficient Markets Hypothesis***

The Efficient Markets Hypothesis (EMH) states that relevant information is instantly and accurately reflected in market share prices.

There are three different levels of efficiency, depending on the information set that is assumed:

1. *Strong form* efficiency is based on all relevant information (including inside or privately available information)

2. *Semi-strong form* efficiency is based on all publicly available information

3. *Weak form* efficiency is based on historical market information (*ie* prices and/or yields and/or trading volumes).

(ii) ***Discuss the observations***

(a) *Annual market returns are negatively correlated*

This observation suggests that, over annual time periods, the market tends to systematically overreact to new information and hence that the market may not be semi-strong form efficient.

In addition, trading rules could be developed based on this information that could generate excess, risk-adjusted returns, which suggests that this observation is inconsistent with the weak form of the EMH.

(b) *The index is higher on Fridays than on Mondays*

The observation suggests that there is a consistent tendency for prices on Fridays to be "inflated", while prices on Mondays are "depressed", *ie* there is a systematic bias present in the prices.

Trading rules could be developed based on this information (*eg* buy on Monday, sell on Friday) that could generate excess, risk-adjusted returns, which suggests that this observation is inconsistent with the weak form of the EMH.

(c) *Announcements take three months to be reflected*

If the semi-strong form of the EMH holds, public dividend announcements should have an immediate effect on the share prices as the market should respond quickly and accurately to new information.

This observation suggests that the market is not semi-strong form efficient.

(d) *Prices fall following a favourable announcement*

The prices are reacting when information is made public. This suggests that the prices have previously been distorted by insider information.

Therefore, this observation contradicts the strong form of the EMH.

Note that, once a particular form of the EMH is contradicted, this also contradicts any of the stronger forms.

Q

Question 2

Discuss the following statement:

The existence of fund managers who sell their services based on their alleged ability to select over-performing sectors and stocks and so add value to portfolios demonstrates that capital markets are not efficient. [6]

Solution 2

This question is taken from Subject 109 September 2000 Question 2.

The Efficient Markets Hypothesis (EMH) suggests that it is not possible to achieve excess risk-adjusted investment returns using investment strategies based only on certain subsets of information. The existence of fund managers who sell their services based on their alleged ability to select over-performing sectors and stocks does not demonstrate that capital markets are inefficient.

In particular, the semi-strong form of the EMH suggests that excess risk-adjusted investment returns cannot be obtained using only publicly available information.
In certain investment markets, it may therefore be possible (and legal) to achieve excess returns using privileged or inside information, which fact would not contradict the semi-strong form of the EMH.

More generally, the EMH does not preclude managers achieving higher investment returns by adopting "riskier" investment strategies and receiving due reward for the risks taken. It says precisely that it is not possible to develop investment strategies that yield excess risk-adjusted returns – though it is difficult to determine exactly how risk should be interpreted in this context.

Some fund managers must necessarily achieve higher than average returns over a given short time period – eg several years. The point of the EMH is that managers cannot consistently achieve above excess returns. Moreover, they cannot guarantee to achieve excess returns over any particular time period.

Finally, rather than reflecting any market inefficiency in contradiction of the EMH, the existence of such managers may instead reflect the following facts:

- Individual investors may be unaware of the EMH or choose not believe it and hence may be inclined to believe the claims of such managers and so place money with them.

- Certain individual investors may choose to believe the claims of such managers, reflecting the fact that investment decisions are often made on the basis of subjective and emotional factors, in addition to, or instead of on the basis of financial theory.

For the above reasons, the existence of such fund managers does not therefore demonstrate that capital markets are inefficient.

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

Chapter 1 Summary

The Efficient Markets Hypothesis (EMH)

In an *efficient* security market the price of every security fully reflects all available and relevant information. The EMH states that security markets *are* efficient.

Three forms of the EMH are commonly distinguished:

1. *Strong form* – market prices incorporate all information, whether or not it is publicly available. If markets are strong form efficient, then insider trading cannot be used to generate excess risk-adjusted returns.

2, *Semi-strong form* – market prices incorporate all publicly available information. If markets are semi-strong form efficient, then fundamental analysis cannot be used to generate excess risk-adjusted returns.

3. *Weak form* – market prices incorporate all of the information contained in historical price data. If markets are weak form efficient, then technical analysis cannot be used to generate excess risk-adjusted returns.

In practice the level of efficiency depends on whether information is freely available, which in turn may depend on the level of disclosure required by regulation.

The importance of market efficiency derives from the fact that if markets are *inefficient* then investors with better information may be able to generate higher investment returns. If, however, they are efficient then active investment management is difficult to justify.

Tests of the EMH

Tests of the EMH are fraught with difficulty. Consequently, the empirical evidence is inconclusive concerning the extent to which security markets are in fact efficient in practice. However:

- studies of directors' share dealings suggest that, even with inside information, it is difficult to out-perform
- studies have failed to identify a difference between the returns on stocks using technical analysis and those from purely random stock selection
- research has concentrated on the semi-strong form of the EMH and in particular tests of informational efficiency and volatility tests:

Informational efficiency

Many studies show that the market over-reacts to certain events and under-reacts to other events. The over/under-reaction is corrected over a long time period. If this is true then traders could take advantage of the slow correction of the market, and efficiency would not hold.

Over-reaction to events

- Past winners tend to be future losers and the market appears to over-react to past performance.
- Certain accounting ratios appear to have predictive powers, an example of the market apparently over-reacting to past growth.
- Firms coming to the market have poor subsequent performance.

Under-reaction to events

- Stock prices continue to respond to earnings announcements up to a year after their announcement.
- Abnormal excess returns for both the parent and subsidiary firms following a de-merger.
- Abnormal negative returns following mergers.

Volatility tests

Shiller first formulated the claim of "excessive volatility" into a testable proposition in 1981. He found strong evidence that the observed level of volatility contradicted the EMH. However, subsequent studies using different formulations of the problem found that the violation of the EMH only had borderline statistical significance. Numerous criticisms were subsequently made of Shiller's methodology. These criticisms covered:

- the choice of terminal value for the stock price
- the use of a constant discount rate
- bias in estimates of the variances because of autocorrelation
- possible non-stationarity of the series, *ie* the series may have stochastic trends that invalidate the measurements obtained for the variance of the stock price.

Chapter 1 Solutions

Solution 1.1

Publicly available information is a subset of all information, whether publicly available or not. Consequently strong form efficiency implies semi-strong form efficiency, in the sense that if a market is strong form efficient, then it must also be semi-strong form efficient.

Similarly as historical price data is a subset of all publicly available information, so a market that is semi-strong form efficient must also be weak form efficient.

Solution 1.2

Such rules would be unnecessary because it would not be possible for senior management to use this information to obtain higher investment returns by trading in the stock of their own company. Thus, senior management would not be at an advantage compared to other investors, who would correspondingly not be disadvantaged by such trades.

Solution 1.3

According to the Efficient Markets Hypothesis, active investment management cannot be justified because it is impossible to exploit the mispricing of securities in order to generate higher expected returns. Even if price anomalies exist, then the costs of identifying them and then trading will outweigh the benefits arising from the additional investment returns.

Solution 1.4

If the market is strong form efficient, then there will be no advantage from insider trading because all the insider knowledge should be reflected in the current share price.

Solution 1.5

If the numbers drawn are truly random, then the market for lottery tickets is weak form efficient. This is because knowledge of the numbers that have been drawn in the past will not help you to predict the numbers that are likely to be drawn in the future, and thereby generate excess returns. It will also be semi-strong and strong form efficient, unless it is operated fraudulently.

Solution 1.6

Although Company A's shares have recently yielded more than Company B's shares, this does not contradict the Efficient Markets Hypothesis (EMH). This is because the EMH implies that it is not possible to identify shares that offer excess *risk-adjusted expected* returns. This is different from the situation described, which refers to actual *past* returns with no allowance being made for the relative riskiness of the two shares involved. Thus Company A may be inherently more risky than Company B. Even if it isn't, then it may not have been possible to predict Company A's success in advance.

Solution 1.7

1. The share price should go up.
2. This should happen immediately.
3. The share price should rise without bias, *ie* the market does not over-react or under-react.

Note that the answer to this question illustrates some of the ways that stock markets tend not to be fully efficient, *ie* information is not fed into share prices immediately and without bias.

Solution 1.8

The laws of probability suggest that some investors will achieve returns in excess of the market even over the long term purely by chance. For example, they might happen to be holding a particular company's shares when some "good" news is announced. However, the Efficient Markets Hypothesis suggests that no one will be able to do so systematically unless:

- they accept a high level of systematic risk – by investing in a portfolio with a beta in excess of one, or
- they have inside information.

Chapter 2

Utility theory

Syllabus objectives

(i) *Describe and discuss the application of utility theory to economic and financial problems.*

1 *Explain the meaning of the term "utility function".*

2 *Explain the axioms underlying utility theory and the expected utility theorem.*

3 *Explain how the following economic characteristics of consumers and investors can be expressed mathematically in a utility function:*

- *non-satiation*
- *risk aversion, risk-neutrality and risk-seeking*
- *declining or increasing absolute and relative risk aversion.*

4 *Discuss the economic properties of commonly used utility functions.*

5 *Discuss how a utility function may depend on current wealth and discuss state-dependent utility functions.*

6 *Explain the concept of utility maximisation and hence explain the traditional theory of consumer choice.*

7 *Perform calculations using commonly used utility functions to compare investment opportunities.*

0 Introduction

This chapter focuses on utility theory, as applied first to consumer choices and more importantly to investment choices. Note that this material used to be covered in Subject CT7, but was moved to the CT8 Course for the April 2010 sitting.

"Utility" is the satisfaction that a consumer obtains from a particular course of action.

In Section 1 we begin by considering consumer choice – how consumption can be allocated between alternative goods and services in a certain, deterministic, framework.

In practice, however, many economic situations involve choices between uncertain outcomes and hence an element of risk. In the remainder of this chapter, we therefore consider choice under uncertainty and in particular between risky assets that offer different returns to an investor according to the investment outcome that prevails.

Section 2 introduces utility functions and the *expected utility theorem*. This provides a means by which to model the way individuals make investment choices.

Section 3 describes the properties that are normally considered desirable in utility functions so as to ensure that they reflect the actual behaviour of investors. Chief amongst these are:

- *non-satiation*, a preference for more over less, and
- *risk aversion*, a dislike of risk.

These ideas underlie the rest of the CT8 course.

Section 4 considers various methods of measuring risk aversion and the way in which risk aversion might vary with wealth. The concepts of *absolute risk aversion* and *relative risk aversion* are discussed.

Section 5 introduces some commonly used examples of utility functions, namely the *quadratic*, *log*, and *power* utility functions and discusses the properties of each.

Section 6 describes how to deal with situations in which a single utility function is inappropriate. In such instances, it may be necessary to vary either the parameters or the form of the utility function according to the particular situation to be modelled. This leads to the idea of *state-dependent* utility functions.

In order to use the expected utility theory, we need an explicit utility function. In Section 7 we look at how we might go about constructing such a utility function.

Expected utility theory can be useful, but it is not without problems. In Section 8, we therefore consider the limitations of the expected utility theory for risk-management purposes – in particular, the need to know the precise form and shape of the individual's utility function.

1 Consumer choice theory

1.1 Introduction

In this section we describe the traditional economic theory of consumer choice. We develop this theory using a scenario in which there are only two goods, apples and bananas. However, the theory is easily generalised to any number of goods or services of any kind. **A given combination of goods (*eg* two apples and five bananas) is called a "consumption bundle".**

The consumer's problem is to decide which of the available consumption bundles to choose. This has to be done subject to the constraint that the consumer's income is fixed.

The traditional theory of consumer choice has three main elements:

1. **the consumer's preferences** (which allow us to compare different consumption bundles in terms of the relative utility that the consumer expects to obtain from their consumption)

2. **the budget constraint** which defines the different consumption bundles that the consumer can afford

3. **how the consumer decides which consumption bundle to choose**, given 1 and 2.

1.2 Consumer preferences

The assumptions

Economists make the following three assumptions about consumers' preferences:

1. **A consumer *can rank* any two bundles.**

So, for any two consumption bundles, X and Y, a consumer can tell us that they either:

prefer X to Y

or prefer Y to X

or are indifferent between X and Y.

Equivalently, in terms of utility, where $U(\mathrm{X})$ denotes the utility from X, a consumer can tell us that:

$$U(\mathrm{X}) > U(\mathrm{Y}) \quad \text{or} \quad \mathrm{U}(\mathrm{X}) < U(\mathrm{Y}) \quad \text{or} \quad U(\mathrm{X}) = U(\mathrm{Y})$$

If a consumer can rank *any* two consumption bundles, then in theory the consumer can fully rank *all* bundles in terms of the utility they provide.

2. Consumers *prefer more* of a good to less of it.

So if a consumption bundle X, contains more of one good *and no less of the other good* than an alternative bundle, Y, then X is always preferred. In terms of utility:

$$U(\mathrm{X}) > U(\mathrm{Y})$$

3. Consumer preferences exhibit diminishing marginal rates of substitution.

The amount of one good that a consumer is prepared to swap for one extra unit of another good is known as the "marginal rate of substitution".

The marginal rate of substitution of bananas for apples is the number of apples the consumer is willing to give up to gain another banana without changing total satisfaction. A diminishing marginal rate of substitution of bananas for apples implies that the consumer is willing to give up diminishing quantities of apples to gain each additional banana. **This means that if it takes, say, *n* extra apples to persuade a consumer to give up one banana, it will take *more* than another *n* extra apples to persuade her to give up yet another banana.**

This stems from the *law of diminishing marginal utility*, which states explicitly that *marginal utility* of a good decreases as more units of that good are consumed.

The *marginal utility* of a good is the increase in total utility gained by consuming one more unit of that good, for a given level of consumption of other goods.

Implications

A consumer is able to rank bundles. It is therefore possible to join together all the bundles that give the consumer equal satisfaction.

An indifference curve joins all the consumption bundles of equal utility.

Consumers are indifferent because they do not care which of the bundles on the indifference curve they have. All bundles on the curve give equal utility.

Figure 2.1 shows indifference curves for different quantities of apples and bananas. Along each curve, utility is equal. The position of the indifference curve reflects the level of utility – the further from the origin, the higher the utility.

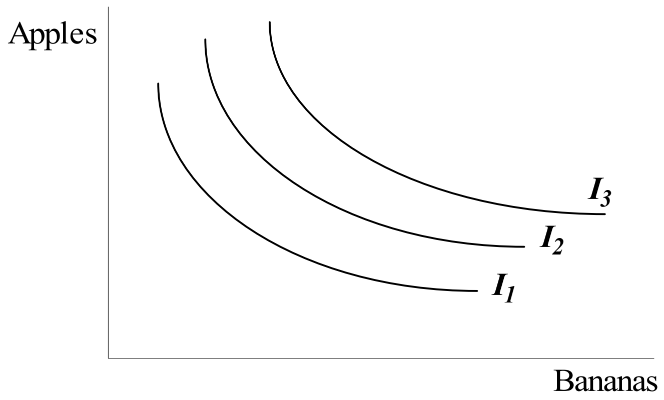


Figure 2.1 – Indifference curves

> ***Question 2.1***
>
> Explain why a consumer's indifference curves drawn for two goods cannot cross.

A consumer *can* rank different bundles, and therefore can pick a set of consumption bundles that give the same utility. Indifference curves further from the origin give higher utility. Indifference curves slope downwards from left to right. Indifference curves are "convex to the origin".

The slope of a consumer's indifference curves will depend on his or her individual preferences and is equal to the marginal rate of substitution. For example, the diagram below illustrates the preferences of a consumer who really likes bananas (or really dislikes apples).

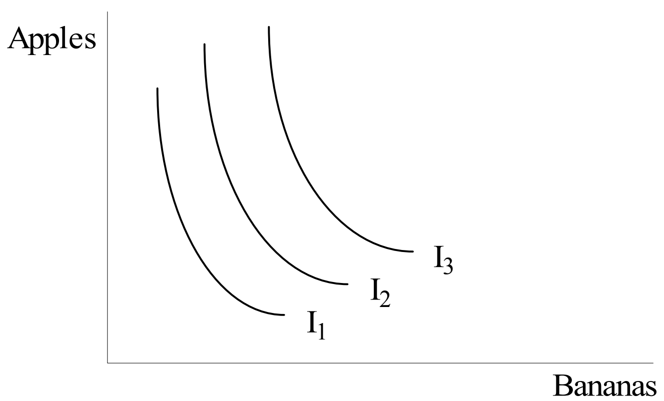


Figure 2.2 – A banana-lover's indifference curves

The *marginal rate of substitution* is equal to the absolute value of the *gradient* of an indifference curve. Both the gradient, and the marginal rate of substitution, fall as we move along the indifference curve. This is due to our assumption of a diminishing marginal rate of substitution. This assumption therefore gives the indifference curve its convex shape.

1.3 The budget constraint

In order to define the different bundles that a consumer can afford, we need to introduce two more assumptions into our theory of consumer choice.

The assumptions

We assume that:

1. The prices of the goods are fixed.

By this, we mean that the prices of the two goods do *not* depend upon which consumption bundle the consumer chooses. Thus, for example, we are ruling out the possibility of bulk discounts.

2. The consumer's income is fixed.

Again, by "fixed" we mean that the consumer's income does not depend upon the choice of consumption bundle.

These two assumptions determine which consumption bundles are affordable.

The budget line joins all points that a consumer can afford, assuming that all income is spent. As the prices of goods are assumed to be fixed in consumer choice theory, the budget line shows a straight-line trade-off between apples and bananas.

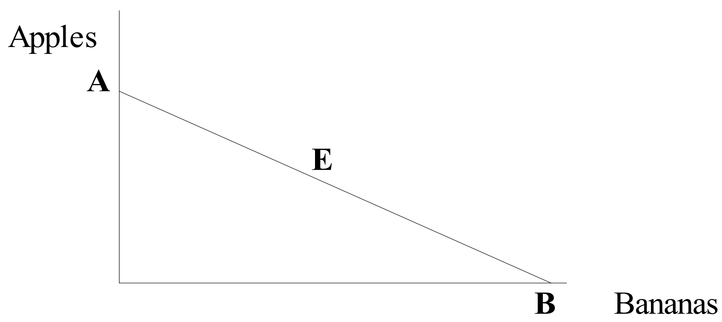


Figure 2.3 – The budget line

With a given income and prices, if a consumer spends all his or her income on apples, he or she can afford to buy A apples. Alternatively, the consumer could buy B bananas, or an intermediate bundle such as E.

1.4 How consumers choose

The assumption

To see how consumers choose which bundle to consume, we need one further assumption:

Economists assume that consumers' choices exhibit *rational behaviour*.

A rational consumer will choose the consumption bundle that maximises his own utility. This is the concept of utility maximisation.

The implications

Combining the budget line with indifference curves, we can determine the consumption bundle which a consumer will choose.

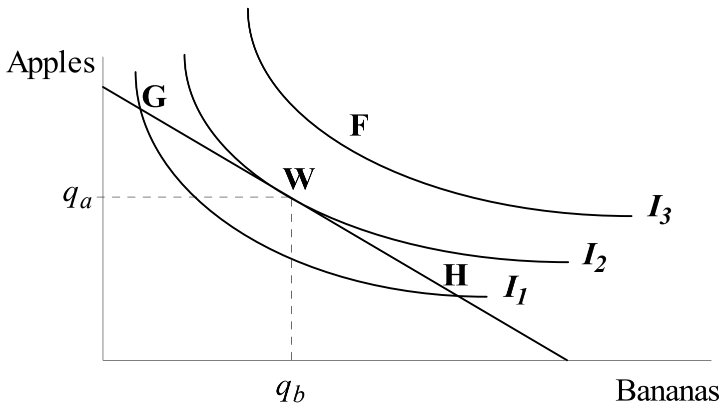


Figure 2.4 – Utility maximisation

Assuming that all income is spent, bundle W will be chosen. This is because W lies on the highest *obtainable* indifference curve. The consumer chooses to buy q_a apples and q_b bananas. A bundle such as F is unaffordable. Bundles such as G and H are affordable, but give lower utility so will not be chosen. W maximises a consumer's utility, given the income represented by the budget line.

Question 2.2

List the assumptions underlying the theory of consumer choice.

Question 2.3

Explain why a consumer will choose the bundle of goods which occurs where the budget line is tangential to an indifference curve. (Hint: consider the gradients of the budget line and the indifference curve.)

A rational consumer will choose a consumption bundle such that the marginal rate of substitution is equal to the slope of the budget line – that is, where the ratios of marginal utilities equal the ratios of prices.

Notice the last point. A consumer maximises utility when the marginal utilities are proportional to the prices.

$$\frac{MU_A}{MU_B} = \frac{p_A}{p_B} \quad \text{or} \quad \frac{MU_A}{p_A} = \frac{MU_B}{p_B}$$

If the price of Good A is twice the price of Good B the consumer must receive twice the marginal utility from Good A compared with Good B.

Question 2.4

The price of Good A is £3 and the price of Good B is £10. The marginal utility from the last unit of Good A is 30 and the marginal utility from the last unit of Good B is 50.

(i) Is the consumer in a utility-maximising equilibrium position?

(ii) If not, what changes in consumption would a rational consumer make to increase total utility?

2 Utility theory

2.1 Introduction

In this section we generalise utility theory to consider situations that involve uncertainty, as will normally be the case where investment choices are concerned. We will refer to a "decision maker" rather than a "consumer" and the utility associated with different levels of wealth rather than the utility associated with different bundles of two goods.

Uncertainty

In what follows, we assume any asset that yields uncertain outcomes or returns, *ie* any *risky* asset, can be characterised as a set of objectively known probabilities defined on a set of possible outcomes. For example, Equity A might offer a return to Investor X of either -4% or +8% in the next time period, with respective probabilities of ¼ and ¾.

Question 2.5

Each year, Mr A is offered the opportunity to invest £1,000 in a risk fund. If successful, at the end of the year he will be given back £2,000. If unsuccessful, he will be given back only £500. There is a 50 per cent chance of either outcome. What is the expected rate of return per annum on the investment to Mr A?

Given the uncertainty involved, the rational investor can no longer maximise his utility with complete certainty. We shall see that he will instead attempt to maximise his *expected* utility by choosing between the available risky assets.

Utility functions

In the application of utility theory to finance it is assumed that a numerical value called the utility can be assigned to each possible value of the investor's wealth by what is known as a *preference function* or *utility function*.

Utility functions show the level of utility associated with different levels of wealth. For example, in a certain world Investor X might have a utility function of the form:

$$U(w) = \log(w)$$

where w is his current or future wealth.

2.2 The expected utility theorem

Introduction

The theorem has two parts.

1. **The expected utility theorem states that a function, *U*(*w*) can be constructed representing an investor's utility of wealth, *w*, at some future date.**

2. **Decisions are made on the basis of maximising the expected value of utility under the investor's particular beliefs about the probability of different outcomes.**

In situations of uncertainty it is impossible to maximise utility with complete certainty. For example, suppose that Investor X invests a proportion a of his wealth in Equity A and places the rest in a non-interest-bearing bank account. Then his wealth in the next period cannot be predicted with complete certainty and hence neither can his utility.

It is possible, however, to say what his *expected* wealth equals. Likewise if the functional form of $U(w)$ is known, then it is possible to calculate his *expected* utility. The expected utility theorem says that when making a choice an individual should choose the course of action that yields the highest expected *utility* – and *not* the course of action that yields the highest expected wealth, which will usually be different.

> ***Question 2.6***
>
> Derive an expression for the expectation of Investor X's next-period wealth if he invests a proportion a of his current wealth w in Equity A (which pays –4% or +8%, with respective probabilities ¼ and ¾) and the rest in a non-interest-bearing bank account.

Calculating the expected utility

Suppose a risky asset has a set of N possible outcomes for wealth $(w_1, ..., w_N)$, each with associated probabilities of occurring of $(p_1, ..., p_N)$, then the *expected utility* yielded by investment in this risky asset is given by:

$$E(U) \quad = \quad \sum_{i=1}^{N} p_i \, U(w_i)$$

ie a weighted average of the utilities associated with each possible individual outcome.

Question 2.7

State an expression for the expectation of the next-period utility of Investor X, again assuming that he invests a proportion a in Equity A and the rest in a non-interest-bearing bank account. He has the utility function $U(w) = \log(w)$.

Note that a *risk-free asset* is a special case of a risky asset that has a probability of one associated with the *certain* outcome, and zero probability associated with all other outcomes.

By combining his beliefs about the set of available assets with his utility function, the investor can determine the optimal investment portfolio in which to invest, *ie* that which maximises his expected utility in that period.

Question 2.8

Investor A has an initial wealth of \$100 and a utility function of the form:

$$U(w) = \log(w)$$

where w is her wealth at any time.

Investment Z offers her a return of -18% or +20% with equal probability.

(i) What is her expected utility if she invests nothing in Investment Z?

(ii) What is her expected utility if she invests entirely in Investment Z?

(iii) What proportion a of her wealth should she invest in Investment Z to maximise her expected utility? What is her expected utility if she invests this proportion in Investment Z?

2.3 Derivation of the expected utility theorem

The expected utility theorem can be derived formally from the following four axioms. In other words, an investor whose behaviour is consistent with these axioms will always make decisions in accordance with the expected utility theorem.

1. Comparability

An investor can state a preference between all available certain outcomes. In other words, for any two certain outcomes A and B, either:

A is preferred to B,
B is preferred to A,
or the investor is indifferent between A and B.

These preferences are sometimes denoted by:

$$U(\mathrm{A}) > U(\mathrm{B}),\ U(\mathrm{B}) > U(\mathrm{A}) \text{ and } U(\mathrm{A}) = U(\mathrm{B})$$

Note that A and B are examples of what we previously referred to as the w_i.

2. Transitivity

If A is preferred to B and B is preferred to C, then A is preferred to C.

ie $U(\mathrm{A}) > U(\mathrm{B})$ and $U(\mathrm{B}) > U(\mathrm{C}) \quad \Rightarrow \quad U(\mathrm{A}) > U(\mathrm{C})$

Also:

$$U(\mathrm{A}) = U(\mathrm{B}) \text{ and } U(\mathrm{B}) = U(\mathrm{C}) \quad \Rightarrow \quad U(\mathrm{A}) = U(\mathrm{C})$$

This implies that investors are consistent in their rankings of outcomes.

3. Independence

If an investor is indifferent between two certain outcomes, A and B, then he is also indifferent between the following two gambles (or lotteries):

(i) A with probability *p* and C with probability (1 – *p*); and

(ii) B with probability *p* and C with probability (1 – *p*).

Hence, if $U(\mathrm{A}) = U(\mathrm{B})$ (and of course $U(\mathrm{C})$ is equal to itself), then:

$$p\,U(\mathrm{A}) + (1-p)\,U(\mathrm{C}) = p\,U(\mathrm{B}) + (1-p)\,U(\mathrm{C}).$$

Thus, the choice between any two certain outcomes is independent of all other certain outcomes.

4. *Certainty equivalence*

Suppose that A is preferred to B and B is preferred to C. Then there is a unique probability, *p*, such that the investor is indifferent between B and a gamble giving A with probability *p* and C with probability (1 – *p*).

Thus if:

$$U(\mathrm{A}) > U(\mathrm{B}) > U(\mathrm{C})$$

Then there exists a unique p ($0 < p < 1$) such that:

$$p\,U(\mathrm{A}) + (1-p)\,U(\mathrm{C}) = U(\mathrm{B}).$$

B is known as the *certainty equivalent* of the above gamble.

It represents the certain outcome or level of wealth that yields the same certain utility as the expected utility yielded by the gamble or lottery involving outcomes A and C. B can also be interpreted as the maximum price that an investor would be willing to pay to accept a gamble.

The four axioms listed above are not the only possible set, but they are the most commonly used.

Question 2.9

Suppose that an investor is asked to choose between various pairs of strategies and responds as follows:

Choose between:	*Response*
B and D	B
A and D	D
C and D	indifferent
B and E	B
A and C	C
D and E	indifferent

Assuming that the investor's preferences satisfy the four axioms discussed above, how does he rank the five investments A to E?

3 The expression of economic characteristics in terms of utility functions

In mainstream finance theory, investors' preferences are assumed to be influenced by their attitude to risk. We need to consider, therefore, how an investor's risk-return preference can be described by the form of his utility function.

The mathematical form of utility functions is normally assumed to satisfy desirable properties that accord with everyday observation about how individuals typically act in the face of uncertainty.

3.1 Non-satiation

It is usually assumed that people prefer more wealth to less. This is known as the principle of non-satiation and can be expressed as:

$$U'(w) > 0$$

This is clearly analogous to individuals preferring more to less of a good or service in the standard choice between different bundles of goods in situations of certainty.

The derivative of utility with respect to wealth is often referred to as the *marginal utility of wealth*. Non-satiation is therefore equivalent to an assumption that the marginal utility of wealth is strictly positive.

3.2 Risk aversion

Attitudes to risk can also be expressed in terms of the properties of utility functions. In particular, we can choose the form of the utility function that we use to model an individual's preferences according to whether or not the individual concerned likes, dislikes or is indifferent to risk.

Risk-averse investor

A *risk-averse* investor values an incremental increase in wealth less highly than an incremental decrease and will reject a fair gamble.

A *fair gamble* is one that leaves the expected wealth of the individual unchanged. Equivalently, it can be defined as a gamble that has an overall expected value of zero.

Example

Suppose that an unbiased coin is tossed once. A gamble in which you receive \$1 if it lands heads up but lose \$1 if it lands tails up is fair, as your expected gain from accepting the gamble is zero and your expected wealth remains unchanged (though your actual wealth will of course change by \$1). Equivalently, the overall expected value of the gamble is equal to:

$$\tfrac{1}{2} \times (+\$1) + \tfrac{1}{2} \times (-\$1) = 0$$

A risk-averse investor derives less additional utility from the prospect of a possible gain than he loses from the prospect of an identical loss with the same probability of occurrence. Consequently he will not accept a fair gamble. He may, however, be willing to trade off lower expected wealth in return for a reduction in the variability of wealth. This is the basic principle underlying insurance.

It is normally assumed that investors are risk-averse and consequently that they will accept additional risk from an investment only if it is associated with a higher level of expected return. Hence, the importance of the trade-off between risk and return that is a feature of most investment decisions.

For a risk-averse investor, **the utility function condition is:**

$$U''(w) < 0$$

In other words, for a risk-averse investor, utility is a (strictly) concave function of wealth, as shown in Figure 2.5.

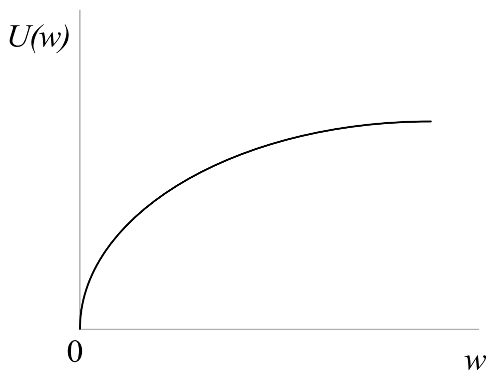


Figure 2.5 – A concave utility function for a risk-averse investor

This concavity condition means that the marginal utility of wealth (strictly) *decreases* with the level of wealth and consequently each additional dollar say, adds less satisfaction to the investor than the previous one.

Question 2.10

Suppose that a risk-averse investor with wealth w is faced with the gamble described in the example above. Show that he will derive less additional utility from the possible gain than he loses from the possible loss and hence that risk aversion is consistent with the condition $U''(w) < 0$.

Risk-seeking investor

A *risk-seeking* investor values an incremental increase in wealth more highly than an incremental decrease and will seek a fair gamble. The utility function condition is:

$$U''(w) > 0$$

A risk-seeking or risk-loving investor will accept any fair gamble and may even accept some unfair gambles (that reduce expected wealth) because the potential increase in utility resulting from the possible gain exceeds the potential decrease in utility associated with the corresponding loss.

Question 2.11

What is the shape of the utility function of a risk-seeking investor?

Risk-neutral investor

A *risk-neutral* investor is indifferent between a fair gamble and the *status quo*. In this case:

$$U''(w) = 0$$

Question 2.12

What can we say about the marginal utility of wealth of a risk-neutral investor?

The utility function of a risk-neutral investor is a linear function of wealth. Assuming non-satiation, then $U'(w) > 0$ and so $U(w)$ increases with w for all w. Thus, the maximisation of expected utility is equivalent to the maximisation of expected *wealth*, in the sense that it will always lead to the same choices being made.

Question 2.13

(i) Ignoring any pleasure derived from gambling, a risk-averse person will:

A never gamble
B accept fair gambles
C accept fair gambles and some gambles with an expected loss
D none of the above

(ii) Ignoring any pleasure derived from gambling, a risk-neutral person will:

I always accept fair gambles
II always accept unfair gambles
III always accept better than fair gambles

A I and II are true
B II and III are true
C I only is true
D III only is true

(iii) A risk-loving person will:

I always accept a gamble
II always accept unfair gambles
III always accept fair gambles

A I and II are true
B II and III are true
C I only is true
D III only is true

4 Measuring risk aversion

4.1 Introduction

In practice, we normally assume that an investor is risk-averse and by looking at the sign of $U''(w)$ we can deduce whether or not this is in fact the case.

The way risk aversion changes with wealth may also be of interest. The degree of risk aversion is likely to vary with the investor's existing level of wealth. For example, we might imagine that wealthy investors are less concerned about risk.

4.2 Risk aversion and the certainty equivalent

Consider the certainty equivalent of a fair gamble. For a risk-averse investor this is negative, *ie* the investor would have to be paid to accept the gamble.

The absolute value of the certainty equivalent of a *gamble* is equal to the maximum sum that the investor would pay to avoid the gamble. It therefore represents the (maximum) premium that the individual would be prepared to pay for insurance against the gamble.

Alternative definitions of the certainty equivalent

Note that we can distinguish between two different types of certainty equivalents depending upon the situation that we are considering, namely:

- The certainty equivalent of the portfolio consisting of the combination of the existing wealth *w and* the gamble *x*, which we can denote c_w. This corresponds to the notion of certainty equivalent in the certainty equivalence axiom.

- The certainty equivalent of the gamble *x* alone, c_x, which will also depend upon the existing level of wealth.

Thus, for a fair gamble and a risk-averse investor, it must be the case that $c_w < w$ and $c_x < 0$.

1. "Additive" or "absolute" gamble

Consider a gamble with outcomes represented by a random variable *x*, in which the sums won or lost are *fixed absolute amounts*. Thus, should the investor accept the gamble, he ends up with a total wealth of $w + x$. The actual sums won or lost are therefore independent of the value of initial wealth *w*.

The certainty equivalent of the combined portfolio of initial wealth plus gamble, c_w, is then defined as the certain level of wealth that solves:

$$U(c_w) = E\left[U\left(w+x\right)\right]$$

The certainty equivalent of the gamble itself is equal to:

$$c_x = c_w - w$$

because we require that $U(w+c_x) = U(c_w)$ and U is a strictly increasing function.

c_x is negative for a fair gamble and its absolute value represents the maximum sum that the risk-averse investor would pay to avoid the risk.

Question 2.14

Recall the example of the fair gamble in the previous section and assume that:

- the investor has initial wealth of 10 and
- a utility function of the form $U(w) = \sqrt{w}$.

Determine the investor's certainty equivalent for this gamble.

2. "Multiplicative" or "proportional" gamble

This is a gamble, with outcomes represented by a random variable y, in which the sums won or lost are all expressed as *proportions* of the initial wealth. If the investor accepts the gamble he therefore ends up with a final wealth of $w \times y$. For example, in a fair gamble of this type, the investor might win 15% of his initial wealth ($y = 1.15$) with probability ¼ and lose 5% ($y = 0.95$) with a probability of ¾. Note that in this case, the actual sums won or lost therefore depend directly upon the value of initial wealth w, *ie* a larger w produces larger wins or losses.

In this instance, the certainty equivalent of total wealth including the proceeds of the gamble can be defined as the level of wealth that satisfies:

$$U(c_w) = E\left[U\left(w \times y\right)\right]$$

The certainty equivalent of the gamble alone is defined as before and is again negative for a fair gamble for a risk-averse investor.

The certainty equivalent and absolute risk aversion

If the absolute value of the certainty equivalent decreases with increasing wealth, the investor is said to exhibit declining *absolute risk aversion*. If the absolute value of the certainty equivalent increases, the investor exhibits increasing *absolute risk aversion*.

Here we have in mind:

- an *additive* gamble
- the certainty equivalent of the gamble alone, c_x.

If the investor's preferences exhibit decreasing (increasing) absolute risk aversion (ARA), then the absolute value of c_x decreases (increases) and the investor is prepared to pay a smaller (larger) absolute amount in order to avoid the risk associated with the gamble.

ie increasing / decreasing ARA $\Leftrightarrow$ increasing / decreasing $|c_x|$

The certainty equivalent and relative risk aversion

If the absolute value of the certainty equivalent decreases (increases) as a *proportion* of total wealth as wealth increases the investor is said to exhibit declining (increasing) *relative risk aversion*.

Here we are looking at:

- a *multiplicative* gamble
- the certainty equivalent of the gamble alone as a *proportion* of initial wealth, *ie* $\dfrac{c_x}{w}$.

If the investor's preferences exhibit decreasing (increasing) relative risk aversion (RRA), then c_x/w decreases (increases) with wealth *w*.

ie increasing / decreasing RRA $\Leftrightarrow$ increasing / decreasing $\left|\dfrac{c_x}{w}\right|$

Example

We shall see below that the log utility function $U(w) = \log(w)$ exhibits:

- decreasing absolute risk aversion
- constant relative risk aversion.

To see these results, consider an individual with an initial wealth of \$100, who faces a fair gamble that offers an equal chance of winning or losing \$20. In this case:

$$U(c_w) = E\left[U(w+x)\right]$$

$$= \frac{1}{2}\left[\ln 120 + \ln 80\right]$$

$$= 4.5848$$

$$\Rightarrow \; c_w = e^{4.5848} = 97.980 \text{ and } c_x = 97.980 - 100 = -2.020$$

If instead his initial wealth is \$200, then:

$$U(c_w) = \frac{1}{2}\left[\ln 220 + \ln 180\right] = 5.2933$$

$$\Rightarrow \quad c_w = e^{5.2933} = 198.997$$

and: $c_x = 198.997 - 200 = -1.003$

The absolute value of c_x, the certainty equivalent of the (fair) gamble alone has decreased with wealth (for a gamble with *fixed absolute proceeds*), as is the case with decreasing absolute risk aversion.

Let us now consider the case of a multiplicative gamble. Suppose the investor is offered an equal chance of winning or losing 20% of his initial wealth. If his initial wealth is £100, then he could win or lose £20. This is equivalent to our first example. We have seen that $c_w = 97.980$ and that $c_x = -2.020$. We can also find that:

$$\frac{c_x}{w} = \frac{-2.020}{100} = -0.0202$$

Example (continued)

If his initial wealth is \$200, the investor could win or lose 20%, *ie* \$40. Then:

$$U(c_w) = \frac{1}{2}[\ln 240 + \ln 160] = 5.2779$$

$$\Rightarrow \qquad c_w = e^{5.2779} = 195.959$$

And: $c_x = 195.959 - 200 = -4.041$

Thus:

$$\frac{c_x}{w} = \frac{-4.041}{200} = -0.0202$$

ie the absolute value of c_x/w, the certainty equivalent of the (fair) gamble as a proportion of initial wealth, is invariant to wealth (for a gamble with *fixed percentage proceeds*), corresponding to constant relative risk aversion.

Question 2.15

Suppose that the investor in Question 2.14 instead had an initial wealth of 20. What would be his revised certainty equivalent for the gamble? What does this suggest about his absolute aversion to risk?

4.3 Risk aversion and the utility function

Absolute and relative risk aversion can be expressed in terms of the utility function as follows.

Absolute risk aversion is measured by the function

$$A(w) = \frac{-U''(w)}{U'(w)}$$

Relative risk aversion is measured by the function

$$R(w) = -w\frac{U''(w)}{U'(w)}$$

These are often referred to as the *Arrow-Pratt* measures of absolute risk aversion and relative risk aversion.

The above results concerning the relationship between the certainty equivalent and the measures of risk aversion arise because it can be shown that the:

- absolute value of the certainty equivalent of a fair gamble is proportional to $\dfrac{-U''(w)}{U'(w)}$

- value of the certainty equivalent of a fair gamble expressed as a proportion of the investor's wealth is proportional to $-w\dfrac{U''(w)}{U'(w)}$.

The following table shows the relationships between the first derivatives of the above functions and declining or increasing absolute and relative risk aversion.

	Absolute risk aversion	***Relative risk aversion***
Increasing	**A'(w) > 0**	**R'(w) > 0**
Constant	**A'(w) = 0**	**R'(w) = 0**
Decreasing	**A'(w) < 0**	**R'(w) < 0**

4.4 Risk aversion and the investment choice

The way that risk aversion changes with wealth can be expressed in terms of the amount of wealth held as risky assets.

Investors who hold an increasing *absolute* amount of wealth in risky assets as they get wealthier exhibit declining absolute risk aversion. Investors who hold an increasing *proportion* of their wealth in risky assets as they get wealthier are exhibiting declining relative risk aversion.

In practice, it is often assumed that as the wealth of a typical investor increases, so the absolute amount that he is willing to invest in risky assets will increase, *ie* that absolute risk aversion decreases with wealth.

It is not so clear cut as to whether we would expect the *proportion* of risky assets to increase or decrease. Consequently the assumption of constant relative risk aversion is sometimes made.

Question 2.16

What can we say about the absolute risk aversion and relative risk aversion of Investor X, whose utility function is $U(w) = \log(w)$?

5 Some commonly used utility functions

5.1 The quadratic utility function

The general form of the quadratic utility function is

$$U(w) = a + bw + cw^2$$

Since adding a constant to a utility function, or multiplying it by a constant will not affect the decision making process, we can write the general form simply as

$$U(w) = w + dw^2$$

Thus:

$$U'(w) = 1 + 2dw$$

and $$U''(w) = 2d$$

Therefore, if the quadratic utility function is to satisfy the condition of diminishing marginal utility of wealth (risk aversion), we must have $d < 0$.

The consequence of this is that the quadratic utility function can only satisfy the condition of non-satiation over a limited range of w:

$$-\infty \le w < -\frac{1}{2d}$$

This constraint on the range of possible values for w is a significant limitation of using quadratic utility functions.

The absolute and relative risk aversion measures are given by:

$$A(w) = \frac{-U''(w)}{U'(w)} = \frac{-2d}{1+2dw}$$

$$A'(w) = \frac{4d^2}{(1+2dw)^2} > 0$$

and:

$$R(w) = w\frac{-U''(w)}{U'(w)} = \frac{-2dw}{1+2dw}$$

$$R'(w) = \frac{-2d}{1+2dw} + \frac{4d^2w}{(1+2dw)^2} = \frac{-2d}{(1+2dw)^2} > 0$$

Thus the quadratic utility function exhibits both increasing absolute and relative risk aversion.

> ***Question 2.17***
>
> Draw the quadratic utility function over the range $0 \le w < -1/2d$ and show why it is valid only for $w < -1/2d$, for a risk-averse investor.

5.2 The log utility function

The form of the log utility function is:

$$U(w) = \ln(w) \qquad (w > 0)$$

Thus:

$$U'(w) = \frac{1}{w}$$

and:

$$U''(w) = -\frac{1}{w^2}$$

Thus the log utility function satisfies the principle of non-satiation and diminishing marginal utility of wealth.

> ***Question 2.18***
>
> Why?

The absolute and relative risk aversion measures are given by:

$$A(w) = \frac{-U''(w)}{U'(w)} = \frac{1}{w}$$

$$A'(w) = -\frac{1}{w^2} < 0$$

and:

$$R(w) = w\frac{-U''(w)}{U'(w)} = 1$$

$$R'(w) = 0$$

Thus the log utility function exhibits declining absolute risk aversion and constant relative risk aversion. This is consistent with an investor who keeps a constant proportion of wealth invested in risky assets as they get richer, in addition to an increasing *absolute* amount of wealth invested in risky assets.

Utility functions exhibiting constant relative risk aversion are said to be iso-elastic.

Iso-elastic means that the elasticity of the marginal utility of wealth is constant with respect to wealth.

The use of iso-elastic utility functions simplifies the determination of an optimal strategy for a multiperiod investment decision because it is possible to make a series of so-called "myopic" decisions. That is the decision at the start of each period only has to consider the possible outcomes at the end of that period and can ignore subsequent periods. Thus, the individual's utility maximisation choice in each period is independent of all subsequent periods. The decision is said to be "myopic" because it is short-sighted, *ie* it does not need to look to future periods.

5.3 The power utility function

The form of the power utility function is:

$$U(w) = \frac{w^{\gamma} - 1}{\gamma} \qquad (w > 0)$$

Thus:

$$U'(w) = w^{\gamma - 1}$$

and:

$$U''(w) = (\gamma - 1)w^{\gamma - 2}$$

Thus for the power utility function to satisfy the principle of non-satiation and diminishing marginal utility of wealth (risk aversion) we require $\gamma < 1$.

The absolute and relative risk aversion measures are given by:

$$\boldsymbol{A(w) = \frac{-U''(w)}{U'(w)} = -\frac{(\gamma-1)}{w}}$$

$$\boldsymbol{A'(w) = \frac{(\gamma-1)}{w^2} < 0}$$

and:

$$\boldsymbol{R(w) = w\frac{-U''(w)}{U'(w)} = -(\gamma-1)}$$

$$\boldsymbol{R'(w) = 0}$$

Thus, like the log utility function, the power utility function exhibits declining absolute risk aversion and constant relative risk aversion. It is therefore also iso-elastic.

The power utility function, in the form given above, is one of a wider class of commonly used functions known as HARA (hyperbolic absolute risk aversion) functions. γ is the risk aversion coefficient.

This is because for such functions the absolute risk aversion is a hyperbolic function of wealth w. For example, in the case of the log utility function:

$$w \times A(w) = constant$$

Hence, a plot of $A(w)$ against w describes a rectangular hyperbola.

Question 2.19

Suppose Investor A has a *power* utility function with $\gamma = 1$, whilst Investor B has a power utility function with $\gamma = 0.5$.

(i) Which investor is more risk-averse (assuming that $w > 0$)?

(ii) Suppose that Investor B has an initial wealth of 100 and is offered the opportunity to buy Investment X for 100, which offers an equal chance of a payout of 110 or 92. Will she choose to buy Investment X?

5.4 Other utility functions

Many different utility functions have appeared in the literature. However, none of those described above allow much freedom in calibrating the function used to reflect a particular investor's preferences.

Question 2.20

Consider the following utility function:

$$U(w) = -e^{-aw}, \qquad a > 0$$

Derive expressions for the absolute risk aversion and relative risk aversion measures. What does the latter indicate about the investor's desire to hold risky assets?

6 The variation of utility functions with wealth

6.1 Introduction

It may not be possible to model an investor's behaviour over all possible levels of wealth with a single utility function. An obvious example is the quadratic utility function described above which only satisfies the non-satiation condition over a limited wealth range. Sometimes this problem can be dealt with by using utility functions with the same functional form but different parameters over different ranges of wealth.

For example, the power utility function could be used to model preferences over all wealth levels, but with the value of the risk aversion coefficient γ changing with wealth. Sometimes, however, it may be necessary to go even further and use different functional forms over different ranges – by constructing *state-dependent utility functions*.

6.2 State-dependent utility functions

State-dependent utility functions can be used to model the situation where there is a discontinuous change in the state of the investor at a certain level of wealth.

They reflect the reality that the usefulness of a good or service to an individual, including wealth, may vary according to the circumstances of the individual. For example, the value of an umbrella depends upon whether or not we believe that it is going to rain over the next few hours or days. In a similar way, the utility that we derive from wealth may also reflect both our existing financial state and our more general circumstances in a way that cannot be captured by a simple functional form. We may therefore need to model preferences using a sophisticated utility function constructed by combining one or more of the standard functions discussed above – so that a different utility function effectively applies over different levels of wealth. A utility function of this kind may involve discontinuities and/or kinks.

Such a situation arises when we consider an insurance company that will become insolvent if the value of its assets falls below a certain level. At asset levels just above the insolvency position, the company will be highly risk-averse and this can be modelled by a utility function that has a discontinuity at the insolvency point.

However, the consequence of applying the same utility function when the company has just become insolvent would be that the company would be prepared to accept a high probability of losing its remaining assets for a chance of regaining solvency.

In other words, at the point of becoming (technically) insolvent the company is very risk-averse, being very keen to avoid this happening. Should it become (technically) insolvent, however, given that the damage (to its reputation or otherwise) has already been done, it may then be willing to take more risks in order to regain solvent status.

This is unlikely to reflect reality and so a different form of utility function would be required to model the company's behaviour in this state.

Question 2.21

Draw the utility function of the above company.

Utility functions can also depend on states other than those, such as insolvency, which are determined by the level of wealth. Obvious examples for an individual include the differences between being healthy or sick, married or single. The state of an individual can also be affected by the anticipation of future events, *eg* if a legacy is expected.

Thus, under some circumstances an individual's utility might accurately be described by a function of the form $U = U(h, w)$ where h is an indicator of health.

Question 2.22

Complete the following sentences by filling in the blank spaces:

Darrell has a logarithmic utility function. Darrell is therefore risk-_________. His preferences also exhibit __________ absolute risk aversion, which means that as his wealth increases, so he will hold _________ wealth in risky assets.

Fiona's utility function, however, is of the form $U(w) = 3w^{2/3}$. Consequently, she is said to be risk-_________ and her preferences exhibit __________ relative risk aversion, which means that as her wealth increases, so the proportion of wealth that she holds in risky assets will _________.

7 Construction of utility functions

7.1 Introduction

In order to use a particular utility function, we need to calibrate the function so that it is appropriate to the particular individual to whom they apply. In other words, we need to find the values of the parameters, for example the value of γ in the power utility function, that apply to the individual.

One approach that has been proposed is to devise a series of questions that allow the shape of an individual's utility function to be roughly determined. A utility curve of a predetermined functional form can then be fitted by a least squares method to the points determined by the answers to the questions. The curve fitting is constrained by the requirement that the function has the desired economic properties (non-satiation, risk aversion and, perhaps, declining absolute risk aversion).

The student may be expected to show how a utility function can be constructed when there is a discrete set of outcomes and the axioms in Section 2.3 of this chapter apply.

7.2 Construction of utility functions by direct questioning

In theory, to determine an individual's utility function we could simply ask the individual what his utility function is. However, in practice it is most unlikely that someone will be able to describe the mathematical form of their utility function.

7.3 Construction of utility functions by indirect questioning

An alternative procedure involves firstly, fixing two values of the utility function for the two extremes of wealth being considered. Secondly, the individual is asked to identify a certain level of wealth such that he or she would be indifferent between that certain level of wealth and a gamble that yields either of the two extremes with particular probabilities. The process is repeated for various scenarios until a sufficient number of plots is found.

Example

Suppose that we wish to determine the nature of an individual's utility function over the range of wealth $0 < w < 4$. One possible approach is to first fix $U(0) = 0$ and $U(4) = 1$. These are the first two points on the individual's utility function.

We could then ask the individual to identify the certain level of wealth, w', such that they would be indifferent between w' for certain and a gamble that yields each of 0 and 4 with equal probability, *ie* w' is the certainty equivalent of the gamble. The expected utility of the gamble is:

$$E(U) = \frac{1}{2}\left[U(0) + U(4)\right] = \frac{1}{2}[0+1] = \frac{1}{2}$$

If $w' = 1.8$ say, then we know that $U(1.8) = 0.5$ and thus have a third point on the utility function.

We could then repeat the exercise for a gamble involving equal probabilities of producing 1.8 and 4, which yields an expected utility of:

$$E(U) = \frac{1}{2}\left[U(1.8) + U(4)\right] = \frac{1}{2}[0.5+1] = 0.75$$

If the certainty equivalent of this gamble is, say, 2.88, then we know that $U(2.88) = 0.75$, giving us a fourth point on the individual's utility function.

This process can be repeated until a sufficient number of points along the individual's utility curve have been identified and a plot of those points produced. Ordinary least squares regression or maximum likelihood methods can then be used to fit an appropriate functional form to the resulting set of values.

Another form of indirect questioning uses information on the premiums that a person is prepared to pay in order to gain an idea of the certainty equivalent of a particular risk.

Thus, we could ask a person what is the maximum that he would be prepared to pay for insurance with a given level of initial wealth and a given potential insurance situation. Points can then be derived on the utility function, which would give rise to the answers given. By repeating this questioning for different initial wealth levels, all the points on the person's utility function could be found.

Consider an example of a person with a house worth £100,000. Suppose that the owner is considering insurance against a variety of perils, each of which would destroy the house completely. These perils have different probabilities of occurring, and the owner has assessed the amount that he is prepared to pay in insurance to insure against each peril, as shown in the following table:

Peril	A	B	C	D
Loss (£K)	100	100	100	100
Probability	0.05	0.15	0.3	0.5
Premium owner is prepared to pay (£K)	20	40	60	80

We can use this table to find out some information about the owner's utility function:

- Fix two values of the utility function. For example, let us suppose that the owner derives utility of zero if he has no wealth, and utility of 1 if he suffers no loss at all, *ie* $U(100)=1$, $U(0)=0$, working in units of £1,000. This is legitimate, because by fixing two points we are just choosing a level and a scale for our measure of utility.

- Consider Peril A. The owner's utility of wealth with insurance will equal his expected utility of wealth without insurance, if he has paid the maximum premium he is prepared to pay. With insurance, his wealth is certain to be 80. Without it, it may be 100 with probability 0.95, or 0 with probability 0.05. So:

  $$U(80)=0.05\times U(0)+0.95\times U(100)=0.95,$$

  and we have a point on the utility function.

Question 2.23

Show, using a similar argument, that $U(60)=0.85$, and find two more points on the owner's utility function. Draw a rough sketch of the graph of the utility function.

The complete utility function can be constructed by considering a large number of different scenarios, each of which contributes a point to the curve. Note that this particular function does appear to satisfy both the usual conditions $U'(x) > 0$ and $U''(x) < 0$.

8 Limitations of utility theory

The expected utility theorem is a very useful device for helping to condition our thinking about risky decisions, because it focuses attention on the types of tradeoffs that have to be made. However, the expected utility theorem has several limitations that reduce its relevance for risk management purposes:

1. **To calculate expected utility, we need to know the precise form and shape of the individual's utility function. Typically, we do not have such information.** Even using the questioning techniques described in the last section, it is still optimistic to assume that it will be possible to construct a utility function that accurately reflects an individual's preferences.

   Usually, the best we can hope for is to identify a general feature, such as risk aversion, and to use the rule to identify broad types of choices that might be appropriate.

2. **The theorem cannot be applied separately to each of several sets of risky choices facing an individual.**

3. **For corporate risk management, it may not be possible to consider a utility function for the firm as though the firm was an individual.**

   The firm is a coalition of interest groups, each having claims on the firm. The decision process must reflect the mechanisms with which these claims are resolved and how this resolution affects the value of the firm. Furthermore, the risk management costs facing a firm may be only one of a number of risky projects affecting the firm's owners (and other claimholders). The expected utility theorem is not an efficient mechanism for modelling the *interdependence* of these sources of risk.

Alternative decision rules that can be used for risky choices include the mean-variance rule (which will be considered in Chapter 5) and stochastic dominance, which is covered in the next chapter**.**

In addition, **new theories of non-rational investment behaviour, known as Behavioural Finance,** are also covered in the next chapter.

9 Exam-style question

We finish this chapter with an exam-style question on utility theory.

Question

An investor can invest in two assets, A and B:

	A	B
expected return	6%	8%
variance	4%%	25%%

The correlation coefficient of the rate of return of the two assets is denoted by ρ and is assumed to take the value 0.5.

The investor is assumed to have an expected utility function of the form:

$$E_\alpha(U) = E(r_p) - \alpha\, Var(r_p)$$

where α is a positive constant and r_p is the rate of return on the assets held by the investor.

(i) Determine, as a function of α, the portfolio that maximises the investor's expected utility. [8]

(ii) Show that, as α increases, the investor selects an increasing proportion of Asset A. [1]
[Total 9]

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:

$$\begin{aligned} E_P &= x_A E_A + (1-x_A)E_B \\ &= 6x_A + 8(1-x_A) \\ &= 8-2x_A \end{aligned}$$

and:

$$\begin{aligned} V_P &= {x_A}^2 V_A + {x_B}^2 V_B + 2x_A x_B \sigma_A \sigma_B \rho_{AB} \\ &= 4{x_A}^2 + 25(1-x_A)^2 + 10x_A(1-x_A) \\ &= 19x_A^2 - 40x_A + 25 \end{aligned}$$

Therefore the investor's expected utility is:

$$\begin{aligned} E_\alpha(U) &= E(r_p) - \alpha Var(r_p) \\ &= 8-2x_A - \alpha\left(19x_A^2 - 40x_A + 25\right) \end{aligned}$$

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

$$\frac{dE}{dx_A} = -2 - \alpha\left(38x_A - 40\right) = 0$$

$$\Leftrightarrow \qquad x_A = \frac{20\alpha - 1}{19\alpha}$$

or:

$$x_A = \frac{20}{19} - \frac{1}{19\alpha}$$

NB The second-order derivative is:

$$\frac{d^2E}{dx_A^2} = -38\alpha < 0$$

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:

$$\frac{dx_A}{d\alpha} = \frac{1}{19\alpha^2} > 0$$

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

Chapter 2 Summary

Consumer choice theory

Consumer preferences are represented by *indifference curves*. They are constructed assuming that:

- consumers can rank different consumption bundles
- consumers prefer more to less
- preferences exhibit diminishing marginal rate of substitution.

Budget constraints are represented by *budget lines* which are determined by:

- the relative prices of goods
- the consumers' incomes.

Consumers are assumed to act rationally by choosing the affordable consumption bundle that maximises utility, *ie* such that the marginal rate of substitution is equal to the slope of the budget line.

The expected utility theorem

The *expected utility theorem* states that:

- a function, $U(w)$, can be constructed representing an investor's utility of wealth, w
- the investor faced with uncertainty makes decisions on the basis of maximising the *expected value* of utility.

Utility functions

The investor's risk-return preference is described by the form of his utility function. It is usually assumed that investors both:

- prefer more to less (*non-satiation*)
- are *risk-averse*.

Additionally, investors are sometimes assumed to exhibit decreasing *absolute risk aversion* – *ie* the *absolute* amount of wealth held in risky assets increases with wealth. In contrast, *relative risk aversion* indicates how the *proportion* of wealth held as risky assets varies with wealth.

Absolute and relative risk aversion are measured by the functions:

$$A(w) = \frac{-U''(w)}{U'(w)} \qquad R(w) = \frac{-wU''(w)}{U'(w)}.$$

Amongst the utility functions commonly used to model investors' preferences are the:

- quadratic utility function
- log utility function
- power utility function.

State-dependent utility functions

Sometimes it may be inappropriate to model an investor's behaviour over all possible levels of wealth with a single utility function. This problem can be overcome either by using:

- utility functions with the same functional form but different parameters over different ranges of wealth, or by using
- *state-dependent utility functions*, which model the situation where there is a discontinuous change in the state of the investor at a certain level of wealth.

Construction of utility functions

One approach to constructing utility functions involves questioning individuals about their preferences. From the answers to these questions, a utility function may be fitted by a least squares method.

The questioning may be direct or indirect. Indirect questioning may be framed in terms of how much an individual would be prepared to pay for insurance against various risks.

Limitations of utility theory

1. We need to know the precise form and shape of the individual's utility function.

2. The expected utility theorem cannot be applied separately to each of several sets of risky choices facing an individual.

3. For corporate risk management, it may not be possible to consider a utility function for the firm as though the firm was an individual

Chapter 2 Solutions

Solution 2.1

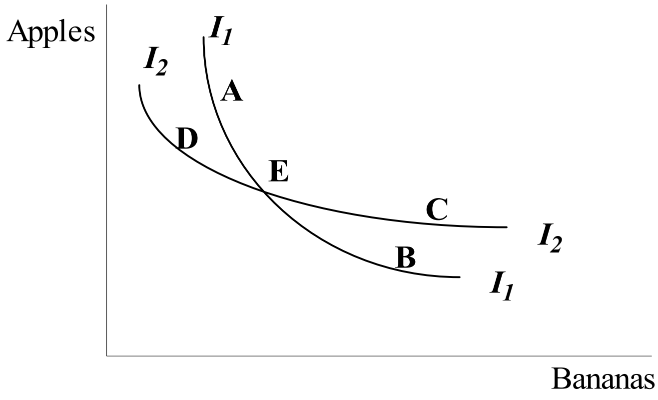


Figure 2.6 – A consumer's indifference curves cannot cross

By definition, a consumer is indifferent between bundles of goods along the same indifference curve. Consider the situation in this diagram, where we have curves that cross. This would mean that the consumer is indifferent between bundles A, E and B, and also between bundles C, E and D. However, if this were the case, the consumer is indifferent between bundles D and A. This violates one of our other assumptions, that consumers prefer more to less. Therefore, a consumer's indifference curves can never cross.

Solution 2.2

1. A consumer can rank any two bundles of goods.
2. Consumers prefer more of a good to less.
3. Consumer preferences exhibit diminishing marginal rate of substitution.
4. The prices of the two goods are not affected by the consumer's consumption choice.
5. The consumer's income is not affected by the consumer's consumption choice.
6. Consumers act rationally to maximise their utility.

Solution 2.3

The assumptions of consumer choice theory mean that indifference curves are downward-sloping and convex to the origin. Therefore, if an indifference curve is *not* tangential to the budget line then either:

1. it must lie entirely to the top-right of the budget line: in this case the consumer is unable to choose a bundle on this curve because his income is insufficient;

or:

2. at least part of it must lie to the bottom-left of the budget line: in this case there must be another, higher, indifference curve which lies on or within the budget line which a consumer would choose to be on instead.

Thus, to avoid a contradiction, the consumer will choose a point of tangency.

Solution 2.4

(i) A consumer is in a utility-maximising equilibrium position when:

$$\frac{MU_A}{p_A} = \frac{MU_B}{p_B}$$

In this case:

$$\frac{30}{£3} > \frac{50}{£10}$$

The consumer is not maximising utility.

(ii) The consumer is receiving more satisfaction per £ spent on Good A than on Good B. The rational consumer will therefore increase consumption of Good A and reduce consumption of Good B. As more units of Good A are consumed, the marginal utility from Good A will fall, and as fewer units of Good B are consumed the marginal utility from Good B will rise. Thus an equilibrium position will be found.

Solution 2.5

We can calculate the expected rate of return as follows:

$$\frac{(0.5\times 2,000)+(0.5\times 500)}{1000}-1=25\%$$

Solution 2.6

$$E(w)=(1-a)\,w+aw[0.25\times 0.96+0.75\times 1.08]$$

$$=(1-a)\,w+1.05aw$$

$$=(1+0.05a)w$$

The answer can also be arrived at directly by noting that the expected next-period wealth will be the initial wealth *w*, plus the expected return of 5% on the investment *aw*.

Solution 2.7

$$E\big[U(w)\big]=0.25\big\{\log\big[(1-0.04a)w\big]\big\}+0.75\big\{\log\big[(1+0.08a)w\big]\big\}$$

Solution 2.8

(i) The expected utility of Investor A is:

$$= \log(100)$$

$$= 4.605$$

(ii) The expected utility of Investor A is:

$$= 0.5\times\log(0.82\times 100)+0.5\times\log(1.2\times 100)$$

$$= 4.597$$

(iii) The expected utility of Investor A is given by:

$$E\big[U(w)\big]=\sum_{i=1}^{n}p_iU(w_i)$$

So:

$$E\left[U\left(w\right)\right] = 0.5\left\{\log\left[(1-0.18a)100\right]\right\} + 0.5\left\{\log\left[(1+0.2a)100\right]\right\}$$
$$= 0.5\left\{\log\left[100-18a\right]\right\} + 0.5\left\{\log\left[100+20a\right]\right\}$$

We differentiate with respect to a to find a maximum.

$$\frac{dE\left[U\left(w\right)\right]}{da} = 0.5 \times \frac{-18}{100-18a} + 0.5 \times \frac{20}{100+20a}$$
$$= \frac{-9}{100-18a} + \frac{10}{100+20a}$$

We then set equal to zero:

$$\frac{9}{100-18a} = \frac{10}{100+20a}$$

Solving, we find $a = 0.2777$.

Checking to see if this gives a maximum:

$$\frac{d^2E\left[U\left(w\right)\right]}{da^2} = \frac{+9(-18)}{(100-18a)^2} + \frac{-10(20)}{(100+20a)^2}$$

This gives a negative value so it is a maximum.

Finding the expected utility from investing 27.77% in Investment Z:

$$E\left[U\left(w\right)\right] = 0.5\left\{\log\left[(1-0.18(0.2777)100\right]\right\} + 0.5\left\{\log\left[(1+0.2(0.2777)100\right]\right\}$$
$$= 0.5\left\{\log\left[100-18(0.2777)\right]\right\} + 0.5\left\{\log\left[100+20(0.2777\right]\right\}$$
$$= 4.6065$$

Solution 2.9

From the responses we can note immediately that:

B > D, D > A, C = D, B > E, C > A, D = E

Hence, transitivity then implies that:

B > D > A
C = D = E

And so we have that:

B > C = D = E > A

Solution 2.10

Whereas the investor's certain utility if the gamble is rejected is $U(w)$, that obtained by accepting the gamble is given by:

$$E(U) = \frac{1}{2}U(w-1) + \frac{1}{2}U(w+1)$$

The gamble is therefore rejected if:

$$\tfrac{1}{2}U(w-1) + \tfrac{1}{2}U(w+1) < U(w)$$

$$\Leftrightarrow \quad U(w-1) + U(w+1) < 2\,U(w)$$

$$\Leftrightarrow \quad U(w+1) - U(w) < U(w) - U(w-1)$$

ie if the additional utility from winning the gamble is less than the loss of utility from losing the gamble. This will be the case if $U''(w) < 0$, which is by definition the case for a risk-averse investor.

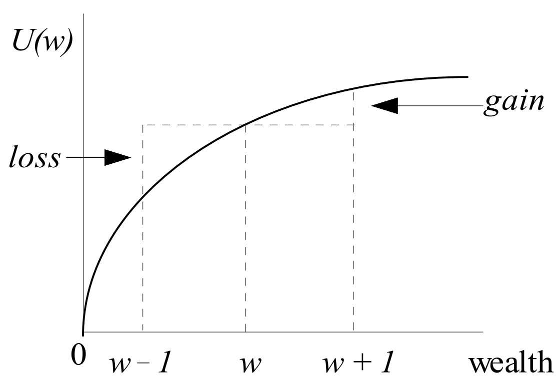


Figure 2.7 – Losses and gains in utility

Solution 2.11

A risk-seeking investor has a convex utility function, because $U''(w) > 0$. The utility function looks as follows:

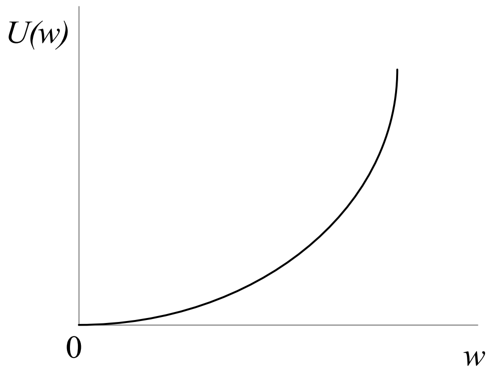


Figure 2.8 – A convex utility function

Solution 2.12

For a risk-neutral investor, $U''(w) = 0$. Thus, $U'(w)$ is constant and so the marginal utility of wealth must itself be constant (and positive assuming non-satiation), so that each additional \$1 is worth the same regardless of wealth.

Solution 2.13

(i) *A risk-averse person will not accept fair gambles*. However, they might accept a gamble where they expected, on average, to win. This would happen if the expected profit from gambling was sufficient to compensate them for taking on the risk. Therefore the correct answer is D.

(ii) *A risk-neutral person will be indifferent to accepting a fair gamble*, but will accept better than fair gambles. Therefore statement III is always true. Statement II is false. Statement I is false (we can't be certain that a person who is indifferent to the gamble will always accept it). The answer is thus D.

(iii) *A risk-loving person will be happy to accept fair gambles*. A risk-loving person will also accept *some* unfair gambles. However, if the odds are very unfair, even a risk-loving person will not accept a gamble. Thus the correct answer is D.

NB This solution shows how risk aversion/neutrality/loving can be defined in terms of a person's attitude towards a fair gamble.

Solution 2.14

With an initial wealth of 10, the expected utility of total wealth is given by:

$$½ \times \left[\sqrt{11} + \sqrt{9}\right] = 3.1583$$

The certainty equivalent, c_w, of the initial wealth plus the gamble satisfies:

$$U(c_w) = 3.1583$$

$$\Rightarrow \qquad \sqrt{c_w} = 3.1583$$

$$\Rightarrow \qquad c_w = 3.1583^2 = 9.9749$$

Hence, the certainty equivalent of the gamble itself is:

$$c_x = c_w - w = 9.9749 - 10 = -0.0251$$

Note that c_x is negative. This means we would have to pay the investor to accept the gamble. Equivalently, the investor would be prepared to pay 0.0251 to avoid the gamble.

Solution 2.15

With an initial wealth of 20, the expected utility of total wealth is given by:

$$½ \times \left[\sqrt{21} + \sqrt{19}\right] = 4.4707$$

The same level of a certain utility is yielded by a certain wealth of:

$$c_w = 4.4707^2 = 19.9875$$

Hence, the certainty equivalent of the gamble is:

$$c_x = c_w - w = 19.9875 - 20 = -0.0125$$

As the absolute value of this certainty equivalent is less than the corresponding certainty equivalent when $w = 10$, we can deduce that the investor's preferences must exhibit decreasing absolute risk aversion.

Solution 2.16

Investor X has a log utility function, which is specified only for $w > 0$. Hence:

$$U'(w) = \frac{1}{w}$$

$$U''(w) = \frac{-1}{w^2}$$

Thus:

$$A(w) = \frac{1}{w} \quad \Rightarrow A'(w) = \frac{-1}{w^2} < 0, \quad \textit{for } w > 0$$

$$R(w) = 1 \quad \Rightarrow R'(w) = 0$$

ie he exhibits decreasing absolute risk aversion and constant relative risk aversion.

Solution 2.17

For non-satiation we require:

$$U'(w) > 0, \ \textit{ie}\ 1 + 2dw > 0$$

$$\Leftrightarrow \quad 2dw > -1$$

$$\Leftrightarrow \quad dw > -½$$

$\Leftrightarrow$ $w < -1/2d$, as $d < 0$ for a risk-averse investor.

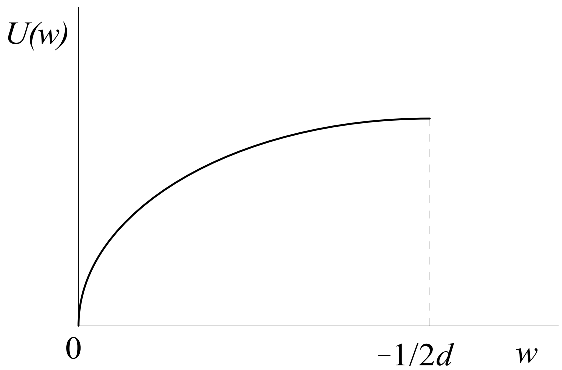


Figure 2.9 – The range of the quadratic utility function

Solution 2.18

Because we have assumed that the log utility function is defined only for positive values of w and $U'(w) = \dfrac{1}{w} > 0$, *ie* non-satiation and $U''(w) = \dfrac{-1}{w^2} < 0$, *ie* diminishing marginal utility of wealth, for $w > 0$.

Solution 2.19

(i) ***Which investor is more risk-averse?***

Investor B is more risk-averse because she has a lower risk aversion coefficient γ. We can show this by deriving the absolute risk aversion and relative risk aversion measures for each investor. For Investor A:

$$A(w) = R(w) = 0$$

ie Investor A is *risk-neutral*.

For Investor B:

$$A(w) = \frac{1}{2w} > 0, \quad R(w) = \frac{1}{2} > 0$$

Hence, Investor B is *strictly risk-averse* for all $w > 0$.

(ii) ***Will Investor B buy Investment X?***

If Investor B buys X, then she enjoys an expected utility of:

$$0.5\left[2\left(\sqrt{110}-1\right)+2\left(\sqrt{92}-1\right)\right] = 18.08$$

If, however, she does not buy X, then her expected (and certain) utility is:

$$2\left(\sqrt{100}-1\right) = 18$$

Thus, as buying X yields a higher expected utility, she ought to buy it.

Solution 2.20

The utility function $U(w) = -e^{-aw}$ is such that:

$$U'(w) = ae^{-aw} \quad \text{and} \quad U''(w) = -a^2 e^{-aw}$$

Thus:

$$A(w) = \frac{-U''(w)}{U'(w)} = a > 0 \quad \text{and} \quad A'(w) = 0$$

and:

$$R(w) = \frac{-wU''(w)}{U'(w)} = aw > 0 \quad \text{and} \quad R'(w) = a > 0$$

Hence, as the absolute risk aversion is constant and independent of wealth the investor must hold the same *absolute* amount of wealth in risky assets as wealth increases. Both this, and the fact that the relative risk aversion increases with wealth, are consistent with a *decreasing proportion* of wealth being held in risky assets as wealth increases.

Solution 2.21

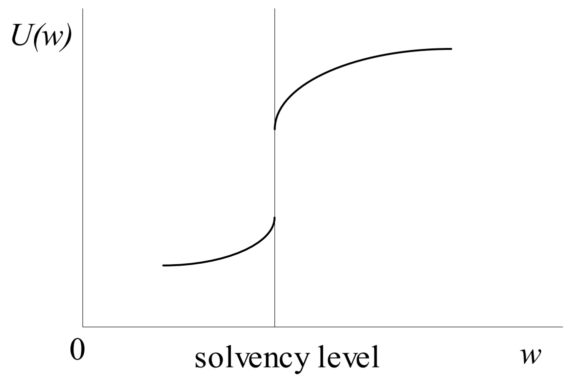


Figure 2.10 – A state-dependent utility function

ie the company is extremely risk-averse when just solvent, so the curve has a rapidly changing gradient. At the solvency level the curve is vertical as any slight increase in wealth leads to a large jump in utility.

Solution 2.22

Darrell has a logarithmic utility function. Darrell is therefore risk-***averse***. His preferences also exhibit ***decreasing*** absolute risk aversion, which means that as his wealth increases, so he will hold ***more (absolutely)*** wealth in risky assets.

Fiona's utility function, however, is of the form $U(w) = 3w^{2/3}$. Consequently, she is said to be risk-***averse*** and her preferences exhibit ***constant*** relative risk aversion, which means that as her wealth increases, so the proportion of wealth that she holds in risky assets will ***not change***.

Solution 2.23

Consider Peril B. With insurance against this peril, the owner's wealth will be 60. Without it, his wealth will either be 100 with probability 0.85 or 0 with probability 0.15. Equating the utilities of these two possibilities, we have

$$U(60) = 0.85 \times U(100) + 0.15 \times U(0) = 0.85 .$$

Similarly, considering Perils C and D, we obtain $U(40) = 0.7$ and $U(20) = 0.5$. So the utility function will look something like this:

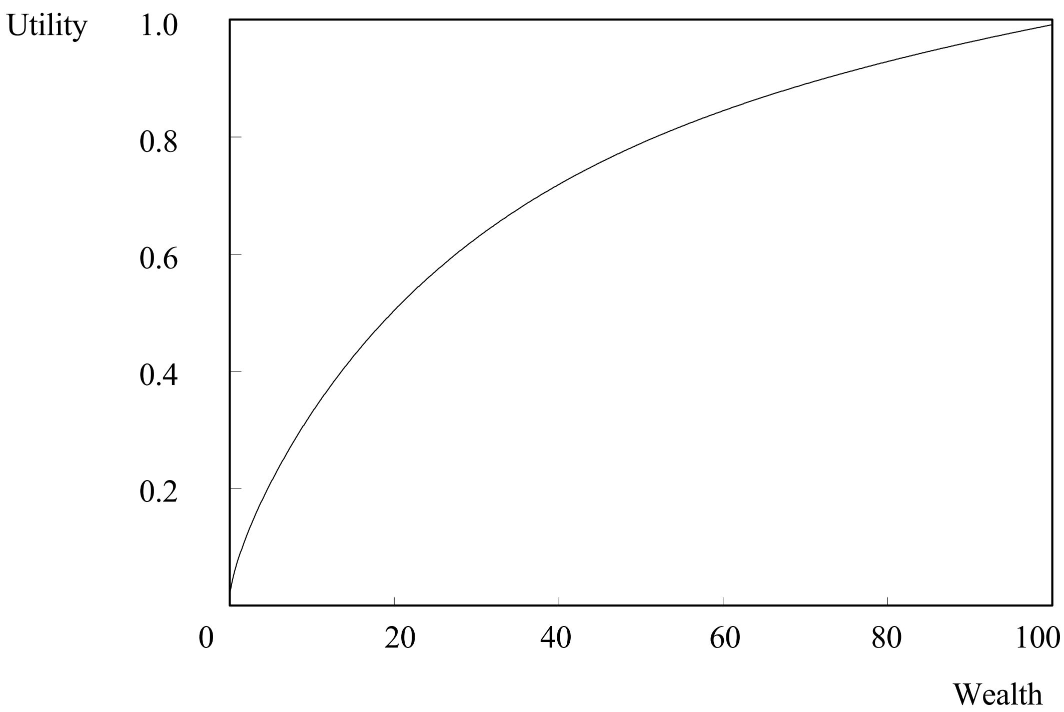


Figure 2.11 – The owner's utility function

Chapter 3

Stochastic dominance and behavioural finance

Syllabus objectives

(i) 8 *State conditions for absolute dominance and for first and second-order dominance and discuss their relationship with utility theory.*

9 *Discuss the key findings in behavioural finance.*

0 Introduction

This chapter focuses on the topics of *stochastic dominance* and *behavioural finance*. Note that:

- Stochastic dominance used to be covered in Subject CT7, but was moved to the CT8 Course for the April 2010 sitting.

- Behavioural finance was added to Subject CT8 for the 2014 exams. Prior to then it appeared in Subject ST5 – where it is now repeated.

Recall that the expected utility theorem suggests that a rational investor will aim to maximise his or her expected utility. In practice, however, it is not always possible to model choices with utility functions. Consequently, alternative approaches, such as *stochastic dominance*, may then be used to say something about an investor's choices without knowing the exact specification of the investor's utility function. Instead, the use of stochastic dominance relies on assumptions regarding the general form of the utility function.

In addition, there is much experimental and empirical evidence to suggest that in practice investors are often *irrational*, in the sense that their actions are inconsistent with the predictions of utility theory. Instead, they appear to be subject to a number of biases and errors. Behavioural finance analyses these biases and their implications for financial decision making.

1 Stochastic dominance

1.1 Background

Absolute dominance exists when one investment portfolio provides a higher return than another in all possible circumstances. Clearly, this situation will rarely occur so we usually need to consider the relative likelihood of out-performance; *ie* stochastic dominance.

We consider two investment portfolios, A and B, with cumulative probability distribution functions of returns F_A and F_B respectively.

Assume that the probability that portfolio i yields an end of period wealth, x, less than or equal to L is given by:

$$P(x \le L) = F_i(L) = \int_{-\infty}^{L} f_i(x)\, dx, \text{ or}$$

$$P(x \le L) = F_i(L) = \sum_{x=-\infty}^{L} p_i(x)$$

depending upon whether or not the distribution of possible outcomes is continuous or discrete. $f_i(x)$ or $p_i(x)$ is therefore the corresponding probability (density) function.

1.2 First-order stochastic dominance

The first-order stochastic dominance theorem states that, assuming an investor prefers more to less, A will dominate B (*ie* the investor will prefer Portfolio A to Portfolio B) if:

$$F_A(x) \le F_B(x), \text{ for all } x, \text{ and}$$

$$F_A(x) < F_B(x) \text{ for some value of } x.$$

In words, this means that the probability of Portfolio B producing a return below a certain value is never less than the probability of Portfolio A producing a return below the same value and exceeds it for at least some value of *x*.

For example, if two normal distributions have the same variance but different means, the one with the higher mean displays first-order stochastic dominance over the other.

Q

Question 3.1

Assuming that Portfolio A first-order stochastically dominates Portfolio B, draw a diagram illustrating the relationship between $F_A(x)$ and $F_B(x)$, the respective cumulative probability distribution functions.

Using first-order stochastic dominance to make investment decisions is similar to basing choices on the "more to less" criterion discussed in the previous chapter. With this in mind, consider Asset X and Asset Z, which offer returns as follows according to whether or not there is a "good" or "poor" investment outcome.

	Asset X	*Asset Z*
Good outcome	6%	10%
Poor outcome	5%	8%

In this instance, it is clear that an investor who prefers more to less should choose Asset Z, which produces a higher return under both possible outcomes. Asset Z is said to *absolutely dominate* Asset X.

Suppose instead, however, that Assets X and Z offer the following possible outcomes with associated probabilities:

Asset X		*Asset Z*	
return	*Probability*	*return*	*probability*
7%	½	8%	½
5%	½	6%	½

In this case the investor's choice is not quite as clear-cut because it is possible that Asset X *may* produce a higher actual return, although we suspect that the investor who prefers more to less should choose Asset Z, which offers a higher expected return. The first-order stochastic dominance theorem formalises the intuition behind the choice.

Consider the table below, which shows the cumulative probabilities of obtaining a return equal to or less than any particular value for each of these two assets.

	cumulative probability	
return	*Asset X*	*Asset Z*
5%	½	0
6%	½	½
7%	1	½
8%	1	1

Asset Z offers a (cumulative) probability of receiving any amount L or less that is never greater, and sometimes strictly less, than that offered by Asset X,

ie $P(x < L \text{ for Asset Z}) \leq P(x < L \text{ for Asset X})$ for $L = 5, 6, 7, 8$

with the inequality being strict for $L = 5, 7$. Equivalently:

$F_Z(x) \leq F_X(x)$ for each of $x = 5, 6, 7, 8$, with

$F_Z(5) < F_X(5)$ and $F_Z(7) < F_X(7)$.

Hence, Asset Z (first-order stochastically) dominates Asset X. Consequently, an investor who prefers "more to less" should choose Asset Z.

On the diagram below, the cumulative probability function for Asset Z, $F_Z(x)$, is never above that of Asset X, $F_X(x)$.

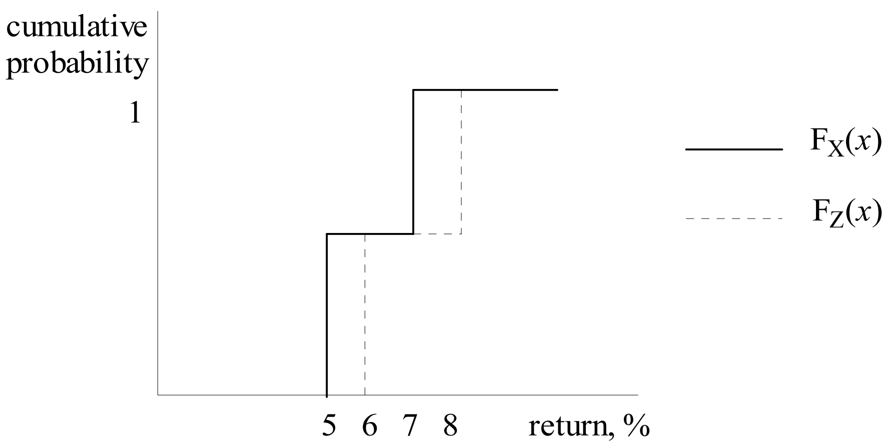


Figure 3.1 – Assessing first-order stochastic dominance for Assets X and Z

A slightly different way of expressing the same idea is to say that Z (first-order stochastically) dominates X if Z can be obtained from X by shifting probability from lower to higher outcome levels.

Question 3.2

Consider two risky assets A and B. A yields \$1 with probability ¼ and \$2 with probability ¾. B yields \$1 with probability ½ and \$2 with probability ½. Why does A dominate B?

What can we say about Asset C, which yields \$1 with probability ¼ and \$3 with probability ¾?

Often first-order stochastic dominance will not be a sufficiently strong criterion by which to choose between assets. On such occasions, we need to call upon the stronger criterion of second-order stochastic dominance.

1.3 *Second-order stochastic dominance*

The second-order stochastic dominance theorem applies when the investor is risk-averse, as well as preferring more to less.

In this case, the condition for A to dominate B is that

$$\int_a^x F_A(y)\,dy \le \int_a^x F_B(y)\,dy, \text{ for all } x,$$

with the strict inequality holding for some value of *x*, where *a* is the lowest return that the portfolios can possibly provide.

The condition for second-order stochastic dominance is similar to that for first-order stochastic dominance, except that it is expressed in terms of the *sums* of the cumulative probabilities, rather than simply the cumulative probabilities themselves.

The interpretation of the inequality above is that a risk-averse investor will accept a lower probability of a given extra return at a low absolute level of return in preference to the same probability of extra return at a higher absolute level. In other words, a potential gain of a certain amount is not valued as highly as a loss of the same amount.

For example, if two normal distributions have the same mean but different variances, the one with the lower variance displays second-order stochastic dominance over the other.

Consider Assets U and V, which offer returns according to the table below:

	probability	
return	*Asset U*	*Asset V*
6%	¼	0
7%	¼	¾
8%	¼	0
9%	¼	¼

Question 3.3

Calculate $F_U(x)$ and $F_V(x)$ and explain why an investor cannot choose between Assets U and V on the basis of first-order stochastic dominance alone.

Question 3.4

Fill in the spaces below, which refer to Assets U and V which offer returns as indicated in the table above.

Assets U and V both offer the same chance of a ____ % return. In addition, U offers a wider spread of returns about ____% – *ie* a greater chance of ____ % return, but at the risk of a greater chance of obtaining only ____%. Assets U and V therefore both offer the same expected return of _____%, but the variance of return is greater for Asset ___. Thus, a risk-averse investor should choose Asset _____.

An investor who bases his choices upon the second-order stochastic dominance theorem will make identical choices to those implied by non-satiation and risk aversion, assuming that he is able to make a choice.

We shall now show that Asset V (second-order stochastically) dominates Asset U.

Consider the following table that shows the *sums* of the cumulative probability functions.

	sum of cumulative probabilities	
return	*U*	*V*
6%	¼	0
7%	¾	¾
8%	1½	1½
9%	2½	2½

In this case, V (second-order stochastically) dominates U because the sum of its cumulative probabilities is never greater than that of U and for one outcome is strictly less, *ie* :

$$\int_{a=6}^{x} F_V(y)\,dy \le \int_{a=6}^{x} F_U(y)\,dy, \text{ for } x = 6, 7, 8, 9$$

with the strict inequality holding for $x = 6$.

According to the second-order stochastic dominance theorem, the investor should therefore always choose Asset V – which offers the same expected return as Asset U but with a lower variance.

Question 3.5

Draw a diagram showing $F_U(x)$ and $F_V(x)$ and use it to explain why Asset V (second-order stochastically) dominates Asset U.

2 Relationship between dominance concepts and utility theory

2.1 Dominance concepts and utility theory

The above stochastic dominance theorems can be used to choose between portfolios given very general assumptions about the investor's utility function.

First-order stochastic dominance

If Security A exhibits first-order stochastic dominance over Security B, then the first-order stochastic dominance theorem states that Security A will be preferred by any investor that prefers more to less, *ie* for whom $U'(x) > 0$.

Thus, if we don't know the exact form of the investor's utility function, but are willing to assume that the investor prefers more to less, then first-order stochastic dominance may enable us to say something about the investor's investment choices between two assets. If not, then we need to make the additional assumption that the investor is risk-averse and then try to apply the second-order stochastic dominance theorem.

Second-order stochastic dominance

If Security A exhibits second-order stochastic dominance over Security B, then the second-order stochastic dominance theorem states that Security A will be preferred by any investor that prefers more to less *and* is risk-averse, *ie* for whom $U'(x) > 0$ and $U''(x) < 0$.

Thus, if we don't know the exact form of the investor's utility function, but are willing to assume both that the investor prefers more to less *and* is risk-averse, then second-order stochastic dominance may enable us to say something about the investor's investment choices between two assets.

2.2 Proof of the theorems

The theorems can be proved by application of the expected utility principle. For example, the assumption that an investor prefers more to less can be stated in utility function terms as:

$$U'(w) > 0$$

The expected utility of A is:

$$E[U_A] = \int_a^b U(w)\, dF_A(w)$$

and the expected utility of an investment in Portfolio B is:

$$E[U_B] = \int_a^b U(w)\, dF_B(w)$$

Note that a and b are the smallest and largest return respectively that could be provided by either portfolio.

Thus, if A is preferred to B:

$$\int_a^b U(w)\, dF_A(w) - \int_a^b U(w)\, dF_B(w) > 0$$

The left-hand side can be written as:

$$\int_a^b U(w)[dF_A(w) - dF_B(w)]$$

and integrating by parts yields:

$$\left[U(w)(F_A(w) - F_B(w))\right]_a^b - \int_a^b U'(w)[F_A(w) - F_B(w)]\, dw$$

Now, $F_A(a) = F_B(a) = 0$ by definition and $F_A(b) = F_B(b) = 1$ by definition, so for the expression to be positive we require the value of the integral to be negative.

This is because:

$$\left[U(w)(F_A(w) - F_B(w))\right]_a^b = 0.$$

$U'(w) > 0$ by assumption, so for the integral to be negative, no matter what the exact form of $U'(w)$, $F_A(w) - F_B(w)$ must be less than or equal to zero for all values of w with $F_A < F_B$ for at least one value of w if the value is not to be zero.

Note that the proof did not require explicit specification of the utility function – beyond the very general requirement that $U'(w) > 0$.

A similar proof of the second-order stochastic dominance theorem can be constructed by bringing in the additional assumption that $U''(w) < 0$.

Recall that risk aversion, $U''(w) < 0$, is a necessary requirement for second-order stochastic dominance

Question 3.6

What are the main advantage and disadvantages of using stochastic dominance to make investment decisions?

3 Behavioural finance

3.1 Introduction

The field of behavioural finance looks at how a variety of mental biases and decision-making errors affect financial decisions. It relates to the psychology that underlies and drives financial decision-making behaviour.

Although traditional economic theory assumes that investors always act rationally, *ie* with the aim of maximising expected utility, experimental and actual evidence suggests that this may not always be entirely the case.

Much of the work to date has concentrated on the impact on prices in capital markets (indeed, some "contrarian" investment funds are run on the basis of taking advantage of errors made by other investors).

A *contrarian fund* is one that tends to take the opposite view to the rest of the market – *ie* it will tend to buy shares when most people in the market are recommending sell and vice versa. For example, it might sell shares when the market is "high" or rising on the basis that the market tends to overreact to positive news and so is likely to be overvalued.

In investment consultancy, behavioural arguments can be applied to trustees and used to justify proposed investment management structures.

In practice, the management of an investment portfolio in accordance with an investment objective will require both:

- an investment strategy
- a choice of investment management structure.

Once these have been determined, the implementation of the chosen investment strategy will then require the selection of the individual investments and investment managers.

If it is recognised that the trustees responsible for the direction of investment policy are subject to the types of mental bias identified by behavioural finance, then the recommended investment management structure may be chosen to reflect those biases.

Likewise, if capital markets are indeed influenced by behavioural factors, then those investors who recognise this may be able to exploit this knowledge to the disadvantage of investors who don't.

Common themes found in research on behavioural finance include:

- anchoring and adjustment
- prospect theory
- framing (and question wording)
- myopic loss aversion
- estimating probabilities
- overconfidence
- mental accounting
- the effect of options.

3.2 *Anchoring and adjustment*

Anchoring is a term used to explain how people will produce estimates. They start with an initial idea of the answer ("the anchor"). They then adjust away from this initial anchor to arrive at their final judgement.

Thus, people base perceptions on past experience or "expert" opinion, which they amend to allow for evident differences to the current conditions. The effects of anchoring are pervasive and robust and are extremely difficult to ignore, even when people are aware of the effect and aware that the anchor is ridiculous. Even patently ridiculous anchor values have been shown to influence post-anchor estimates.

The effect of anchoring and adjustment grows with the size of the difference between the *anchor value* – the original estimate provided – and the *pre-anchor estimate* – the mean estimate people make before being exposed to an explicit anchor. In other words, the bigger this difference, the greater the influence of the anchor value on the post-anchor estimate.

Thus, for example, we could ask an estate agent to value a house without giving them any clues (such as a listed price) or anchor values. The resulting estimate would be a *pre-anchor* value. We could then tell them the listed price/anchor value and get them to re-estimate the value of the house with this additional information. The resulting estimate would be the *post-anchor* estimate.

The following example is taken from the paper

- Northcraft G B, and M A Neale , Experts, amateurs and real estate: an anchoring and adjustment perspective on property pricing decisions, *Organizational behaviour and human decision processes*, *39*.

Example

An experiment was conducted in which a large number of real estate agents were asked to value a property and come up with a recommended selling price. They were each provided with an information booklet containing a large volume of information concerning the property. The booklet was identical for all of the agents, except that four different versions were used, each with a different listed (*ie* suggested) price for the property. It turned out that the average selling price recommended by the agents increased with the listed price as shown in the following table.

Listed price ($)	Average recommended selling price ($)
119,900 (Version 1)	117,745
129,900 (Version 2)	127,836
139,900 (Version 3)	128,530
149,900 (Version 4)	130,981

In the above example, the anchor value is the listed price. Thus, the agents' estimates were influenced by the "anchor" or benchmark given to them in the form of the listed price. In addition it turns out that the further the anchor value gets from the "true" value, then the more it will pull people's estimates away from the true value.

3.3 Prospect theory

Prospect theory is a theory of how people make decisions when faced with risk and uncertainty. It replaces the conventional risk-averse / risk-seeking decreasing marginal utility theory based on total wealth with a concept of value defined in terms of gains and losses relative to a reference point. This generates utility curves with a point of inflexion at the chosen reference point.

Prospect theory attempts to explain why people may make asymmetric choices when faced with similar possible gains and losses. It was originally developed by Kahneman and Tversky in their paper:

- Kahneman, D and Tversky, A (1979), Prospect theory: an analysis of decision under risk, *Econometrica 47*.

In this paper, Kahneman and Tversky describe an experiment in which they asked people to choose between two alternatives:

- Alternative 1: an 80% chance of winning \$4,000 and a 20% chance of winning nothing
- Alternative 2: a 100% chance of winning \$3,000.

Although the first alternative offers higher expected winnings (\$3,200 *v* \$3,000 for certain), 80% of people chose Alternative 2. This choice is consistent with the assumption of *risk aversion* that underpins expected utility theory. A risk-averse person may prefer a more certain outcome, even if the expected gains are lower (because the additional value derived from the extra certainty outweighs the additional value of the higher possible return.

The same people were then offered the following choice:

- Alternative 3: an 80% chance of losing \$4,000 and a 20% chance of losing nothing
- Alternative 4: a 100% chance of losing \$3,000.

Here 92% of people chose Alternative 3, even though the expected losses are greater (expected losses of \$3,200 *v* a certain loss of \$3,000). This evidence suggests that rather than being risk-averse, people may actually become *risk-seeking* when facing losses.

Prospect theory suggests that:

- Value is based on gains and losses relative to some *reference point.*

- The loss in value from a loss is typically twice as much as the gain in value from the same size monetary gain.

- People are typically *risk-averse* when considering gains relative to the reference point and *risk-seeking* when considering losses relative to the reference point.

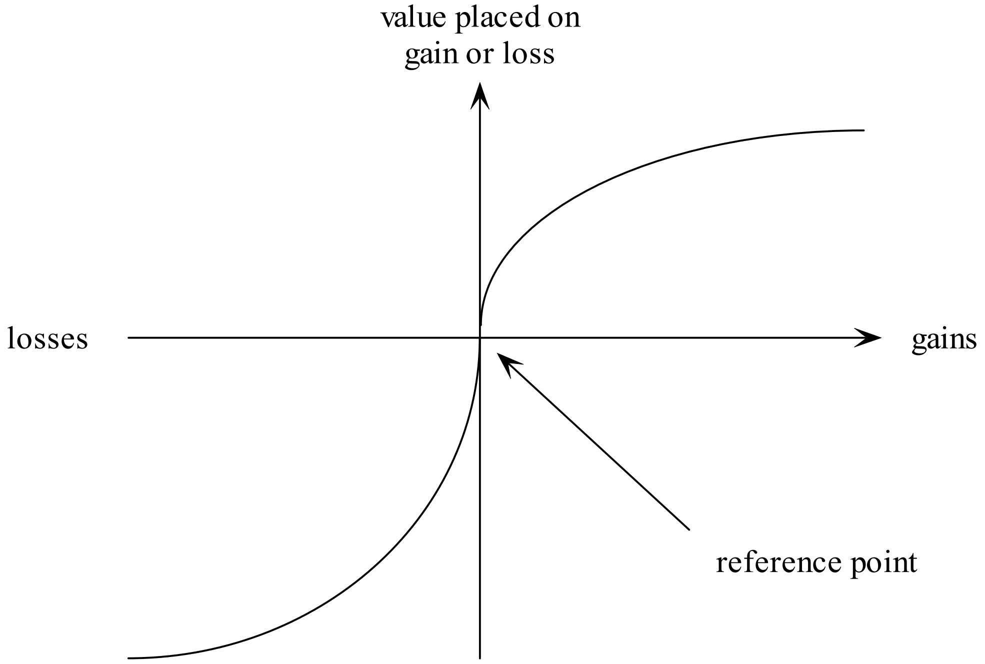


Figure 3.2: Prospect theory value function

Prospect theory suggests that the decision made depends on how a problem is presented or "framed", *ie* whether the available choices are presented as gains or losses relative to the chosen reference point. If the alternative choices are presented as possible gains, then the value function is concave, reflecting the risk-averse nature of individuals. Conversely if they are presented as possible losses, then the value function is convex reflecting the risk-seeking nature of individuals. So, the value function will be as shown on the diagram above.

Prospect theory is therefore associated with the concept of framing.

3.4 Framing (and question wording)

The way a choice is presented ("framed") and, particularly, the wording of a question in terms of gains and losses, can have an enormous impact on the answer given or the decision made. Changes in the way a question is framed of only a word or two can have a profound effect.

Question 3.7

Which of the following two alternatives would you prefer?

1. $30 for certain plus a 50% chance of losing $20
2. $10 for certain plus a 50% chance of winning $20

In the same way, "structured response" questions are found to convey an implicit range of acceptable answers.

In the book:

- Plous S, (1993), *The psychology of judgement and decision making*, McGraw-Hill Inc

Plous describes an experiment in which people were asked the following questions about the length of a film (the same one) they had all recently watched:

- Question 1: How long was the movie?
- Question 2: How short was the movie?

The mean answer to the first question was 2 hours and 10 minutes, whereas that to the second question was 1 hour and 40 minutes!

3.5 Myopic loss aversion

This is similar to prospect theory, but considers repeated choices rather than a single "gamble". It may therefore be relevant when considering investment choices, which can often be thought of as a series of repeated gambles. For example, if an investor reviews and possibly changes its investment strategy on an annual basis, then the investment strategy decision can be thought of as a series of repeated one-year gambles.

Research suggests that investors are less "risk-averse" when faced with a multi-period series of "gambles", and that the frequency of choice / length of reporting period will also be influential.

As its name suggests, myopic loss aversion relates to investors' aversion to short-term losses. The basic idea is that investors have been shown to be less "risk-averse" when faced with a repeated series of "gambles" than when faced with a single gamble.

Thus, if the investor recognises that the investment strategy decision is in fact a series of repeated short-term gambles and consequently takes a long-term view when determining strategy, then they are likely to be less risk-averse than if they instead consider only the immediate short-term gamble and so take too short-term a view. In this latter case, they will tend to focus more on the short-term risk of loss than is necessarily in their best interests, the consequence being that their resulting portfolio ends up being overweight in less risky assets.

In addition, the extent of investors' short-sightedness may be influenced by the frequency with which they review their investment choices and/or the length of the reporting period. For example, a pension fund that has to report its financial position every year may be more averse to very short-term investment losses than one that has to report only every three years.

Question 3.8

What might be the consequence on its investment strategy of requiring a pension fund to report its financial position annually rather than triennially?

3.6 Estimating probabilities

In many situations, investors need to estimate probabilities, for example, in order to assess the riskiness of a particular security. However, research again suggests that investors' estimates may be influenced by a number of biases.

Issues (other than anchoring) which might affect probability estimates include:

- ***Dislike of "negative" events* – the "valence" of an outcome (the degree to which it is considered as negative or positive) has an enormous influence on the probability estimates of its likely occurrence.**

  In particular, experiments suggest that individuals are prone to underestimate the probability that a negative event may occur.

- ***Representative heuristics* – people find more probable that which they find easier to imagine. As the amount of detail increases, its *apparent* likelihood may increase (although the *true* probability can only decrease steadily).**

  As the amount of detail increases, the more specific (less generalised) the event becomes and the less probable the occurrence of such a specific event must become.

- ***Availability* – people are influenced by the ease with which something can be brought to mind. This can lead to biased judgements when examples of one event are inherently more difficult to imagine than examples of another.**

  An example here concerns the incidences of deaths due to car crashes and cancer. When asked to estimate the relative numbers of deaths due to each, people tend to overestimate the number of deaths due to car crashes perhaps because they receive more publicity and are easier to imagine.

3.7 Overconfidence

People tend to overestimate their own abilities, knowledge and skills. For example, if you ask 100 people if they are better than average drivers, then you might not be surprised if more than 50% of them reply "yes"!

Question 3.9

Do you think you will pass the Subject CT8 exam?

Moreover, studies show that the discrepancy between accuracy and overconfidence increases (in all but the simplest tasks) as the respondent is more knowledgeable! (Accuracy increases to a modest degree but confidence increases to a much larger degree.) If this is true then it may not be wise to pass the Subject CT8 exam!

Overconfidence could therefore be a potentially serious problem in fields such as investment where most of the participants are likely to be highly knowledgeable. Moreover, the available evidence suggests that even when people are aware that they are overconfident they remain so.

This may be a result of:

- ***Hindsight bias* – events that happen will be thought of as having been predictable prior to the event; events that do not happen will be thought of as having been unlikely prior to the event.**

  A possible example of the first type of event is the credit crunch of 2007/09. A possible example of the second type of event is when an underdog is heavily beaten in a sporting event. Although supporters may have had high hopes of an upset prior to the event, after the event a heavy defeat will always have seemed inevitable.

- ***Confirmation bias* – people will tend to look for evidence that confirms their point of view (and will tend to dismiss evidence that does not justify it).**

  For example, failing an exam might reinforce an existing preconception that the exam system is a lottery – even if the real reason for failure was a lack of preparation. In contrast, passing an exam might reinforce a view of the exam system as a lottery, even if it was in fact due reward for months of long, hard study!

3.8 Mental accounting

People show a tendency to *separate* related events and decisions and find it difficult to *aggregate* events. Thus, rather than netting out all gains and losses (as standard financial economic theory would suggest) **people set up a series of "mental accounts" and view individual decisions as relating to one or another of these accounts.**

Question 3.10

Consider the following scenarios:

- Scenario 1: You lose \$10 today and find a \$10 note on the street tomorrow.
- Scenario 2: Neither of the above events occurs.

Which scenario would you prefer?

An example of when mental accounting might be of relevance relates to the paying off of a mortgage early. Consider the following two situations:

1. have a 100*k* mortgage and 10*k* of savings in the bank
2. use 5*k* of savings to pay off part of your mortgage, so that you end up with a 95*k* mortgage and 5*k* of savings in the bank.

In theory (*ie* ignoring issues such as tax and the interest rates applicable) you should be indifferent between these two situations because either way you have a *net* indebtedness of 90*k* (and somewhere to live!).

In practice, however, this may not be the case. Some people may actually feel happier in the second situation because they *feel* that a smaller mortgage means that they have less debt. This will be the case if they use separate mental accounts for savings and debt.

In practice, of course, the choice of course of action will be driven to some extent by issues such as tax, the interest rates applicable and the need to hold some liquid assets.

3.9 Effect of options

Experimental evidence also suggests that the range of options or choices presented to people may influence their choices.

In addition to the "framing" effect discussed in Section 3.4 above, other issues include:

- **the *primary effect* – people are more likely to choose the first option presented**

- **the *recency effect* – in some instances, the final option that is discussed may be preferred! The gap in time between the presentation of the options and the decision may influence this dichotomy.**

  More specifically, the sooner/later the decision is made, the more likely it is that the first/last option will be chosen.

- **other research suggesting that people are more likely to choose an intermediate option than one at either end!**

- **a greater range of options tends to discourage decision-making. On the other hand, a higher probability is attributed to options explicitly stated than when included in a broader category.**

- ***status quo bias* – people have a marked preference for keeping things as they are.**

- ***regret aversion* – by retaining the existing arrangements, people minimise the possibility of regret (the pain associated with feeling responsible for a loss).**

- ***ambiguity aversion* – people are prepared to pay a premium for rules.**

  Much standard financial economic theory (such as portfolio theory, which is the subject of Chapter 5) assumes that investors know the actual *distribution* of future investment returns when making their investment choices – although they don't know what the actual investment return will turn out to be in any future period.

The terms *ambiguity* and *uncertainty* are both used to refer to a situation in which the investor is uncertain about the distribution of future investment returns. An investor who dislikes such ambiguity will tend to err on the side of caution when making an investment choice. The investor will also be prepared to pay for further information that reduces the degree of uncertainty faced.

Q

Question 3.11

List the eight common themes found in research on behavioural finance.

3.10 Further reading

The material in this section is based on the paper:

- Nigel Taylor (2000), *Making actuaries less human: lessons from behavioural finance*, Staple Inn Actuarial Society (SIAS).

This is an easy-to-read paper, which is well worth studying if you are interested in these ideas. It can be downloaded from the SIAS website: **www.sias.org.uk**.

4 Exam-style question

We finish this chapter with an exam-style question on stochastic dominance.

Question

(i) Define first-order and second-order stochastic dominance. Illustrate the definitions by sketching cumulative distribution functions of two random variables which represent the returns on two investments, one of which dominates the other. [5]

(ii) Explain how an investor's economic characteristics will affect his choice of an investment that:

(a) first-order stochastically dominates another

(b) second-order stochastically dominates another. [3]

[Total 8]

Solution

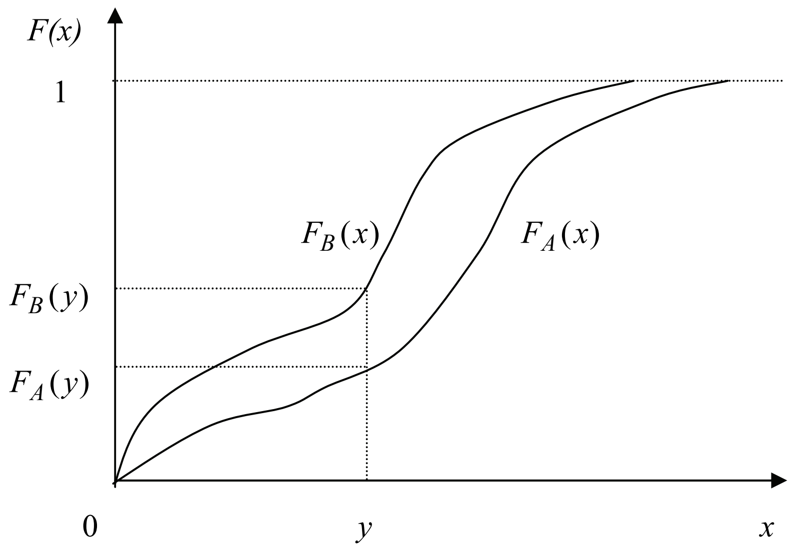
This is Subject 109, September 2000, Question 9.

(i) ***Definitions***

A distribution of investment returns A is said to exhibit *first-order stochastic dominance* over a distribution of investment returns B if:

$$F_A(x) \le F_B(x) \qquad \text{for all } x\text{, and}$$
$$F_A(x) < F_B(x) \qquad \text{for some value of } x.$$

In other words, the probability of B producing a return below a certain value is never less than the probability of A producing a return below the same value and exceeds it for at least some value of x.

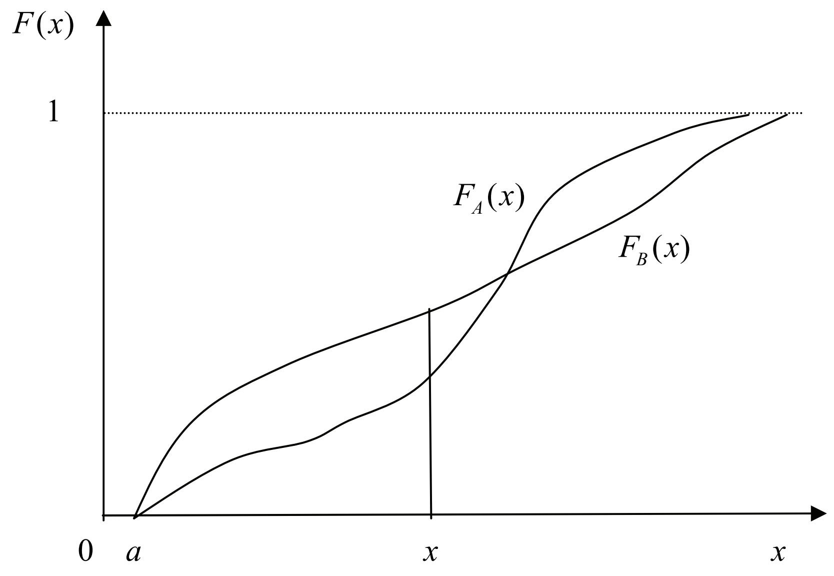


Here $F_A(x)$ must never be above $F_B(x)$. Note, however, that the lowest possible value of x may be non-zero and even negative.

A distribution of investment returns A is said to exhibit *second-order stochastic dominance* over a distribution of investment returns B if:

$$\int_a^x F_A(y)\,dy \le \int_a^x F_B(y)\,dy$$

for all x, with the strict inequality holding for some value of x, where a is the lowest return that the portfolios can possibly provide.



Here the area under $F_B(x)$ must never be less than the area under $F_A(x)$ for any value of x.

(ii) ***Economic characteristics***

(a) An investor who prefers more to less – *ie* who is not satiated – will always choose an investment whose distribution of returns first-order stochastically dominates another.

(b) An investor who prefers more to less and is also risk-averse will always choose an investment whose distribution of returns second-order stochastically dominates another.

Chapter 3 Summary

Stochastic dominance

Stochastic dominance offers an approach to modelling choices under uncertainty that does not require the use of explicit utility functions.

Given two investment portfolios, A and B, with cumulative probability distribution functions of returns F_A and F_B respectively:

The *first-order stochastic dominance theorem* states that A will be preferred to B if:

- the investor prefers more to less, $U'(x) > 0$ and
- $F_A(x) \le F_B(x)$ for all x, with $F_A(x) < F_B(x)$ for at least one x.

The *second-order stochastic dominance theorem* states that A will be preferred to B if:

- the investor prefers more to less, $U'(x) > 0$,
- he is risk-averse, $U''(x) < 0$, and
- $\int_a^x F_A(y)\,dy \le \int_a^x F_B(y)\,dy$, for all x, with the strict inequality holding for at least one x.

Behavioural finance

The field of behavioural finance relates to the psychology that underlies and drives financial decision-making behaviour. Eight common themes include:

1. anchoring and adjustment
2. prospect theory
3. framing (and question wording)
4. myopic loss aversion
5. estimating probabilities
6. overconfidence
7. mental accounting
8. effect of options.

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

Chapter 3 Solutions

Solution 3.1

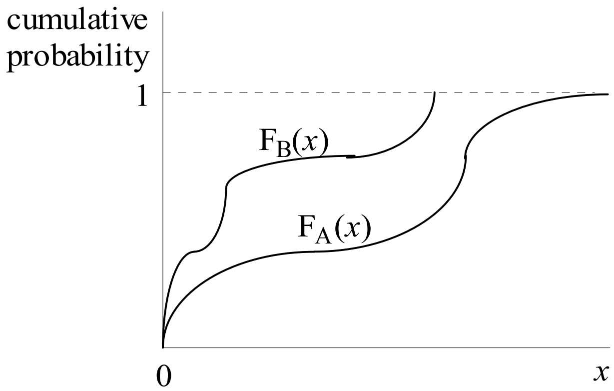


Figure 3.3 – First-order stochastic dominance

The important points to note are that:

- $F_A(x)$ and $F_B(x)$ are both monotonically increasing functions of x
- $F_A(x)$ is never above (to the left of) $F_B(x)$.

Solution 3.2

A can be obtained from B by shifting a probability of ¼ from the $1 outcome to the higher $2 outcome. Consequently A dominates B.

C can be obtained from A by shifting ¾ from $2 to $3 and so C dominates A. C also dominates B because first-order stochastic dominance is *transitive*.

These results (C > A > B) are confirmed by probabilities shown in the table below.

return	*probability*			*cumulative probability*		
	A	*B*	*C*	*A*	*B*	*C*
$1	¼	½	¼	¼	½	¼
$2	¾	½	0	1	1	¼
$3	0	0	¾	1	1	1

Solution 3.3

The cumulative probabilities are as follows.

	cumulative probability	
return	$F_U(x)$	$F_V(x)$
6%	¼	0
7%	½	¾
8%	¾	¾
9%	1	1

The choice between Assets U and V cannot be based upon first-order stochastic dominance alone, because neither asset dominates the other, *ie*:

$F_U(6) > F_V(6)$, but $F_U(7) < F_V(7)$.

Alternatively, you could say that neither asset first-order stochastically dominates the other because we cannot obtain U from V by shifting probability from lower to higher outcome levels; nor can we obtain V from U by shifting probability from lower to higher outcome levels.

Solution 3.4

Assets U and V both offer the same chance of a ***9*** % return. In addition, U offers a wider spread of returns about ***7½*** % – *ie* a greater chance of ***an 8*** % return, but at the risk of a greater chance of obtaining only ***6*** %. Assets U and V therefore both offer the same expected return of ***7½*** %, but the variance of return is greater for Asset ***U***. Thus, a risk-averse investor should choose Asset ***V***.

Solution 3.5

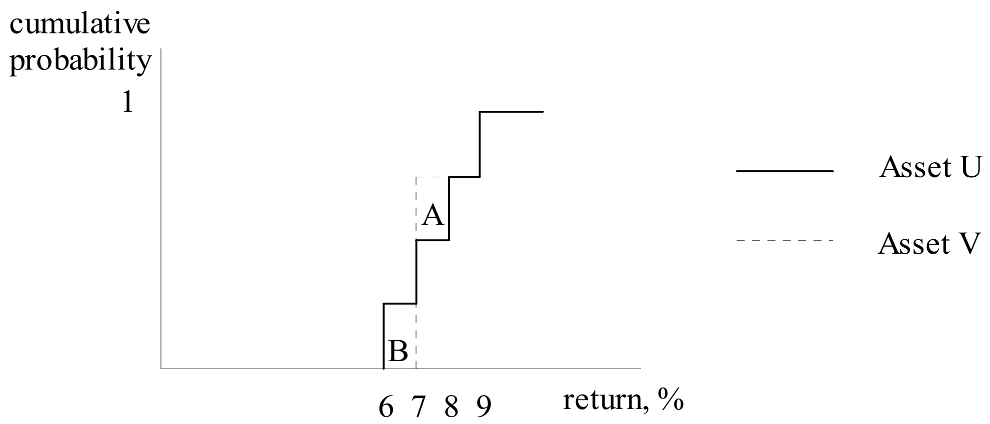


Figure 3.4 – Assessing the stochastic dominance of Assets U and V

In this case, first-order stochastic dominance is insufficient to choose between the assets, because the cumulative probability graphs cross. U is second-order stochastically dominated by V because the extra possibility of obtaining 8% (represented by the box marked A) is of less value to the investor than the possibility of avoiding 6% (represented by Box B).

Solution 3.6

The main advantage of stochastic dominance is that it does not require explicit formulation of the investor's utility function, but can instead be used to make investment decisions for a wide range of utility functions.

The main disadvantages are that it:

- may be unable to choose between two investments and
- generally involves pair-wise comparisons of alternative investments, which may be problematic if there is a large number of investments between which to choose.

Solution 3.7

The two alternatives are of course identical in that they both offer an equal chance of winning either \$30 or \$10. However, experimental evidence based on similar choices suggests that people can and do view identical alternatives differently depending on how they are framed or worded. Thus, the proportion of people selecting each choice typically differs greatly from 50%.

Solution 3.8

The consequence of requiring a pension fund to report its financial position annually rather than triennially might be to force the fund to invest in less volatile assets in order to reduce the risk of having to report a poor financial position. It might therefore force the fund to invest less heavily in equities, and possibly also less heavily in long-term bonds.

This assumes that assets are valued using market values.

Solution 3.9

Only you will know how you answered this. However your answer probably reflected behavioural issues as much as an objective assessment of your true chances. This is not an entirely satisfactory example as some people may in fact be under-confident as regards their chances of passing the exam. Either way, it is quite possible that the percentage of you answering "yes" is very different to the 50-60% or so which may turn out to be the case in practice!

Solution 3.10

Standard financial economic theory suggests that you should be indifferent between the two outcomes because they both result in no net gain or loss. In practice, however, you might have chosen one or the other. When asked about scenarios of this nature in experiments, there is typically a very strong majority preference for one or other of the scenarios described.

Solution 3.11

The eight common themes found in research on behavioural finance are:

1. anchoring and adjustment
2. prospect theory
3. framing (and question wording)
4. myopic loss aversion
5. estimating probabilities
6. overconfidence
7. mental accounting
8. effect of options.

Chapter 4

Measures of investment risk

Syllabus objectives

(ii) Discuss the advantages and disadvantages of different measures of investment risk.

1. Define the following measures of investment risk:
- *variance of return*
- *downside semi-variance of return*
- *shortfall probabilities*
- *Value at Risk (VaR) / Tail VaR*

2. Describe how the risk measures listed in (ii)1 above are related to the form of an investor's utility function.

3. Perform calculations using the risk measures listed above to compare investment opportunities.

4. Explain how the distribution of returns and the thickness of tails will influence the assessment of risk.

0 Introduction

In financial economics, it is often assumed that the key factors influencing investment decisions are "risk" and "return". In practice, return is almost always interpreted as the *expected* investment return. However, there are many possible interpretations and different ways of measuring investment risk, of which the variance is just one, each of which corresponds to a different utility function.

This chapter outlines a small number of such measures, together with their relative merits.

1 Measures of risk

1.1 Introduction

Most mathematical investment theories of investment risk use variance of return as the measure of risk. Examples include (mean-variance) portfolio theory and the capital asset pricing model, both of which are discussed later in this course.

However, it is not obvious that variance necessarily corresponds to investors' perception of risk and other measures have been proposed as being more appropriate.

Some investors might not be concerned with the mean and variance of returns, but simpler things such as the maximum possible loss. Alternatively, some investors might be concerned not only with the mean and variance of returns, but also more generally with other higher moments of returns, such as the *skewness* of returns. For example, although two risky assets might yield the same expectation and variance of future returns, if the returns on Asset A are positively skewed, whilst those on Asset B are symmetrical about the mean, then Asset A might be preferred to Asset B by some investors.

1.2 Variance of return

For a continuous distribution, **variance of return is defined as:**

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

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

"Return" here means the proportionate increase in the market value of the asset, *eg* $x = 0.05$ if the asset has increased by 5% over the period.

The units of variance are "%%", which means "per 100 per 100".

eg $(4\%)^2 = 16\%\% = 0.16\% = 0.0016$

Question 4.1 (CT3 revision)

Investment returns (% *pa*), X, on a particular asset are modelled using a probability distribution with density function:

$$f(x) = 0.00075\left(100 - (x-5)^2\right) \qquad \text{where } -5 \le x \le 15$$

Calculate the mean return and the variance of return.

For a discrete distribution, variance of return is defined as:

$$\sum_x (\mu - x)^2 P(X = x)$$

where μ is the mean return at the end of the chosen period.

Question 4.2 (CT3 revision)

Investment returns (% *pa*), X, on a particular asset are modelled using the probability distribution:

X	*probability*
-7	0.04
5.5	0.96

Calculate the mean return and variance of return.

Variance has the advantage over most other measures in that it is mathematically tractable and the mean-variance framework discussed in Chapter 5 below leads to elegant solutions for optimal portfolios. This ease of use should not be lightly disregarded and to justify using a more complicated measure it would have to be shown that it was both more theoretically correct and that it lead to significantly different choices than the use of variance. In fact the use of mean-variance theory has been shown to give a good approximation to several other proposed methodologies.

Mean-variance portfolio theory can be shown to lead to optimum portfolios if investors can be assumed to have quadratic utility functions or if returns can be assumed to be normally distributed.

In an earlier chapter we discussed how the aim of investors is to maximise their expected utility. The mean-variance portfolio theory discussed in the next chapter assumes that investors base their investment decisions solely on the mean and variance of investment returns. This assumption is consistent the maximisation of expected utility provided that the investor's expected utility depends only the mean and variance of investment returns. It can be shown that this is the case if:

- the investor has a quadratic utility function, and/or
- investment returns follow a distribution that is characterised fully by its first two moments, such as the normal distribution.

If, however, neither of these conditions holds, then we cannot assume that investors make choices solely on the basis of the mean and variance of return. For example, with more complex utility functions and non-normal return distributions investors may need to consider other features of the distribution of returns, such as skewness and kurtosis.

Question 4.3 (CT3 revision)

Define both the skewness and the fourth central moment (called the *kurtosis*) of a continuous probability distribution.

1.3 Semi-variance of return

The main argument against the use of variance as a measure of risk is that most investors do not dislike uncertainty of returns as such; rather they dislike the possibility of low returns.

For example, all investors would choose a security that offered a chance of either a 10% or 12% return in preference to one that offered a certain 10%, despite the greater uncertainty associated with the former.

One measure that seeks to quantify this view is downside semi-variance (also referred to as simply semi-variance). For a continuous random variable, **this is defined as:**

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

For a discrete random variable, the downside semi-variance is defined as:

$$\sum_{x<\mu} (\mu - x)^2 P(X = x)$$

Semi-variance is not easy to handle mathematically and it takes no account of variability above the mean. Furthermore if returns on assets are symmetrically distributed semi-variance is proportional to variance.

Question 4.4

What is the relationship between semi-variance and variance for the normal distribution?

Question 4.5

Calculate the downside semi-variance of return for the asset modelled in Question 4.1.

Question 4.6

Calculate the downside semi-variance of return for the asset modelled in Question 4.2.

1.4 Shortfall probabilities

A shortfall probability measures the probability of returns falling below a certain level. For continuous variables, the risk measure is given by:

$$\textbf{Shortfall probability} = \int_{-\infty}^{L} f(x)\, dx$$

where L is a chosen benchmark level.

For discrete random variables, the risk measure is given by:

$$\text{Shortfall probability} = \sum_{x<L} P(X = x)$$

The benchmark level can be expressed as the return on a benchmark fund if this is more appropriate than an absolute level. In fact any of the risk measures discussed can be expressed as measures of the risk relative to a suitable benchmark which may be an index, a median fund or some level of inflation.

L could alternatively relate to some pre-specified level of surplus or fund solvency.

The main advantages of the shortfall probability are that it is easy to understand and calculate.

Question 4.7

Calculate the shortfall probability for the asset modelled in Question 4.1 where the benchmark return is 0% *pa*.

Question 4.8

Calculate the shortfall probability for the asset modelled in Question 4.2 where the benchmark return is 0% *pa*.

Question 4.9

What is the main drawback of the shortfall probability as a measure of investment risk?

1.5 Value at risk

***Value at Risk (VaR)* generalises the likelihood of under-performing by providing a statistical measure of downside risk.**

For a continuous random variable, **Value at Risk can be determined as:**

$$VaR(X) = -t \quad \text{where} \quad P(X < t) = p$$

VaR assesses the potential losses on a portfolio over a given future time period with a given degree of confidence.

For example, if we adopt a 99% confidence limit, the VaR is the amount of loss that will be exceeded only one time in a hundred over a given time period and we would need to find t such that $P(X < t) = 0.01$.

Note that Value at Risk is a "loss amount". Therefore:

- a positive Value at Risk (a negative t) indicates a loss
- a negative Value at Risk (a positive t) indicates a profit
- Value at Risk should be expressed as a *monetary* amount and not as a percentage.

Question 4.10

Calculate the VaR over one year with a 95% confidence limit for a portfolio consisting of £100*m* invested in the asset modelled in Question 4.1.

For a discrete random variable, VaR is defined as:

$$VaR(X) = -t \quad \text{where } t = \max\left\{x : P(X < x) \leq p\right\}$$

Question 4.11

Calculate the 95% VaR over one year with a 95% confidence limit for a portfolio consisting of £100*m* invested in the asset modelled in Question 4.2.

Question 4.12

Calculate the 97.5% VaR over one year for a portfolio consisting of £200*m* invested in shares. You should assume that the return on the portfolio of shares is normally distributed with mean 8% *pa* and standard deviation 8% *pa*.

VaR can be measured either in absolute terms or relative to a benchmark. Again, VaR is based on assumptions that may not be immediately apparent.

The problem is that in practice VaR is usually calculated assuming that investment returns are normally distributed.

Portfolios exposed to credit risk, systematic bias or derivatives may exhibit non-normal distributions. The usefulness of VaR in these situations depends on modelling skewed or fat-tailed distributions of returns, either in the form of statistical distributions (such as the Gumbel, Frechet or Weibull distributions) or via Monte Carlo simulations. However, the further one gets out into the "tails" of the distributions, the more lacking the data and, hence, the more arbitrary the choice of the underlying probability becomes.

Hedge funds are a good example of portfolios exposed to credit risk, systematic bias and derivatives. These are private collective investment vehicles that often adopt complex and unusual investment positions in order to make high investment returns. For example, they will often short-sell securities and use derivatives.

If the portfolio in Question 4.12 was a hedge fund then modelling the return using a normal distribution may no longer be appropriate. A different distribution could be used to assess the lower tail but choosing this distribution will depend on the data available for how hedge funds have performed in the past. This data may be lacking or include survivorship bias, *ie* hedge funds that do very badly may not be included.

The Gumbel, Frechet and Weibull distributions are three examples of extreme value distributions, which are used to model extreme events.

The main weakness of VaR is that it does not quantify the size of the "tail". Another useful measure of investment risk therefore is the *Tail Value at Risk*.

1.6 Tail value at risk (TailVar) and expected shortfall

Closely related to both shortfall probabilities and VaR are the TailVaR and Expected Shortfall measures of risk.

The risk measure can be expressed as the expected shortfall below a certain level.

For a continuous random variable, the expected shortfall is given by:

$$\textbf{Expected shortfall} = E\left[\max(L - X, 0)\right] = \int_{-\infty}^{L} (L - x) f(x)\, dx$$

where L is the chosen benchmark level.

For a discrete random variable, the expected shortfall is given by:

$$\text{Expected shortfall} = E\left[\max(L - X, 0)\right] = \sum_{x<L} (L - x) P(X = x)$$

If L is chosen to be a particular percentile point on the distribution, then the risk measure is known as the TailVaR.

The $(1-p)$ TailVaR is the expected shortfall in the p^{th} lower tail. So, for the 99% confidence limit, it represents the expected loss *in excess of* the 1% lower tail value.

Question 4.13

Calculate the 95% TailVaR over one year for a portfolio consisting of £100*m* invested in the asset modelled in Question 4.1.

Question 4.14

Calculate the 95% TailVaR over one year for a portfolio consisting of £100*m* invested in the asset modelled in Question 4.2.

However, Tail VaR can also be expressed as the Expected Shortfall conditional on there being a shortfall.

To do this, we would need to take the expected shortfall formula and divide by the shortfall probability.

Other similar measures of risk have been called:

- **expected tail loss**
- **tail conditional expectation**
- **conditional VaR**
- **tail conditional VaR**
- **worst conditional expectation.**

They all measure the risk of underperformance against some set criteria. It should be noted that the characteristics of the risk measures may vary depending on whether the variable is discrete or continuous in nature.

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

In other words, for continuous random variables, we could use a measure of the form:

$$\int_{-\infty}^{L} g(L-x)f(x)\,dx$$

Two particular cases of note are when:

1. $g(L-r)=(L-r)^2$ – this is the so-called *shortfall variance*
2. $g(L-r)=(L-r)$ – the *average or expected shortfall* measure defined above.

Note also that if $g(x) = x^2$ and $L = \mu$, then we have the semi-variance measure defined above.

Shortfall measures are useful for monitoring a fund's exposure to risk because the expected underperformance relative to a benchmark is a concept that is apparently easy to understand. As with semi-variance, however, no attention is paid to the distribution of outperformance of the benchmark, *ie* returns in excess of L are again completely ignored**.**

Question 4.15

Consider an investment whose returns follow a continuous uniform distribution over the range 0% to 10% *pa*.

(i) Write down the probability density function for the investment returns.

(ii) What is the mean investment return?

(iii) Calculate the variance and semi-variance measures of investment risk.

(iv) Calculate the shortfall probability and the expected shortfall based on a benchmark level of 3% *pa*.

Question 4.16

Consider two investments: Investment A offers an investment return of either –3%, +2% or +7% *pa*, whereas Investment B offers –5%, +2% or +3% *pa*. In each case the probability of each possible outcome is one third.

Which is the riskier investment according to each of the following measures of investment risk?

- variance
- shortfall probability (with a benchmark of 0% *pa*)
- expected shortfall (with a benchmark of 0% *pa*)

2 Relationship between risk measures and utility functions

An investor using a particular risk measure will base his decisions on a consideration of the available combinations of risk and expected return. Given a knowledge of how this trade off is made it is possible, in principle, to construct the investor's underlying utility function. Conversely, given a particular utility function, the appropriate risk measure can be determined.

For example, if an investor has a quadratic utility function, the function to be maximised in applying the expected utility theorem will involve a linear combination of the first two moments of the distribution of return. Thus variance of return is an appropriate measure of risk in this case.

Question 4.17

(i) State the expected utility theorem.

(ii) Draw a typical utility function for a non-satiated, risk-averse investor.

(iii) Explain the last two paragraphs of Core Reading.

If expected return and semi-variance below the expected return are used as the basis of investment decisions, it can be shown that this is equivalent to a utility function that is quadratic below the expected return and linear above.

Thus, this is equivalent to the investor being risk-averse below the expected return and risk-neutral for investment return levels above the expected return. Hence, no weighting is given to variability of investment returns above the expected return.

Use of a shortfall risk measure corresponds to a utility function that has a discontinuity at the minimum required return. This therefore corresponds to the state-dependent utility functions discussed in Chapter 2.

Question 4.18

What is meant by a state-dependent utility function?

Question 4.19

What are the advantages of variance of return as the definition of risk compared with semi-variance and shortfall probabilities? What is the main disadvantage?

3 Exam-style questions

We finish this chapter with two exam-style questions.

Question 1

Define the following measures of investment risk:

(i) variance of return [2]

(ii) downside semi-variance of return [2]

(iii) shortfall probability. [2]

(iv) value at risk [1]

[Total 7]

Solution 1

Most of this question is taken from Subject 109 April 2000 Question 3.

In the following we assume an investment return is given by a continuous random variable X with density function $f(x)$. This is the return over a chosen time period. Analogous formulae could be given for discrete or mixed cases.

(i) ***Variance***

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

where $\mu = E[X]$ is the mean return for the chosen period.

(ii) ***Downside semi-variance***

The downside semi-variance only takes into account returns below the mean return:

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

(iii) ***Shortfall probability***

A shortfall probability measures the probability of returns falling below a certain chosen benchmark level L:

$$P(X < L) = \int_{-\infty}^{L} f(x)\,dx$$

(iv) ***Value at risk***

Value at risk assesses the potential loss in value on a portfolio over a given future time period with a given degree of confidence.

Alternatively, for a given confidence limit $(1-p)$:

$$VaR(X) = -t \quad \text{where} \quad P(X < t) = p$$

Question 2

This question is taken from Subject CT8 September 2005 Question 1.

An investor is contemplating an investment with a return of £ R , where:

$$R = 300,000 - 500,000U$$

where U is a uniform $[0,1]$ random variable.

Calculate each of the following four measures of risk:

(a) variance of return

(b) downside semi-variance of return

(c) shortfall probability, where the shortfall level is £100,000

(d) Value at Risk at the 5% level. [8]

Solution 2

(a) ***Variance***

R is defined by $R = 300{,}000 - 500{,}000U$, where U is $U[0,1]$. So R has a uniform distribution on the range from –200,000 to 300,000.

The variance of R can be calculated directly from the formula $\frac{1}{12}(b-a)^2$ on page 13 of the *Tables*:

$$\begin{aligned}\text{var}(R) &= \tfrac{1}{12}[300{,}000 - (-200{,}000)]^2 \\ &= \tfrac{1}{12} \times 500{,}000^2 \\ &= (£144{,}338)^2\end{aligned}$$

An alternative approach is to evaluate the integral:

$$\int_{-200{,}000}^{300{,}000} (\mu - x)^2 f(x)dx$$

where $\mu = \frac{1}{2}(-200{,}000 + 300{,}000) = 50{,}000$ *and* $f(x) = \dfrac{1}{500{,}000}$.

(b) ***Downside semi-variance***

Since the uniform distribution is symmetrical, the semi-variance is just half the "full" variance:

$$\text{semi-variance} = \tfrac{1}{2} \times \tfrac{1}{12} \times 500{,}000^2 = (£102{,}062)^2$$

Alternatively, you can evaluate the integral:

$$\int_{-200{,}000}^{50{,}000} (\mu - x)^2 f(x)dx$$

(c) ***Shortfall probability***

The shortfall probability can be evaluated using the formula $\frac{x-a}{b-a}$ on page 13 of the *Tables* for the distribution function of the uniform distribution:

$$P(R < 100,000) = \frac{100,000-(-200,000)}{300,000-(-200,000)} = \frac{300,000}{500,000} = 0.6$$

Alternatively, you can evaluate the integral:

$$\int_{-200,000}^{100,000} f(x)dx$$

(d) ***Value at Risk***

We need to find the (lower) 5% percentile of the distribution of values of R. We can do this using the same formula we used in part (c):

$$\frac{x-a}{b-a} = 0.05$$

ie $$\frac{x-(-200,000)}{500,000} = 0.05$$

$\Rightarrow$ $$x = 0.05 \times 500,000 - 200,000 = -175,000$$

Therefore, the Value at Risk at the 5% level is $-(-175,000) = £175,000$.

Chapter 4 Summary

Measures of investment risk

Many investment models use *variance* of return as the measure of investment risk.

For a continuous random variable: $V = \int_{-\infty}^{\infty} (\mu - x)^2 f(x)\,dx$.

For a discrete random variable: $V = \sum_x (\mu - x)^2 P(X = x)$.

Variance has the advantage over most other measures that it:

- is mathematically tractable
- leads to elegant solutions for optimal portfolios, within the context of mean-variance portfolio theory.

The main argument against the use of variance as a measure of risk is that most investors do not dislike uncertainty of returns as such; rather they dislike the *downside risk* of low investment returns. Consequently, alternative measures of downside risk sometimes used include (in the continuous and then discrete cases):

- semi-variance of return: $\int_{-\infty}^{\mu} (\mu - x)^2 f(x)\,dx$ $\sum_{x<\mu} (\mu - x)^2 P(X = x)$
- shortfall probability: $\int_{-\infty}^{L} f(x)\,dx$ $\sum_{x<L} P(X = x)$

each of which ignores upside risk.

Value at Risk (VaR) generalises the likelihood of under-performing by providing a statistical measure of downside risk.

VaR assesses the potential losses on a portfolio over a given future time period with a given degree of confidence $(1-p)$. It is often calculated assuming that investment returns follow a normal distribution, which may not be an appropriate assumption.

For a continuous random variable, $VaR(X) = -t$, where $P(X < t) = p$.

For a discrete random variable, $VaR(X) = -t$, where $t = \max\{x : P(X < x) \le p\}$.

The expected shortfall, relative to a benchmark L is given by $E[\max(L-X,0)]$.

For a continuous random variable, expected shortfall = $\int_{-\infty}^{L}(L-x)f(x)dx$.

For a discrete random variable, expected shortfall = $\sum_{x<L}(L-x)P(X=x)$.

When L is the VaR with a particular confidence interval, the expected shortfall is known as *TailVaR*. TailVaR measures the expected loss in excess of the VaR.

It is also possible to calculated the expected shortfall and TailVar *conditional* on a shortfall occurring by dividing through by the shortfall probability.

Relationship between risk measures and utility functions

If *expected return* and *variance* are used as the basis of investment decisions, it can be shown that this is equivalent to a quadratic utility function.

If *expected return* and *semi-variance* below the expected return are used as the basis of investment decisions, it can be shown that this is equivalent to a utility function that is quadratic below the expected return and linear above.

Use of a *shortfall risk measure* corresponds to a utility function that has a discontinuity at the minimum required return.

Note that some of the past questions on this topic have required you to evaluate integrals using techniques such as integration by parts.

Chapter 4 Solutions

Solution 4.1

The density function is symmetrical about $x = 5$. Hence the mean return is 5%. Alternatively, this could be found by integrating as follows:

$$E[X] = 0.00075\int_{-5}^{15} x\left(100-(x-5)^2\right)dx$$

$$= 0.00075\int_{-5}^{15} 75x + 10x^2 - x^3 dx$$

$$= 0.00075\left[\frac{75}{2}x^2 + \frac{10}{3}x^3 - \frac{1}{4}x^4\right]_{-5}^{15}$$

$$= 0.00075\left[7031.25 - 364.5833\right]$$

$$= 5$$

ie 5% *pa*.

The variance is given by:

$$\text{var}[X] = 0.00075\int_{-5}^{15} (5-x)^2\left(100-(x-5)^2\right)dx$$

$$= 0.00075\int_{-5}^{15} 100(x-5)^2 - (x-5)^4)dx$$

$$= 0.00075\left[\frac{100}{3}(x-5)^3 - \frac{1}{5}(x-5)^5\right]_{-5}^{15}$$

$$= 0.00075\left[13{,}333.33 - (-13{,}333{,}33)\right]$$

$$= 20$$

ie 20%% *pa*.

Alternatively, you may have calculated the variance using the formula: $\text{var}[X] = E[X^2] - E[X]^2$, where $E[X^2]$ can be found by integration to be 45%%.

Solution 4.2

The mean return is given by:

$$E[X] = -7 \times 0.04 + 5.5 \times 0.96 = 5$$

ie 5% *pa*.

The variance of return is given by:

$$\text{var}[X] = (5-(-7))^2 \times 0.04 + (5-5.5)^2 \times 0.96 = 6$$

ie 6%% *pa*.

Alternatively, you may have calculated the variance using the formula: $\text{var}[X] = E[X^2] - E[X]^2$, where $E[X^2]$ is 31%%.

Solution 4.3

The skewness of a continuous probability distribution is defined as the third central moment:

$$Skew = \int_{-\infty}^{\infty} (x-\mu)^3 f(x)\, dx$$

It is a measure of the extent to which a distribution is asymmetric about its mean. For example, the normal distribution is symmetric about its mean and therefore has zero skewness, whereas the lognormal distribution is positively skewed.

The kurtosis of a continuous probability distribution is defined as the fourth central moment:

$$K = \int_{-\infty}^{\infty} (x-\mu)^4 f(x)\, dx$$

It is a measure of the "peakedness" or "pointedness" of a distribution.

Solution 4.4

The normal distribution is symmetrical. Hence the semi-variance is half of the variance.

Solution 4.5

The continuous distribution in Question 4.1 is symmetrical. Therefore, the downside semi-variance is half the variance, *ie* 10%%.

Solution 4.6

For the discrete distribution in Question 4.2, the downside semi-variance is given by:

$$\sum_{x<5}(5-x)^2 P(X=x)=(5-(-7))^2\times 0.04=5.76$$

ie 5.76%% *pa*.

Solution 4.7

The shortfall probability is given by:

$$\begin{aligned}P(X<0)&=0.00075\int_{-5}^{0}100-(x-5)^2\,dx\\&=0.00075\left[100x-\frac{1}{3}(x-5)^3\right]_{-5}^{0}\\&=0.00075\left[41.6667-\left(-166.6667\right)\right]\\&=0.15625\end{aligned}$$

Solution 4.8

The shortfall probability is given by:

$$P(X<0)=0.04$$

Solution 4.9

The shortfall probability gives no indication of the magnitude of any shortfall (being independent of the extent of any shortfall).

For example, consider two securities that offer the following combinations of returns and associated probabilities:

Investment A: 100% with probability of 0.9 and 9.9% with probability of 0.1

Investment B: 10.1% with probability of 0.91 and 0% with probability of 0.09

An investor who chooses between them purely on the basis of the shortfall probability based upon a benchmark return of 10% would choose Investment B!

Solution 4.10

We start by finding t, where $P(X < t) = 0.05$:

$$\Rightarrow \qquad 0.00075\int_{-5}^{t} 100-(x-5)^2\,dx = 0.05$$

$$\Rightarrow \qquad 0.00075\left[100x-\frac{1}{3}(x-5)^3\right]_{-5}^{t} = 0.05$$

Since the equation in the brackets is a cubic in t, we are going to need to solve this equation numerically, by trial and error.

$$t=-3 \Rightarrow 0.00075\left[100x-\frac{1}{3}(x-5)^3\right]_{-5}^{-3} = 0.028$$

and $$t=-2 \Rightarrow 0.00075\left[100x-\frac{1}{3}(x-5)^3\right]_{-5}^{-2} = 0.06075$$

Interpolating between the two gives $t=-3+\dfrac{0.05-0.028}{0.06075-0.028}=-2.3$.

In fact, the true value is $t=-2.293$. Since t is a percentage investment return per annum, the 95% value at risk over one year on a £100m portfolio is £$100m \times 2.293\% =$ £$2.293m$. This means that, we are 95% certain that we will not lose more than £2.293m over the next year.

Solution 4.11

We start by finding t, where $t = \max\left\{x : P(X < x) \leq 0.05\right\}$.

Now $P(X < -7) = 0$ and $P(X < 5.5) = 0.04$. Therefore $t = 5.5$.

Since t is a percentage investment return per annum, the 95% value at risk over one year on a £100*m* portfolio is $£100m \times -5.5\% = -£5.5m$. This means that, we are 95% certain that we will not make profits of less than £5.5*m* over the next year.

Solution 4.12

We start by finding t, where:

$$P(X < t) = 0.025\text{, where } X \sim N(8, 8^2)$$

Standardising gives:

$$P\left(Z < \frac{t-8}{8}\right) = \Phi\left(\frac{t-8}{8}\right) = 0.025$$

But $\Phi(-1.96) = 0.025$ from page 160 of the *Tables*, so $t = -7.68$.

Since t is a percentage investment return per annum, the 97.5% value at risk over one year on a £200*m* portfolio is $£200m \times 7.68\% = £15.36m$. This means that, we are 97.5% certain that we will not lose more than £15.36*m* over the next year.

Solution 4.13

The expected shortfall in returns below –2.293% is given by:

$$E[\max(-2.293 - X, 0)] = 0.00075\int_{-5}^{-2.293} (-2.293 - x)\left(100 - (x-5)^2\right) dx$$

$$= 0.00075\int_{-5}^{-2.293} \left(-171.975 - 97.93x - 7.707x^2 + x^3\right) dx$$

$$= 0.00075\left[-171.975x - 48.965x^2 - 2.569x^3 + 0.25x^4\right]_{-5}^{-2.293}$$

$$= 0.0462$$

On a portfolio of £100*m*, the 95% TailVaR is $£100m \times 0.000462 = £0.0462m$. This means that the expected loss *in excess of* £2.293*m* is £46,200.

Solution 4.14

The expected shortfall in returns below 5.5% is given by:

$$E[\max(5.5 - X, 0)] = \sum_{x<5.5} (5.5 - x)P(X = x)$$

$$= 12.5 \times 0.04 = 0.5$$

On a portfolio of £100*m*, the 95% TailVaR is $£100m \times 0.005 = £0.5m$. This means that the expected reduction in profits *below* £5.5*m* is £0.5*m*.

Solution 4.15

(i) ***Probability density function***

The probability density function is simply $f(x) = \dfrac{1}{10}$ for $0\% \le x \le 10\%$ and zero otherwise.

We've assumed here that you're working in % units. If you've converted the returns to decimals, $f(x) = 10$.

(ii) ***Mean***

The mean investment return is 5% *pa*.

(iii) ***Variance and semi-variance***

The variance and semi-variance are given by:

$$V = \int_0^{10} \frac{(5-x)^2}{10}\,dx = 8.33\,\%\%\ pa$$

$$SV = \int_0^{5} \frac{(5-x)^2}{10}\,dx = 4.17\,\%\%\ pa$$

(iv) ***Shortfall probability and expected shortfall***

The shortfall probability is given by:

$$SP = \int_0^{3} \frac{1}{10}\,dx = 0.3$$

The expected shortfall is given by:

$$ES = \int_0^{3} \frac{(3-x)}{10}\,dx = 0.45\,\%$$

Solution 4.16

Using the usual formula, the *variances* of investment returns are 16.67%% *pa* and 12.67%% *pa* for A and B respectively.

The *shortfall probability* is in each case equal to one third.

The *expected shortfall* is:

$(0-(-3))\times\frac{1}{3}=1\%\ pa$ for Investment A

and $(0-(-5))\times\frac{1}{3}=\frac{5}{3}\%\ pa$ for Investment B.

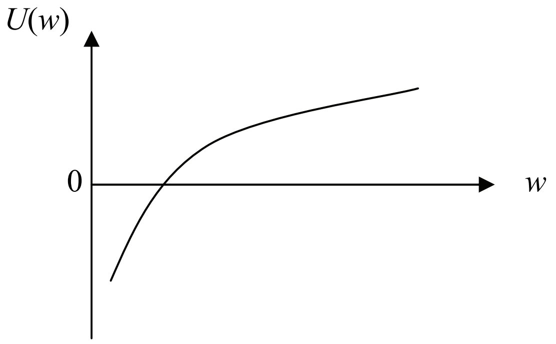
Thus, whereas A is riskier according to the variance, B is riskier according to the expected shortfall and they are equally risky using the shortfall probability as the measure of investment risk.

Solution 4.17

(i) The expected utility theorem states that, assuming some basic conditions apply:

- then an investor can construct a utility function *U(w)* for each level of wealth *w* such that
- the investor will base their investment decisions upon maximising their expected utility.

(ii)



Utility functions are drawn in $(w, U(w))$ space. They show the investor's utility of wealth *U(w)* at various points of possible wealth *w*.

Non-satiated investors prefer more wealth to less and so the graph slopes upwards, *ie* $U'(w) > 0$.

Risk-averse investors have diminishing marginal utility of wealth and so the slope of the graph decreases with *w*, *ie* $U''(w) < 0$.

(iii) An investor's utility function describes their attitude towards risk and return and as such can be translated into other ways of expressing this attitude. For example, if an investor has a quadratic utility function then their attitude towards risk and return can be expressed purely in terms of the mean and variance of investment opportunities. For this investor, it would be possible to calculate indifference curves in expected return-standard deviation space. Such indifference curves are discussed further in Chapter 5.

Solution 4.18

Sometimes it may be inappropriate to model an investor's behaviour over all possible levels of wealth with a utility function with a single formula. This problem can be overcome by using *state-dependent* utility functions, which model the situation where there is a discontinuous change in the state of the investor at a certain level of wealth.

Solution 4.19

Advantages

- Mathematical manipulation is easier using variance of return.
- The variance is a familiar summary statistic.
- It has not been shown that other measures of risk give "better" results.
- Variance is the measure of investment risk used in mean-variance portfolio theory (see Chapter 5).

Main disadvantage

Arguably, it is the downside risk, rather than the total uncertainty, that is the more important measure of risk. Hence measures that concentrate more on the "bad" outcomes, such as semi-variance or shortfall probability, may be more appropriate.

Chapter 5

Portfolio theory

Syllabus objectives

(iii) Describe and discuss the assumptions of mean-variance portfolio theory and its principal results.

1. *Describe and discuss the assumptions of mean-variance portfolio theory.*

2. *Discuss the conditions under which application of mean-variance portfolio theory leads to the selection of an optimum portfolio.*

3. *Calculate the expected return and risk of a portfolio of many risky assets, given the expected return, variance and covariance of returns of the individual assets, using mean-variance portfolio theory.*

4. *Explain the benefits of diversification using mean-variance portfolio theory.*

0 Introduction

This chapter introduces *mean-variance portfolio theory (also called modern portfolio theory*). As well as being very important in its own right, mean-variance portfolio theory (MPT) forms the basis of the capital asset pricing model discussed in Chapter 7. A very detailed description of mean-variance portfolio theory is given in the book by Elton & Gruber listed in the Study Guide.

MPT assumes that investment decisions are based solely upon risk and return – more specifically the mean and variance of investment return – and that investors are willing to accept higher risk in exchange for higher expected return. This can be consistent with the maximisation of expected utility as outlined in Chapter 2, if the investor is assumed to have a utility function that only uses the mean and variance of investment returns, such as the quadratic utility function. It can also be consistent if the distribution of investment returns is a function only of its mean and variance.

Based upon these and other assumptions, MPT specifies a method for an investor to construct a portfolio that gives the maximum return for a specified risk (variance), or the minimum risk for a specified return. Such portfolios are described as *efficient*.

Section 1 of the chapter focuses upon the means by which to determine:

- the *opportunity set* of available portfolios and the corresponding risk-return combinations between which the investor must choose from and

- the set of efficient portfolios, the *efficient frontier*.

A rational investor who prefers more to less and is risk-averse will always choose an efficient portfolio.

Section 2 introduces the benefits of diversification and shows that if the returns from all the assets in a portfolio are statistically independent then the variance of the portfolio return tends towards the market variance when the number of assets in that portfolio is increased.

Note carefully that:

- Figures 5.1 to 5.3 included within this chapter are all part of the Core Reading for this topic.

- Within the context of mean-variance portfolio theory, risk is defined very specifically as the *variance* – or equivalently standard deviation – of investment returns. Elsewhere in this and other courses we discuss other possible measures or types of risk that might be relevant depending upon the exact context considered.

Question 5.1

Can you think of three other possible types of risk that might be relevant in an investment context?

1 Portfolio theory

1.1 Introduction

Mean-variance portfolio theory, sometimes called modern portfolio theory (MPT), specifies a method for an investor to construct a portfolio that gives the maximum return for a specified risk, or the minimum risk for a specified return.

If the investor's utility function is known, then MPT allows the investor to choose the portfolio that has the optimal balance between return and risk, as measured by the variance of return, and consequently maximises the investor's expected utility.

However, the theory relies on some strong and limiting assumptions about the properties of portfolios that are important to investors. In the form described here the theory ignores the investor's liabilities although it is possible to extend the analysis to include them.

By ignoring actuarial risk – the risk that the investor fails to meet his or her liabilities – the theory as presented here is severely limited. In its defence, MPT was the first real attempt to use statistical techniques to show the benefit of diversification for investors. The extension to include reference to the investor's liabilities is discussed in a later subject.

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

1. **First the definition of the properties of the portfolios available to the investor – the opportunity set.** Here we are looking at the risk and return of the possible portfolios available to the investor.

2. **Second, the determination of how the investor chooses one out of all the feasible portfolios in the opportunity set,** *ie* the determination of the investor's *optimal* portfolio from those available.

1.2 *Assumptions of mean-variance portfolio theory*

The application of mean-variance portfolio theory is based on some important assumptions:

- 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, *ie* short-selling is possible, we can have infinitely divisible holdings, and there are no maximum investment limits.

You will meet these assumptions again at various points throughout this chapter.

1.3 *Specification of the opportunity set*

In specifying the opportunity set it is necessary to make some assumptions about how investors make decisions. Then the properties of portfolios can be specified in terms of relevant characteristics. It is assumed that investors select their portfolios on the basis of:

- **the expected return and**
- **the variance of that return**

over a single time horizon. Thus all the relevant properties of a portfolio can be specified with just two numbers — the mean return and the variance of the return. The variance (or standard deviation) is known as the risk of the portfolio.

So, according to MPT, variance of return and expected return are all that matter – this is a key assumption. Other key factors that might influence the investment decision in practice are ignored. These include:

- the suitability of the asset(s) for an investor's liabilities
- the marketability of the asset(s)
- higher moments of the distribution of returns such as skewness and kurtosis
- taxes and investment expenses
- restrictions imposed by legislation
- restrictions imposed by the fund's trustees.

Finally, we should note that the length of the time horizon, which is likely to vary between investors, is not specified.

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

The variance/covariance matrix shows the covariance between each pair of the variables. So, if there are three variables, 1, 2 and 3 say, then the matrix has the form:

$$\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$

where c_{ij} is the covariance between variables i and j.

It follows that:

- $c_{ij} = c_{ji}$ and so the matrix is symmetric about the leading diagonal
- c_{ii} is the variance of variable i.

Note that this means that with N different securities an investor must specify:

- N expected returns
- N variances of return
- $\frac{N(N-1)}{2}$ covariances.

Whilst estimates of the required parameters can be obtained using historical data, these are unlikely to prove reliable predictors of the future behaviour of investment returns and it may be necessary to adjust the historical estimates in the light of the other factors.

Question 5.2

If you assume that there are 350 shares in an equity index (as there are in the FTSE 350), how many items of data need to be specified for an investor to apply MPT?

This requirement for an investor to make thousands of estimates of covariances is potentially a major limitation of mean-variance portfolio theory in its general form.

However, we will see in Chapters 6 and 7 that:

- multifactor models and single-index models have been developed in an attempt to reduce the data requirements
- the capital asset pricing model gets round this problem.

Efficient portfolios

Two further assumptions about investor behaviour allow the definition of efficient portfolios. You have already met these in the assumptions of mean-variance portfolio theory above.

The assumptions are:

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

This means that investors prefer "more to less" and that they are risk averse.

A portfolio is *inefficient* if the investor can find another portfolio with the same (or higher) expected return and lower variance, or the same (or lower) variance and higher expected return.

A portfolio is *efficient* if the investor cannot find a better one in the sense that it has either a higher expected return and the same (or lower) variance or a lower variance and the same (or higher) expected return, *ie* an *efficient* portfolio is one that isn't *inefficient*.

Thus, every portfolio – including those that consist of a single asset – is either efficient or inefficient.

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

This is because an investor who is risk-averse and prefers more to less will never choose an inefficient portfolio. The set of efficient portfolios is known as the *efficient frontier*.

However, an investor may be able to rank efficient portfolios by using a utility function, as shown in Section 1.4 below – thereby determining the investor's optimal portfolio. Recall from Chapter 2 that if we know an individual's utility function then we can describe their attitude towards risk and return. If the assumption that investors make their decisions purely on the basis of expected return and variance holds, then this attitude towards risk and return can equally be described by indifference curves. Indifference curves join points of equal expected utility in expected return-standard deviation space, *ie* portfolios that an individual is indifferent between.

Suppose an investor can invest in any of the N securities, $i = 1,\ldots N$. A proportion x_i is invested in security S_i.

Note that:

- x_i is a proportion of the total sum to be invested
- given infinite divisibility, x_i can assume any value along the real line, subject to the restriction that $\sum x_i = 1$
- we have not specified the nature of the N securities.

The return on the portfolio R_P is:

$$R_P = \sum_i x_i R_i$$

where R_i is the return on security S_i, *ie* the portfolio return is a weighted average of the individual security returns.

The expected return on the portfolio is:

$$E = E[R_P] = \sum_i x_i E_i$$

where E_i is the expected return on security S_i.

The variance is:

$$V = \text{var}[R_P] = \sum_j \sum_i x_i x_j C_{ij}$$

where C_{ij} is the covariance of the returns on securities S_i and S_j and we write $C_{ii} = V_i$.

So, the lower the covariance between security returns, the lower the overall variance of the portfolio. This means that the variance of a portfolio can be reduced, by investing in securities whose returns are uncorrelated, *ie* by diversification.

Q

Question 5.3

How can the covariance of the returns on securities S_i and S_j be expressed in terms of the correlations of the returns on securities S_i and S_j?

The case of two securities

We can derive expressions for the mean and the variance of portfolio returns in simple cases.

If there are just two securities, S_A and S_B, the above expressions reduce to:

$$E = x_A E_A + x_B E_B$$

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

Question 5.4

Consider a portfolio consisting of equal holdings of two securities, S_x and S_y. You have the following information about the securities:

- the return on S_x is equally likely to be 5% or 10% *pa*
- the return on S_y is equally likely to be 10% or 20% *pa.*

(i) Calculate the means and variances of returns on each individual security.

(ii) Calculate the mean and variance of the return on the portfolio as a whole, given that the correlation coefficient of the two securities is:

(a) 1
(b) 0
(c) –1
(d) 0.7

(iii) Comment on your results.

Question 5.5

Repeat (ii)(c) of the previous question for a portfolio consisting of $\frac{2}{3}S_x$ and $\frac{1}{3}S_y$ and again comment on your result.

As the proportion invested in S_A is varied a curve is traced in $E-V$ space. The minimum variance can easily be shown to occur when:

$$x_A = \frac{V_B - C_{AB}}{V_A - 2C_{AB} + V_B}$$

Question 5.6

Prove the above result.

As an example consider the case where:

$$E_A = 4\% \qquad V_A = 4\%\% \qquad (\sigma_A = 2\%)$$
$$E_B = 8\% \qquad V_B = 36\%\% \qquad (\sigma_B = 6\%)$$

We now let the correlation coefficient between the two securities vary by considering ρ_{AB} equal to –0.75, 0, and +0.75 in turn. The results are plotted in Figure 5.1 where the vertical axis represents expected values of return and the horizontal axis represents standard deviation of return. In this space $(E-\sigma)$ the curves representing possible portfolios of the two securities are hyperbolae. It is possible to plot the same results in $E-V$ space, where the lines would be parabolas.

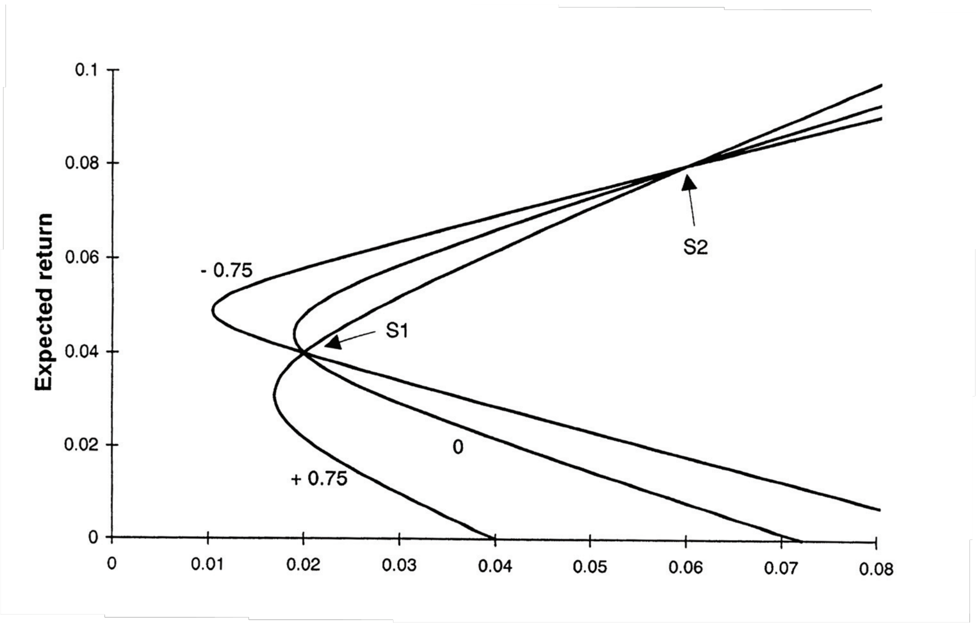


Figure 5.1

Note that S_A and S_B are denoted by S1 and S2 respectively in Figures 5.1 and 5.2, *ie* the points S1 and S2 represent the expected return and standard deviation for portfolios consisting entirely of S_A and S_B respectively.

Question 5.7

What are the co-ordinates of the point of minimum variance in Figure 5.1 in the case when $\rho_{AB} = 0$? Comment on this result.

Figure 5.2 shows combinations of securities with correlation coefficients of +1, 0 and –1. In such cases it is possible to obtain risk-free portfolios, with zero standard deviation of return, as shown in Question 5.5. Note that in each case, the set of efficient portfolios consists only of those portfolios above the point of minimum variance.

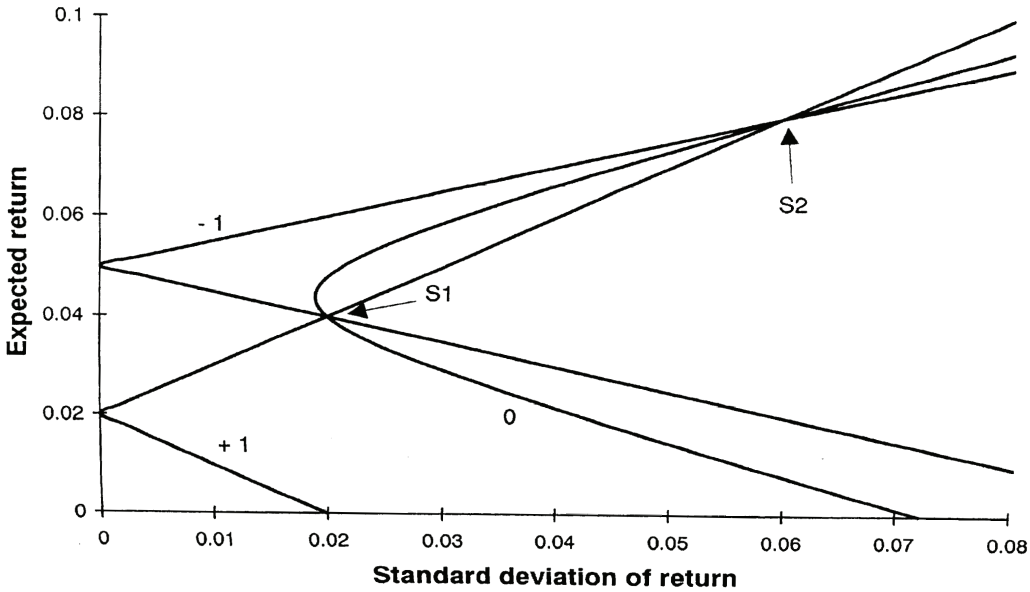


Figure 5.2

Note that if ρ is equal to +1 or –1, then there exists a *risk-free* portfolio with $V = 0$. If $\rho = -1$, it involves positive holdings of both securities, *ie* $x_1, x_2 > 0$. If $\rho = +1$, it involves a negative holding of S_B (S2) and a positive holding of S_A (S1).

Question 5.8

Show that with 2 assets, in the case where $\rho = 1$, the efficient frontier is a straight line, as shown in Figure 5.2.

When there are *N* securities the aim is to choose x_i to minimise *V* subject to the constraints:

$$\sum_i x_i = 1$$

and $E = E_P$**, say,**

in order to plot the minimum variance curve.

The aim is to choose the proportions to invest in each possible security in a way that minimises risk subject to the constraints that:

- all the investor's money is invested somewhere and
- the expected return on the portfolio is set equal to the desired level.

An alternative approach would be to *maximise* E subject to:

$$\sum_i x_i = 1 \text{ and } V = V_P \text{ say}$$

However, the first approach is usually easier.

Note carefully that E and V without the subscripts are the *portfolio* expected return or variance, *ie* the quantities that we are optimising, and that E_P and V_P are the specified values used in the *constraints*.

Lagrangian multipliers

One way of solving such a minimisation problem is the method of Lagrangian multipliers.

The Lagrangian approach – in which we set up and optimise a suitable Lagrangian function – can be used to solve constrained optimisation problems such as we are considering here, provided that:

- they are static, *ie* involve only a single time period
- the constraints are all strict equalities.

The basic idea is that if we wish to:

- maximise or minimise some function $f(x_1, ..., x_N)$
- subject to a set of M constraints $g_j(x_1, ..., x_N) = c_j$, $j = 1, \dots, M$
- by choice of N variables x_i, $i = 1, \dots, N$

then we can set up a Lagrangian function of the form:

$$W = f(x_1, ..., x_N) - \sum_{j=1}^{M} \lambda_j \left[g_j(x_1, ..., x_N) - c_j \right]$$

It turns out that maximising/minimising $f(x_1, ..., x_N)$ subject to the constraints $g_j(x_1, ..., x_N) = c_j$, $j = 1, \ldots, M$ is the same as maximising/minimising W with respect to $x_1, \ldots, x_N$ and $\lambda_1, \ldots, \lambda_M$.

Thus we set the derivatives of W with respect to each of the x_i and the λ_j's to zero. This gives us a set of $M + N$ equations in $M + N$ unknowns. Under suitable conditions this first-order conditions can be solved simultaneously to find the optimal values of the x_i – and also the λ_j. Note that the equations $\dfrac{\partial W}{\partial \lambda_j} = 0$ are just the M constraint equations themselves.

In this instance we wish to:

- minimise the portfolio variance, V
- subject to the two constraints $\sum_i x_i = 1$ and $E = E_P$
- by choice of the securities x_i, $i = 1, \ldots, N$.

The Lagrangian function is:

$$\boldsymbol{W = V - \lambda(E - E_P) - \mu\left(\textstyle\sum_i x_i - 1\right)}$$

or $$W = \sum_i \sum_j C_{ij} x_i x_j - \lambda\left(\sum_i E_i x_i - E_P\right) - \mu\left(\sum_i x_i - 1\right)$$

where:

- *V, E* and x_i are defined as above
- E_P and 1 are the constraining constants and
- λ and μ are known as the *Lagrangian multipliers*. Remember that we are trying to *minimise* the variance *V*, subject to the expected return and "all money invested" constraints.

Question 5.9

Recall the example in the Core Reading involving Securities S_A and S_B illustrated in Figure 5.1, in which:

$$E_A = 4\%, \quad V_A = 4\%\%$$

$$E_B = 8\%, \quad V_B = 36\%\%$$

Write down the Lagrangian function W in the case where the correlation coefficient is $\rho_{AB} = 0.75$.

To find the minimum we set the partial derivatives of W with respect to all the x_i and λ and μ equal to zero. The result is a set of linear equations that can be solved.

Question 5.10

Why are the equations linear?

The partial derivative of W with respect to x_i is:

$$\frac{\partial W}{\partial x_i} = 2\sum_j x_j C_{ij} - \lambda E_i - \mu$$

Question 5.11

Explain where the $2\sum_{j=1}^{N} x_j C_{ij}$ comes from in the above expression.

The partial derivative of W with respect to λ is:

$$\frac{\partial W}{\partial \lambda} = -\left(\sum_i E_i x_i - E_P\right)$$

and with respect to μ is:

$$\frac{\partial W}{\partial \mu} = -\left(\sum_i x_i - 1\right)$$

Setting each of these to zero gives:

$$2\sum_j C_{ij}x_j - \lambda E_i - \mu = 0 \quad \textbf{(one equation for each of } N \textbf{ securities)}$$

$$\sum_i x_i E_i = E_P$$

$$\sum_i x_i = 1$$

These $N+2$ equations in $N+2$ unknowns (first-order conditions) can be solved to find the optimal values of the security proportions, *ie* the x_i's, as functions of the portfolio expectation E_P. These functions can then be substituted into the expression for the portfolio variance, the resulting expression for the portfolio variance as a function of the portfolio expectation being the equation of the minimum variance curve. It is the top half of this curve, *ie* above the point of global minimum variance, that is the efficient frontier.

Question 5.12

Write down the above conditions for the example in Question 5.9, where we had:

$$E_A = 4\%, \quad V_A = 4\%\%$$

$$E_B = 8\%, \quad V_B = 36\%\%$$

and $\rho_{AB} = 0.75$.

Matrix notation

These $N+2$ equations are best represented using matrix notation as:

$$\boldsymbol{Ay} = \boldsymbol{b}$$

where

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{2C} & -\boldsymbol{E} & -\boldsymbol{I} \\ \boldsymbol{E^T} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{I^T} & \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$$

For example, in the 2-security case:

$$A = \begin{bmatrix} 2C_{11} & 2C_{12} & -E_1 & -1 \\ 2C_{21} & 2C_{22} & -E_2 & -1 \\ E_1 & E_2 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{y^T} = \left\{\boldsymbol{x^T}\,\lambda\;\mu\right\}$$

$$\boldsymbol{b^T} = \{\boldsymbol{0\,0\,...\,0\,E_P\,1}\} \quad \textbf{(}\boldsymbol{N}\textbf{ zeros)}$$

$$\boldsymbol{x^T} = \{\boldsymbol{x_1\,x_2\,...\,x_N}\}$$

$$\boldsymbol{E^T} = \{\boldsymbol{E_1\,E_2\,...\,E_N}\}$$

$$\boldsymbol{I^T} = \{\boldsymbol{1\,1\,...\,1}\} \quad \textbf{(}\boldsymbol{N}\textbf{ ones)}$$

Note that a "T" superscript denotes a transpose, which means that, although it has been written on the page as a row vector, it is actually a column vector. C is the covariance matrix.

Question 5.13

What are the dimensions of A, y and b? Hence write out the equation $A.y = b$ in full in the case in which there are just 2 securities.

The solution is then:

$$\boldsymbol{y} = \boldsymbol{A}^{-1}\boldsymbol{b}$$

from which we can see that x_i is linear in E_P . Note that the x_i's are contained in the vector y and E_P appears only once on the right-hand side of the solution equation, *ie* in b but not in A^{-1}. This assumes that A^{-1} exists.

By substituting for x in the equation for $V = x^T Cx$, we can see that the corresponding V_P is quadratic in E_P, *ie* the minimum portfolio variance is a quadratic function of the portfolio expected return. This follows immediately from the result that x_i is linear in E_P. Hence, if V_P is quadratic in x, then it must be quadratic in E_P.

We now generalise to any E and V rather than the specific values of E_P and V_P . In other words we now look at (E, V) as E is allowed to vary.

The solution to the problem shows that the minimum variance V is a quadratic in E and each x_i is linear in E.

The usual way of representing the results of the above calculations is by plotting the minimum standard deviation for each value of E_P as a curve in expected return–standard deviation $(E - \sigma)$ space.

Recall that this is done for the two-security case in Figures 5.1 and 5.2 earlier in this chapter. As V is quadratic in E, so the resulting minimum standard deviation curve is a hyperbola in $E - \sigma$ space and a parabola in $E - V$ space. Hereafter we always consider $E - \sigma$ space unless stated otherwise.

In this space, with expected return on the vertical axis, the *efficient frontier* is the part of the curve lying above the point of the global minimum of standard deviation. All other possible portfolios are inefficient and, according to MPT, should never be chosen by a rational, non-satiated, risk-averse investor.

Question 5.14

Solve the set of equations you wrote down in Question 5.12 to derive an expression for the efficient frontier when S_A and S_B are the only two securities available and they have a correlation coefficient equal to 0.75.

Recall that investors make their decisions purely on the basis of expected return and variance. This means that, if the returns on securities are normally distributed, the returns can be characterised by just the mean and the variance. In this case MPT could be used to select optimal portfolios.

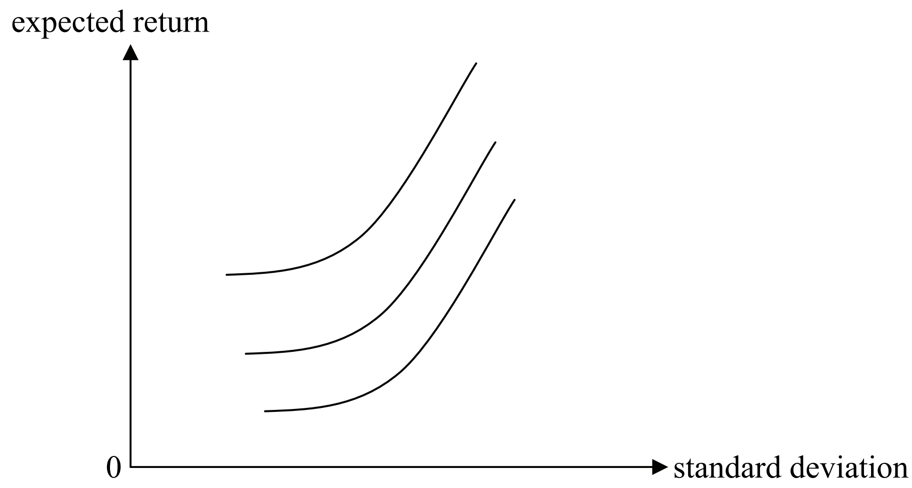
In fact, it can be shown that normality of returns is not a necessary condition for the selection of optimal portfolios. There is a more general class of distributions called the elliptically symmetrical family, which also result in optimality. All the distributions in this class have the property that the higher order moments can be expressed in terms of just their mean and variance.

1.4 Choosing an efficient portfolio

Recall from Section 1.3 that indifference curves join points of equal expected utility in expected return-standard deviation space, *ie* portfolios that an individual is indifferent between. Note that it is *expected utility* because we are considering situations involving uncertainty.

In return–standard deviation space a series of indifference curves can be plotted. Portfolios lying along a single curve all give the same value of expected utility and so the investor would be indifferent between them.

The following diagram shows typical indifference curves in $E-\sigma$ space.



By combining the investor's indifference curves with the efficient frontier of portfolios, we can determine the investor's optimal portfolio, *ie* the portfolio that maximises the investor's expected utility.

Utility is maximised by choosing the portfolio on the efficient frontier at the point where the frontier is at a tangent to an indifference curve.

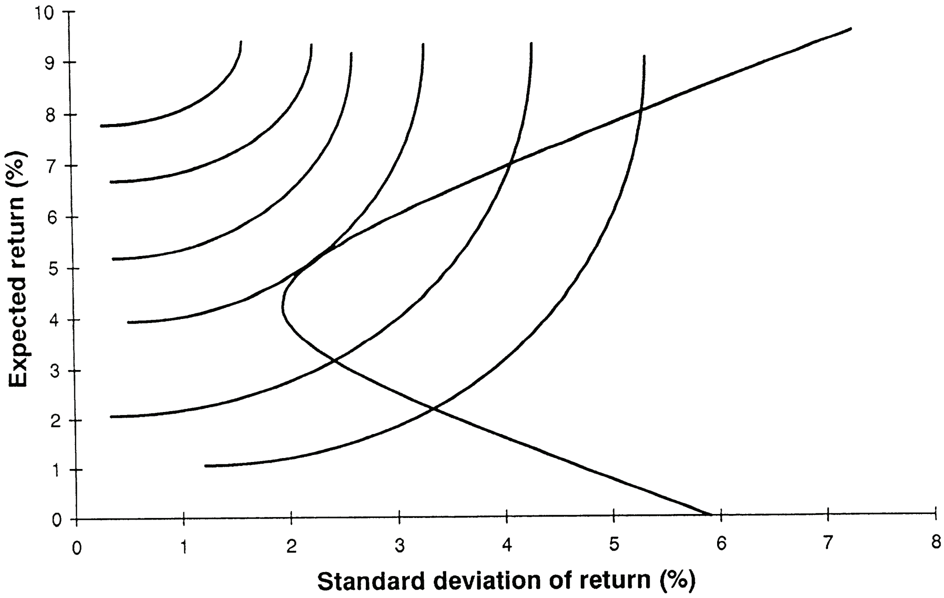


Figure 5.3

Question 5.15

Explain why the optimal portfolio on the efficient frontier is at the point where the frontier is at a tangent to an indifference curve.

Question 5.16

Why do the investor's indifference curves slope upward? What determines their gradient?

For quadratic utility functions the process described above produces optimal portfolios whatever the distribution of returns, because expected utility is uniquely determined if we know the mean and variance of the distribution.

Recall from Chapter 2 that, if the investor has a quadratic utility function, their attitude towards risk can be fully characterised by just the mean and the variance of return. Hence, when maximising expected utility by the choice of portfolio, the investor is concerned only with the first two moments of the investment returns yielded by the available portfolios and ignores all other factors.

Optimal portfolios are also produced for any utility function if investment returns are assumed to be normally distributed. This is very important because investment returns are often modelled using a normal distribution – although, as we will see in Chapter 10, there is some debate concerning the appropriateness of their use.

If it is felt that the assumptions leading to a two-dimensional mean-variance type portfolio selection model are inappropriate, it is possible to construct models with higher dimensions. For example, skewness could be used in addition to expected return and a dispersion measure. It would then be necessary to consider an efficient surface in three dimensions rather than an efficient frontier in two. Clearly, the technique can be extended to more than three dimensions.

Although such models have been constructed, they do not appear to be widely used. It is doubtful whether the additional mathematical complexity, input data requirements and difficulty of interpretation in a pragmatic way, are compensated by real improvements in value added.

2 Benefits of diversification

Recall from Section 1.3 that the variance of the portfolio is:

$$V = \text{var}[R_P] = \sum_i \sum_j x_i x_j C_{ij}$$

This expression can be rewritten as:

$$V = \sum_i x_i^2 V_i + \sum_j \sum_{i \neq j} x_i x_j C_{ij}$$

Recall also from Section 1.3 that, the lower (*ie,* closer to zero) the covariance between security returns, the lower the overall variance of the portfolio. This means that the variance of a portfolio can be reduced, by investing in securities whose returns are uncorrelated or, equivalently, investing in independent assets.

Where all assets are independent, the covariance between them is zero and the formula for variance becomes:

$$V = \sum_i x_i^2 V_i$$

If we assume that equal amounts are invested in each asset, then with *N* assets the proportion invested in each is $1/N$. Thus:

$$\sum_i (1/N)^2 V_i = (1/N) \sum_i (1/N) V_i = \frac{V^*}{N}$$

where V^* represents the average variance of the stocks in the portfolio. As *N* gets larger and larger, the variance of the portfolio approaches zero. This is a general result – if we have enough *independent* assets, the variance of a portfolio of these assets approaches zero.

In other words, a lower variance, *ie* lower risk, can be achieved by diversification.

In general, we are not so fortunate. In most markets, the correlation coefficient and the covariance between assets is positive. If you read the financial press or watch the city news, you'll usually find that the commentators think that the market as a whole is either doing well or badly. This suggests that investment returns tend to move together, *ie they are positively correlated*.

In these markets, the risk on the portfolio cannot be made to go to zero, but can still be much less than the variance of an individual asset. With equal investment, the proportion invested in any one asset x_i is $1/N$ and the formula for the variance of the portfolio becomes:

$$V = \sum_i (1/N)^2 V_i + \sum_j \sum_{i \neq j} (1/N)(1/N) C_{ij}$$

Factoring out $1/N$ from the first summation and $(N-1)/N$ from the second yields:

$$V = (1/N) \sum_i (1/N) V_i + \frac{(N-1)}{N} \sum_j \sum_{i \neq j} \frac{C_{ij}}{N(N-1)}$$

Replacing the variances and covariances in the summation by their averages V* and C*, we have:

$$V = \frac{V^*}{N} + \frac{N-1}{N} C^*$$

The contribution to the portfolio variance of the variances of the individual securities goes to zero as *N* gets very large. However, the contribution of the covariance terms approaches the average covariance as *N* gets large.

So, as the number of assets in the portfolio is increased, the variance of the return on the portfolio gets closer to the average covariance of return between the pairs of assets in that portfolio.

The individual risk of securities can be diversified away, but the contribution to the total risk caused by the covariance terms cannot be diversified away.

Some sections of this Chapter have been adapted from lecture notes originally written by David Wilkie.

David Wilkie has also produced a stochastic model for the economy, which is mentioned briefly in Chapter 10. He once received many nominations for the greatest actuary of all time in *The Actuary* magazine.

3 Exam-style questions

We finish this chapter with two exam-style questions.

Question 1

(i) Using mean-variance portfolio theory, prove that the efficient frontier becomes a straight line in the presence of a risk-free asset. [3]

Consider two independent assets, Asset A and Asset B, with expected returns of 6% *pa* and 11% *pa* and standard deviations of returns of 5% *pa* and 10% *pa* respectively.

(ii) If only Assets A and B are available, calculate the equation of the efficient frontier in expected return-standard deviation space. [3]

(iii) A third Asset, Asset C, is risk-free and has an expected return of 4% *pa*. A Lagrangian function is to be used to calculate the equation of the new efficient frontier. Write down, but do not solve, the five simultaneous equations that result from this procedure. [3]

(iv) Use your simultaneous equations to derive the relationship between x_A and x_B, the holdings of Assets A and B, on the new efficient frontier. [2]

(v) Hence derive the equation of the new efficient frontier in expected return-standard deviation space. [4]

[Total 15]

Solution 1

(i) ***Efficient frontier with a risk-free asset***

The risk-free asset sits on the efficient frontier because you cannot get a higher return with no risk and you can not have less risk than zero.

Consider any efficient portfolio of risky assets with mean E_P, a variance V_P and a portfolio composition x_P.

The variance of a portfolio on the efficient frontier can be expressed as:

$$V = x_r^2\sigma_r^2 + x_P^2\sigma_P^2 + 2x_r x_P \sigma_r \sigma_P \rho$$

This simplifies because the risk-free asset has $\sigma_r = 0$:

$$V = 0 + x_P^2\sigma_P^2 + 0 = x_P^2\sigma_P^2 \iff \sigma = x_P\sigma_P$$

Next, we can use the formula for the expected return on the portfolio to show that:

$$E = x_r E_r + x_P E_P = \left(1 - x_P\right) E_r + x_P E_P$$

$$\iff \qquad x_P = \frac{E - E_r}{E_P - E_r}$$

Substituting this into the above expression for the standard deviation we see that σ is linear in E:

$$\sigma = \left(\frac{E - E_r}{E_P - E_r}\right)\sigma_P = \frac{\sigma_P}{E_P - E_r}E - \frac{E_r\sigma_P}{E_P - E_r}$$

Hence the efficient frontier is a straight line in expected return-standard deviation space.

(ii) ***Efficient frontier of A and B***

We can express the portfolio proportions as functions of the portfolio expected return:

$$E = x_A E_A + x_B E_B = 6x_A + 11(1 - x_A)$$

$$\Leftrightarrow \qquad x_A = \frac{E-11}{6-11} = \frac{11-E}{5} \quad \text{and} \quad x_B = 1 - x_A = \frac{E-6}{5}$$

The two assets are independent and so the portfolio variance is:

$$V = x_A^2 V_A + x_B^2 V_B$$

Substituting in the portfolio proportions we get:

$$\begin{aligned} V &= 5^2\left(\frac{11-E}{5}\right)^2 + 10^2\left(\frac{E-6}{5}\right)^2 \\ &= 121 - 22E + E^2 + 4\left(E^2 - 12E + 36\right) \\ &= 5E^2 - 70E + 265 \end{aligned}$$

Finally, we square root to get the equation for the opportunity set in expected return-standard deviation space:

$$\sigma = \sqrt{5E^2 - 70E + 265}$$

The efficient frontier is the part of this line above the point at which the variance is minimised. To find this point we differentiate:

$$\frac{dV}{dE} = 10E - 70 = 0$$

So the efficient frontier is the part of the opportunity set where $E \geq \frac{70}{10} = 7\%$

ie, the efficient frontier is $\sigma = \sqrt{5E^2 - 70E + 265}$, $E \geq 7\%$

(iii) ***Simultaneous equations***

The Lagrangian function is given by:

$$V = 5^2 x_A^2 + 10^2 x_B^2$$
$$-\lambda\left(6x_A + 11x_B + 4x_C - E\right)$$
$$-\mu\left(x_A + x_B + x_C - 1\right)$$

We now differentiate the function with respect to its five parameters and set to zero:

$$\frac{\partial V}{\partial x_A} = 50x_A - 6\lambda - \mu = 0 \qquad (1)$$

$$\frac{\partial V}{\partial x_B} = 200x_B - 11\lambda - \mu = 0 \quad (2)$$

$$\frac{\partial V}{\partial x_C} = -4\lambda - \mu = 0 \quad (3)$$

$$\frac{\partial V}{\partial \lambda} = E - 6x_A - 11x_B - 4x_C = 0 \qquad (4)$$

$$\frac{\partial V}{\partial \mu} = 1 - x_A - x_B - x_C = 0 \quad (5)$$

(iv) ***Relationship between holdings in A and B***

From Equation (3) in part (iii), we have:

$$\mu = -4\lambda$$

Substituting this into Equations (1) and (2) gives:

$$50x_A - 6\lambda + 4\lambda = 0 \Rightarrow x_A = \frac{2\lambda}{50} = \frac{8\lambda}{200}$$

and $\quad 200x_B - 11\lambda + 4\lambda = 0 \Rightarrow x_B = \dfrac{7\lambda}{200} = \dfrac{7}{8}x_A$

(v) ***New efficient frontier***

Again, the risk-free asset is part of the efficient frontier. Given this, we can use the result from part (i) to deduce that the new efficient frontier is a straight line.

This straight line must be a combination of the risk-free Asset C and a risky portfolio consisting of Assets A and B in the ratio determined above.

Part (iv) establishes that $x_B = \frac{7}{8} x_A$ along this line. The efficient portfolio consisting entirely of risky assets must have $x_C = 0$ and so at this point we have:

$$x_A = \frac{8}{15} \text{ and } x_B = \frac{7}{15}$$

The expected return of this portfolio is:

$$\begin{aligned} E &= x_A E_A + x_B E_B \\ &= \frac{8}{15} \times 6 + \frac{7}{15} \times 11 = 8.33\% \end{aligned}$$

The standard deviation of this portfolio is:

$$\sigma = \sqrt{x_A^2 V_A + x_B^2 V_B} = \sqrt{\left(\frac{8}{15}\right)^2 5^2 + \left(\frac{7}{15}\right)^2 10^2} = 5.375\%$$

Therefore, the efficient frontier is a straight line in expected return-standard deviation space joining the points $(\sigma = 0, E = 4)$ and $(\sigma = 5.375, E = 8.333)$:

$$E = 4 + \frac{8.333 - 4}{5.375 - 0} \sigma = 4 + 0.806\sigma$$

Alternatively, $\sigma = 1.240E - 4.961$.

Question 2

Two assets, A and B with independent returns R_A and R_B, are available to investors. The return on Asset A is assumed to be normally distributed:

$$R_A \sim N\left(8\%, 36\%\%\right)$$

The return on Asset B can be described by the following probability distribution:

$$R_B = \begin{cases} -15\% & \text{with probability } 0.04 \\ 1\% & \text{with probability } 0.2 \\ 10\% & \text{with probability } 0.6 \\ 20\% & \text{with probability } 0.16 \end{cases}$$

(i) Calculate the expected return, E, and the standard deviation, σ, of the return on Asset B. [2]

(ii) For both assets, find the value at risk (VaR) at the 5% confidence level on a £1m portfolio invested entirely in that asset. [3]

(iii) Compare assets A and B in the light of your answers to parts (i) and (ii) and comment on the use of VaR as a measure of investment risk. [3]

(iv) Specify the equation of the efficient frontier in $\sigma - E$ space for portfolios invested in Assets A and B. [4]

(v) An investor's risk preferences can be adequately described by indifference curves, in $E - V$ space, of the form:

$$E = e^{0.01V + \alpha} - 1$$

for some $\alpha > 0$.

Find the expected return of this investor's optimal portfolio. [5]

[Total 17]

Solution 2

(i) ***Expected return and standard deviation of return on Asset B***

The expected return on Asset B (working in percentage units) is:

$$E[R_B] = -15 \times 0.04 + 1 \times 0.2 + 10 \times 0.6 + 20 \times 0.16$$
$$= 8.8\%$$

The standard deviation of return on Asset B is:

$$\sigma[R_B] = \sqrt{(-15)^2 \times 0.04 + 1^2 \times 0.2 + 10^2 \times 0.6 + 20^2 \times 0.16 - (8.8\%)^2}$$
$$= \sqrt{55.76\%\%} = 7.467\%$$

(ii) ***Value at risk (VaR)***

The 5% tail for R_B is at 1%. This corresponds to a VaR of:

$$£1m \times -1\% = -£10,000$$

So, VaR can actually be negative. Here it means that the likely worst outcome at the 5% level of confidence is a profit rather than a loss.

The 5% tail for R_A is at $8\% - 1.645 \times 6\% = -1.87\%$.

This corresponds to a VaR of:

$$£1m \times 1.87\% = £18,700$$

(iii) ***Compare assets A and B***

Asset B is riskier because it has a higher standard deviation of return (7.467% versus 6%). However, it has a higher expected return, as one would expect for a riskier asset.

The VaR is lower for Asset B than it is for Asset A, indicating that Asset B is less risky in this sense.

However, this highlights a weakness of the VaR methodology more than it suggests a meaningful comparison between the two assets.

The problem with VaR, as highlighted here, is that it does not give us any idea about the potential magnitude of the loss should the 5% confidence level be exceeded.

Although Asset B only has a 4% chance of making a loss, the loss it would make is very high. For this reason, we should concentrate on other measures of investment risk to compare these two assets, eg variance, semi-variance and expected shortfall.

In fact, one useful accompaniment to VaR is "tail value at risk", which quantifies the expected value of the 5% tail.

(iv) ***Efficient frontier***

Let x_A be the proportion invested in Asset A. Then, the expected return on a portfolio consisting entirely of Assets A and B is:

$$\begin{aligned} E &= x_A E[R_A] + (1 - x_A) E[R_B] \\ &= 8x_A + 8.8(1 - x_A) \end{aligned}$$

Rearranging, we get:

$$x_A = \frac{8.8 - E}{0.8} = 11 - 1.25E$$

We are told that the assets are independent so that there is no covariance term in the formula for the variance of this portfolio:

$$\begin{aligned} V &= x_A^2 \operatorname{var}[R_A] + (1 - x_A)^2 \operatorname{var}[R_B] \\ &= 36(11 - 1.25E)^2 + 55.76(1.25E - 10)^2 \\ &= 143.375E^2 - 2{,}384E + 9{,}932 \end{aligned}$$

But V is a quadratic function of E and the efficient frontier is only the "top" half of this curve. So we must find the minimum by differentiating and setting to zero:

$$\frac{dV}{dE} = 286.75E - 2{,}384 = 0$$

$$\Leftrightarrow \qquad E = 8.314\%$$

Square-rooting V gives the equation in $\sigma - E$ space:

$$\sigma = \sqrt{143.375E^2 - 2{,}384E + 9{,}932}$$

So, the efficient frontier is:

$$\sigma = \sqrt{143.375E^2 - 2{,}384E + 9{,}932}$$

valid for $E \geq 8.314\%$.

We can check this answer by finding the proportion of Asset A in the minimum variance portfolio:

$$x_A = \frac{V_B - C_{AB}}{V_A + V_B - 2C_{AB}} = \frac{55.76 - 0}{36 + 55.76 - 2(0)} = 0.607672$$

The corresponding value of E is then:

$$E = 8 \times 0.607672 + 8.8 \times (1 - 0.607672) = 8.313(862)\%$$

(v) ***The investor's optimal portfolio***

In $V - E$ space, the efficient frontier and indifference curve equations are:

$$V = 143.375E^2 - 2{,}384E + 9{,}932$$

and $$V = 100\ln\left(E+1\right) - 100\alpha$$

The optimal portfolio will be where an indifference curve runs tangential to the efficient frontier. So, we differentiate both equations and set them equal to each other:

$$\frac{dV}{dE} = 286.75E - 2{,}384 = \frac{100}{E+1}$$

$$\Leftrightarrow \qquad 286.75E^2 - 2{,}097.25E - 2{,}484 = 0$$

Using the formula for solving a quadratic, we get:

$$E = \frac{2,097.25 \pm \sqrt{(-2,097.25)^2 + 4 \times 286.75 \times 2,484}}{2 \times 286.75}$$

$$= 8.351\%, -1.037\%$$

Only the first of these solutions lies within the range determined in part (iv). So, the expected return at the tangent point, and hence the expected return of the optimal portfolio for this investor, is $E = 8.35\%$.

Chapter 5 Summary

Assumptions underlying mean-variance portfolio theory

- 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, *ie* short-selling, infinitely divisible holdings, no maximum investment limits.

Definitions

- The *opportunity set* is the set of points in $E-V$ space that are attainable by the investor based on the available combinations of securities.
- A portfolio is *efficient* if there is no other portfolio with either a higher mean and the same or lower variance, or a lower variance and the same or higher mean.
- The *efficient frontier* is the set of efficient portfolios in $E-V$ space.
- *Indifference curves* join points of equal expected utility in $E-V$ space.
- The *optimal portfolio* is the portfolio that maximises the investor's expected utility.

Derivation of efficient frontier – the case of two securities

The three equations used to derive the equation of the efficient frontier when only two securities are available are:

1. $1 = x_A + x_B$
2. $E_P = x_A E_A + x_B E_B$
3. $V_P = x_A{}^2 V_A + x_B{}^2 V_B + 2x_A x_B C_{AB}$

V_p is minimised when $x_A = \dfrac{V_B - C_{AB}}{V_A + V_B - 2C_{AB}}$.

Derivation of the efficient frontier – the case of N securities

Here we use the Lagrangian function:

$$W = \sum_i \sum_j C_{ij} x_i x_j - \lambda\left(\sum_i x_i E_i - E\right) - \mu\left(\sum_i x_i - 1\right)$$

This is differentiated with respect to the x_i's, λ and μ to give:

$$\frac{\partial W}{\partial x_i} = 2\sum_j C_{ij} x_j - \lambda E_i - \mu$$

$$\frac{\partial W}{\partial \lambda} = E - \sum_i x_i E_i$$

$$\frac{\partial W}{\partial \mu} = 1 - \sum_i x_i$$

Setting these derivatives to zero and solving $N+2$ equations in $N+2$ unknowns then gives the equation of the efficient frontier in $E-V$ space.

Benefits of diversification

Assuming an equal holding of each asset within a portfolio then, as the number of assets in the portfolio is increased, the variance of the return on the portfolio gets closer to the average covariance of return between the pairs of assets in that portfolio.

$$V = \frac{V^*}{N} + \left(\frac{N-1}{N}\right)C^*$$

This means that if all the assets are independent then the variance of return on the portfolio tends towards zero, *ie* the risk can be diversified away completely.

$$V = \frac{V^*}{N} \to 0$$

Chapter 5 Solutions

Solution 5.1

There are *many* other types of investment-related risk, which are discussed in detail in Subjects CA1, ST5 and ST9. Amongst the more important are:

- *default* or *credit risk* – the other party to an investment deal fails to fulfil their obligations
- *inflation risk* – inflation is higher than anticipated, so reducing real returns
- *exchange rate* or *currency risk* – exchange rate moves in an unanticipated way
- *reinvestment risk* – stems from the uncertainty concerning the terms on which investment income can be reinvested
- *marketability risk* – the risk that you might be unable to realise the true value of an investment if it is difficult to find a buyer.

Solution 5.2

The required number of items of data is:

$$350+350+\frac{350\times 349}{2}=61,775$$

Note that this ignores all the other available investments that are not included in the FTSE 350 Index, *eg* FTSE Smallcap, FTSE Fledgling, unquoted equities, non-UK equities, property, bonds *etc*.

Solution 5.3

The covariance C_{ij} is equal to $\rho_{ij}\sigma_i\sigma_j$,

where:

σ_i = standard deviation of security i returns

σ_j = standard deviation of security j returns

ρ_{ij} = correlation coefficient between security i returns and security j returns

Solution 5.4

(i) The mean returns on each security are:

$$E(S_x) = 0.5\times0.05+0.5\times0.10 = 0.075, \quad \textit{ie}\ 7.5\%$$

$$E(S_y) = 0.5\times0.10+0.5\times0.20 = 0.15, \quad \textit{ie}\ 15\%$$

The variances are given by:

$$V(S_x) = E(S_x^2)-[E(S_x)^2]$$

$$E(S_x^2) = 0.5\times0.05^2+0.5\times0.10^2 = 0.00625$$

$$E(S_y^2) = 0.5\times0.10^2+0.5\times0.20^2 = 0.025$$

So,

$$V(S_x) = 0.00625-0.075^2 = 0.025^2, \quad \textit{ie}\ (2.5\%)^2$$

$$V(S_y) = 0.025-0.15^2 = 0.05^2, \quad \textit{ie}\ (5\%)^2$$

(ii) The mean return on the portfolio in each case is given by:

$$E(R_p) = 0.5\times E(S_x)+0.5\times E(S_y)$$

ie $E(R_p) = 0.5\times0.075+0.5\times0.15 = 0.1125$, *ie* 11.25%

To find the variance of return we need to use the relationship:

ie
$$\begin{aligned} V(R_p) &= (½)^2V(S_x)+(½)^2V(S_y)+2Cov(½S_x,½S_y) \\ &= ¼V(S_x)+¼V(S_y)+2(½)(½)Cov(S_x,S_y) \end{aligned}$$

(a) When the correlation coefficient = 1, we have:

$$Cov(S_x,S_y) = 1\times0.025\times0.05 = 0.00125$$

ie
$$V(R_p) = ¼V(S_x)+¼V(S_y)+2(½)(½)Cov(S_x,S_y)$$

$$= ¼\times0.000625+¼\times0.0025+2\times½\times½\times0.00125 = 0.00140625$$

ie $(3.75\%)^2$

(b) When the correlation coefficient = 0, we have:

$$Cov(S_x, S_y) = 0 \times 0.025 \times 0.05 = 0$$

$$V(R_p) = ¼ \times 0.000625 + ¼ \times 0.0025 = 0.00078125 \text{ , } ie \ (2.795\%)^2$$

(c) When the correlation coefficient = –1, we have:

$$Cov(S_x, S_y) = -1 \times 0.025 \times 0.05 = -0.00125$$

ie $V(R_p) = ¼ \times 0.000625 + ¼ \times 0.0025 - 2 \times ½ \times ½ \times 0.00125 = 0.00015625$

ie $(1.25\%)^2$

(d) When the correlation coefficient = 0.7, we have:

$$Cov(S_x, S_y) = 0.7 \times 0.025 \times 0.05 = 0.000875$$

$$V(R_p) = ¼ \times 0.000625 + ¼ \times 0.0025 + 2 \times ½ \times ½ \times 0.000875 = 0.00121875 ,$$

ie $(3.491\%)^2$

(iii) The more closely correlated the investments in the portfolio, the larger the variance. The highest result was obtained when the securities were assumed to be perfectly correlated (*ie* the correlation coefficient was +1) and the lowest result when the securities were assumed to be perfectly negatively correlated (*ie* the correlation coefficient was –1).

This is to be expected. For example, with negative correlation, the potential deviations from the expected return of each security separately will tend to cancel each other out, giving a smaller overall portfolio deviation from the overall expected return of the portfolio.

Solution 5.5

$$V(R_p) = \left(\tfrac{2}{3}\right)^2 \times 0.000625 + \left(\tfrac{1}{3}\right)^2 \times 0.0025 - 2 \times \tfrac{2}{3} \times \tfrac{1}{3} \times 0.00125 = 0$$

$$E(R_p) = \tfrac{2}{3} \times 0.075 + \tfrac{1}{3} \times 0.15 = 0.10$$

This shows that a portfolio consisting of two risky assets can be risk-free if the correlation coefficient is –1 and the appropriate proportion is invested in each asset.

Solution 5.6

The variance of the portfolio return is:

$$\begin{aligned} V_P &= x_A{}^2 V_A + x_B{}^2 V_B + 2x_A x_B C_{AB} \\ &= x_A{}^2 V_A + (1-x_A)^2 V_B + 2x_A(1-x_A)C_{AB} \end{aligned}$$

We want to choose the value for x_A that minimises the variance V_P. To do this, we differentiate and set to zero:

$$\frac{dV_P}{dx_A} = 2x_A V_A + 2(1-x_A)V_B(-1) + 2(1-x_A)C_{AB} - 2x_A C_{AB} = 0$$

$$\Leftrightarrow \quad 2x_A V_A - 2V_B + 2x_A V_B + 2C_{AB} - 4x_A C_{AB} = 0$$

$$\Leftrightarrow \quad x_A(V_A + V_B - 2C_{AB}) = V_B - C_{AB}$$

$$\Leftrightarrow \quad x_A = \frac{V_B - C_{AB}}{V_A - 2C_{AB} + V_B}$$

Solution 5.7

The minimum variance occurs when:

$$x_A = \frac{V_B - C_{AB}}{V_A - 2C_{AB} + V_B}$$

If $\rho = 0$, then this reduces to $x_A = \dfrac{V_B}{V_A + V_B}$, so that in this instance:

$$x_A = \frac{0.0036}{0.0004 + 0.0036} = 0.9 \text{ and so } x_B = 0.1 .$$

Thus, the resulting portfolio has expected return and variance given by:

$$\begin{aligned} E_P &= x_A E_A + x_B E_B \\ &= 0.9 \times 0.04 + 0.1 \times 0.08 \\ &= 0.044 \end{aligned}$$

and:

$$\begin{aligned} \sigma_P^2 &= x_A^2 V_A + x_B^2 V_B + 2 x_A x_B C_{AB} \\ &= 0.81 \times 0.0004 + 0.01 \times 0.0036 + 2 \times 0.9 \times 0.1 \times 0 \\ &= 0.01897^2 \end{aligned}$$

ie $\quad \sigma_P = 0.019$ or 1.9%

This is less than both of the individual security standard deviations of 2% and 6%, illustrating the benefits of diversification.

Solution 5.8

We can use the formula for the expected return of the portfolio to express x_A in terms of E_P.

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

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

The variance of the portfolio return is:

$$\begin{aligned} V_P &= {x_A}^2 V_A + {x_B}^2 V_B + 2 x_A x_B C_{AB} \\ &= {x_A}^2 \sigma_A^2 + {x_B}^2 \sigma_B^2 + 2 x_A x_B \sigma_A \sigma_B \rho \end{aligned}$$

If there is perfect correlation, $\rho = 1$, then this simplifies:

$$V_P = {x_A}^2\sigma_A^2 + {x_B}^2\sigma_B^2 + 2x_A x_B \sigma_A \sigma_B$$
$$= \left(x_A\sigma_A + x_B\sigma_B\right)^2$$

$$\Leftrightarrow \qquad \sigma_P = x_A\sigma_A + x_B\sigma_B = x_A\sigma_A + \left(1 - x_A\right)\sigma_B$$
$$= \frac{E_P - E_B}{E_A - E_B}\sigma_A + \frac{E_A - E_P}{E_A - E_B}\sigma_B$$
$$= \frac{\sigma_A - \sigma_B}{E_A - E_B}E_P + \frac{\sigma_B E_A - \sigma_A E_B}{E_A - E_B}$$

This is a straight line in $\left(E_P, \sigma_P\right)$ space.

Solution 5.9

Here the Lagrangian function is:

$$W = x_A^2\sigma_A^2 + x_B^2\sigma_B^2 + 2x_A x_B \rho_{AB}\sigma_A\sigma_B - \lambda\left(E_A x_A + E_B x_B - E_P\right) - \mu\left(x_A + x_B - 1\right)$$

$$W = 4x_A^2 + 36x_B^2 + 2x_A x_B \times 0.75 \times 2 \times 6 - \lambda\left(4x_A + 8x_B - E_P\right) - \mu\left(x_A + x_B - 1\right)$$

ie $$W = 4x_A^2 + 36x_B^2 + 18x_A x_B - \lambda\left(4x_A + 8x_B - E_P\right) - \mu\left(x_A + x_B - 1\right)$$

Solution 5.10

The equations are linear because:

- the variance of portfolio returns is a quadratic function of the x_i and
- the constraint terms are linear in each of x_i, λ and μ.

Hence, the first-order conditions contain powers in the x_i no higher than one.

Solution 5.11

For an N-security portfolio, the variance is equal to:

$$V = \sum_{i=1}^{N}\sum_{j=1}^{N} x_i x_j C_{ij}$$

Writing out the covariance matrix in full gives:

$$\begin{bmatrix} x_1^2 C_{11} & x_1 x_2 C_{12} & \cdots & x_1 x_i C_{1i} & \cdots & x_1 x_N C_{1N} \\ x_2 x_1 C_{21} & x_2^2 C_{22} & \cdots & x_2 x_i C_{2i} & \cdots & x_2 x_N C_{2N} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_i x_1 C_{i1} & x_i x_2 C_{i2} & \cdots & x_i^2 C_{ii} & \cdots & x_i x_N C_{iN} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_N x_1 C_{N1} & x_N x_2 C_{N2} & \cdots & x_N x_i C_{Ni} & \cdots & x_N^2 C_{NN} \end{bmatrix}$$

Thus, if we differentiate with respect to x_i, the ith row of the matrix gives us the following terms:

(1) $x_1 C_{i1} + x_2 C_{i2} + \ldots + 2x_i C_{ii} + \ldots + x_n C_{in}$

Also from the ith column, each of the other j = 1, … , i–1, i+1, …, n rows, yields an $x_j C_{ji}$ term, *ie* the following terms:

(2) $x_1 C_{1i} + x_2 C_{2i} + \ldots + x_{i-1} C_{i-1,i} + x_{i+1} C_{i+1,i} + \ldots + x_n C_{ni}$

Hence, if we sum (1) and (2) we obtain:

$$2x_1 C_{1i} + 2x_2 C_{2i} + \ldots + 2x_i C_{ii} + \ldots + 2x_n C_{in}$$

which we can write as:

$$2\sum_{j=1}^{N} x_j C_{ij}$$

Solution 5.12

Differentiating the Lagrangian function given in the solution to Question 5.9 gives the first-order conditions:

$$\frac{\partial W}{\partial x_A} = 8x_A + 18x_B - 4\lambda - \mu = 0$$

$$\frac{\partial W}{\partial x_B} = 18x_A + 72x_B - 8\lambda - \mu = 0$$

$$\frac{\partial W}{\partial \lambda} = -\left(4x_A + 8x_B - E_P\right) = 0$$

$$\frac{\partial W}{\partial \mu} = -\left(x_A + x_B - 1\right) = 0$$

Solution 5.13

The dimension of:

- A is $(N+2) \times (N+2)$. The submatrix C is $N \times N$, *ie* $C = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1N} \\ C_{21} & C_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ C_{N1} & \cdots & \cdots & C_{NN} \end{bmatrix}$.

  The N rows represent the equations for each of the N variables x_i, $i = 1, \ldots, N$.

- y is $(N+2) \times 1$
- b is $(N+2) \times 1$.

When $N = 2$, we can derive this set of equations from first principles using the Lagrangian function for the optimisation problem, which is:

$$W = x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\rho\sigma_1\sigma_2 - \lambda(x_1E_1 + x_2E_2 - E_p) - \mu(x_1 + x_2 - 1)$$

where ρ is the correlation coefficient between the two security returns.

Thus, the first-order conditions are:

$$\frac{\partial W}{\partial x_1} = 2x_1\sigma_1^2 + 2x_2\rho\sigma_1\sigma_2 - \lambda E_1 - \mu = 0$$

$$\frac{\partial W}{\partial x_2} = 2x_2\sigma_2^2 + 2x_1\rho\sigma_1\sigma_2 - \lambda E_2 - \mu = 0$$

$$\frac{\partial W}{\partial \lambda} = x_1 E_1 + x_2 E_2 - E_p = 0$$

$$\frac{\partial W}{\partial \mu} = x_1 + x_2 - 1 = 0$$

Hence, $A.y = b$ can be written as:

$$\begin{bmatrix} 2\sigma_1^2 & 2\rho\sigma_1\sigma_2 & -E_1 & -1 \\ 2\rho\sigma_1\sigma_2 & 2\sigma_2^2 & -E_2 & -1 \\ E_1 & E_2 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ E_p \\ 1 \end{bmatrix}$$

Solution 5.14

In order to minimise the portfolio variance V for a given level of portfolio expected return E we need to solve the equations derived in the solution to Question 5.12 above. Recall that these were:

$$\frac{\partial W}{\partial x_A} = 8x_A + 18x_B - 4\lambda - \mu = 0$$

$$\frac{\partial W}{\partial x_B} = 18x_A + 72x_B - 8\lambda - \mu = 0$$

$$\frac{\partial W}{\partial \lambda} = -\left(4x_A + 8x_B - E\right) = 0$$

$$\frac{\partial W}{\partial \mu} = -\left(x_A + x_B - 1\right) = 0$$

Here we need to find x_A and x_B as (linear) functions of E. In general terms, this requires us to solve the above 4 simultaneous equations in 4 unknowns x_A, x_B, λ and μ. In fact, in the 2-security case we have here we can simply use the last two equations to obtain x_A and x_B (as we are not interested in solving for λ and μ).

These give:

$$x_A = \frac{8-E}{4}, \; x_B = \frac{E-4}{4}$$

Substituting these back into the expression for the variance (and simplifying) then gives:

$$V = \frac{1}{16}\left(22E^2 - 136E + 256\right) = \frac{1}{8}\left(11E^2 - 68E + 128\right)$$

Note that V is a quadratic function of E – the actual minimum variance curve is shown in Figure 5.1 on page 10. The efficient frontier is that part of the curve above the point of minimum variance.

Solution 5.15

The optimal portfolio occurs at the point where the indifference curve is tangential to the efficient frontier for the following two reasons.

1. The indifference curves that correspond to a higher level of expected utility are unattainable as they lie strictly above the efficient frontier.
2. Conversely, lower indifference curves that cut the efficient frontier are attainable, but correspond to a lower level of expected utility.

The highest attainable indifference curve, and corresponding level of expected utility, is therefore the one that is tangential to the efficient frontier. The optimal portfolio occurs at the tangency point, which is in fact the only attainable point on this indifference curve, which is why it is optimal.

Solution 5.16

The investor's indifference curves slope upwards because the investor is assumed to be risk-averse and prefer more to less. Consequently, additional expected return yields extra utility, whereas additional risk reduces utility. Thus, any increase (decrease) in risk/standard deviation must be offset by an increase (decrease) in expected return in order to maintain a constant level of expected utility.

The gradient of the indifference curves is determined by the degree of the investor's risk aversion. The more risk-averse the investor, the steeper the indifference curves – as the investor will require a greater increase in expected return in order to offset any extra risk.

Chapter 6

Models of asset returns

Syllabus objectives

(iv) Describe and discuss the properties of single and multifactor models of asset returns.

1. *Describe the three types of multifactor models of asset returns:*
   - *macroeconomic models*
   - *fundamental factor models*
   - *statistical factor models*
2. *Discuss the single index model of asset returns.*
3. *Discuss the concepts of diversifiable and non-diversifiable risk.*
4. *Discuss the construction of the different types of multifactor models.*
5. *Perform calculations using both single and multifactor models.*

0 Introduction

In the previous chapter, we looked at how mean-variance portfolio theory can be used to determine the efficient frontier, from which the investor can then determine the optimal portfolio. Unfortunately, this approach can be difficult to implement in practice due to computational difficulty and the amount and type of data required.

Much subsequent research has therefore been directed at developing means of simplifying the implementation process. This has lead to the development of the *multifactor* or *multi-index models* and the *single-index models* that we discuss in this chapter. These facilitate the determination of the efficient frontier with substantially less information than the standard mean-variance portfolio theory described in Chapter 5. In addition, they can be used to characterise the sensitivities of a security's returns to various factors. Consequently, they are very important tools for portfolio management, being used both to model and predict the future investment returns yielded by different assets.

Multifactor models also form the basis of the arbitrage pricing theory discussed in Chapter 7 of this course.

This chapter also discusses the important ideas of *specific risk* – risk that can be diversified away – and *systematic risk* – which cannot. These ideas are discussed further in Subjects CT2 and CA1.

1 Multifactor models

1.1 Definition

A multifactor model of security returns attempts to explain the observed historical return by an equation of the form:

$$R_i = a_i + b_{i,1}I_1 + b_{i,2}I_2 + ... + b_{i,L}I_L + c_i \ ,$$

where:

- R_i **is the return on security** i
- a_i, c_i **are the constant and random parts respectively of the component of return unique to security** i
- $I_1 ... I_L$ **are the changes in a set of** L **factors which explain the variation of** R_i **about the expected return** a_i
- $b_{i,k}$ **is the sensitivity of security** i **to factor** k**.**

Here the interpretation of the $b_{i,k}$'s is that if $b_{i,k} = 1.5$ say, then an increase (decrease) of 1 in factor k is expected to produce an increase (decrease) of 1.5 in the return provided by security i. Conversely, a_i represents the expected value of that element of the investment return that is independent of the set of L factors and hence unique to security i.

The L factors are therefore the *systematic* factors that influence the returns on every security and the corresponding part of the total return which is equal to:

$$b_{i,1}I_1 + b_{i,2}I_2 + ... + b_{i,L}I_L$$

is referred to as the *systematic return* of security i. Conversely, the remaining element of the total return, which is specific to each individual security and independent of the returns on all other securities, namely:

$$a_i + c_i$$

is referred to as the *specific return* of security i.

When applying the multifactor model, other assumptions are usually made, similar to those described below in Section 2 in respect of single-index models. For example, it is usual to assume that:

- $E[c_i] = 0$
- $Cov[c_i, c_j] = 0$ for all $i \neq j$.
- $Cov[c_i, I_k] = 0$ for all stocks and indices.

Note that:

- although the above equation for the model assumes that asset returns are linearly related to the factors or indices, this requirement is not as restrictive as it might first appear, as the factors themselves may be non-linear functions of the underlying variables, *eg* the log of inflation
- a_i and c_i are sometimes combined into a single parameter, with a non-zero expectation.

Question 6.1

Consider a two-factor model. If:

- the mean specific return is 1.0%
- the expected values of the two factors are 3.0% and 2.2% and
- the sensitivities of investment returns to each of the factors are 0.8 and –0.3 respectively,

what is the expected return predicted by the model?

The goal of the builders of such a model is to find a set of factors which explain as much as possible of the observed historical variation, without introducing too much "noise" into predictions of future returns.

Multifactor models can be classified into three categories, depending on the type of factors used. These are *macroeconomic*, *fundamental* and *statistical factor models*. We discuss each of these in the following sections.

1.2 *Macroeconomic factor models*

These use observable economic time series as the factors. They could include factors such as the annual rates of inflation and economic growth, short-term interest rates, the yields on long-term government bonds, and the yield margin on corporate bonds over government bonds.

The rationale here is that the price of, and hence the returns obtained from, a security should reflect the discounted present value to investors of the cashflows that it is expected to produce in the future. The macroeconomic variables mentioned above are the factors that we might typically expect to influence both the size of the future cashflows from security *i* and the discount rate used to value them. Note, however, that the more factors that are included, the more complex the model and so the harder it is to handle mathematically.

A related class of model uses a market index plus a set of industry indices as the factors.

Here, "industry" refers to the company sectors, such as banking, energy, food, support services *etc*.

So, security returns are assumed to reflect the influence of both market-wide and industry-specific effects.

Once the set of factors has been decided on, a time series regression is performed to determine the sensitivities for each security in the sample.

This kind of time series regression is a natural multivariable extension of the one-variable case considered in the single-index models discussed later in this chapter.

A common method used to determine how many factors to include is to start with relatively few, perform the regression and measure the residual (unexplained) variance. An extra factor is then added and the regression repeated. The whole process is repeated until the addition of an extra factor causes no significant reduction in the residual variance.

An alternative approach is to start with a more general model containing a large number of possible factors and then to remove those whose elimination does not significantly affect the explanatory power of the model – *ie* the size of the residual variance.

1.3 Fundamental factor models

Fundamental factor models are closely related to macroeconomic models but instead of (or in addition to) macroeconomic variables the factors used are company-specific variables. These may include such fundamental factors as:

- **the level of gearing**
- **the price earnings ratio**
- **the level of R&D (research & development) spending**
- **the industry group to which the company belongs.**

Again, the models are constructed using regression techniques.

Commercial fundamental factor models are available which use many tens of factors. They are used for risk control by comparing the sensitivity of a portfolio to one of the factors with the sensitivity of a benchmark portfolio.

Suppose that you can find a portfolio that has similar sensitivities to similar factors as a benchmark portfolio. Then by holding that portfolio you should be able to closely replicate the performance of the benchmark. This technique could, for example, be used to construct a portfolio that follows or tracks the performance of an investment index without the investor needing to hold every individual constituent security of the index. Index tracking is discussed in detail in Subjects CA1, ST5 and SA6.

1.4 Statistical factor models

Statistical factor models do not rely on specifying the factors independently of the historical returns data. Instead a technique called principal components analysis can be used to determine a set of indices which explain as much as possible of the observed variance. However, these indices are unlikely to have any meaningful economic interpretation and may vary considerably between different data sets.

Question 6.2 (Revision)

Explain the distinction between systematic return and specific return.

1.5 Construction of models

Principal components analysis is a technique used to investigate the relationship between a set of endogenous variables, such as the factors determining the investment return in a multifactor model. Within this context it can be used to:

- determine the relative significance of the various factors by analysing the variance-covariance matrix (between them) to determine which factors have the most influence upon the total variance of security returns
- combine groups of highly correlated factors into single factors or *principal components* that are much less highly correlated with each other – thereby reducing the number of factors in the model and improving the efficiency of the model.

For applications of multi-index models to portfolio selection problems it is convenient if the factors used are uncorrelated (or orthogonal). Principal components analysis automatically produces a set of uncorrelated factors.

Question 6.3

Why is it convenient for the factors used to be uncorrelated?

Where the factors are derived from a set of market indices or macroeconomic variables it is possible to transform the original set into an orthogonal set which retain a meaningful economic interpretation.

Suppose, for example, we have two indices I_1 and I_2. I_1 could be a market index and I_2 an industry sector index. Two new, uncorrelated factors, I_1^* and I_2^*, can be constructed as follows:

First, let $I_1^* = I_1$.

This is done solely to keep our notation consistent in the final equation below.

We then carry out a linear regression analysis to determine the parameters γ_1 and γ_2 in the equation:

$$I_2 = \gamma_1 + \gamma_2 I_1^* + d_2$$

So γ_1 and γ_2 represent the intercept and the slope of the regression line and d_2 is the "error" term, which by definition is independent of $I_1^* = I_1$.

We then set:

$$\boldsymbol{I_2^* = d_2 = I_2 - \left(\gamma_1 + \gamma_2 I_1^*\right).}$$

By construction I_2^* is uncorrelated with I_1.

This is because $I_2^* = d_2$, which was the residual term in the previous equation. Mathematically:

$$\text{cov}(I_2^*, I_1) = \text{cov}\left(I_2 - (\gamma_1 + \gamma_2 I_1^*), I_1\right)$$

$$= \text{cov}(d_2, I_1) = \text{cov}(d_2, I_1^*) = 0$$

Changes in I_2^* can be interpreted as the change in the observed value of I_2 that cannot be explained by the observed change in I_1.

If there were a third index, a regression would be performed to determine the component of that index which could not be explained by the observed values of I_1 and I_2, and so on.

Q

Question 6.4

A modeller has developed a two-factor model to explain the returns obtained from security *i*. It has the form:

$$R_i = 2 + 1.3I_1 + 0.8I_2 + c_i$$

However, he is concerned that the two indices 1 and 2 may be correlated and so decides to re-express the model in terms of orthogonal factors. He therefore regresses Index 1 on Index 2 and obtains the following equation for the line of best fit:

$$I_1 = 0.8 + 0.3I_2$$

Use this information to re-express the two-factor model in terms of two orthogonal factors I_1^* and I_2^*.

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:

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

and $\quad \hat{\alpha}_i = \bar{E}_i - \hat{\beta}_i \bar{E}_M$

where $\bar{E}_i = \frac{1}{N}\sum_{t=1}^{N} E_{it}$ and $\bar{E}_M = \frac{1}{N}\sum_{t=1}^{N} E_{Mt}$.

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:

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

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

and $\quad \boldsymbol{C_{ij} = \beta_i \beta_j V_M}$ (6.3)

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

Question 6.6

Derive the first of the above results.

Question 6.7

In no more than seven words, interpret the second of the above results.

Let's now look at these equations in a little more detail.

Equation 6.1

The first of these three equations will be seen to be identical in form to the main result of the Capital Asset Pricing Model (CAPM) discussed in Chapter 7. However, it should be emphasised that the single-index model is purely empirical and is not based on any theoretical relationships between β_i and the other variables. This is in contrast to CAPM, which can be derived using economic theory.

Equation 6.2

The second equation models the variance of the return on security *i* as the sum of a term related to the variance of the return on the market and a term specific to security *i*. These two terms are usually called systematic and specific risk respectively.

***Systematic risk* can be regarded as relating to the market as a whole while *specific risk* depends on factors peculiar to the individual security.**

It can be shown that in a diversified portfolio, consisting of a large number of securities, the contribution of the specific risk on each security, to the total risk of the portfolio, becomes very small. In this case, the contribution of each security to the portfolio's total risk is then only the systematic risk of that security.

So, diversification:

- can be used to reduce and ultimately eliminate specific risk
- leads to an averaging of systematic risk.

Thus it is only the systematic risk, measured by β_i of a security, that should be expected to be rewarded by increased return since this is non-diversifiable. For this reason systematic risk is also sometimes referred to as non-diversifiable or market risk.

Investors can diversify away specific risk and do not therefore require compensation for accepting it. Specific risk is sometimes also referred to as alpha, unsystematic, diversifiable or residual risk.

Recall that α_i in the single-index model equation is the expected value of the component of security i's return that is *independent* of the market's performance and *specific* to that particular security.

Question 6.8

Show that the variance of portfolio returns can be written as:

$$V_P = \sum_{i=1}^{N} x_i^2 V_{\varepsilon i} + \beta_p^2 V_M$$

and use this expression to show that:

- the contribution of the specific risk on each security to the total risk of the portfolio becomes very small as the number of securities increases and
- the contribution of each security to the portfolio's total risk is only the systematic risk of that security, *ie*

$$\sigma_P \to \beta_p \sigma_M = \sigma_M \sum_{i=1}^{\infty} x_i \beta_i \text{ , as } N \to \infty$$

Equation 6.3

The third equation shows that, in this particular model, any correlation between the returns on two securities comes only from their joint correlation with the market as a whole. In other words, the only reason that securities move together is a common response to market movements, there are no other possible common factors.

In reality, however, we might expect shares within the same sector (*eg* banks) to tend to move together because of factors influencing that particular sector which might not affect other sectors (*eg* oil companies).

Data requirements

Although many studies have found that incorporating more factors into the model (for example industry indices) leads to a better explanation of the historical data, correlation with the market is the largest factor in explaining security price variation. Furthermore, there is little evidence that multi-factor models are significantly better at forecasting the future correlation structure.

The use of the single-index model dramatically reduces the amount of data required as input to the portfolio selection process. For N securities, the number of data items needed has been reduced from $N(N+3)/2$ to $3N+2$.

> ***Question 6.9***
>
> Derive these expressions for the number of data items.

A primary use of both single-index models and multifactor models is to determine the expected return, variance and covariance of security returns, thereby enabling the investor to determine the efficient frontier. The single-index model in particular allows this to be done with a lot less information than is required with the standard mean-variance portfolio theory as described in the previous chapter.

Furthermore, the nature of the estimates required from security analysts conforms much more closely to the way in which they traditionally work.

Traditional investment analysis concentrates on estimating future performance, which in practice has normally meant the expected returns of securities, although increasing emphasis has also been placed on the risk or volatility of individual securities. Mean-variance portfolio theory, however, also requires estimates of each security's pairwise correlation with all other securities that may be included in the portfolio.

In addition, considerably simplified methods for calculating the efficient frontier have been developed under the single-index model although, with increasing computer power, this is of considerably less importance than it was at the time the model was first published.

Question 6.10 (Revision)

State the main uses of multifactor models and single-index models. What is their main limitation?

Question 6.11 (Revision)

If β_i doubles (with everything else remaining unchanged) then so does the expected return on security *i*. True or false?

Further reading

More detail on single-index models and multifactor models can be found in the book by Elton & Gruber listed in the Study Guide.

3 Exam-style question

We finish this chapter with an exam-style question on the single-index model.

Question

(i) A portfolio P consists of n assets, with a proportion x_i invested in asset i, $i = 1, 2, \ldots, n$ (so that $\sum_{i=1}^{n} x_i = 1$). Derive a formula for the portfolio beta, β_P, in terms of the individual betas for each asset. [2]

(ii) The annual returns R_P on this portfolio can be assumed to conform to the single-index model of asset returns. Write down an equation defining this model and show that:

$$\text{var}(R_P) = \beta_P^2 \,\text{var}(R_M) + \text{var}(\varepsilon_P)$$

where ε_P denotes the component of the portfolio return that is independent of movements in the market. [3]

(iii) Explain why the specific risk $\text{var}(\varepsilon_P)$ is sometimes referred to as the "diversifiable risk", giving an algebraic justification for your answer. [4]

(iv) Discuss the following statement:

"A portfolio with a beta of zero is equivalent to a risk-free asset." [2]

[Total 11]

Solution

(i) ***Formula for portfolio beta***

The (random) return on the portfolio is:

$$R_P = \sum_{i=1}^{n} x_i R_i$$

Using the definition given for β from page 43 of the *Tables*, the portfolio beta is:

$$\beta_P = \frac{\operatorname{cov}(R_P, R_M)}{\operatorname{var}(R_M)} = \operatorname{cov}\left(\sum_{i=1}^{n} x_i R_i, R_M\right) \Big/ \operatorname{var}(R_M)$$

Using the properties of covariances, we can simplify the numerator to get:

$$\beta_P = \sum_{i=1}^{n} x_i \operatorname{cov}(R_i, R_M) \Big/ \operatorname{var}(R_M) = \sum_{i=1}^{n} x_i \frac{\operatorname{cov}(R_i, R_M)}{\operatorname{var}(R_M)} = \sum_{i=1}^{n} x_i \beta_i$$

(ii) ***Equations for model and variance***

According to the single-index model:

$$R_P = \alpha_P + \beta_P R_M + \varepsilon_P$$

Taking variances of both sides gives:

$$\begin{aligned} \operatorname{var}(R_P) &= \operatorname{var}(\alpha_P + \beta_P R_M + \varepsilon_P) \\ &= \beta_P^2 \operatorname{var}(R_M) + \operatorname{var}(\varepsilon_P) \end{aligned}$$

We have used the facts that:

- α_P is a constant
- ε_P and R_M are uncorrelated.

(iii) ***Why specific risk is "diversifiable"***

Modelling portfolio returns using the single-index model is usually based on the underlying assumption that each of the individual assets i also follows a single-index model of the form:

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

where ε_i and ε_j are uncorrelated when $i \neq j$.

So: $$R_P = \sum_{i=1}^{n} x_i R_i = \sum_{i=1}^{n} x_i \alpha_i + \beta_P R_M + \sum_{i=1}^{n} x_i \varepsilon_i$$

So ε_P corresponds to $\sum_{i=1}^{n} x_i \varepsilon_i$ and it follows (since ε_i and ε_j are uncorrelated) that:

$$\text{var}(\varepsilon_P) = \text{var}\left(\sum_{i=1}^{n} x_i \varepsilon_i\right) = \sum_{i=1}^{n} x_i^2 \text{ var}(\varepsilon_i)$$

If we select a portfolio with equal proportions of each asset, *ie* $x_i = \frac{1}{n}$, then:

$$\text{var}(\varepsilon_P) = \sum_{i=1}^{n} \frac{1}{n^2} \text{var}(\varepsilon_i) = \frac{1}{n} \times \left(\frac{1}{n} \sum_{i=1}^{n} \text{var}(\varepsilon_i)\right) = \frac{1}{n} \times \overline{V}$$

where $\overline{V}$ denotes the average specific risk of the assets in the portfolio.

So, as $n \to \infty$, the variance $\text{var}(\varepsilon_P)$ will tend to zero.

So this component of the variance of returns can be reduced to a very small level by selecting a sufficiently diversified portfolio. It is therefore often called the "diversifiable risk".

(iv) ***Zero beta versus risk-free***

The statement is not correct.

A zero beta indicates no systematic risk.

A portfolio will only be totally risk-free if it has zero variance, *ie* if $\text{var}(R_P) = 0$.

From the equation $\text{var}(R_P) = \beta_P^2 \text{var}(R_M) + \text{var}(\varepsilon_P)$, we see that, even if $\beta_P = 0$, the overall variance will only be zero if $\text{var}(\varepsilon_P) = 0$, as well.

But, any (non-trivial) portfolio of risky assets will have a non-zero specific risk, *ie* $\text{var}(\varepsilon_P) > 0$.

However, it is theoretically true that a well-diversified portfolio with a beta of zero will be approximately risk-free.

4 End of Part 1

What next?

1. Briefly **review** the key areas of Part 1 and/or re-read the **summaries** at the end of Chapters 1 to 6.
2. Attempt some of the questions in Part 1 of the **Question and Answer Bank**. If you don't have time to do them all, you could save the remainder for use as part of your revision.
3. Attempt **Assignment X1**.

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

Chapter 6 Summary

The multifactor model

A multifactor model of security returns attempts to explain the observed historical return by an equation of the form:

$$R_i = a_i + b_{i,1}I_1 + b_{i,2}I_2 + \ldots + b_{i,L}I_L + c_i \ ,$$

where:

- R_i is the return on security i,
- a_i, c_i are the constant and random parts respectively of the component of return unique to security i
- $I_1, \ldots, I_L$ are the changes in a set of L factors which explain the variation of R_i about the expected return a_i
- $b_{i,k}$ is the sensitivity of security i to factor k.

Types of multifactor model

Macroeconomic – the factors are the main macroeconomic variables such as interest rates, inflation, economic growth and exchange rates.

Fundamental – the factors will be company specifics such as P/E ratios, liquidity ratios and gearing measurements.

Statistical – the factors are not specific items initially. The method uses *principal component analysis* and historical returns on stocks to decide upon the factors.

Single-index model

This has just a single factor, which is usually the return on the market, R_M .

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

where:

- $$\beta_i = \frac{Cov(R_i, R_M)}{V_M}$$

It can then be shown that for any security *i*:

- $E_i = \alpha_i + \beta_i E_M$
- $V_i = \beta_i^2 V_M + V_{\varepsilon_i}$ *ie* total risk is the sum of *systematic risk* and *specific risk*
- $C_{ij} = \beta_i \beta_j V_M$ *ie* securities "covary" only through their covariance with the market

Likewise, the variance of portfolio returns is equal to:

$$V_P = \sum_{i=1}^{N} x_i^2 V_{\varepsilon i} + \beta_p^2 V_M$$

where:

- x_i is the weight in security *i*
- $$\beta_P = \sum_i^N x_i \beta_i$$

Chapter 6 Solutions

Solution 6.1

In a two-factor model:

$$R_i = a_i + b_{i,1} I_1 + b_{i,2} I_2 + c_i$$

In this case:

$$a_i = 1.0 \quad b_{i,1} = 0.8 \quad b_{i,2} = -0.3$$

Thus:

$$\begin{aligned} E(R_i) &= E\left[a_i + b_{i,1}I_1 + b_{i,2}I_2 + c_i\right] \\ &= a_i + b_{i,1}E\left[I_1\right] + b_{i,2}E\left[I_2\right] + E\left[c_i\right] \\ &= 1.0 + 0.8 \times 3 - 0.3 \times 2.2 + 0 \\ &= 2.74\% \end{aligned}$$

Solution 6.2

The *systematic return* of security i is the element of the total return that arises due to the influences of the L factors that influence the returns on every security. It is equal to:

$$b_{i,1}I_1 + b_{i,2}I_2 + ... + b_{i,L}I_L$$

The *specific return* of security i is the element of the total return that is independent of the L factors and hence independent of the returns on all other securities. It is therefore specific or unique to security i and is equal to:

$$a_i + c_i$$

Solution 6.3

Intuitively, the less correlated the factors are, the easier it is to disentangle the influences of each upon security returns. If they are highly correlated, then they act in unison and have insufficient independent variation to enable the model to isolate their separate influences.

Within a regression context, this problem is known as *multicollinearity* and it has the effect of making the coefficient estimates less efficient.

Solution 6.4

In this case it is easier to carry out the process the other way round, because he's regressed I_1 on I_2, not I_2 on I_1 as before.

So, first define two new variables:

(1) $I_2^* = I_2$

(2) $I_1^* = I_1 - 0.8 - 0.3I_2$

where I_1^* is equal to the residuals from the regression of I_1 on I_2, which by definition are independent.

We now need to re-express R_i in terms of the new variables I_1^* and I_2^*. It follows from (1) and (2) that:

$$I_1^* = I_1 - 0.8 - 0.3I_2^*$$

from which:

(3) $I_1 = I_1^* + 0.8 + 0.3I_2^*$

Using (1) and (3) to substitute for I_1 and I_2 in the original model then gives:

$$R_i = 2 + 1.3\left(I_1^* + 0.8 + 0.3I_2^*\right) + 0.8I_2^* + c_i$$

ie $R_i = 3.04 + 1.3I_1^* + 1.19I_2^* + c_i$

Note that in general the process can be carried out either way around, although a different, but equally acceptable, answer will be obtained in each case.

Solution 6.5

α_i can be interpreted as the expected value of the component of security i's return that is independent of the market's performance and specific to that particular security.

β_i quantifies the component of the security return that is directly related to movements in the market – so that if $\beta_i = x$, then security i's return is expected to increase (decrease) by x% when the market return increases (decreases) by 1%.

Solution 6.6

The expected return on security i is given by:

$$E(R_i) = E[\alpha_i + \beta_i R_M + \varepsilon_i]$$

By the linear additivity of expected values this can be written as:

$$E(R_i) = E(\alpha_i) + E(\beta_i R_M) + E(\varepsilon_i)$$

And since α_i and β_i are constants and α_i is chosen so that $E(\varepsilon_i) = 0$, we have:

$$E(R_i) = \alpha_i + \beta_i E_M$$

as required.

Solution 6.7

Risk equals systematic risk plus specific risk.

Solution 6.8

The variance of portfolio returns can be found as follows:

$$V_P = \sum_{i=1}^{N} x_i^2 V_i + \sum_{i=1}^{N} \sum_{j \neq i}^{N} x_i x_j C_{i,j}$$

where the x's are the portfolio weightings and the second summation is over $j \neq i$.

ie $$V_P = \sum_{i=1}^{N} x_i^2 (\beta_i^2 V_M + V_{\varepsilon i}) + \sum_{i=1}^{N} \sum_{j \neq i}^{N} x_i x_j (\beta_i \beta_j V_M)$$

Combining all of the variance and covariance terms for the systematic risk then gives:

$$V_P = \sum_{i=1}^{N} x_i^2 V_{\varepsilon i} + \sum_{i=1}^{N} \sum_{j=1}^{N} x_i x_j \beta_i \beta_j V_M$$

Note that the second summation now includes the case $i = j$. Collecting together the i and j terms in this summation then gives:

$$V_P = \sum_{i=1}^{N} x_i^2 V_{\varepsilon i} + \left(\sum_{i=1}^{N} x_i \beta_i \right) \left(\sum_{j=1}^{N} x_j \beta_j \right) V_M$$

$$V_P = \sum_{i=1}^{N} x_i^2 V_{\varepsilon i} + \beta_P^2 V_M$$

where $\beta_P = \sum_{i=1}^{N} x_i \beta_i$ which is the required expression.

Now, suppose that equal amounts of money are invested in each of the N securities, so that $x_i = 1/N$ for all $i = 1, \ldots, N$. Then the first term in the expression for V_P, which represents the contribution of the specific risk to the total portfolio variance, can be written as:

$$\frac{1}{N} \sum_{i=1}^{N} \frac{1}{N} V_{\varepsilon i}$$

ie $\frac{1}{N}$ times the average specific risk associated with each individual security V^*. As N increases and the average specific risk remains unchanged, so this term will rapidly become smaller. Hence, as $N \to \infty$, the specific risk tends to zero and:

$$V_P \;\to\; \beta_P^2 V_M$$

ie $$\sigma_P \;\to\; \beta_P \sigma_M \;=\; \sigma_M \sum_{i=1}^{N} x_i \beta_i$$

In other words, the contribution of the specific risk on each security to the total risk of the portfolio becomes very small and the contribution of each security to the portfolio's total risk is therefore only the systematic risk of that security.

Solution 6.9

In order to apply mean-variance portfolio theory with N securities, we need estimates of the following:

- N expected returns
- N variances
- $\frac{1}{2} \times N \times (N-1)$ correlation coefficients or covariances.

So, in total we need $\frac{1}{2} \times N \times (N+3)$ items of data. For example, if $N = 200$, then $\frac{1}{2} \times N \times (N+3) = 20{,}300$.

Under the single-index model we need:

- N values for the α_i's
- N values for the β_i's
- N values for the $V_{\varepsilon i}$'s
- the expected return E_M and variance V_M for the market.

Thus, in total we need $3N+2$ items of data, which with 200 securities amounts to 602 items of data.

Solution 6.10

The main uses include:

- Determination of the investor's efficient frontier, as part of the derivation of the investor's optimal portfolio.

- Risk control – by enabling the investor to forecast the variability of portfolio returns both absolutely and relative to some benchmark. For example, by constructing a portfolio whose sensitivities to the relevant factors are the same as the benchmark, it is possible to reduce the risk of under- or over-performance compared to that benchmark.

- Performance analysis – by comparing the actual performance of the portfolio to that predicted by the model and based on the portfolio's actual exposure to the relevant factors over the period considered.

- Categorisation of investment styles – according to the extent of the exposure to particular factors.

The main limitation is that the construction of factor models is based on historical data that reflect conditions that may not be replicated in the future. Moreover, a model that does produce good predictions in one time period, may not produce good predictions in subsequent time periods.

Note that this particular limitation applies equally to many of the other investment models in common use, including the capital asset pricing model and the arbitrage pricing theory discussed in the next chapter, whose parameters are typically estimated using past data.

Solution 6.11

This is *false*. If β_i doubles then, the expected return on security *i* will not also double, as the specific risk, α_i and ε_i, will be unchanged. Thus, if:

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

then:

$$R_i' = \alpha_i + 2\beta_i R_M + \varepsilon_i = R_i + \beta_i R_M \neq 2R_i$$

Chapter 7

Asset pricing models

Syllabus objectives

(v) Describe asset pricing models, discussing the principal results and assumptions and limitations of such models.

1. *Describe the assumptions and the principal results of the Sharpe-Lintner-Mossin Capital Asset Pricing Model (CAPM).*

2. *Discuss the limitations of the basic CAPM and some of the attempts that have been made to develop the theory to overcome these limitations.*

3. *Discuss the assumptions, principal results and limitations of the Ross Arbitrage Pricing Theory model (APT).*

4. *Perform calculations using the CAPM.*

0 Introduction

Mean-variance portfolio theory and the multifactor and single-index models, discussed previously, showed how an individual investor can characterise the relationship between risk and return for a particular security. The *asset pricing models* described in this chapter extend the ideas discussed previously, in an attempt to characterise the entire investment market on the assumption that investors behave exactly as predicted by those models. In this chapter, we introduce two such models, namely the *capital asset pricing model* and *arbitrage pricing theory*.

The capital asset pricing model tells us about the relationship between risk and return in the security market as a whole, assuming that investors act in accordance with mean-variance portfolio theory and that the market is in equilibrium. The arbitrage pricing theory model is likewise based upon the foundation provided by multifactor models.

Note that such asset pricing models are also sometimes referred to as *equilibrium models*, because they characterise the equilibrium outcome in investment markets.

1 The capital asset pricing model (CAPM)

1.1 Introduction

A single investor given his own estimates of security returns, variances and covariances can apply portfolio theory. The capital asset pricing model developed by Sharpe, Lintner and Mossin introduces additional assumptions regarding the market and the behaviour of other investors to allow the construction of an equilibrium model of prices in the whole market.

Hence, the assumption here is that all investors select their investments by applying the ideas and assumptions underlying mean-variance portfolio theory.

Question 7.1

List the assumptions underlying mean-variance portfolio theory.

If this is the case, then the introduction of some additional assumptions enables us to characterise how investors may act in aggregate and thereby construct an equilibrium model of security prices. The resulting capital asset pricing model (CAPM) tells us about the relationship between risk and return for security markets as a whole.

1.2 Assumptions

The extra assumptions of CAPM are:

- **All investors have the same one-period horizon.**
- **All investors can borrow or lend unlimited amounts at the same risk-free rate.**
- **The markets for risky assets are perfect. Information is freely and instantly available to all investors and no investor believes that they can affect the price of a security by their own actions.**
- **Investors have the same estimates of the expected returns, standard deviations and covariances of securities over the one-period horizon.**
- **All investors measure in the same "currency" *eg* pounds or dollars or in "real" or "money" terms.**

A number of conditions need to be met for an investment market to be perfect – basically there must be no anomalies or distortions in the pricing of assets. The following are the basic requirements for a perfect market:

- There are many buyers and sellers, so that no one individual can influence the market price.
- All investors are perfectly informed.
- Investors all behave rationally.
- There is a large amount of each type of asset.
- Assets can be bought and sold in very small quantities, *ie* perfect divisibility.
- There are no taxes.
- There are no transaction costs.

Not all of the above assumptions are 100% realistic. However, the fact that the assumptions do not hold does not necessarily invalidate the CAPM, as it may nevertheless yield useful insight into the operation of security prices and returns. We must recognise, however, that it will be only an approximation.

1.3 Consequences of the extra assumptions

Given the extra assumptions above, we can build on the results of mean-variance portfolio theory to develop the standard form of the capital asset pricing model as follows.

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

> ***Question 7.2***
>
> Why?

2. **If in addition they are all subject to the same risk-free rate of interest, the efficient frontier collapses to the straight line in $E-\sigma$ space which passes through the risk-free rate of return on the *E*-axis and is tangential to the efficient frontier for risky securities** (See Figure 7.1 on the next page).

Question 7.3

Suppose there are only two portfolios A and B available to invest in. A is a portfolio of risky assets and B is a portfolio consisting of just one risk-free asset. Show that the efficient frontier must be a straight line.

So, if there is a risk-free asset, the efficient frontier must be a straight line.

Question 7.4

Explain why the new efficient frontier must be at a tangent to the old efficient frontier of risky assets.

All investors face the *same* straight-line efficient frontier in expected return-standard deviation space.

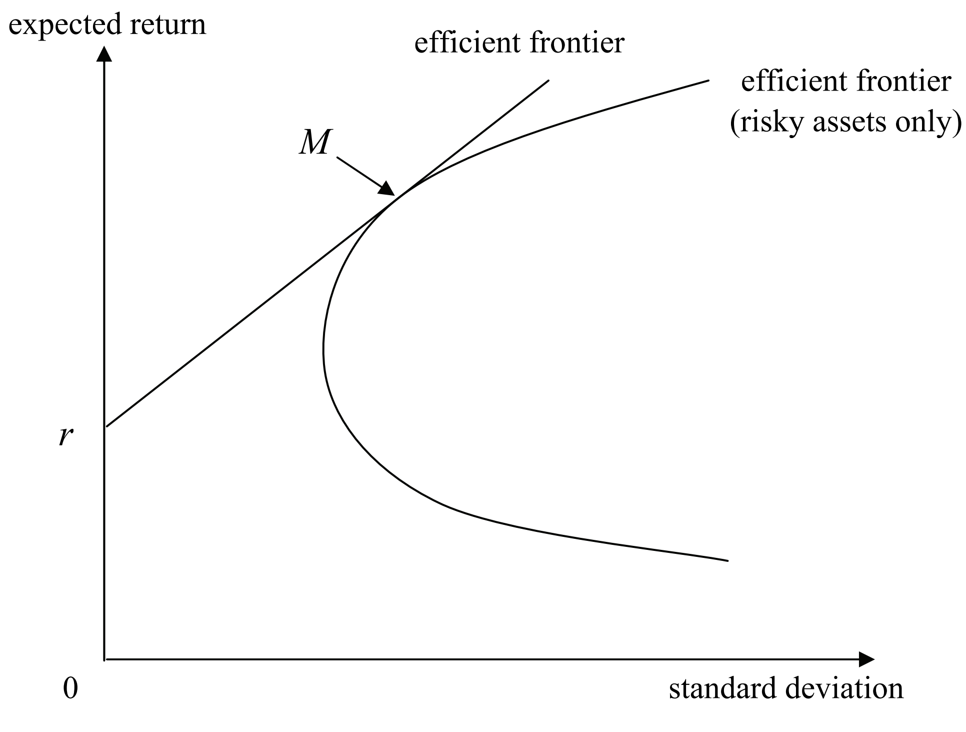


Figure 7.1: The efficient frontier with risk-free lending and borrowing

Note that:

- r represents the risk-free asset (which has zero standard deviation).

- The efficient frontier with a risk-free asset is *parabolic* in *expected return-variance* space. The straight line above in *expected return-standard deviation space* is a degenerate case of a hyperbola.

- The model does not require that investors all have the same attitude to risk, only that their views of the available securities are the same – and hence that the opportunity set is identical for all investors.

A rigorous derivation of the capital asset pricing model also exists.

3. All rational investors will hold a combination of the risk-free asset and *M*, the portfolio of risky assets at the point where the straight line through the risk-free return touches the original efficient frontier.

This is because if investors are rational they should only invest in efficient portfolios, which are located along the straight-line efficient frontier. Every investor should choose a portfolio of the form:

a% of the risk-free asset + $(100–a)$% of M

The portfolio of risky assets is shown as M on Figure 7.1. The choice of a depends on the investor's level of aversion to risk. Note that it could be negative for an investor who has low risk aversion, so that the investor's optimal portfolio lies on the efficient frontier to the right of M. This would mean that the investor is borrowing the risk-free asset and investing in the portfolio of risky assets M.

4. Because this is the portfolio held in different quantities by all investors it must consist of all risky assets in proportion to their market capitalisation. It is commonly called the "market portfolio". The proportion of a particular investor's portfolio consisting of the market portfolio will be determined by their risk-return preference.

This is a key result of the CAPM. It is important to realise that M *is* the market portfolio, rather than the market portfolio just being a name we give to M.

Example

Suppose that the market of risky assets being considered is that consisting of FTSE 100 companies only and that there are 15 million investors. The market portfolio is then the FTSE 100 index itself. If Investor 1 has 5% of their portfolio of risky assets in Vodafone shares then Investor 2 must also have the same percentage and so on up to Investor 15,000,000. CAPM says that every single investor holds 5% of their portfolio of risky assets in Vodafone shares and so, because these investors cover the whole market, the market must hold 5% of its risky assets in Vodafone, *ie* Vodafone's market capitalisation would be 5% of all FTSE 100 shares.

If you apply this logic across all companies in the FTSE 100 then it becomes clear that every investor must hold shares in proportion to their market capitalisation. In this case every investor's portfolio of risky assets would be the FTSE 100 index, *ie* the market portfolio.

1.4 The separation theorem

The fact that the optimal combination of risky assets for an investor can be determined without any knowledge of their preferences towards risk and return (or their liabilities) is often known as the *separation theorem*.

The separation theorem doesn't tell us whereabouts on the new efficient frontier an individual investor's portfolio will be. For this we need to know the investor's attitude towards risk and return, or equivalently the investor's utility function, in order to determine their preferred split between risky assets and the risk-free asset, *ie* the value of a.

However, we no longer have to make thousands of estimates of covariances in order to determine the portfolio of risky assets, because we know that it is always the market portfolio, M.

1.5 The capital market line

The straight line denoting the new efficient frontier is called the *capital market line* and its equation is:

$$E_P - r = (E_M - r)\sigma_P / \sigma_M$$

where:

- E_P **is the expected return of any portfolio on the efficient frontier**
- σ_P **is the standard deviation of the return on portfolio *P***
- E_M **is the expected return on the market portfolio**
- σ_M **is the standard deviation of the return on the market portfolio**
- r **is the risk-free rate of return.**

Thus the expected return on any efficient portfolio is a linear function of its standard deviation. The factor $(E_M - r)/\sigma_M$ is often called the *market price of risk*. (The market price of risk is also considered in Chapter 17.)

Question 7.5

Derive the above equation of the capital market line.

Note that:

- The expected return on any efficient portfolio P can be written as:

$$E_P = r + \left(\frac{E_M - r}{\sigma_M}\right)\sigma_P$$

ie expected return = risk-free rate + (market price of risk) × (amount of risk)

The "(price of risk) × (amount of risk)" term is sometimes referred to as the *risk premium*.

- The market price of risk is equal to the gradient of the capital market line in expected return-standard deviation space.

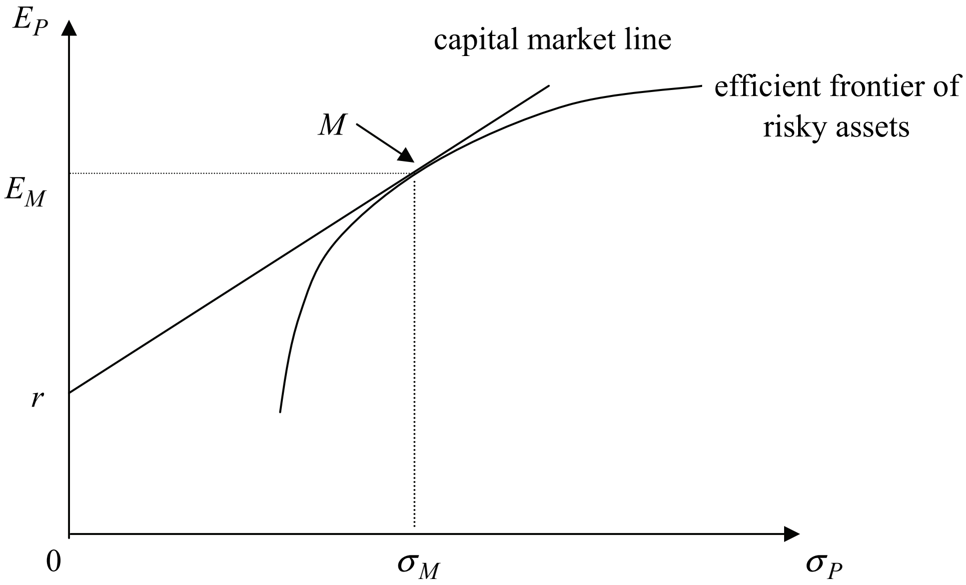


Figure 7.2: The capital market line

Hence, if the assumptions underlying the capital asset pricing model are true, then rational behaviour by investors should result in equilibrium security prices such that the expected return on any efficient portfolio is a *linear* function of its standard deviation. It is important to note that this result applies to efficient portfolios only and not to inefficient portfolios.

1.6 The security market line

It is also possible to develop an equation relating the expected return on *any* asset to the return on the market:

$$E_i - r = \beta_i \left(E_M - r\right)$$

where:

- E_i **is the expected return on security *i*,**
- r **is the return on the risk-free asset,**
- E_M **is the expected return on the market portfolio,**
- β_i **is the beta factor of security *i* defined as** $\text{cov}\left[R_i, R_M\right]/V_M$ **.**

The equation is often written as:

$$E_i = r + \beta_i \left(E_M - r\right)$$

This presentation emphasises that the expected return on any asset is again equal to the risk-free rate plus a risk premium, which here is derived from the beta of the security.

This is the equation of a straight line in $E-\beta$ space called the *security market line*. It shows that the expected return on any security can be expressed as a linear function of the security's covariance with the market as a whole. Since the beta of a portfolio is the weighted sum of the betas of its constituent securities, the security market line equation applies to portfolios as well as individual securities.

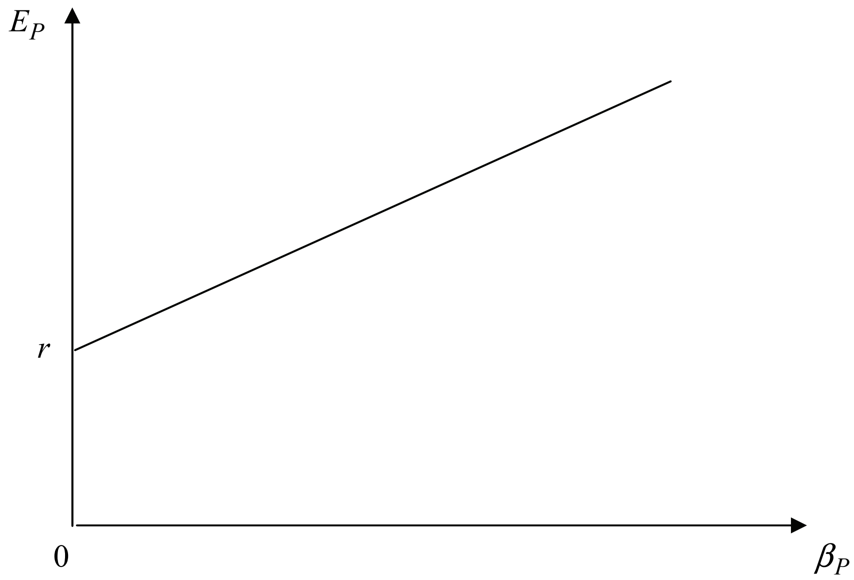


Figure 7.3: The security market line

Note that all portfolios, including those comprising a single security or asset, lie on the security market line whether or not they are efficient. Thus, according to the capital asset pricing model, the security market line relationship can be used to determine the expected return of *any* asset or portfolio from its beta. The expected return of a portfolio depends linearly upon its beta, which measures systematic risk and is independent of other non-systematic risk. Consequently, investors are rewarded only for systematic risk and not for non-systematic risk, precisely because they are able to diversify it away.

The security market line result has many diverse applications, a number of which go beyond the subject of investment. For example, in a life office you could calculate and use the market price of risk when choosing the risk discount rate for profit testing new contracts.

Question 7.6

Consider Security A, which has a standard deviation of investment returns of 4%. If:

- the standard deviation of the market return is 5%
- the correlation between A's return and that of the market is 0.75
- the risk-free rate is 5%
- and the expected return on the market is 10%

then calculate:

(i) the beta of Security A

(ii) Security A's expected return.

Derivation of the security market line

One way to derive the security market line is to note that according to the capital asset pricing model, all investors will hold the market portfolio of risky assets as part of their overall investment portfolio. Consequently, each investor will hold a *very* well-diversified portfolio.

When discussing systematic and specific risk earlier in the course we saw that, as a portfolio becomes very well-diversified:

- the systematic risk of the portfolio tends towards a weighted average of the systematic risks of the constituent securities
- the non-systematic or specific risk tends to zero.

So, an investor who holds a well-diversified portfolio, should be concerned only with expected return and systematic risk. Hence portfolio decisions should be based only upon the expected return and the beta of the portfolio.

Consider two well-diversified portfolios, 1 and 2, with expected returns E_1 and E_2 and betas β_1 and β_2. If we construct a third portfolio consisting of equal holdings of Portfolios 1 and 2, then its expected return and beta will be equal to:

$$E_3 = \tfrac{1}{2} \times (E_1 + E_2)$$

and $$\beta_3 = \tfrac{1}{2} \times (\beta_1 + \beta_2)$$

Question 7.7

Why?

Hence, if we plot Portfolios 1, 2 and 3 in expected return – beta space then they must all lie on the same straight line. A similar argument applies to all portfolios and individual securities, the straight line in question being the security market line. The general form of the equation of a straight line in expected return-beta space is (for any security *i*):

$$E_i = a_0 + a_1 \beta_i$$

Thus, to complete the derivation we need to determine the values of a_0 and a_1.

Question 7.8

What are the betas of the market portfolio and the risk-free asset? Use these beta values to complete the derivation of the security market line equation.

2 Limitations of CAPM

2.1 Limitations and empirical evidence

The limitations of the basic CAPM are well known and attempts have been made to overcome them since the model was first published. Most of the assumptions of the basic model can be attacked as unrealistic and, furthermore, empirical studies do not provide strong support for the model.

Models with less restrictive or unrealistic assumptions may therefore provide better predictions of the actual behaviour of security markets than the basic capital asset pricing model described above. A further subtle yet important issue is that if the model assumes that a particular factor such as tax does not exist, then it cannot tell us about the influence of taxation upon security markets. We might therefore wish to develop a model that explicitly introduces taxes.

There are basic problems in testing the model since, in theory, account has to be taken of the entire investment universe open to investors, not just capital markets.

An important asset of most investors, for example, is their human capital (*ie* the value of their future earnings).

In other words tests of the model should allow for every single asset that an investor could purchase and that yields an uncertain return to an investor, *eg* houses, works of art, training courses *etc*. It is therefore extremely difficult to test the validity of the capital asset pricing model if the returns on many assets cannot be observed. Moreover, even if actual returns are observable, the capital asset pricing model is expressed in terms of *expected* returns, which typically are not.

Nevertheless the studies that have been carried out do provide some evidence:

- in support of a linear relationship between return and systematic risk over long periods of time
- suggesting that return is *not* related to unsystematic or residual risk.

2.2 Extensions of the basic CAPM

Models have been developed which allow for decisions over multiple periods and for the optimisation of consumption over time to take account of this.

Other versions of the basic CAPM have been produced which allow for taxes and inflation, and also for a situation where there is no riskless asset.

Multi-period models

The basic capital asset pricing model assumes that investors are concerned only with a single-period time horizon. Various multi-period models have, however, been developed that consider how investors with preferences defined over a time horizon of more than one period make investment decisions over that time horizon. These models have shown that under certain assumptions, the investment decisions taken by investors with multi-period time horizons are consistent with those made by investors with just a single-period time horizon. Consequently, each single period can be considered in isolation and the results of the single period capital asset pricing model are equally valid (or not) in the multi-period context.

Amongst the multi-period models are:

- the *consumption capital asset pricing model* – which relates investment returns to the growth rate of per capita consumption
- one that allows for the uncertain *inflation* that will be present in a multi-period context – so that investors are concerned with *real* returns.

Model with taxes

The absence of taxes in the basic model means that investors are indifferent between income and capital gains. One extension of the basic capital asset pricing model therefore derives a somewhat more complicated equilibrium relationship allowing for differential taxation and based upon the means and variances of post-tax returns.

Zero-beta model

It can be shown that the absence of a risk-free asset does not alter the form of the security market line, the role of the risk-free asset simply being replaced by a "zero-beta" portfolio, *ie* a portfolio of risky assets with a beta equal to zero. The equation of the security market line is then:

$$E_i = E_Z + \left(E_M - E_Z\right)\beta_i$$

where E_Z is the expected return of the zero-beta portfolio. This is sometimes referred to as the *zero-beta* version of the capital asset pricing model.

In the international situation there is no asset which is riskless for all investors (due to currency risks) so a model has been developed which allows for groups of investors in different countries, each of which considers their domestic currency risk-free.

3 Arbitrage pricing theory

3.1 Introduction

Arbitrage pricing theory (APT), originally developed by Ross, is an equilibrium market model that does not rely on the strong assumptions of the CAPM.

For example, strong assumptions are no longer required concerning the form of investors' preferences – though an assumption of homogeneous expectations is still required. Nor is it necessary to identify all risky assets or a "market portfolio" in order to test it. As a consequence it is both more general and more easily testable than the CAPM, although there *are* still difficulties in testing its applicability.

Instead it is based upon the assumption of an investment world that is *arbitrage-free*.

Question 7.9

Define what is meant by the term *arbitrage*.

Before we look at the algebra underlying APT, you may find it helpful to consider the model in terms of basic concepts. In essence, there are two basic concepts to grasp:

1. The process of arbitrage ensures that when a market is in equilibrium there should be just one price for a given asset. Two items that are identical cannot have different prices – if they did, the forces of supply and demand created by arbitrageurs would quickly remove the price anomaly. This is the *no-arbitrage principle*, which we will use extensively in Part 3 of this course to value derivatives.

2. The return provided by each asset is a linear function of various factors or indices. The same factors apply for all assets, but different assets have different exposures or weightings to the various factors. In other words, asset returns are generated by multifactor models such as we discussed earlier in the course. As was the case then, the theory does not attempt to say what the factors may be, or even how many factors there are.

The no-arbitrage principle is relevant because any asset – or more precisely the returns generated by that asset – can in theory at least be replicated by appropriate combinations and weightings of other assets. This means that the price of each asset should be consistent with the price of the combinations of the other, replicating assets. The following simple question helps to illustrate this point.

Question 7.10

Suppose that just two factors, I_1 and I_2, are needed to describe the returns from assets. Asset A is heavily influenced by I_1, whereas Asset B is heavily influenced by I_2. Suppose also that the expected returns from each of A and B are given by the following equations:

$$E(R_A) = 4 + 2I_1 + I_2$$

$$E(R_B) = 5 + I_1 + 2I_2$$

Asset C has equal weighting on the two factors. Write down the equation for the expected return on Asset C.

Having summarised the general principles underlying APT, we can study the more formal description of the theory.

3.2 The theory

APT requires that the returns on any stock be linearly related to a set of indices as shown below:

$$R_i = a_i + b_{i,1}I_1 + b_{i,2}I_2 + \dots + b_{i,L}I_L + c_i$$

where:

- R_i **is the return on security** i
- a_i **and** c_i **are the constant and random parts respectively of the component of return unique to security** i
- $I_1, \dots, I_L$ **are the returns on a set of** L **indices**
- $b_{i,k}$ **is the sensitivity of security** i **to index** k**.**

We also have:

- $E[c_i] = 0$
- $E[c_i c_j] = 0$ **for all** $i \neq j$,
- $E[c_i(I_j - E[I_j])] = 0$ **for all stocks and indices.**

This is exactly the same as the multi-index model for returns on individual securities. The contribution of APT is to describe how we can go from a multi-index model for individual security returns to an equilibrium market model.

3.3 *Proof of the theory*

Non-mathematically, the argument can be made as follows. Consider a two-index model. The return on the *i*th security is given by:

$$R_i = a_i + b_{i,1}I_1 + b_{i,2}I_2 + c_i$$

For investors who hold well-diversified portfolios the specific risk of each security, represented by c_i, can be diversified away, so an investor need only be concerned with expected return, $b_{i,1}$ and $b_{i,2}$ in choosing his portfolio.

By diversifying their investment across a wide enough range of securities, the investor can eliminate any risk specific to individual securities. Consequently, they should be concerned only with expected return and systematic risk, as indicated by the b_i's, when choosing between different available portfolios. This is reasonable if there is a sufficiently large number of securities available.

Question 7.11 (Revision)

What is meant by systematic risk?

Suppose we hypothesise the existence of three widely-diversified portfolios, represented by the points $(E_i, b_{i,1}, b_{i,2})$ in $E - b_1 - b_2$ space where $i = 1, 2, 3$.

The three portfolios are therefore represented by the three points $P_1 = (E_1, b_{1,1}, b_{1,2})$, $P_2 = (E_2, b_{2,1}, b_{2,2})$ and $P_3 = (E_3, b_{3,1}, b_{3,2})$ in three-dimensional space, with E, b_1 and b_2 as the axes.

Question 7.12

Why is this representation of the three points valid?

These three portfolios define a plane in $E - b_1 - b_2$ space with equation:

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

Question 7.13

Explain this equation.

A portfolio having any combination of b_1 and b_2 can be formed by combining portfolios 1, 2 and 3 in the correct proportions. For example the portfolio *P*, obtained by taking one-third of each of 1, 2 and 3 would have:

$$b_{P,1} = \left(b_{1,1} + b_{2,1} + b_{3,1}\right)/3$$

$$b_{P,2} = \left(b_{1,2} + b_{2,2} + b_{3,2}\right)/3$$

and $$E\left[R_P\right] = \lambda_0 + \lambda_1 b_{P,1} + \lambda_2 b_{P,2}$$

So, P is a point on the plane defined by P_1, P_2 and P_3 that lies between the points P_1, P_2 and P_3.

Now, consider what would happen if another portfolio *Q* existed, with exactly the same values of b_1 and b_2 but a higher expected return. Both portfolios would have the same degree of systematic risk but *Q* would have a higher expected return than *P*.

Rational investors would therefore sell *P* and buy *Q*, and this would continue until the forces of supply and demand had brought portfolio *Q* onto the same plane as portfolios 1, 2 and 3. Thus, in equilibrium, all securities and portfolios must lie on a plane in $E - b_1 - b_2$ space.

The crucial concept underpinning the arbitrage pricing theory is therefore the no-arbitrage principle, hence the name of the theory.

The general result

The more general result of APT, that all securities and portfolios have expected returns described by the *L*-dimensional hyperplane:

$$E_i = \lambda_0 + \lambda_1 b_{i,1} + \lambda_2 b_{i,2} + \ldots + \lambda_L b_{i,L}$$

can be derived by a more rigorous mathematical argument.

Note that:

- According to the arbitrage pricing theory you can calculate the expected return on a security without knowing about the future values of the indices. Instead you need to know the values of each $b_{i,k}$ – the sensitivity of security i's investment return to index k – and each λ_k

- Arbitrage pricing theory does not specify what the indices should be nor does it say anything about the values of the λ_k's. In testing the theory, therefore, both the indices and the λ_k's have to be extracted from the data, making interpretation of the tests difficult.

Question 7.14

What does each λ_k represent?

Question 7.15

An investor is trying to predict the expected return on a security using the arbitrage pricing theory. She decides that a two-factor model is appropriate and estimates that:

$$\lambda_0 = 3.0\% \quad \lambda_1 = 0.9\% \quad \lambda_2 = 0.8\%$$

$$b_{i,1} = 0.5 \quad b_{i,2} = 1.5$$

(i) $b_{i,2}$ is numerically greater than $b_{i,1}$ – what does this imply?

(ii) Estimate the expected return on the security.

3.4 Conclusion

The principal strength of the APT approach is that it is based on the no-arbitrage conditions. Another important characteristic is that it is extremely general, but this 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. Furthermore, APT tells us nothing about the size or the signs of the λ's. Until we actually fit the model to a set of data, it's difficult to guess what values we'll get for the λ parameters.

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.

3.5 Further reading

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

4 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.

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:

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

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:

$$\begin{aligned} E_A &= r + \beta_A(E_M - r) \\ &= 0.05 + 0.6(0.10 - 0.05) \\ &= 0.08 \end{aligned}$$

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:

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

Chapter 8

Brownian motion and martingales

Syllabus objectives

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

1. Explain the definition and basic properties of standard Brownian motion (or Wiener process).

0 Introduction

Essentially, a stochastic process is a sequence of values of some quantity where the future values cannot be predicted with certainty. This and the following chapter are concerned with continuous-time stochastic processes that have applications in financial economics. These chapters are of a very mathematical nature and you may find some of it hard-going. It is more important that you gain a higher-level understanding than that you master the pure maths that underlies it. For example, you may find it useful to learn Ito's lemma as a procedure rather than trying to understand the pure mathematical concepts.

The most important process studied here is Brownian motion, which is the subject of Section 1. We define this process as having stationary, independent and normally distributed increments. It turns out that this implies that its sample paths are continuous. In fact, a Brownian motion is the continuous-time version of a random walk, as we will see. The graph at the start of Section 1 shows a typical sample path.

If security prices can be modelled in some way in terms of Brownian motion, this will be useful for pricing certain types of options. This is discussed further in Parts 3 and 4 of the course.

Section 2 of this chapter introduces martingales. A martingale is a process whose current value is the best estimate of its future values. We will see later that martingale theory has important applications in relation to financial derivatives.

In order to specify precisely what we mean by the "best estimate", we need to work from the definition of a conditional expectation. Unfortunately when we are dealing with processes operating in continuous-time, this gets a bit "mathematical" and involves some technical definitions. This is another area where it is the overall approach that is important, rather than the technical details.

For an alternative presentation of some of these ideas, you *may* find it useful to consult the book by Baxter and Rennie listed in the Study Guide.

The notation used in financial economics generally is not standardised and similar notation can refer to different quantities: readers should check the definitions provided in each section.

1 Introduction to Brownian motion

1.1 Introduction

The phenomenon of Brownian motion is named after the nineteenth century botanist Robert Brown (who observed the random movement of pollen particles in water) because the path of a two-dimensional Brownian motion process bears a resemblance to the track of a pollen particle. The stochastic process was used by Bachelier to model the movements of the Paris stock exchange index.

Brownian motion is a continuous-time stochastic process with a continuous state space.

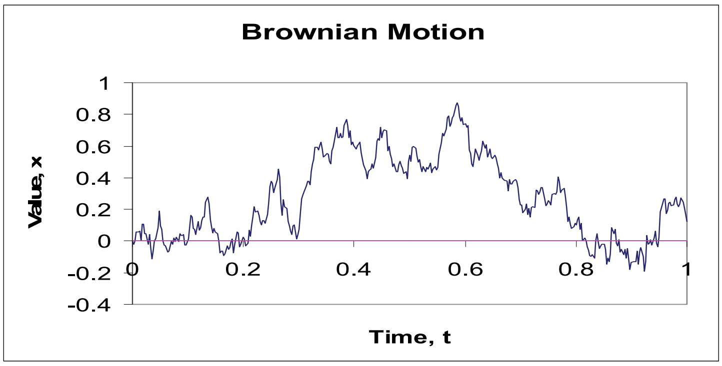


Figure 8.1: a typical trajectory (sample path) **of Brownian motion**

The simplest form of Brownian motion is *standard* Brownian motion. **It can be seen as the continuous version of a simple symmetric random walk.**

1.2 Definition of standard Brownian motion

Standard Brownian motion (also called the Wiener process) is a stochastic process $\{B_t, t \geq 0\}$ with state space $S = R$ (the set of real numbers) **and the following defining properties:**

(i) **B_t has independent increments, *ie* $B_t - B_s$ is independent of $\{B_r, r \leq s\}$ whenever $s < t$**

It follows that the changes in the value of the process over any two non-overlapping periods are statistically independent.

(ii) **B_t has stationary increments, *ie* the distribution of $B_t - B_s$ depends only on $t - s$**

(iii) B_t **has Gaussian increments, *ie* the distribution of** $B_t - B_s$ **is** $N(0, t-s)$

"Gaussian" is the official name for the normal distribution.

(iv) B_t **has continuous sample paths** $t \to B_t$

This means that the graph of B_t as a function of t doesn't have any breaks in it.

(v) $B_0 = 0$

1.3 Brownian motion in general

Standard Brownian motion is a special case of the more general form of Brownian motion.

The term Brownian motion refers to a process $\{W_t, t \geq 0\}$ **that satisfies (i), (ii) and (iv) above and also (iii) the distribution of** $W_t - W_s$ **is** $N[\mu(t-s), \sigma^2(t-s)]$**.**

Here "$\mu(t-s)$" means "μ multiplied by $(t-s)$", not that μ is a function of $t-s$.

Here μ **is the drift coefficient and** σ **the diffusion coefficient** (or volatility).

Standard Brownian motion is obtained when $\mu = 0$ and $\sigma = 1$.

It is not an easy task to prove that the above properties are mutually compatible. A surprising fact (which shows how tight the constraints are) is that either condition (iii) or condition (iv) can be dropped from the definition as it can be shown to be a consequence of the other three properties. In particular, Brownian motion is the *only process with stationary independent increments and continuous sample paths*.

This is far from obvious and we won't prove it here.

The relationship between standard Brownian motion and Brownian motion is the same as the relationship between a standard normal distribution, $N(0,1)$**, and a general** $N(\mu, \sigma^2)$ **distribution. A Brownian motion with given diffusion and drift coefficients can be constructed out of a standard Brownian motion** $\{B_t, t \geq 0\}$ **by setting:**

$$W_t = W_0 + \sigma B_t + \mu t$$

Question 8.1

Let B_t be a standard Brownian motion. Prove that $W_t = W_0 + \sigma B_t + \mu t$ is a Brownian motion with diffusion coefficient σ and drift μ.

Question 8.2

How can a Brownian motion W_t that has drift μ and diffusion parameter σ and a starting value of W_0 be converted into a standard Brownian motion?

1.4 Properties of Brownian motion

Standard Brownian motion has a number of other properties inherited from the simple symmetric random walk.

You may recall the simple symmetric random walk if you have previously studied Subject CT4. A simple symmetric random walk is a discrete-time stochastic process:

$$X_n = \sum_{i=1}^{n} Z_i \text{ where } Z_i = \begin{cases} +1 & \text{with probability } \frac{1}{2} \\ -1 & \text{with probability } \frac{1}{2} \end{cases}$$

The value of the process increases or decreases randomly by 1 unit (= "simple") with equal probability (= "symmetric").

We will see an example in Section 2 of a simple symmetric random walk (the man on the "never-ending" ladder).

If we reduce the step size progressively from 1 unit until it is infinitesimal (and rescale the X values accordingly), the simple symmetric random walk becomes standard Brownian motion.

Many of the properties of standard Brownian motion can be demonstrated using the following decomposition. For $s < t$:

$$\boldsymbol{B_t = B_s + (B_t - B_s)} \tag{6.1}$$

a decomposition in which the first term is known at time s and the second is independent of everything up until time s. In calculations involving Brownian motion, we often need to split up B_t in this way, so that we are working with *independent* increments.

(vi) $\boldsymbol{\text{cov}(B_s, B_t) = \min(s,t)}$

since, for $s < t$, $\text{cov}(B_s, B_s) = s$ and $\text{cov}(B_s, B_t - B_s) = 0$

Here we've used the decomposition just described:

$$\text{cov}(B_s, B_t) = \text{cov}[B_s, B_s + (B_t - B_s)] = \text{cov}(B_s, B_s) + \text{cov}(B_s, B_t - B_s)$$

(vii) $\boldsymbol{\{B_t, t \geq 0\}}$ **is a Markov process: this follows directly from the independent increment property**

If you have studied Subject CT4 you will be familiar with the concept of a Markov process. Intuitively, a Markov process is one where, if you know the latest value of the process, you have all the information required to determine the probabilities for the future values. Knowing the historical values of the process as well would not make any difference.

(viii) $\boldsymbol{\{B_t, t \geq 0\}}$ **is a martingale: (6.1) demonstrates that $E(B_t \mid F_s) = B_s$**

We will discuss martingales in detail later in this chapter.

(ix) $\boldsymbol{\{B_t, t \geq 0\}}$ **returns infinitely often to 0, or indeed to any other level**

Standard Brownian motion possesses sample path properties of its own:

(x) If $\{B_1(t), t \geq 0\}$ is defined by $B_1(t) = \frac{1}{\sqrt{c}} B_{ct}$ then $\{B_1(t), t \geq 0\}$ is also a standard Brownian motion. (This is the *scaling property* of Brownian motion.)

Question 8.3

Show that $B_1(t) - B_1(s) \sim N(0, t-s)$ for $s < t$.

(xi) If $\{B_2(t), t \geq 0\}$ is defined by $B_2(t) = tB_{1/t}$ then $\{B_2(t), t \geq 0\}$ is also a standard Brownian motion. (This is the time inversion property of Brownian motion.)

The law of any Gaussian stochastic process is completely determined by its expectation and its covariance function.

So if we know that a stochastic process has normally distributed ("Gaussian") increments and we know the first two moments of those increments, then we can determine all the statistical properties of that process.

To prove (x) and (xi) it is therefore sufficient to demonstrate that B_1 and B_2 have the correct first and second moments. Both evidently have expectation zero:

For $s < t$, we have:

$$\text{Cov}(B_1(s), B_1(t)) = \frac{1}{c}\text{Cov}(B_{cs}, B_{ct}) = \frac{1}{c}(cs) = s$$

and $$\text{Cov}(B_2(s), B_2(t)) = st\,\text{Cov}(B_{1/s}, B_{1/t}) = st(1/t) = s$$

since $1/t < 1/s$.

1.5 Non-differentiability of sample paths

A result frequently quoted in books on Brownian motion is that (with probability 1) the sample path of a Brownian motion is not differentiable anywhere.

A function is "differentiable" at a particular point if its graph has a well-defined slope at that point. Brownian motion, however, is too "squiggly" for this to be true.

The proof of this is beyond the scope of the syllabus, but a weaker result may be proved to illustrate the usefulness of the time inversion property.

Result 8.1

For any given time t_0, the probability that the sample path of a standard Brownian motion is differentiable at t_0 is zero.

At first sight, you might think that these two results are the same. If we can prove it's not differentiable at any chosen point t_0, then doesn't this mean that it's not differentiable *anywhere*? But, in fact, because of the way probabilities work with infinite sets, these are not the same. The following example illustrates the point – and shows how careful you need to be when dealing with processes such as Brownian motion.

Example

If I choose a value at random from the $N(0,1)$ distribution and you choose a *particular* number (t_0, say), then the probability I have chosen your number will always be zero. But the probability I have chosen *some* number is 1.

Although the probability might be zero for each individual value t_0, this doesn't necessarily mean that it will be zero when we consider *all* possible values of t_0.

Proof

We may assume that $t_0 = 0$, as otherwise we need only consider $B_t^* = B_{t+t_0} - B_{t_0}$, which is a standard Brownian motion.

If the graph is differentiable at this point, then if we look at a small enough region just after this time (times in the range $(0,h)$, say), then the slopes formed by joining points on the graph in this region must be very nearly constant. We can express this mathematically as follows.

If B is to be differentiable at 0, with derivative a, say, it must certainly be the case that there is some positive h and some positive δ such that:

$$(a-\delta)s < B_s < (a+\delta)s \quad \text{for all } 0 < s < h$$

If we divide through by s, this is the same as saying:

$$a-\delta < \frac{B_s - B_0}{s} < a+\delta$$

In other words, the slope of the line from the origin to the point (s, B_s) can be "sandwiched" between $a-\delta$ and $a+\delta$.

Writing $t = 1/s$, this statement becomes:

$$a - \delta < t\,B_{1/t} < a + \delta \text{ for all } t > 1/h$$

By the time-inversion property, however, $\{t\,B_{1/t}, t > 0\}$ is a standard Brownian motion, and we know that the probability that a standard Brownian motion remains confined to the region $(a-\delta, a+\delta)$ for sufficiently large t is zero.

Question 8.4

Express the probability that a standard Brownian motion is confined to the region $(a-\delta, a+\delta)$ at time t in terms of the standard normal distribution.

1.6 Geometric Brownian motion

As mentioned at the start of this chapter, Brownian motion was used by Bachelier to model the movements of the Paris stock exchange index.

However successful the Brownian motion model may be for describing the movement of market indices in the short run, it is useless in the long run, if only for the reason that a standard Brownian motion is certain to become negative eventually. It could also be pointed out that the Brownian model predicts that daily movements of size 100 or more would occur just as frequently when the process is at level 100 as when it is at level 10,000.

A more useful model is:

$$S_t = e^{W_t}$$

where W is the Brownian process $W_t = W_0 + \sigma B_t + \mu t$. Thus S_t, which is called geometric Brownian motion, is lognormally distributed with parameters $W_0 + \mu t$ and $\sigma^2 t$.

In other words, the values of $\log S_t$ are normally distributed with mean $W_0 + \mu t$ and variance $\sigma^2 t$.

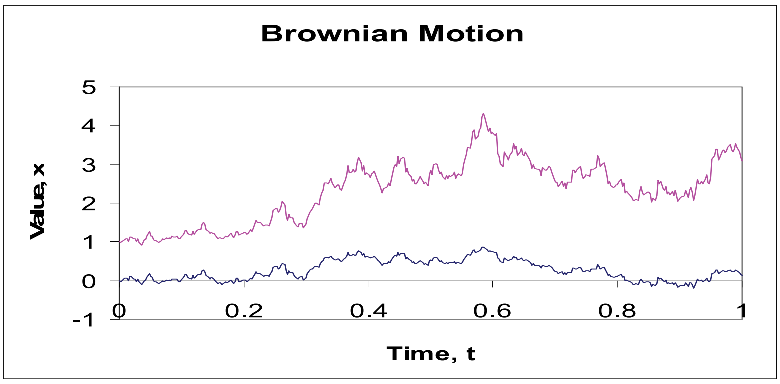


Figure 8.2: a Brownian motion B_t and a geometric BM, $S_t = e^{B_t}$

The most important property of S is:

$$S_t \geq 0 \text{ for all } t$$

From the properties of the lognormal distribution we also have:

$$E(S_t) = \exp[(W_0 + \mu t) + \tfrac{1}{2}\sigma^2 t] \quad \text{and} \quad \text{var}(S_t) = [E(S_t)]^2\{\exp(\sigma^2 t) - 1\}$$

Geometric Brownian motion features heavily in Subject CT8. We will discuss it further in this chapter and you will also meet it again in Chapter 9 and extensively in Parts 3 and 4 of the course. For example, Black and Scholes' Nobel prize-winning formula for pricing European options assumes that the price of the underlying asset is a geometric Brownian motion. This is discussed in detail in Parts 3 and 4 of the course.

The properties of S are less helpful than those of Brownian motion. For example, S has neither independent increments nor stationary increments.

The increments of S are of the form $S_t - S_s = e^{W_t} - e^{W_s}$.

But this is not so important because B does possess these desirable properties. Analysis of path properties of S should involve two steps: first, take the logarithm of the observations; second, perform the analysis using techniques appropriate to Brownian motion.

The log-return $\log \dfrac{S_t}{S_s}$ from time s to time t is given by $\log \dfrac{S_t}{S_s} = \log \dfrac{e^{W_t}}{e^{W_s}} = W_t - W_s$.

It follows by the independent increments property of Brownian motion that the log-returns, and hence the returns themselves, are independent over disjoint time periods.

2 Martingales

2.1 Introduction

In simple terms, a martingale is a stochastic process for which its current value is the optimal estimator of its future value.

So, the expected future value is the current value. Other ways of thinking of a martingale are that the expected change in the process is zero or that the process has "no drift".

> ***Note***
>
> Throughout this course we will be using the word "expected" in its statistical sense, rather than in the everyday sense.
>
> Consider a man standing on a "never-ending" ladder. Every minute he moves up or down the ladder one step, depending on whether a coin that his friend tosses, comes up heads or tails.
>
> In the everyday sense of the word, after the next toss of the coin, we "expect" him to move up or down (but we don't know which way). However, in the statistical sense, because $\frac{1}{2}\times(+1)+\frac{1}{2}\times(-1)=0$, we "expect" him to stay on the step he is currently standing on, even though there's no way that that can happen!

The idea of martingales is consistent with the original equestrian term "martingale", meaning a holster used to keep a horse "pointing straight ahead".

Their importance for modern financial theory cannot be overstated. In fact the whole theory of pricing and hedging of financial derivatives is formulated in terms of martingales.

For this reason, it may be best to think of a martingale as being a random process that has "no drift" because the idea of drift is more consistent with the way we think about real financial assets. We will see shortly that it is possible to model the log of a share price, $\log S_t$, using Brownian motion with a drift μ. You can think of μ as being the rate of the long-term drift of the log of the share price. It is the underlying non-random trend. It should not come as a great surprise that when we remove this underlying non-random trend (or drift) and look at $\log S_t - \mu t$, we obtain a martingale.

The features of martingales rely on the application of conditional expectations.

2.2 Conditional expectation

Recall from Subject CT3, that the conditional expectation of a random variable X, given $Y = y$, is defined to be the mean of the conditional distribution of X given $Y = y$, denoted by $E\left[X|Y = y\right]$. As y varies, so too will $E\left[X|Y = y\right]$ and we get the *random variable* $E\left[X \mid Y\right]$. This random variable is just a function of the random variable Y.

As $E\left[X \mid Y\right]$ is a random variable, we can take expectations in the usual way and this leads to the following fundamental result:

$$E\left[E\left[X \mid Y\right]\right] = E\left[X\right]$$

However, this definition of a conditional expectation is quite restrictive. It doesn't define what the conditional expectation means if we want to condition on the entire past history of any type of stochastic process.

The conditional expectation $E\left[X \mid Y\right]$ defined in Subject CT3 can be extended to conditioning on several random variables, yielding $E\left[X \mid Y_1, Y_2, \ldots, Y_n\right]$. The important property:

$$E\left[E\left[X \mid Y_1, Y_2, \ldots, Y_n\right]\right] = E\left[X\right]$$

is retained. In what follows the notation $\underline{Y}$ will be used to denote the vector $Y_1, Y_2, \ldots, Y_n$.

Question 8.5

Prove the result $E\left[E\left[X \mid Y_1, Y_2, \ldots, Y_n\right]\right] = E\left[X\right]$.

The importance of conditional expectations comes from the following property:

Result 8.2

$E[X \mid \underline{Y}]$ is the optimal estimator of X based on $Y_1, Y_2, \ldots, Y_n$ in the sense that for every function h:

$$E\left\{\left(X - E[X \mid \underline{Y}]\right)^2\right\} \leq E\left\{\left(X - h(\underline{Y})\right)^2\right\}$$

This result can be proved using techniques covered in Subject CT3, but students will not be required to prove the result in the Subject CT8 examination.

2.3 Martingales in discrete time

A discrete-time stochastic process $X_0, X_1, X_2, \ldots$ is said to be a martingale if:

- **$E[|X_n|] < \infty$ for all n**
- **$E[X_n \mid X_0, X_1, \ldots, X_m] = X_m$ for all $m < n$**

The real content of this definition is the second condition. The first condition is just a technicality that ensures that the mathematics makes sense. In most questions we are only concerned with the second condition and we'll assume the first condition holds.

In words, the current value X_m of a martingale is the optimal estimator of its future value X_n.

Note that the expectations here must be calculated based on a particular probability distribution and we will get different answers if we assume different distributions. So a process may only be a martingale if we assume a particular probability distribution.

If the discounted price of a financial asset is a martingale when calculated using a particular probability distribution, the probability distribution can be described as *risk-neutral*.

We will have a lot more to say about this later in the course when we discuss risk-neutral probability measures.

"Discounted" is used in the sense of Subject CT1, *ie* it is reduced to allow for the time value of money.

Example

Consider a share, whose value is assumed to change each day according to the following probability distribution:

$$C = \begin{cases} +10p & \text{with probability } \frac{1}{4} \\ +8p & \text{with probability } \frac{1}{4} \\ -9p & \text{with probability } \frac{1}{2} \end{cases}$$

Show that C is a martingale under this probability distribution.

Solution

The expected change in the share price over one day is:

$$E[C] = \tfrac{1}{4} \times 10 + \tfrac{1}{4} \times 8 + \tfrac{1}{2} \times (-9) = 0$$

So on average the share price stays the same each day. This means that C is a martingale under the probability distribution $\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right)$, which we have assumed for the three possible states.

Another (more complicated) example

Suppose a share has price S_t at time t, and assume for simplicity that it does not pay dividends. Let $D_t = e^{-rt} S_t$ be the discounted price of the share, discounting at the risk-free rate r. If, according to a particular probability distribution we assume for S_t, $D_t = e^{-rt} S_t$ is a martingale, then for $m < n$:

$$E\left[e^{-rn} S_n \middle| S_0, S_1, \ldots, S_m \right] = e^{-rm} S_m$$

Since we are conditioning on S_m, we can treat it as a constant and take it inside the expectation. It follows that:

$$E\left[\frac{S_n}{S_m}\middle|S_0, S_1, \ldots, S_m\right] = e^{r(n-m)}$$

In other words, the expected return on the share from time m to time n is just the risk-free rate. This is why this set of probabilities is described as risk-neutral. In reality, investing in a share involves risk and investors will require some extra reward for investing in the share. So the real-world probability distribution will be different from the risk-neutral one.

We will meet the discounted share price process again in Part 4 of this course.

Of all the properties of martingales, the most useful is also the simplest.

Result 8.3

A martingale has constant mean, *ie* $E[X_n] = E[X_0]$ for all n.

Proof

$$E[X_m] = E\{E[X_n \mid X_0, X_1, \ldots, X_m]\} = E[X_n] \text{ for all } m < n$$

The conclusion follows.

Note that if you know the value of X_0 then $E[X_n] = E[X_0] = X_0$ for all n.

2.4 Martingales in continuous time

The definition of martingales given in Section 2.3 is too narrow for many applications. In particular it does not lend itself to extension to continuous time because of the difficulty of conditioning on a continuum of random variables. A more general approach to conditional expectations and martingales is described below.

Mathematical framework

As you will see, there is very little difference intuitively, between a martingale in discrete time and a martingale in continuous time.

Question 8.6
In words, what is a martingale?

Pure mathematicians do not like to rely on intuition because it can lead to the wrong conclusions. They prefer "rigour" and mathematical proof. This section outlines the mathematical framework used to set up the theory of martingales in continuous time.

The key concept here is that of a *filtration*. Intuitively, the filtration F_t represents the "history of the process up until time t".

The following structures underlie any stochastic process X_t :

- **a sample space Ω,** *ie* a set of all possible outcomes. **Each outcome ω in Ω determines a sample path $X_t(\omega)$.**
- **a set of events F. This is a collection of events, by which is meant subsets of Ω, to which a probability can be attached.**
- **for each time t, a smaller collection of events $F_t \subset F$. This is the set of those events whose truth or otherwise are known at time t. In other words an event A is in F_t if it depends only on X_s, $0 \le s \le t$.**

As t increases, so does F_t, *ie* $F_t \subset F_u$, $t < u$. Taken collectively, the family $(F_t)_{t \ge 0}$ is known as the *filtration* associated with the stochastic process $X_t, t \ge 0$. It describes the information gained by observing the process.

Some random variables will be known by time t. We say that Y is F_t*-measurable* if the event $\{Y \le y\}$ belongs to F_t for all values of y.

A stochastic process $Y_t, t \ge 0$ is said to be *adapted* to the filtration F_t if Y_t is F_t-measurable for all t. If F_t is the filtration generated by X_t, then any function of X_t is adapted to F_t.

To formulate our new definition of conditional expectations, we simply insist on the key property of Result 8.2.

The *conditional expectation* $E[X \mid F_t]$ is by definition the optimal estimator of X among all F_t-measurable random variables with finite expectation, or equivalently:

$$E\{(X - E[X \mid F_t])Y\} = 0$$

for all F_t-measurable bounded random variables Y.

The following key properties of conditional expectations are a relatively easy consequence of the above geometric definition:

Result 8.4

(i) $E\left\{E\left[X \middle| F_t\right]\right\} = E\left[X\right]$

(ii) **if X is F_t-measurable, then** $E\left[X \middle| F_t\right] = X$

(iii) **if Y is F_t-measurable and bounded, then**

$$E\left[XY \middle| F_t\right] = YE\left[X \middle| F_t\right]$$

(iv) **if X is independent of F_t, then** $E\left[X \middle| F_t\right] = E\left[X\right]$

Question 8.7

In words, what does $P\left[X_t = x \middle| F_s\right]$ mean?

Martingales

In this setting, a martingale is a stochastic process such that:

- $E[|X_t|] < \infty$ **for all** t
- $E[X_t \mid F_s] = X_s$ **for all** $s < t$.

Once again, the real content of this definition is the second condition. In most questions, we'll simply assume that the first condition holds.

Question 8.8

In words, what does the $E[X_t \mid F_s]$ in the second condition mean?

Question 8.9

An asset's value at time t (in pence and measured in years) is denoted by A_t and fluctuates in value from day to day. Within these random fluctuations, there appears to be an underlying long-term trend, in that the asset's value is increasing by 2 pence on average each week. Assuming that there are exactly 52 weeks in a year, suggest a process based on A_t that you think might be a martingale.

Question 8.10

Suppose that the price increments have a continuous uniform distribution such that $A_t - A_s \sim U[-(t-s), 5(t-s)]$. Construct a martingale out of A_t.

Using the independence property of the increments, one can prove that Brownian motion is a martingale.

Also if N_t is the Poisson process with rate λ, then $N_t - \lambda t$ is a martingale.

You met the Poisson process in Subjects CT4 and CT6. We won't be using it in CT8.

Note that a process can be a martingale with respect to the filtration generated by another process. For instance, if B_t is Brownian motion, $B_t^2 - t$ is martingale for the filtration generated by $B_t, t \geq 0$.

Example

Let $X_t = e^{\lambda B_t - ½\lambda^2 t}$, where λ is any constant. Show that X_t is a martingale with respect to F_t, the filtration associated with B_t.

Solution

We need to show that, when $s<t$:

$$E[X_t \mid F_s] = X_s$$

The left-hand side is:

$$E[e^{\lambda B_t - \frac{1}{2}\lambda^2 t} \mid F_s]$$

Writing B_t as $B_s + (B_t - B_s)$, this becomes:

$$E[e^{\lambda B_s + \lambda(B_t - B_s) - \frac{1}{2}\lambda^2 t} \mid F_s]$$

Since we are conditioning on F_s, the value of B_s is not random, and can be treated as a constant. The time t is also a fixed number. So this becomes:

$$e^{\lambda B_s - \frac{1}{2}\lambda^2 t} E[e^{\lambda(B_t - B_s)} \mid F_s]$$

The increment $B_t - B_s$ is independent of the past history up to time s. So this is just:

$$e^{\lambda B_s - \frac{1}{2}\lambda^2 t} E[e^{\lambda(B_t - B_s)}]$$

We know that $B_t - B_s \sim N(0, t-s)$. So the expectation is of the form $E[e^{\lambda X}]$, where λ is a fixed number and $X \sim N(0, t-s)$. This is just the MGF of a normal distribution, and the formula on page 11 of the Tables (with t replaced by λ) tells us that this equals $e^{\mu\lambda + \frac{1}{2}\sigma^2\lambda^2}$. Here $\mu = 0$ and $\sigma^2 = t-s$. So the expectation is $e^{\frac{1}{2}\lambda^2(t-s)}$ and we have:

$$e^{\lambda B_s - \frac{1}{2}\lambda^2 t} e^{\frac{1}{2}\lambda^2(t-s)} = e^{\lambda B_s - \frac{1}{2}\lambda^2 s}$$

But this is just the formula for X_s. So the result is proved.

Q

Question 8.11

Show that $B_t^2 - t$ is a martingale with respect to F_t, the filtration associated with B_t.

3 Exam-style question

We finish this chapter with an exam-style question on Brownian motion and martingales.

Question

This question is based on Subject 103 September 2004 Question 10(i) and (ii).

Assume that the spot rate of interest at time t, $S(t)$, can be modelled by $S(t) = e^{-2\mu W(t)}$, where $W(t)$ is a Brownian motion with drift coefficient μ and volatility coefficient 1 such that $W(0) = 0$.

(i) Write down an expression for $W(t)$ in terms of a standard Brownian motion, $B(t)$. [1]

(ii) Show that $\{S(t) : t > 0\}$ is a continuous-time martingale. [4]

[Total 5]

Solution

(i) ***General Brownian motion***

A general Brownian motion can be defined as:

$$W(t) = W(0) + \mu t + \sigma B(t)$$

where $B(t)$ is standard Brownian motion, μ is the drift, σ is the volatility coefficient and $W(0)$ is the value of general Brownian motion at time 0.

Since $\sigma = 1$ and $W(0)=0$:

$$W(t) = \mu t + B(t)$$

(ii) ***$S(t)$ is a martingale***

We write $S(t)$ as S_t and $B(t)$ as B_t for neatness. Let $t > s$, then:

$$\begin{aligned}
E\left[S_t \middle| F_s\right] &= E\left[e^{-2\mu(\mu t + B_t)} \middle| F_s\right] \\
&= E\left[e^{-2\mu^2 t - 2\mu B_t} \middle| F_s\right] \\
&= e^{-2\mu^2 t} E\left[e^{-2\mu B_t} \middle| F_s\right] \\
&= e^{-2\mu^2 t} E\left[e^{-2\mu\left(B_t - B_s + B_s\right)} \middle| F_s\right] \\
&= e^{-2\mu^2 t} e^{-2\mu B_s} E\left[e^{-2\mu\left(B_t - B_s\right)} \middle| F_s\right] \\
&= e^{-2\mu^2 t} e^{-2\mu B_s} E\left[e^{-2\mu\left(B_t - B_s\right)}\right]
\end{aligned}$$

The filtration F_s can be left out because of the independent increments property. Now $B_t - B_s \sim N\left(0, t-s\right)$ and so the expectation is of the form $E[e^{aX}]$ where $X \sim N\left(0, t-s\right)$ and $a = -2\mu$. Therefore, we can use the MGF of a normal distribution calculated at point -2μ to determine the expectation.

So, using the MGF formula, from page 11 of the *Tables*, we get:

$$E\left[S_t \middle| F_s\right] = e^{-2\mu^2 t} e^{-2\mu B_s} e^{0(-2\mu)+\frac{1}{2}(t-s)(-2\mu)^2}$$

$$= e^{-2\mu^2 s - 2\mu B_s}$$

$$= S_s$$

Finally we check that $E\left[\left|S_t\right|\right] < \infty$ for all values of t.

$$E[|S_t|] = E\left[e^{-2\mu^2 t - 2\mu B(t)}\right]$$

$$= e^{-2\mu^2 t} E\left[e^{-2\mu B_t}\right]$$

$$= e^{-2\mu^2 t} M_{B_t}(-2\mu)$$

$$= e^{-2\mu^2 t} e^{0(-2\mu)+\frac{1}{2}t(-2\mu)^2} = 1 < \infty$$

where: $M_{B_t}(-2\mu)$ is the MGF of B_t at -2μ.

Chapter 8 Summary

Standard Brownian motion

Standard Brownian motion is a stochastic process with continuous state space $\mathbb{R}$ and the continuous-time set $\mathbb{R}^+$. Its defining properties are:

- $B_0 = 0$
- it has independent increments
- it has stationary increments
- it has Gaussian increments, *ie* $B_t - B_s \sim N(0, t-s)$
- it has continuous sample paths.

Brownian motion is the continuous-time analogue of a random walk.

Other properties of Standard Brownian motion include:

- $\text{cov}(B_s, B_t) = \min(s,t)$
- $\{B_t, t \geq 0\}$ is a *Markov process*
- $\{B_t, t \geq 0\}$ is a *martingale*, *ie* $E(B_t \mid F_s) = B_s$
- $\{B_t, t \geq 0\}$ returns infinitely often to 0, or indeed to any other level
- If $\{B_1(t), t \geq 0\}$ is defined by $B_1(t) = \dfrac{1}{\sqrt{c}} B_{ct}$ then $\{B_1(t), t \geq 0\}$ is also a standard Brownian motion. (*scaling property*)
- If $\{B_2(t), t \geq 0\}$ is defined by $B_2(t) = tB_{1/t}$ then $\{B_2(t), t \geq 0\}$ is also a standard Brownian motion. (*time inversion property*)
- For any given time t_0, the probability that the sample path of a standard Brownian motion is differentiable at t_0 is zero.

Brownian motion with drift

Brownian motion with drift is related to standard Brownian motion by the equation:

$$W_t = W_0 + \sigma B_t + \mu t$$

where σ is the *volatility* or *diffusion coefficient* and μ is the *drift*.

Geometric Brownian motion (lognormal model)

For modelling purposes a Brownian motion may have to be transformed, for example by taking logarithms. A useful model for security prices is *geometric Brownian motion*:

$$S_t = e^{W_t}$$

where W is the Brownian process $W_t = W_0 + \sigma B_t + \mu t$. Thus S_t, which is called geometric Brownian motion, is lognormally distributed with parameters $W_0 + \mu t$ and $\sigma^2 t$.

Martingales

A *martingale* is a stochastic process such that:

- $E\left[|X_t|\right] < \infty$ for all t
- $E[X_t \mid F_s] = X_s$ for all $s < t$

Martingales are processes with no drift. In fact, it can be shown that a martingale has constant mean, *ie*:

$$E[X_n] = E[X_0] \text{ for all } n$$

Martingales constructed from Brownian motion

Various martingales can be constructed from standard Brownian motion, for example, B_t, $B_t^2 - t$ and $e^{\lambda B_t - \lambda^2 t/2}$.

Chapter 8 Solutions

Solution 8.1

The first two properties of a Brownian motion – that the increments are (i) independent of the past and (ii) stationary – follow for W_t because:

$$W_t - W_s = \sigma\left(B_t - B_s\right) + \mu\left(t-s\right)$$

and we know that the increments $B_t - B_s$ have these properties.

The third property we require is that:

$$W_t - W_s \sim N\left(\mu\left(t-s\right), \sigma^2\left(t-s\right)\right)$$

This follows because $B_t - B_s \sim N\left(0, \left(t-s\right)\right)$ and therefore:

$$\begin{aligned} W_t - W_s &= \sigma\left(B_t - B_s\right) + \mu\left(t-s\right) \\ &\sim \sigma N\left(0, t-s\right) + \mu\left(t-s\right) \\ &\sim N\left(0, \sigma^2\left(t-s\right)\right) + \mu\left(t-s\right) \\ &\sim N\left(\mu\left(t-s\right), \sigma^2\left(t-s\right)\right) \end{aligned}$$

Solution 8.2

Just invert the relationship given:

$$B_t = \frac{W_t - W_0 - \mu t}{\sigma}$$

This is analogous to converting an observed value x from a general normal distribution, $N(\mu, \sigma^2)$, into a value from the standard normal distribution, $N(0,1)$, by calculating the standardised value $z = \dfrac{x-\mu}{\sigma}$.

Solution 8.3

We have the definition $B_1(t) = \frac{1}{\sqrt{c}} B_{ct}$. It follows that:

$$B_1(t) - B_1(s) = \frac{1}{\sqrt{c}}\left(B_{ct} - B_{cs}\right) \sim \frac{1}{\sqrt{c}} N\left(0, ct - cs\right) \sim N\left(0, t - s\right)$$

Solution 8.4

The distribution of B_t (the value at time t) is $N(0,t)$.

So we have:

$$\begin{aligned} P[a-\delta < B_t < a+\delta] &= P[a-\delta < N(0,t) < a+\delta] \\ &= P\left[\frac{a-\delta}{\sqrt{t}} < N(0,1) < \frac{a+\delta}{\sqrt{t}}\right] \end{aligned}$$

As $t \to \infty$, this range will get narrower and narrower, and the probability will tend to zero.

Solution 8.5

Let $y_1, y_2, \ldots, y_n$ be denoted by $\underline{y}$. Then the right-hand side is:

$$\begin{aligned} E[X] &= \sum_x xP(X = x) \\ &= \sum_x \sum_{\underline{y}} xP\left(X = x \mid \underline{Y} = \underline{y}\right) P\left(\underline{Y} = \underline{y}\right) \\ &= \sum_{\underline{y}} \left(\sum_x xP\left(X = x \mid \underline{Y} = \underline{y}\right) \right) P\left(\underline{Y} = \underline{y}\right) \\ &= \sum_{\underline{y}} E\left[X \mid \underline{Y} = \underline{y}\right] P\left(\underline{Y} = \underline{y}\right) \\ &= E\left[E[X \mid \underline{Y}]\right] \end{aligned}$$

The first equation is just the definition of an expectation. The second follows by conditioning on $\underline{Y}$. In the third equation we interchange the order of summation. Then we can apply the definition of conditional expectation to get the fourth equation. The last line just uses the definition of an expectation of a function of $\underline{Y}$.

Solution 8.6

There are many good answers, *eg*:

- The expected future value is the current value.
- The expected change in the process is zero.
- The process has "no drift".

Solution 8.7

$P\left[X_t = x \middle| F_s\right]$ is the probability that the process equals x at time t, given that we are at time s and we know the history of the process up until time s.

Solution 8.8

$E\left[X_t \middle| F_s\right]$ means the expected value of the process at time t, given that we are at time s and we know the history of the process up until time s.

Solution 8.9

The value of the asset is increasing on average by 2 pence a week. Assuming that there are exactly 52 weeks in a year, this means the asset is "drifting" by 104 pence a year. A martingale is a process without drift and so a good suggestion would be to remove this drift and consider the process:

$$A_t - 104t$$

Solution 8.10

Using the formula for the expected value of a uniform distribution from page 13 of the *Tables*, we have:

$$E\left[A_t \middle| F_s\right] = E\left[A_s + (A_t - A_s) \middle| F_s\right] = A_s + \frac{5(t-s)-(t-s)}{2} = A_s + 2(t-s)$$

So for every increase of $t-s$ in the time, the process is "drifting" by $2(t-s)$. So this is the same process as in the previous question.

One way to construct a martingale is to subtract the drift. Mathematically, we can subtract $2t$ from both sides of the last equation to get:

$$E\left[A_t - 2t \middle| F_s\right] = A_s - 2s$$

This now fits in with the definition given for a martingale in continuous time.

So the process $A_t - 2t$ is a martingale.

Solution 8.11

Let $X_t = B_t^2 - t$.

We need to show that, when $s < t$:

$$E[X_t \mid F_s] = X_s$$

The left-hand side is:

$$E[B_t^2 - t \mid F_s]$$

Writing B_t as $B_s + (B_t - B_s)$, this becomes:

$$E[\{B_s + (B_t - B_s)\}^2 - t \mid F_s]$$

We can expand the square to get:

$$E[B_s^2 + 2B_s(B_t - B_s) + (B_t - B_s)^2 - t \mid F_s]$$

Since we are conditioning on F_s, the value of B_s is not random, and can be treated as a constant. The time t is also a fixed number. So this becomes:

$$B_s^2 + 2B_s E[B_t - B_s \mid F_s] + E[(B_t - B_s)^2 \mid F_s] - t$$

The increment $B_t - B_s$ is independent of the past history up to time s. So this is just:

$$B_s^2 + 2B_s E[B_t - B_s] + E[(B_t - B_s)^2] - t$$

We know that $B_t - B_s \sim N(0, t-s)$. So the first expectation is zero and the second equals the variance, which is $t-s$. So we have:

$$B_s^2 + 0 + (t-s) - t = B_s^2 - s$$

But this is just the formula for X_s. So the result is proved.

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 the price of shares. 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.

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

An integral of this type is called an *Ito integral*.

It is worth pointing out, however, that 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 random integrands Y_s by an indirect method that we outline now.

We will simplify the problem for the moment and assume that the integrand is not itself 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: $\int_1^3 f(s)\,dB_s = \int_1^2 1dB_s + \int_2^3 2dB_s = \left[B_s\right]_1^2 + \left[2B_s\right]_2^3 = 2B_3 - B_2 - B_1$

More generally:

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

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.

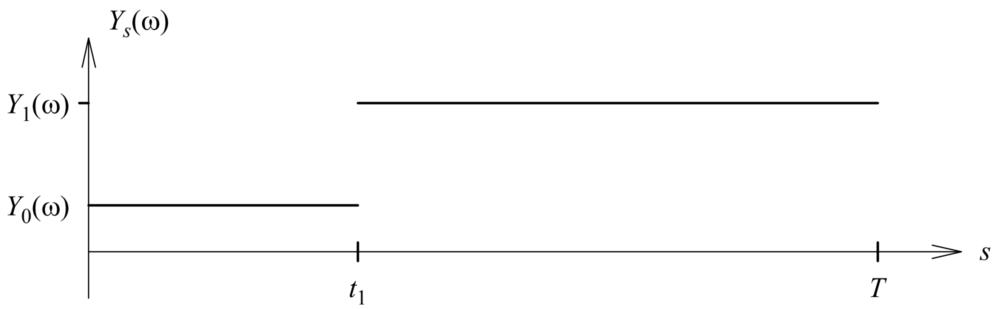
We now move on to integrals of random functions.

Ito integrals for random functions

Start with simple processes adapted to the natural filtration $(F_t)_{t\geq 0}$ of standard Brownian motion:

$$Y_s = \begin{cases} Y_0 & 0 \leq s < t_1 \\ Y_1 & t_1 \leq s \leq T \end{cases}$$

where Y_0 is an F_0-measurable random variable – that is, a random variable whose value is known at time 0 – and Y_1 is an F_{t_1}-measurable random variable – that is, its value is fully determined by $\{B_t : 0 \leq t \leq t_1\}$.



For such processes, we require that the integral has the obvious meaning:

$$\int_0^t Y_s\,dB_s = \begin{cases} Y_0 B_t & \text{if } t < t_1 \\ Y_0 B_{t_1} + Y_1(B_t - B_{t_1}) & \text{if } t \geq t_1 \end{cases}$$

The idea here is the same as with the deterministic functions. We insist that the integral of a piecewise constant function can be split into integrals on the constant pieces.

Then observe that the resulting process 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)

$$E\left[\left(\int_0^t Y_s dB_s \right)^2 \right] = \begin{cases} t_1 E[Y_0^2] + (t - t_1) E[Y_1^2] = E\left[\int_0^t Y_s^2 ds \right] & t \geq t_1 \\ t E[Y_0^2] = E\left[\int_0^t Y_s^2 ds \right] & t < t_1 \end{cases}$$

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.

This is known as the *norm-preservation* property. A *norm* is a measure of size or distance. The norm of a random variable X is often taken to be $E\left[X^2 \right]$. So the above equation is saying that the norm of the integral is the same as the integral of the norm.

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

The above properties are easily checked.

Question 9.1

Consider the integral $\int_0^t Y_s dB_s = \begin{cases} Y_0 B_t & \text{if } t \leq t_1 \\ Y_0 B_{t_1} + Y_1 (B_t - B_{t_1}) & \text{if } t \geq t_1 \end{cases}$

Prove each of the four properties above.

Moreover, the above properties continue to hold when the simple two-step adapted integrand is replaced by an adapted integrand consisting of any finite number of random steps.

Now any adapted integrand can be approximated by a step function with a finite number of random steps. We would like to say: if $\{Y_s^{(1)} : s \geq 0\}$ and $\{Y_s^{(2)} : s \geq 0\}$ are two step functions which are "close" to each other, then $\left\{\int_0^t Y_s^{(1)} dB_s : t \geq 0\right\}$ and $\left\{\int_0^t Y_s^{(2)} dB_s : t \geq 0\right\}$ are also close to each other.

And this is true because of the norm preservation property (iii):

$$E\left[\left(\int_0^t Y_s^{(1)}\, dB_s - \int_0^t Y_s^{(2)}\, dB_s\right)^2\right] = E\left[\left(\int_0^t (Y_s^{(1)} - Y_s^{(2)})\, dB_s\right)^2\right] = E\left[\int_0^t (Y_s^{(1)} - Y_s^{(2)})^2\, ds\right]$$

This means that, if $\left\{Y_s^{(n)}\right\}$ is a sequence of step functions that converges to the adapted integrand $\{Y_s\}$, then the sequence of integrals $\left\{\int_0^t Y_s^{(n)}\, dB_s\right\}$ also converges and the limit does not depend on which sequence of approximating step functions we choose. The limit is formally defined as the integral $\int_0^t Y_s dB_s$.

The method of approximating an integral by means of integrals of step functions works for any adapted square integrable integrand Y (*ie* any Y satisfying $E\left[\int_0^t Y_s^2 ds\right] < \infty$) and ensures that properties (i)–(iv) are preserved.

The last point is saying that for any function such that $E\left[\int_0^t Y_s^2\, ds\right] < \infty$, the Ito integral can be well defined, and furthermore, it satisfies the given properties. This can be proved using limits of sequences of simple functions and applying Property (iii). But this is less important for us. What matters is the fact that it can be done.

It is as well to point out straightaway 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 below) to identify the integrated process.

Note that in the random case the integral does not have to be normally distributed in general since the function Y_s introduces another random source.

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. 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.

Alternatively, we can use the following formula approach, although this is less intuitive.

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.

Alternative solution

The idea is to differentiate the whole expression with respect to *t*. Recall that:

$$\int_0^t y(s)\,df(s) = \int_0^t y(s)\frac{df}{ds}\,ds$$

We also need to know how to differentiate an integral. The full formula is:

$$\frac{d}{dt}\int_{a(t)}^{b(t)} g(x,t)\,dx = b'(t)\,g\big(b(t),t\big) - a'(t)\,g\big(a(t),t\big) + \int_{a(t)}^{b(t)} \frac{\partial}{\partial t} g(x,t)\,dx$$

So: $$\frac{dx}{dt} = y(t)\frac{df}{dt} + a(t)$$

This can be written: $$dx(t) = y(t)\,df(t) + a(t)\,dt$$

Question 9.2

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 rule and the only rule that we will need. (If you haven't come across the word "lemma" before, a lemma is just a minor theorem.)

To be consistent with what is to follow, we will first give a derivation of the chain rule (function-of-a-function rule) for standard calculus.

Suppose we have a function-of-a-function $f(b_t)$. 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 Gold Tables if you are unfamiliar with Taylor series.

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

$$\frac{df(b_t)}{dt} = f'(b_t)\frac{db_t}{dt} + \lim_{\delta t \to 0} \tfrac{1}{2} f''(b_t)\frac{(\delta b_t)^2}{\delta t}$$

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:

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

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:

$$df(B_t) = f'(B_t)\,dB_t + \tfrac{1}{2} f''(B_t)\,dt$$

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:

$$d\left(B_t^2\right) = 2B_t dB_t + ½2dt = 2B_t dB_t + dt$$

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:

$$\int_0^s B_t dB_t = ½\left(B_s^2 - s\right)$$

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 basically 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 $f(X_t)$ we can proceed as with the functions of Brownian motion. Taylor's theorem to second-order is:

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

which we can write in the limit as:

$$df(X_t) = f'(X_t)dX_t + ½ f''(X_t)(dX_t)^2$$

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 $f(t, X_t)$.

Before proceeding to the result itself, we will quickly remind ourselves of the classical chain rule in two variables. If we have $f(x(t), y(t))$, 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 $f(t, x(t))$, 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 (Ito's lemma)

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

$$df(t, X_t) = \frac{\partial f}{\partial x} Y_t dB_t + \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} A_t + ½ \frac{\partial^2 f}{\partial x^2} Y_t^2 \right] dt \tag{9.3}$$

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

- **use Taylor's formula to second-order to write:**

$$df(t, X_t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + ½ \frac{\partial^2 f}{\partial x^2} (dX_t)^2$$

- **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:

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} A_t + ½ \frac{\partial^2 f}{\partial x^2} Y_t^2 \quad \text{and} \quad \frac{\partial f}{\partial x} Y_t$$

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 $f(t, X_t) = B_t^2 - t$. Then:

$$d(B_t^2 - t) = 2B_t dB_t + [-1 + \tfrac{1}{2} \times 2]dt = 2B_t dB_t$$

This can also 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:

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

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

Question 9.3

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 9.4

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.* The unknown process S_t appears on both sides of the equation.

It is of course shorthand for the integral equation:

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

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 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 $\alpha t + \sigma B_t$ for $\log(S_t / S_0)$ and thus to $S_0 \exp(\alpha t + \sigma B_t)$ for S_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, use Ito's Lemma to calculate $d \log S_t$:

$$\begin{aligned} d \log S_t &= \frac{1}{S_t} dS_t + \frac{1}{2}\left(\frac{-1}{S_t^2}\right)(dS_t)^2 \\ &= \left(\alpha - \frac{1}{2}\sigma^2\right) dt + \sigma dB_t \end{aligned}$$

Written in integral form, 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 satisfying the stochastic differential equation 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:

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

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.)

A diffusion model or Ito model 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:**

(i) it is a Markov process

(ii) it has continuous sample paths

(iii) there exist functions $\mu(x)$ and $\sigma^2(x) > 0$ such that as $h \to 0+$

$$E\left[X_{t+h} - X_t \mid X_t = x\right] = h\mu(x) + o(h)$$

$$E\left[\left(X_{t+h} - X_t\right)^2 \mid X_t = x\right] = h\sigma^2(x) + o(h)$$

$$E\left[\left|X_{t+h} - X_t\right|^3 \mid X_t = x\right] = o(h)$$

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 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.

Q

Question 9.5

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:

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

$$\sigma^2(x) = \sigma^2$$

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 $\mu(x)$ and the diffusion function $\sigma^2(x)$. 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 calculus to revisit the Ornstein-Uhlenbeck process. In view of the drift and diffusion coefficients $\mu(x) = -\gamma x$, $\sigma^2(x) = \sigma^2$ we may define the Ornstein-Uhlenbeck process as the solution to the equation:

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

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".

We have:

$$\begin{aligned} dU_t = d\left(e^{\gamma t} X_t\right) &= \gamma e^{\gamma t} X_t dt + e^{\gamma t} dX_t \\ &= \gamma e^{\gamma t} X_t dt + e^{\gamma t}\left(-\gamma X_t dt + \sigma dB_t\right) \\ &= \sigma e^{\gamma t} dB_t \end{aligned}$$

The first line is based on a form of the product rule:

$$d(e^{\gamma t} X_t) = e^{\gamma t} dX_t + X_t d(e^{\gamma t})$$

But, since $\dfrac{d(e^{\gamma t})}{dt} = \gamma e^{\gamma t}$, we know that $d(e^{\gamma t}) = \gamma e^{\gamma t} dt$.

Thus:

$$U_t = U_0 + \sigma \int_0^t e^{\gamma s} dB_s$$

Consequently:

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

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 9.6

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 9.1

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 9.7

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:

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

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

- $\alpha = e^{-\gamma}$
- $\dfrac{\sigma_e^2}{1-\alpha^2} = \dfrac{\sigma^2}{2\gamma}$

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:

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

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:

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

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

Chapter 9 Summary

Ito integrals

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

- 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:

$$df\left(t,X_t\right)=\frac{\partial f}{\partial t}dt+\frac{\partial f}{\partial x}dX_t+½\frac{\partial^2 f}{\partial x^2}\left(dX_t\right)^2$$

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:

$$X_t=e^{-\gamma t}U_t=X_0e^{-\gamma t}+\sigma\int_0^t e^{-\gamma(t-s)}dB_s$$

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

Chapter 9 Solutions

Solution 9.1

Define the process to be the integral:

$$X_t = \int_0^t Y_s dB_s = \begin{cases} Y_0 B_t & \text{if } t \le t_1 \\ Y_0 B_{t_1} + Y_1(B_t - B_{t_1}) & \text{if } t \ge t_1 \end{cases}$$

First we prove that the means are zero. Note that:

$$E\left[X_t\right] = \begin{cases} E\left[Y_0 B_t\right] & \text{if } t \le t_1 \\ E\left[Y_0 B_{t_1} + Y_1(B_t - B_{t_1})\right] & \text{if } t \ge t_1 \end{cases}$$

Since the increments of standard Brownian motion have zero mean, the result follows.

The martingale property now follows as in the deterministic case:

$$E\left[X_t \mid F_u\right] = E\left[\int_0^u Y_s dB_s + \int_u^t Y_s dB_s \,\middle|\, F_u\right] = X_u + E\left[\int_u^t Y_s dB_s \,\middle|\, F_u\right] = X_u$$

where the last equality holds because of the zero expectation property we've just proved.

Equation (iii) requires more work.

Assume first that $t \ge t_1$. For the first equation (on the top line) we multiply out the expression for $\int_0^t Y_s \, dB_s$ as a quadratic and note that, for example, $E\left[B_t^2\right] = \text{var}\left(B_t\right) = t$.

$$E\left[\left(\int_0^t Y_s dB_s\right)^2\right] = E\left[\left(Y_0 B_{t_1} + Y_1\left(B_t - B_{t_1}\right)\right)^2\right]$$

$$= E\left[Y_0^2 B_{t_1}^2 + Y_1^2\left(B_t - B_{t_1}\right)^2 + 2Y_0 Y_1 B_t\left(B_t - B_{t_1}\right)\right]$$

$$= t_1 E\left[Y_0^2\right] + \left(t - t_1\right) E\left[Y_1^2\right]$$

Showing that this equals $E\left[\int_0^t Y_s^2 ds\right]$ is straightforward since the function Y_s^2 is easy to calculate and the integral in the last term is a standard integral with respect to s. It is not a stochastic integral.

The $t < t_1$ case is similar.

Observation (iv) also follows from the fact that $\int_0^t Y_s \, dB_s$ is a sum of terms, each of which has continuous sample paths. Also, we can see that the path of this integral remains continuous at the critical point t_1, since the area under the graph will increase gradually as we pass t_1.

Solution 9.2

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 9.3

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 9.4

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:

$$df(x) = f'(x)dx + \tfrac{1}{2} f''(x)(dx)^2 + O\left((dx)^3\right)$$

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:

$$dX_t^3 = 3X_t^2\left(\sigma(t)\,dB_t + \mu(t)\,dt\right) + 3X_t\sigma^2(t)\,dt$$
$$= 3X_t^2\sigma(t)\,dB_t + \left(3X_t\sigma^2(t) + 3X_t^2\mu(t)\right)dt$$

as Ito's formula gave us.

Solution 9.5

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 9.6

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 9.7

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 9.1(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:

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

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

Chapter 10

Stochastic models of security prices

Syllabus objectives

(vii) Demonstrate a knowledge and understanding of stochastic models of the behaviour of security prices.

1. *Discuss the continuous-time lognormal model of security prices and the empirical evidence for or against the model.*

2. *Discuss the structure of autoregressive models of security prices and other economic variables, such as the Wilkie model, and describe the economic justification for such models.*

3. *Discuss the main alternatives to the models covered in (vii)1. and (vii)2. above and describe their strengths and weaknesses.*

4. *Perform simple calculations involving the models described above.*

5. *Discuss the main problems involved in estimating parameters for asset pricing models:*
    - *data availability*
    - *data errors*
    - *outliers*
    - *stationarity of underlying time series*
    - *the role of economic judgement.*

0 Introduction

Stochastic models of security prices are used to simulate future investment returns as an important part of modelling processes such as *asset liability modelling*, which you will meet in Subject CA1. Several basic types of models are available, each making different assumptions about the processes that generate investment returns. Each will, in general, yield different results. Consequently, the results of any simulation process should be analysed with regard to both the nature of the model and the values of the model parameters employed.

The models described and discussed in this chapter are used primarily for longer-term risk management applications rather than pricing or hedging. In other words we are focused on time horizons of several years rather than months, weeks, days or shorter.

The pricing and hedging models referred to here are models such as the Black-Scholes model, which assumes that share prices follow geometric Brownian motion. This type of model is commonly used for pricing and hedging options, which typically have lives of only a few months.

Many of the underlying issues have a strong overlap with those integral to the pricing and hedging models, but the rather different historical development and application of the two strands means that jargon and areas of focus can be quite different.

Whether or not markets are efficient, the experience of active fund managers shows that it is surprisingly difficult to consistently outperform the market. In an actuarial context, it is imprudent to take advance credit for exploitation of future anomalies. To reflect this it is often considered good practice to construct models that are consistent with markets being efficient.

One such model that is widely used in financial economics is the *continuous-time lognormal model* of security prices, which assumes that the log of security prices follows a continuous-time random walk with drift. Note that this is equivalent to saying that the log of security prices follows a Brownian motion with drift. The share prices themselves will therefore follow geometric Brownian motion as introduced in Chapter 8.

In an efficient market, security prices will change only in response to the arrival of new information, which by definition must be unpredictable and hence random. In other words, since the price of an asset at time s already incorporates all the information up until that time, the subsequent change in the price between times s and t ($s < t$) will be independent of any history up until time s. A (geometric) Brownian motion model is consistent with this because of the property of *independent increments*.

If, however, markets are inefficient then the lognormal model may be inappropriate. As a consequence of this, other security price models have been developed. Amongst these is the family of *autoregressive models*, which assume that future security price movements revert to a long-term mean and hence are not entirely independent of past values. These models are also referred to as *mean-reverting* or error correction models. Foremost amongst the actuarial investment models of this type is the *Wilkie model*.

This chapter discusses in detail:

- the *continuous-time* lognormal model of security prices – this is a statistical model of the process determining share prices

- autoregressive models of security prices – with reference to the *Wilkie model*, which was developed so as to ensure a reasonable fit to historical data and also to be consistent with economic principles and intuition.

It concludes with sections discussing some issues relating to alternative models and the estimation of parameters for asset pricing models.

1 The continuous-time lognormal model

1.1 Definition

The *continuous-time lognormal model* is another name for geometric Brownian motion. Although the basic (non-geometric) Brownian motion model may be good at describing the movement of market indices in the short run, it is not very good in the long run for two reasons:

1. Brownian motion without a positive drift is certain to become negative eventually. Even with a positive drift, there is the possibility of negative security prices in the future, which isn't very realistic.

2. The Brownian motion model predicts that daily movements of size 10 say would occur just as frequently when the process is at a level of 500 as when it is at a level of 5000.

We can remove these two problems by working with logs.

The conventional continuous-time lognormal model of security prices assumes that log prices form a random walk. If S_t denotes the market price of an investment, then the model states that, for $u > t$, log returns are given by:

$$\log(S_u) - \log(S_t) \sim N\left[\mu(u-t), \sigma^2(u-t)\right]$$

where μ is the *drift*, and σ is the *volatility* (or diffusion coefficient)**.**

Note that the μ that appears in the lognormal model (which is described in Section 1.3 as the "drift *parameter*") refers to the drift in the *log* price. This is not quite the same as the rate of drift of the price itself, which is $\mu + \frac{1}{2}\sigma^2$. The corresponding stochastic differential equations for the log price and the price itself are $d(\log S_t) = \mu dt + \sigma dZ_t$ and $dS_t = (\mu + \frac{1}{2}\sigma^2)S_t dt + \sigma S_t dZ_t$.

These parameters are specific to the investment considered. From the expression above we can see that the proportional change in the price is lognormally distributed, so the returns over any interval do not depend on the initial value of the investment S_t.

Question 10.1

Why is the *proportionate* change in the price lognormally distributed?

An equivalent way of specifying the model is the equation:

$$\log(S_u) = \log(S_t) + \mu(u-t) + \sigma\varepsilon_u\sqrt{u-t}$$

where ε_u is a series of independent and identically distributed standard normal random variables. These are sometimes referred as the *innovations* of the stochastic process – *ie* the terms that introduce random variation into the process. Here we have "discretised" time, so that the model is now represented as a random walk. The times u and t might correspond, for example, to the first of January in different years.

1.2 Properties

The lognormal model has the following properties:

- **The mean and variance of the log returns are proportional to the length of the interval considered $(u-t)$, and so the standard deviation of the log returns (often taken to be a measure of volatility) increases with the square root of the interval.**

  The dependence on the length of the interval means that the mean, variance and standard deviation tend to infinity as the length of the interval increases, the exceptions being the simplified cases where there is no drift in the log price $(\mu = 0)$ **or where there is no volatility** $(\sigma = 0)$**.**

In other words, the mean will tend to infinity unless the log security price doesn't drift and the standard deviation will tend to infinity unless there is no volatility. This isn't necessarily as unrealistic as it first sounds.

Q

Question 10.2

Why is not as unrealistic as it first sounds?

Assuming zero volatility would be an extreme simplification because it renders the model deterministic! It means that there is nothing random about movements in the security price process, and in a no-arbitrage world the return on the security should equal the risk-free rate of interest because the security is itself risk-free.

The lognormal model has often been used to model ordinary share prices, in which case $\mu\,(> 0)$ represents the upward drift of log share prices due to growth in company profits, which is linked to other economic factors.

Although the mean and standard deviation of a security price will tend to infinity under this model, the mean and standard deviation of the annual changes in the log of the security price are constant.

Question 10.3

Explain why the mean and standard deviation of the annual changes in the log of the security price are constant under this model.

- **It is assumed that returns over non-overlapping intervals are independent of each other.**

  This is because the standard normal variables generating the random variation in the log of the security price are assumed to be independent.

- **We can write the value of the investment at time u as:**

  $$S_u = S_t \exp(X_{u-t})$$

  where $X_{u-t} \sim N\left[\mu(u-t), \sigma^2(u-t)\right]$

  This format should be familiar to you from Chapter 8.

- **We know that S_u is lognormally distributed and so we can write:**

  $$E[S_u] = S_t \exp\left(\mu(u-t) + \tfrac{1}{2}\sigma^2(u-t)\right)$$

  $$\text{var}[S_u] = (S_t)^2 \exp\left(2\mu(u-t) + \sigma^2(u-t)\right)\left[\exp\left(\sigma^2(u-t)\right) - 1\right]$$

  The mean and variance for the lognormal distribution for a random variable X with parameters μ and σ are given on page 14 of the Tables:

  $$E(X) = \exp\left(\mu + \sigma^2/2\right)$$

  and $$Var(X) = \exp\left(2\mu + \sigma^2\right)\left(\exp(\sigma^2) - 1\right)$$

1.3 *Empirical tests of the log random walk model of security prices*

Empirical results for testing the log random walk model are mixed.

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

Question 10.4 (Revision)

Can you remember what "technical analysis" is?

It should be pointed out that the lognormal model is widely used. For example, it is the assumed underlying process for share prices that we will use to price share options later on in the course. This is largely due to the (relative) mathematical simplicity of the model.

However, more detailed analysis reveals several weaknesses in the log random walk model.

Volatility, σ

The most obvious is that estimates of σ vary widely according to what time period is considered, and how frequently the samples are taken.

For example, the volatility has been found to be greater in recessions and periods of financial crisis. Also σ can be estimated based on daily, monthly or annual share price records. Even if the data used covers the same time period, this usually leads to different numerical estimates of σ.

As an example, the graph below shows the rolling annual volatility of the total returns on the FTSE All-Share index, between 1966 and 1998, and the rolling volatility from a simulated lognormal random walk with the same mean and standard deviation. The lognormal random walk clearly fails to capture the differing volatilities that are observed.

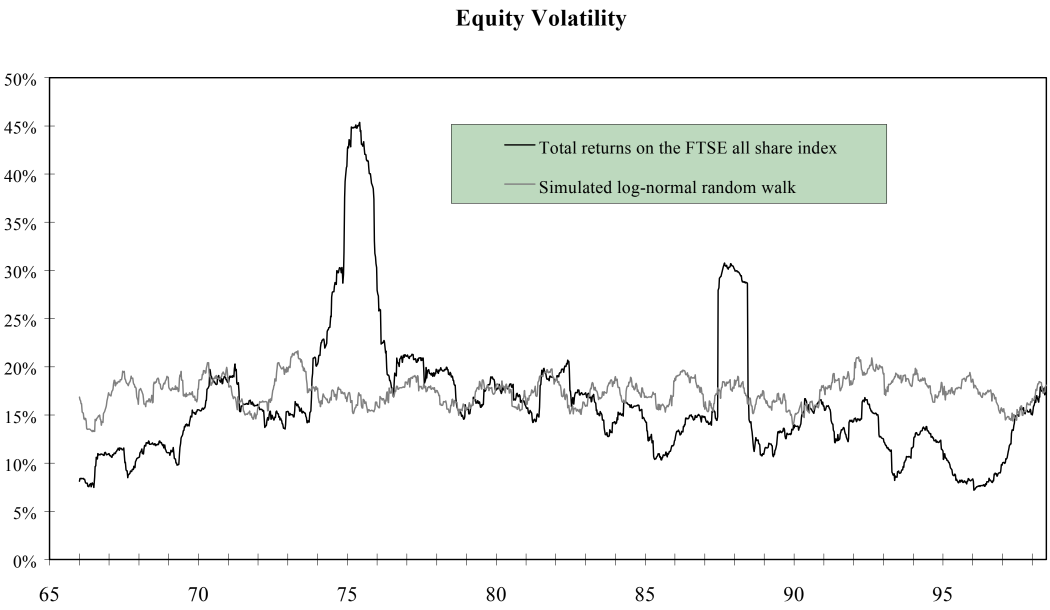


Figure 10.1: Equity volatility

(This chart is part of the Core Reading.)

The returns on the FTSE All-Share index are given by the line with the two tall peaks. The simulated lognormal random walk is much less volatile than the actual returns graph, which has volatility peaks around the market crashes of 1974 and 1987.

We can also take some evidence from option prices. The Black-Scholes formula describes option prices in terms of anticipated values of volatility over the term of the option. Given observed option prices in the market, it is possible to work backwards to the *implied volatility*, that is, the value of σ which is consistent with observed option prices.

Later on in this course, we will discuss the Black-Scholes formula. This formula expresses the price of an option as a function of several variables, one of which is σ. Hence, given the actual price of the option in the marketplace, together with values for all the other variables in the model (which we can actually observe), we can work backwards to derive the implied value for σ that is consistent with that observed price. In other words, the price of an option tells us implicitly what the market believes the volatility of the security price to be.

Examination of historical option prices suggests that volatility expectations fluctuate markedly over time.

The evidence therefore suggests that the assumption of a constant volatility may be inappropriate. Thus, investment returns could be *heteroscedastic* as opposed to *homoscedastic*, *ie* the volatility of the process is not constant. (The random walk model is based upon an assumption of homoscedasticity since σ is a constant.)

One way of modelling this behaviour is to take volatility as a process in its own right. This can explain why we have periods of high volatility and periods of low volatility. One class of models with this feature is known as ARCH – standing for *autoregressive conditional heteroscedasticity*. These models are covered briefly in Subject CT6.

Question 10.5

Explain what is meant by an ARCH model.

Drift parameter, μ

A more contentious area relates to whether the drift parameter μ is constant over time. There are good theoretical reasons to suppose that μ should vary over time. It is reasonable to suppose that investors will require a risk premium on equities relative to bonds.

The risk premium compensates the investor for the extra risk taken – both default risk (if compared to bonds) and volatility of the share price.

As a result, if interest rates are high, we might expect the equity drift, μ, to be high as well.

So, if the expected return on bonds is currently high say, then investors will require a correspondingly higher expected return on equities in order to make it worthwhile to hold them, instead of bonds. If this is not the case, then investors will sell equities and buy bonds until the expected returns are again brought back into line.

Mean reversion

One unsettled empirical question is whether markets are *mean-reverting*, or not. A mean-reverting market is one where rises are more likely following a market fall, and falls are more likely following a rise.

Hence, if returns have recently been above the long-run average level, then we might expect them to be lower than average over the next few periods, so that average returns revert back towards their long-run trend level.

There appears to be some evidence for this, but the evidence rests heavily on the aftermath of a small number of dramatic crashes. After a major crash, we might well expect the market to revert to its former level after sufficient time.

Momentum effects

Furthermore, there also appears to be some evidence of momentum effects, which imply that a rise one day is more likely to be followed by another rise the next day. For example, if returns increase, then everyone may jump on the bandwagon and drive prices even higher.

Normality assumption

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

This would seem to justify the consideration of Levy processes. Levy processes no longer appear in the actuarial syllabus. One of the features of a Levy process is that the sample paths are not necessarily continuous, but can have jump discontinuities.

So, the distribution of actual market returns appears to be more peaked and with fatter tails than is consistent with strict normality.

One way of comparing two distributions, especially if you are interested in their tails, is to draw a *Quantile-Quantile* or *Q-Q plot*. For example, you might wish to compare the actual distribution of the observed changes in the FTSE All-Share index against those that would be expected if the returns were lognormally distributed. This is done as follows:

1. Rank the n observations of the actual log investment returns in ascending order and denote them by $\log y_i$, $i = 1, \ldots, n$.

2. Find the log returns, $\log x_i$, such that the normal distribution function has the value i/n, *ie* $\Phi\left(\log x_i\right) = i/n$.

3. Plot the pairs $(\log x_i, \log y_i)$.

The closer the distribution of the actual returns is to a lognormal, the straighter will be the resulting plot. A straight line therefore indicates a perfect fit, whereas deviations away from a straight line at either end of the plot indicate that the tails of the distributions differ.

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

It is not important to understand the nature of the quintic distribution. This is merely included as an example of a non-normal distribution that provides a closer fit to the observed data.

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

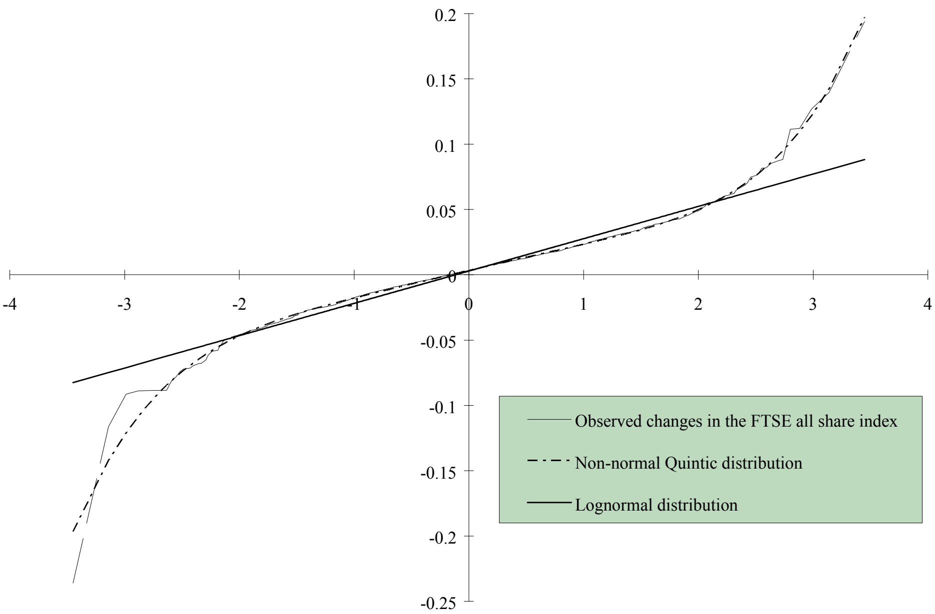


Figure 10.2: Q-Q graph

(This graph is part of the Core Reading.)

On the above graph the horizontal axis is measured in standard deviations from the mean $\left(\text{of } \log x_i\right)$, whilst the vertical axis is simply $\log y_i$.

It is also informative to consider what the volatility of a simulation from the non-normal distribution looks like. The chart below reproduces the volatility of the actual returns on UK Equities, between 1966 and 1998, with a lognormal simulation and also shows the rolling volatility of the non-normal distribution.

So, this is the first chart from above, with the addition of the thick solid line showing the rolling volatility of the non-normal quintic distribution.

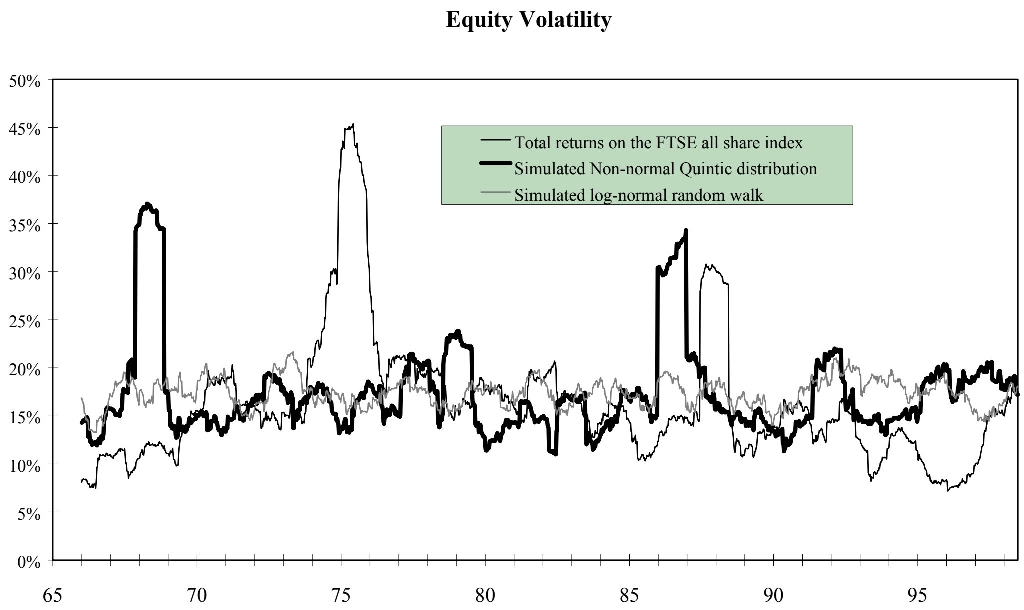


Figure 10.3: Equity volatilities – actual and simulated

(This graph is part of the Core Reading.)

Despite the non-normal quintic distribution having a constant volatility, the rolling volatilities show significant differences over different periods. This process gives rise to volatility, which has the same characteristics as the observed volatility from the equity market. This evidence suggests that modelling capital markets requires the use of distributions that more accurately reflect the returns observed. These distributions also provide an improved description of the varying volatility without requiring volatility to be modelled as a stochastic process.

One measure of these non-normal features is the *Hausdorff fractal dimension* of the price process. A pure jump process (such as a Poisson process) has a fractal dimension of 1. Random walks have a fractal dimension of 1½. Empirical investigations of market returns often reveal a fractal dimension around 1.4.

An understanding of fractal processes and the Hausdorff fractal dimension are beyond the syllabus. The paragraph above is saying that for any process, we can associate a number, the Hausdorff fractal dimension. This is a mathematical way of quantifying how "erratic" a particular stochastic process is. The value for the lognormal model would be 1.5, but the actual value observed in the market is 1.4. Because these differ, one could argue that geometric Brownian motion does not provide a perfect model of security prices.

Question 10.6

The shares of Abingdon Life can be modelled using a lognormal model in which $\mu = 0.104$ *pa* and $\sigma = 0.40$ *pa*. If the current share price is 2.00, derive a 95% confidence interval for the share price in one week's time, assuming that there are exactly 52 weeks in a year.

Summary

So, to summarise Section 1.3, the continuous-time lognormal model may be inappropriate for modelling investment returns because:

- The volatility parameter σ may not be constant over time. Estimates of volatility from past data are critically dependent on the time period chosen for the data and how often the estimate is re-parameterised.

- The drift parameter μ may not be constant over time. In particular, bond yields will influence the drift.

- There is evidence in real markets of mean-reverting behaviour, which is inconsistent with the independent increments assumption.

- There is evidence in real markets of momentum effects, which is inconsistent with the independent increments assumption.

- The distribution of security returns $\log\left(S_u / S_t\right)$ has a taller peak in reality than that implied by the normal distribution. This is because there are more days of little or no movement in financial markets.

- The distribution of security returns $\log\left(S_u / S_t\right)$ has fatter tails in reality than that implied by the normal distribution. This is because there are big "jumps" in security prices.

1.4 Market efficiency

It is important to appreciate that many of the empirical deviations from the random walk do not imply market inefficiency.

This implies that the lognormal model is not the only model that is consistent with market efficiency. If we believe in market efficiency then we must be able to explain how an efficient market can be consistent with the evidence we've given above.

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

Consequently, the hypothesis of market efficiency can be difficult to disprove.

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

2 Cross-sectional and longitudinal properties

We can distinguish two ways of looking at the properties of time series models. For a given quantity (for example, the force of inflation) we can imagine a two-dimensional table in which each row is one simulation and each column corresponds to a future projection date. All the simulations start from the same starting position, which is determined by reference to market conditions on the run date.

An example of a summary table appears below, with the simulations $i = 1, \ldots, 1000$ listed down the side and the future time periods $t = 1, \ldots, 10$ across the top. All the simulations start with a current inflation rate of 2.5%.

	I(t): Force of inflation at time t					
Simulation, i	$t = 0$	$t = 1$	$t = 2$	$t = 3$	$\cdots$	$t = 10$
$i = 1$	2.5	1.9	2.2	1.4	$\cdots$	3.6
$i = 2$	2.5	1.6	2.8	2.6	$\cdots$	4.9
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
$i = 1{,}000$	2.5	2.6	3.1	1.7	$\cdots$	2.6

2.1 Cross-sectional properties

A *cross-sectional property* fixes a time horizon and looks at the distribution over all the simulations.

This kind of property is therefore derived from a column of figures. For example, if we want to know $P\big(I(2) < 3\big)$, we can look at the 1000 values of $I(2)$, the force of inflation at time 2, and work out the proportion that are less than 3.

For example, we might consider the distribution of inflation next year. Implicitly, this is a distribution conditional on the past information that is built into the initial conditions, and is, of course, common to all simulations.

If those initial conditions change, then the implied cross-sectional distribution will also change. As a result, cross-sectional properties are difficult to validate from past data, since each year of past history typically started from a different set of conditions.

If next year's inflation (at time $t = 1$) actually turns out to be 2.4%, then we will get a different distribution of projections for time $t = 2$ if we repeat the simulations in a year's time. Hence, each starting year will lead to a different set of results.

However, the prices of derivatives today should reflect market views of a cross-sectional distribution. Cross-sectional information can therefore sometimes be deduced from the market prices of options and other derivatives.

Recall, for example, the implied volatility we discussed earlier. This is the volatility implicitly assumed by the market, based on the current information.

2.2 *Longitudinal properties*

A *longitudinal property* picks one simulation and looks at a statistic sampled repeatedly from that simulation over a long period of time. It is derived from a row of figures in the table above.

For example, we might consider one simulation and fit a distribution to the sampled rates of inflation projected for the next 1000 years.

For example, we might pick the second simulation. If we want to know about $I(2)$ again, as we did in the cross-sectional case, we can do the following. The first sample value of $I(2)$ is 2.8. The next number in the row, 2.6, can then be thought of as the second value, but with starting point 1.6 (*ie* the time 3 value with respect to the time 1 value is the second sample of $I(2)$). Repeating this, the whole row can be thought of as a sample of values of $I(2)$, with the tth value in the row giving the sample value of $I(2)$ with starting condition given by the $(t-2)$th value.

For some models, this longitudinal distribution will converge to some limiting distribution as the time horizon lengthens. Furthermore this limiting distribution is common to all simulations. Such convergence results are sometimes called ergodic theorems. The resulting distribution is an *ergodic* distribution.

An analysis of when and how this might occur is beyond the syllabus. Suffice it to say that under suitable circumstances, it doesn't matter which simulation you look at, and as long as your time period is long enough, the result will be close to the "true" result (obtained as $t \to \infty$).

Unlike cross-sectional properties, longitudinal properties do not reflect market conditions at a particular date but, rather, an average over all likely future economic conditions. Most statistical properties computed from historical data are effectively longitudinal properties.

A set of historical data effectively forms a single simulation, *ie* a row in the table above. It follows that a statistical property derived from real historical data will generally be a longitudinal one.

In a pure random walk environment, asset returns are independent across years and also (as for any model) across simulations.

By random walk here we mean the lognormal model, *ie* it is log of share price that follows the actual random walk (or Brownian motion). The log returns therefore follow a white noise process consisting of independent values for each year.

As a result, cross-sectional and longitudinal quantities coincide.

In other words, if we have a table similar to the one above giving the log returns for individual years, then whether we move down a column or along a row, each number is independent of the previous one. Any evidence that cross-sectional and longitudinal values of returns are not actually equal in real life is evidence against the lognormal model.

Care must be taken to ensure that a cross-sectional property is not used when a longitudinal one is required, or vice versa.

For example, it is common to see historical volatility used as an input to option pricing models. The historical volatility estimates a longitudinal standard deviation, while option pricing requires cross-sectional inputs.

An estimate of volatility based on historical returns will be a longitudinal property, as discussed in general above. However, when pricing an option, the volatility should be an estimate of the future and should therefore be estimated conditional on the current information, *ie* a cross-sectional value.

To equate the two is valid in a random walk setting, but not for more general models which we now consider.

In the lognormal model, as mentioned above, the returns in each year are all independent. So we don't need to distinguish between the two. However, the main example of an autoregressive model considered in the next section (the Wilkie model) doesn't have the independence property and so it is not valid to equate cross-sectional and longitudinal properties.

Question 10.7

Are the longitudinal properties of log returns simulated using the lognormal model identical to their cross-sectional properties?

3 Autoregressive models

3.1 Introduction

The material in this and the next section provides applications of *autoregressive moving average* (*ARMA)* time series models introduced in Subject CT6. Some familiarity with that material will be helpful.

These processes are discrete-time models, in contrast to the continuous-time lognormal model discussed in the last section. This disparity is of no consequence, however – we can always consider the lognormal model in discrete time for comparison's sake. For example, we can think of the log share prices as following a discrete-time random walk rather than a Brownian motion. It should also be noted that when Core Reading refers to a "random walk" process for share prices, it is the lognormal model that it is referring to. Hence it is the *log* of the share prices that follow the actual random walk – the share prices themselves follow a geometric random walk.

3.2 Mean reversion

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

The log share price process, and hence the share price process itself, is therefore a non-stationary process.

However, there are many quantities that should not behave like this. For example, we do not expect interest rates to jump off to infinity, or to collapse to zero.

On the other hand, the *returns* in the lognormal model are independent of each other from one year to the next. This would not be expected to be the case for all processes, for example, interest rates.

Instead, we would expect some mean-reverting force to pull interest rates back to some normal range. In the same way, while dividend yields can change substantially over time, we would expect them, over the long run, to form some stationary distribution, and not run off to infinity. Similar considerations apply to the annual rate of growth in prices or in dividends. In each case, these quantities are not independent from one year to the next; times of high interest rates or high inflation tend to bunch together, *ie* the models are autoregressive.

In particular then, we do not expect white noise to be a good model of interest rates, dividend yields or inflation. Instead, some dependence on previous values must be built into the model, and this is what autoregressive models do.

Recall from Subject CT6, that a general $ARMA(p,q)$ model X_t is defined by the relation:

$$X_t = \mu + \alpha_1 \left(X_{t-1} - \mu \right) + \cdots + \alpha_p \left(X_{t-p} - \mu \right) + e_t + \beta_1 e_{t-1} + \cdots + \beta_q e_{t-q}$$

where e_t is a zero-mean white noise process. For the simple $AR(1)$ case this becomes:

$$X_t = \mu + \alpha \left(X_{t-1} - \mu \right) + e_t$$

As shown in Subject CT6, this process is stationary in the long-run if and only if $|\alpha| < 1$. In that case, we can rearrange this equation slightly:

$$X_t - \mu = \alpha \left(X_{t-1} - \mu \right) + e_t$$

to show that the process is *mean-reverting.* If we ignore the white noise error term, the distance of X_t from its long-run mean μ is α times the previous distance. If $|\alpha| < 1$ then the distance is decreasing so the process is being pulled back to the mean. It will usually be the case that $0 < \alpha < 1$ in our economic examples.

Q

Question 10.8

How can we describe the process X_t if $\alpha = 1$?

Note that if $|\alpha| \geq 1$, then the process is not mean-reverting. We should be slightly careful about the terminology used here. In the context of stochastic investment models, the term "autoregressive" is often used to describe models in which future price or yield movements tend to a long-run average value. In other words, we assume implicitly that we are talking about stationary, mean-reverting autoregressive processes.

Question 10.9

Which of the following variables are likely to be mean-reverting:

1. the money supply
2. monetary growth rates
3. actuarial salaries?

So, how can we model an economy where some quantities are mean-reverting?

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

It is impossible, for example, to have a negative dividend, and hence a negative dividend yield.

Question 10.10

Which of the following economic variables could take negative values?

(a) inflation
(b) equity dividend yields
(c) real yields on index-linked bonds
(d) nominal yields on conventional bonds
(e) dividends
(f) dividend growth
(g) property prices.

4 The Wilkie model

4.1 Introduction

We will now look briefly at the Wilkie model, which is an example of an autoregressive model that can be used to model various economic variables.

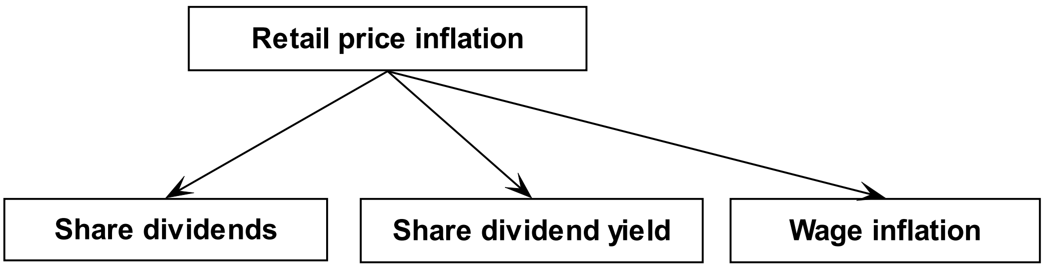
Although it is rarely used in its published form, the Wilkie model (full details of the Wilkie model are set out in his paper "More on a Stochastic Asset Model for Actuarial Use", BAJ Volume 1, pages 777–964) provides a convenient published example of a model designed explicitly for long-term applications.

David Wilkie has also written more recent papers relating to his model.

4.2 Structure of the Wilkie model

The Wilkie model is used to simulate the future values of economic variables in each future year, over a number of years.

The structure of the Wilkie model as originally conceived can be shown diagrammatically as a "cascade" model. This form shows how the parameters of the model can be estimated serially, but also "hides" some of the implicit structure (in particular the lack of correlation between some variables) imposed on the relationships between the variables.



etc.

(This diagram is part of the Core Reading.)

You can see from the diagram that, for example, the rate of inflation in a particular year will feed into the values of the dividend yield for the next year but not vice versa. More generally, under this *cascade structure*, higher variables influence those lower down but not vice versa, *ie* the causality "cascades" down through the model.

The "*etc*" in the diagram indicates that this is just a subset of the full model, which includes many other variables, and that there are further tiers of variables below the ones shown, *eg* **dividend growth, share prices, interest rates (short- and long-term), property returns, index-linked stock and currency exchange rates.**

The Wilkie model can also be recast as a single multivariate model based on a *vector autoregressive moving average (VARMA) process*. The vector representation makes it clear that the parameter count is far higher than is first apparent as a zero value is a parameter value.

In the vector representation, rather than modelling each variable individually, with its own equation, we instead put the variables of interest into a vector and model the vector as an ARMA process, *ie*:

$$X_t = A + BX_{t-1} + C_0 Z_t + C_1 Z_{t-1} + C_2 Z_{t-2}\,.$$

where:

- X_t is a vector of this period's values of the variables of interest, *eg* inflation, the share dividend yield.
- X_{t-1} is a vector of last period's values of the same set of variables.
- $Z_{t-i}, i = 0, 1, 2$, are vectors of random innovation terms (each with a standard normal distribution).
- A, B, C_0, C_1, C_2 are all matrices of constants, estimated using historical data.

In fact, the Wilkie model turns out to be a VARMA(1, 2) process, which means it is:

- *autoregressive of order 1 or AR(1)*, as this year's vector of variable values is assumed to depend upon last year's vector of variable values (but not values before then)
- *moving average of order 2 or MA(2)*, as this year's vector of variable values is assumed to depend upon the random innovation terms from this year, last year and also from two years ago.

In theory, it would be possible to set up the model so that each of the variable values this year is directly dependent on each of the variable values last year, which would result in a set of equations with a large number of parameters to which we could assign values. In practice, however, Wilkie assumed a much simpler *cascade structure*, which drastically reduces the number of parameters actually used. For example, in the Wilkie model, inflation this year is linked only to inflation last year, which is equivalent to saying that the parameters for the links to last year's values of all the other variables are equal to zero.

In fact, Wilkie models price inflation as an AR(1) process using the equation:

$$I_t = QMU + QA\left[I_{t-1} - QMU\right] + QSD.\,QZ_t$$

Here:

- I_t, the force of inflation, is derived from movements in the Retail Price Index Q_t, so that $I_t = \ln Q_t - \ln Q_{t-1}$
- QMU, QA and QSD are parameters to be estimated
- QZ_t is a series of independent, identically distributed (*iid*) standard normal random variables, also called "innovations".

Notice the structure of the equation:

This year's value = long-run mean + $QA \times$(last year's value − long-run mean)
+ a "random shock to the system"

In this structure:

- the long-run mean is QMU
- the *auto-regressive* component is based upon the difference between last year's value and the long-run average value of inflation, *ie* $QA\left[I_{t-1} - QMU\right]$. As $|QA| < 1$, the process is also *mean-reverting*.
- the "shock to the system" (random component) is $QSD.QZ_t$. QSD dictates the typical size of this "shock".

Also, none of the other variables in the model (*eg* dividend streams and dividend yields) appear in the equation because inflation is assumed to influence them, but they are assumed not to influence inflation.

Q

Question 10.11

What do you think might be an appropriate value for QMU ?

The original form was decided upon using a combination of high-level economic expectations, statistical analysis and the requirement to keep the model parsimonious, *ie* to avoid the estimation of "too many" parameters. It should be remembered that there is no over-riding economic reason for the model to have the particular form it was originally published in. Analysis using different (for example more recent) data sets may indicate that the best statistical model has a different form. This may also be true when analysing economies other than the UK.

5 The main alternative models

5.1 Economic constraints and market efficiency

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

Consequently, they are not based directly upon economic theory, although Wilkie did use some judgement regarding how the economy worked in practice when deciding upon the pyramid structure of his model, *eg* that inflation is the key process that influences other variables.

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

Asset	*Prospective income*	*Expected income growth pa*	*Market value*
A	**$60**	**2%**	**$2,000**
B	**$60**	**3%**	**$1,500**

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

Asset	*Prospective income*	*Expected income growth pa*	*Market value*
A	**$60**	**2%**	**$1,500**
B	**$60**	**3%**	**$2,000**

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

Riskier assets should offer a higher expected return to compensate investors for the additional risk. When developing a model, we might therefore wish to ensure that the model parameters are consistent with this prediction.

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

Ultimately, it is the modeller's responsibility to judge the importance or otherwise of consistency with market efficiency, or indeed any other economic constraint, when developing a model.

The decision to adopt an efficient or inefficient market model has important implications for model output. For example, statistical models often produce yields that fluctuate more than would be the case if markets were efficient.

Such models steer investors into tactical switches, from low-yielding assets to high-yielding ones – a strategy that, on the basis of past data, would have been profitable in several countries including the UK. Recommended asset allocations will thus fluctuate substantially between different run dates, according to market conditions at the time.

In practice, re-calibrating the model to give a predetermined answer each time it is run, while ignoring the impact of future re-calibrations on the dynamics of the economy usually mitigates this problem. For example, in the context of UK pension schemes, a statistical model might be calibrated to suggest 75% equities, while a comparable scheme in the US or the Netherlands would be 60% or 40% respectively. These figures are derived in order to reproduce the average scheme asset allocation in each country. This calibration technique is a crude example of an *equilibrium approach*.

5.2 Economic models

An alternative to purely statistical models is to give more weight to economic theory. If we are convinced by the theoretical or empirical arguments for market efficiency or purchasing power parity, we may wish to use these theories to guide the construction of stochastic investment models.

An *economic* model is therefore one that imposes constraints upon both the structure of a model and its parameter variables based upon economic theory. An *econometric* model is a time series model with economic constraints imposed upon the model structure and parameters.

By way of an example, we can note that purchasing power parity suggests that the real exchange rates between two countries ought to remain constant, suggesting the presence of a simple cointegrating relationship between price levels measured in a common currency. If we impose this and other similar economic relationships upon our model then we are fitting an econometric model.

> ***Question 10.12***
>
> What is the main advantage of an economic model?

For example, market efficiency considerations would force us to identify two factors in dividend growth – one component that strongly mean reverts in the short term, and a longer term component which mean reverts more slowly. Statistical analysis of historical dividend growth enables us to be reasonably precise about the short-term effect, but this effect alone is insufficient to explain observed yield volatility. The long-term effect drives longer-term expectations, and hence yields, but is swamped by the short-term effect if dividend statistics are considered in isolation. In other words, we can use the volatility of observed yields, together with notions of market efficiency, to infer the long-term behaviour of income indices. In this way, we can achieve a model which is statistically adequate, but which can also be rationalised in an efficient market framework.

Any stochastic investment model, combined with a suitable business model, can provide simulated distributions of a wide range of future cashflows and capital items. These distributions can be computed for a range of different strategies. Some strategies will turn out to be high risk / high return and others low risk / low return.

Such projections might be used as part of an asset liability modelling process to determine the optimal asset allocation of an investment fund. Alternatively, they could be used as part of a project appraisal process, to help determine whether to proceed with a capital project. The use of such models to assess the viability of a capital project is discussed further in Subject CA1.

Depending on the precise definitions of risk and return, different strategies will appear optimal. So, some form of objective technique involving the use of an objective function is required in order to determine the optimal policy choice.

This causes a problem in using the models to support decisions. Although asset projection is well developed, interpreting the output remains more of an art than a science. As always, much judgement and skill is required, as most problems do not have a single "correct" or definitive answer.

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

Question 10.13

Based on the parameters suggested in his 1995 paper, the Wilkie model assumes that the property asset class has a higher expected total return and a lower variance of return (and a similar covariance) to the equity asset class. In what sense could this be said to be inconsistent with the predictions of mean-variance portfolio theory, given that investors invest in equities in practice?

6 Estimating parameters for asset pricing models

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

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

It is simplest to estimate parameters for a statistical model, such as the lognormal model discussed in Section 1, because no consideration of economic or other non-statistical factors is required. The calibration process itself will be based upon techniques discussed in earlier subjects. Even here though, a degree of judgement will be required from the modeller.

Of course, there may not always be as much data as we would like, and the statistical error in estimating parameters may be substantial. Consequently, the results obtained may lack credibility.

Furthermore, there is a difficulty in interpreting data, which appears to invalidate the model being fitted. For example, what should be done when fitting a Gaussian model in the presence of large outliers in the data? Perhaps the obvious course of action is to reject the hypothesis of normality, and to continue building the model under some alternative hypothesis. After all, in many applications, the major financial risks lie in the outliers, so it seems foolish to ignore them.

The rejection of normality is a big step because non-normal models are generally much more difficult to apply and use than those based on the assumption of normality.

Q

Question 10.14

Why do the major financial risks lie in the outliers?

In practice, a more common approach to outliers is to exclude them from the statistical analysis, and focus attention instead on the remaining residuals that appear more normal. The model standard deviation may be subjectively nudged upwards after the fitting process, in order to give some recognition to the outliers which have been excluded.

It is common practice in actuarial modelling to use the same data set to specify the model structure, to fit the parameters, and to validate the model choice. A large number of possible model structures are tested, and testing stops when a model is found which passes a suitable array of tests. Unfortunately, in this framework, we may not be justified in accepting a model simply because it passes the tests. Many of these tests (for example, tests of stationarity) have notoriously low power, and therefore may not reject incorrect models. Indeed, even if the "true" model was not in the class of models being fitted, we would still end up with an apparently acceptable fit, because the rules say we keep generalising until we find one.

As we add more and more variables to a model, the model will necessarily fit the historical data more closely. Whether it is capable of a meaningful interpretation is another matter. Again, and as always, the modelling process relies heavily on the skill and judgement of the modeller.

This process of generalisation tends to lead to models that wrap themselves around the data, resulting in an understatement of future risk, and optimism regarding the accuracy of out-of-sample forecasts. For example, Huber recently compared the out-of-sample forecasts of the Wilkie model to a naïve "same as last time" forecast over a 10-year period. The naïve forecasts proved more accurate.

An alternative approach to this that is widely used by econometricians and which avoids this problem of "data mining" is the *general-to-specific* approach. Here the modeller starts with a very general model, which includes all the variables (both current values and lagged values) that are thought likely to influence the variable being modelled. The model is then made more specific by eliminating variables (one at a time) that do not materially affect the significance of the fit to past data.

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

Thus, we need to decide which is the most important – fitting past data as accurately as possible or complying with economic theory. In practice, there will often be a trade off between these two objectives.

Q

Question 10.15

Why should the modeller not aim solely to fit the model as accurately as possible to past data?

7 Exam-style question

We finish this chapter with an exam-style question based on the lognormal model of share prices.

Question

An investor has decided to model PPB plc shares using the lognormal model. Using historical data, she has estimated the drift and volatility parameters to be 6% and 25% respectively. PPB's current share price is $2.

(i) Calculate the mean and variance of PPB's share price in one year's time. [3]

(ii) Calculate the probability that:

(a) PPB's shares fall in value over the next year.

(b) PPB's shares yield a return of greater than 30% over the next year.

Assume that no dividends are to be paid over the next year. [4]

[Total 7]

Solution

(i) ***Mean and variance of PPB's share price in one year's time***

Based on the information given in the question:

$$\log S_1 - \log S_0 \sim N\left[\mu, \sigma^2\right]$$

So, using the formulae for the mean and variance of the lognormal distribution from page 14 in the *Tables*, we have:

$$E\left[S_1 \middle| S_0\right] = S_0 \exp\left(\mu + ½\sigma^2\right)$$

$$Var\left[S_1 \middle| S_0\right] = S_0^2 \exp\left(2\mu + \sigma^2\right)\left(\exp\left(\sigma^2\right) - 1\right)$$

So, with $\mu = 0.06, \sigma = 0.25$ and $S_0 = 2$ we have:

$$E\left[S_1 \middle| S_0\right] = 2 \times \exp\left(0.06 + ½ \times 0.25^2\right) = 2.1911$$

$$Var\left[S_1 \middle| S_0\right] = 2^2 \times \exp\left(2 \times 0.06 + 0.25^2\right)\left(\exp\left(0.25^2\right) - 1\right) = 0.30963$$

(ii)(a) ***Probability that PPB's shares fall in value over the next year***

Here we want the probability that $S_1 < S_0$.

$$\begin{aligned}
P\left[S_1 < S_0\right] &= P\left[\frac{S_1}{S_0} < 1\right] \\
&= P\left[\log\left(\frac{S_1}{S_0}\right) < \log(1)\right] \\
&= P\left[\frac{\log\left(\frac{S_1}{S_0}\right) - 0.06}{0.25} < \frac{0 - 0.06}{0.25}\right]
\end{aligned}$$

ie $\quad P\left[S_1 < S_0\right] = P\left[Z \sim N[0,1] < -0.24\right]$

$$= 1 - \Phi(0.24)$$

$$= 1 - 0.59483$$

$$= 0.40517$$

So, there is a probability of almost 41% that the share price will fall over the next year.

(ii)(b) ***Probability that PPB's shares yield more than 30% over the next year***

Here we want the probability that $S_1 > 1.30S_0$.

$$P\left[S_1 > 1.30S_0\right] = P\left[\frac{S_1}{S_0} > 1.30\right]$$

$$= P\left[\log\left(\frac{S_1}{S_0}\right) > \log(1.30)\right]$$

ie $$= P\left[\frac{\log\left(\frac{S_1}{S_0}\right) - 0.06}{0.25} > \frac{\log(1.30) - 0.06}{0.25}\right]$$

$$= P\left[Z \sim N[0,1] > 0.80946\right]$$

$$= 1 - \Phi(0.80946)$$

$$= 1 - 0.79087$$

$$= 0.20913$$

So, there is a 21% or so chance that the share will yield a return of 30% or more over the next year.

The relatively high probabilities in both (a) and (b) reflect the volatility parameter of 25%, which isn't unrealistic for an individual share.

8 End of Part 2

What next?

1. Briefly **review** the key areas of Part 2 and/or re-read the **summaries** at the end of Chapters 7 to 10.
2. Attempt some of the questions in Part 2 of the **Question and Answer Bank**. If you don't have time to do them all, you could save the remainder for use as part of your revision.
3. Attempt **Assignment X2**.

Chapter 10 Summary

The continuous-time lognormal model

$\ln S_u - \ln S_t \sim N\left(\mu(u-t), \sigma^2(u-t)\right)$ where S_t is the share price at time t, μ is the *drift parameter* and σ is the *volatility parameter* and $u > t$.

- $E[S_u | S_t] = S_t \exp\left(\mu(u-t) + \frac{1}{2}\sigma^2(u-t)\right)$
- $V[S_u | S_t] = E^2[S_u | S_t]\left\{\exp\left(\sigma^2(u-t)\right) - 1\right\}$
- The mean and variance of the log returns are proportional to the interval $u-t$
- $\ln S_t$ has independent and stationary increments.
- $\ln S_t$ has continuous sample paths.

The continuous-time lognormal model may be inappropriate for modelling investment returns because:

- The volatility parameter σ may not be constant over time. Estimates of volatility from past data are critically dependent on the time period chosen for the data and how often the estimate is re-parameterised.
- The drift parameter μ may not be constant over time. In particular, interest rates will influence the drift.
- There is evidence in real markets of mean-reverting behaviour, which is inconsistent with the independent increments assumption.
- There is evidence in real markets of momentum effects, which is inconsistent with the independent increments assumption.
- The distribution of security returns $\log\left(S_u / S_t\right)$ has a taller peak in reality than that implied by the normal distribution. This is because there are more days of little or no movement in financial markets.
- The distribution of security returns $\log\left(S_u / S_t\right)$ has fatter tails in reality than that implied by the normal distribution. This is because there are big "jumps" in security prices.

Cross-sectional and longitudinal properties

A *cross-sectional property* fixes a time horizon and looks at the distribution over all the simulations.

A *longitudinal property* picks one simulation and looks at a statistic sampled repeatedly from that simulation over a long period of time.

Cross-sectional and longitudinal properties of asset returns differ unless the asset returns are independent through time, *eg* they follow a random walk.

The Wilkie model

The Wilkie model:

- is a long-term model and its results are only meaningful when looking over the long term
- is a statistical model although the exact form was determined with reference to economic judgement
- has a cascading structure so that individual variables of interest can be analysed without the need to consider the rest of the economy
- can be presented in VARMA form.

Alternative models

There are a number of alternative models in use, some of which are statistically based, whereas others are based upon economic theory.

Models based on economic theory have the advantage that the results produced are:

- capable of an economic interpretation
- consistent with economic principles, such as arbitrage-free markets.

However, a purely statistical model may provide a better fit to past data and hence give more accurate predictions of the future.

Estimating parameters for models

The estimation of parameters is one of the most time-consuming aspects of stochastic asset modelling and requires much skill and judgement.

Chapter 10 Solutions

Solution 10.1

By definition:

$$\log(S_u) - \log(S_t) \sim Normal$$

ie $\log(S_u / S_t) \sim Normal$

ie $S_u / S_t \sim Lognormal$

Thus, the proportionate change is lognormally distributed, *ie* share price (percentage) returns are lognormally distributed.

Solution 10.2

Although the model predicts that the share price will drift higher and higher towards infinity, time is also tending to infinity. For example, is it unrealistic to assume that if we go until the end of planet Earth, share prices in Microsoft, say, will be worth zillions of dollars? No, not really. We cannot accurately predict share prices 100 years into the future, let alone beyond that! This exponential increase in uncertainty is consistent with the volatility tending to infinity as well.

Solution 10.3

The change in the log of the security price is $\log S_t - \log S_u$. This will always have the same distribution for any time period of length $t-u$. So the expected value and standard deviation of annual changes in the log of the security price are constant.

Solution 10.4

Technical analysis is using past data on asset prices to predict the future. Chartism is an example and you should recall from Chapter 1 that, if the market is weak-form efficient, there is no advantage to be gained from technical analysis.

Solution 10.5

An *ARCH (autoregressive conditional heteroscedasticity) model* is a model of the residuals from a time series regression in which the residuals themselves are assumed to be autoregressive. In practice, ARCH models usually specify that the *variance* of the residuals is autoregressive and not constant. For example, an order 1 ARCH model is defined by:

$$X_n = \mu + e_n \sqrt{\alpha_0 + \alpha_1 \left(X_{n-1} - \mu\right)^2}$$

where e_n is a zero-mean white noise with variance 1. The standard deviation of the residual is therefore equal to the square root term. Thus, a large deviation from the mean at time $n-1$ is followed by a period of high volatility, as the variance will be large. Conversely, a small deviation from the mean will tend to be followed by further small change.

Solution 10.6

Here:

$$\mu = 0.104 \quad \sigma = 0.40 \quad S_t = 2.00 \quad u - t = \tfrac{1}{52}$$

Thus, if we denote the share price in one week's time by S_u, then a 95% confidence interval for $\log(S_u)$ is given by:

$$\log(S_u) \;=\; \log(S_t) + \mu(u-t) \pm 1.96 \times \sigma\sqrt{u-t}$$

Substituting in the relevant values gives:

$$\begin{aligned}\log(S_u) \;&=\; \log 2 + \frac{0.104}{52} \pm 1.96 \times 0.40\sqrt{\tfrac{1}{52}} \\ &=\; \log 2 + 0.002 \pm 0.1087 \\ &=\; \left(0.5864, 0.8039\right)\end{aligned}$$

Thus, a 95% confidence interval for next week's share price is $\left(1.80, 2.23\right)$

Solution 10.7

Yes, they are the same because the (log) share price increments are independent over non-overlapping intervals.

Solution 10.8

If $\alpha = 1$, then X_t follows a random walk. This is not mean-reverting since the increments are independent of the past. However, the term "autoregressive" can still be used because the random walk model is a member of the same family of models.

Solution 10.9

Both the money supply and actuarial salaries are unlikely to be mean-reverting. Instead they will tend to grow without limit, *ie* they are generated by non-stationary processes.

The growth rate of money will not generally grow without limit, but will tend to oscillate around some long-term average value – it is therefore mean-reverting.

Solution 10.10

Dividends, dividend yields, and property prices would never take negative values.

Nominal yields are extremely unlikely to take negative values, unless there is high deflation.

It is possible for real yields on index-linked bonds to be negative, and this has occurred in the UK in recent times. However, negative real yields only occur very rarely, so it is usual to model real yields as positive quantities.

The others could take negative values.

Solution 10.11

QMU represents the annual force of inflation that will typically apply in the distant future. Wilkie used historical data from 1923 to 1994 to estimate a value of $QMU = 0.047$. However, most of the major economies have had relatively low inflation for a number of years now and some have introduced inflation targets. So you may have suggested a lower figure than this.

Solution 10.12

An economic model has the advantage that the model output is capable of an economic interpretation. In essence, it is able to tell us "where the results come from" and why the results are as they are.

Solution 10.13

Portfolio theory suggests that investors prefer more to less (higher expected returns) and dislike risk (variance of investment returns). It therefore suggests that an investor should invest only in assets that increase the expected return of its portfolio or decrease the variance of portfolio returns. Hence, based on Wilkie's 1995 parameters, mean-variance portfolio theory suggests that investors should never invest in equities. The fact that they do in practice means that the Wilkie model is inconsistent with the predictions of portfolio theory.

Solution 10.14

The major financial risks lie in the outliers because it is very low investment returns –*ie* returns in the lower tail of the distribution – that will lead to a fall in asset values and possible difficulties with regard to financial solvency.

Solution 10.15

The modeller should not aim solely to fit the model as accurately as possible to past data because the past might not be a very accurate indication of the future – *eg* if there are structural changes in investment and/or economic markets.

Chapter 11

Introduction to the valuation of derivative securities

Syllabus objectives

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

1. State what is meant by arbitrage and a complete market.

2. Outline the factors that affect option prices.

3. Derive specific results for options that are not model-dependent:

- *show how to value a forward contract*
- *develop upper and lower bounds for European and American call and put options*
- *explain what is meant by put-call parity.*

0 Introduction

A *derivative* is a security or contract which promises to make a payment at a specified time in the future, the amount of which depends upon the behaviour of some *underlying* security up to and including the time of the payment.

The value of the derivative contract at any time is therefore *derived* from that of the underlying security – which is itself often referred to simply as the *underlying*. The underlying asset could be a share, bonds, index, interest rate, a currency, or a commodity such as gold or wheat.

The remainder of the Subject CT8 course builds on the background material concerning derivatives introduced in Subjects CT1 and CT2 by focusing on some of the mathematics underlying the valuation of derivatives. After this introductory chapter, the remainder of Part 3 outlines the more detailed analysis and procedures used to value derivatives and, in particular, options.

1 Arbitrage

One of the central concepts in this section of financial economics is that of *arbitrage*. This topic was first introduced in Subject CT1 so you may wish to refresh your mind on this material before reading on.

1.1 Definition

Put in simple terms, an *arbitrage opportunity* is a situation where we can make a sure profit with no risk. This is sometimes described as a *free lunch*. Put more precisely an arbitrage opportunity means that:

(a) we can start at time 0 with a portfolio that has a net value of zero (implying that we are long in some assets and short in others). This is usually called a zero-cost portfolio.

(b) at some future time *T*:

- **the probability of a loss is 0**
- **the probability that we make a strictly positive profit is greater than 0.**

If such an opportunity existed then we could multiply up this portfolio as much as we wanted to make as large a profit as we desired. The problem with this is that all of the active participants in the market would do the same and the market prices of the assets in the portfolio would quickly change to remove the arbitrage opportunity.

So, if we can find a strategy, or an investment portfolio, that gives an arbitrage profit, then we can simply repeat this strategy, or buy this investment portfolio, as many times as possible before asset prices change to "close out" this opportunity for arbitrage profit.

Question 11.1

Imagine you find a horse race involving just three horses and you notice that the odds are as follows:

Nag Portachio	evens
GG Badeye	3/1
Horsey Mullet	7/1

Devise a betting strategy to make an arbitrage profit.

Note that, in gambling, 3/1 means that if you put a £1 bet on and GG Badeye wins then you receive £3 + £1 = £4. "Evens" means 1/1.

Question 11.2

Imagine instead that you are Bookmaker A offering the following odds on the outcome of a boxing match that cannot be drawn:

Tyson	1/2
Bruno	6/4

Bookmaker B is offering a "special" bet, where for a stake of £11.50, you receive £20 in total if Bruno wins and £6 in total if Tyson does.

Devise a betting strategy to make an arbitrage profit.

1.2 *The principle of no arbitrage*

The *Principle of No Arbitrage* states simply that arbitrage opportunities do not exist.

This principle is essential in the pricing of derivative securities. Essentially, any two assets that behave in exactly the same way must have the same price. If this were not true, we could buy the "cheap" one and sell the "expensive" one as many times as we liked, making an unlimited arbitrage profit!

If we assume that there are no arbitrage opportunities in a market, then it follows that any two securities or combinations of securities that give exactly the same payments must have the same price. This is sometimes called the "Law of One Price".

Question 11.3

Use the Bookmaker A odds given in the previous question to replicate the "special" bet being offered by Bookmaker B. Hence explain why the principle of no arbitrage is not applying here.

In real life derivative securities are generally based on financial assets instead of horse races and boxing matches. Instead of having bookmakers setting prices for "assets" dependent on the outcome of a sporting event, we have market-makers setting prices for assets dependent on the outcome of other financial assets, *ie* derivatives. But the ideas and concepts are the same.

2 Preliminary concepts

2.1 Notation

t	**is the current time**
S_t	**is the underlying share price at time *t***
K	**is the strike or exercise price**
T	**is the option expiry date**
c_t	**is the price at time *t* of a European call option**
p_t	**is the price at time *t* of a European put option**
C_t	**is the price at time *t* of an American call option**
P_t	**is the price at time *t* of an American put option**
r	**is the risk-free continuously-compounding rate of interest (assumed constant),** *ie r* here corresponds to the force of interest δ.

As we will see, the descriptions "European" and "American" don't have a geographical meaning here. They are labels used to distinguish two types of option.

2.2 European call options

Recall from Subject CT1 that the buyer of a European call option has the right, but not the obligation, to buy the share from the seller on a set date, usually referred to as the exercise date (or expiry date). Conversely, the seller or *writer* of the option, who has no choice, is obliged to deliver the share should the holder of the option exercise the option.

If at time T the share price S_T is less than the strike price, then the holder would lose money if he exercised the option to buy. So, in fact, the option will not be exercised if $S_T < K$.

This is because the holder of the option would be silly to pay K for the share when he could just as easily go and buy the share for S_T $(< K)$ in the open market. He therefore simply walks away and "loses" the premium that he originally paid for the option itself.

Conversely, the seller or writer of the option gets to keep the premium that he received at the outset and therefore makes an overall profit on the deal equal in value to that premium.

If $S_T > K$ then the holder can buy the share at the strike price and sell immediately at the market price, S_T , making a profit of $S_T - K$.

Here the option should be exercised because the holder can make a profit by doing so.

Question 11.4

What is the profit or loss to the seller or *writer* of the call option contract in this latter case?

The profit made by the holder at the expiry date is called the "payoff". The payoff on a European call option is therefore:

$$f(S_T) = \max\left\{S_T - K, 0\right\} \text{ at time } T.$$

To purchase a traded option the buyer must pay the seller an option *premium* when the contract is made.

As always the writer of the option makes a profit or loss that is equal and opposite to that of the buyer, so that the total profit or loss to the two parties sums to zero.

The writer of the option keeps the premium regardless of whether or not the option is ultimately exercised. It is paid by the buyer to the writer in order to enjoy the *choice* conferred by holding the option. Remember that the writer has no such choice, but must trade if the buyer wishes to do so. Later in this chapter, we will discuss the factors that influence the value of the option, *ie* how much the premium should be.

A call option is described as:

- *in-the-money* if the current price, S_t , is greater than the strike price, K
- *out-of-the-money* if $S_t < K$
- *at-the-money* if $S_t = K$.

Hence, an in-the-money call option is one that would result in an immediate profit – ignoring the premium originally paid – if it could be exercised now, whereas an out-of-the-money option would produce an immediate loss. Recall that if the option is European then it can be exercised only at the expiry date.

Example

Suppose that the current price of Share X is 115 and that a call option is available on Share X with an exercise price of 110.

The option is currently *in-the-money*, as the share price exceeds the exercise price. If the share price remained unchanged until the exercise date, then the holder of the option would exercise it and make a profit of 5. If instead the share price fell to 105 at the exercise date, then it would not be exercised, as the share would then be worth less than the exercise price. In this case, if the option holder actually wanted to buy the share, they would be able to do so more cheaply on the open share market.

2.3 European put options

The holder of the contract has the right to *sell* one share at time T to the issuer at the strike price K.

Again the choice, in this case to sell the share on the specified exercise date, lies with the buyer of the option, who is not obliged to sell if he chooses not to do so. As before, the writer of the put is obliged to buy the share from the option holder should the latter exercise the option to sell.

The option will only be exercised if $S_T < K$ and the payoff is, therefore:

$$f(S_T) = \max\{K - S_T, 0\} \text{ at time } T.$$

Question 11.5

Why is the put option exercised only if $S_T < K$?

A put option is described as:

- *in-the-money* if $S_t < K$
- *out-of-the-money* if $S_t > K$
- *at-the-money* if $S_t = K$.

As before, an in-the-money option is one that would result in an immediate profit – ignoring the premium originally paid – if it could be exercised now, whereas an out-of-the-money option would produce an immediate loss.

2.4 American options

The only difference between an American and a European option is that with an American option the holder can exercise the option before the expiry date, not just *on* the expiry date, as is the case for a European option. As always, the writer of the option is obliged to trade should the holder wish to do so.

Note that the names *European option* and *American option* have arisen for historical reasons. There is no longer any direct link with the place where the contracts are traded.

2.5 Other terminology

Intrinsic value

The *intrinsic value* of a derivative is the value assuming expiry of the contract immediately rather than at some time in the future. For a call option, for example, the intrinsic value at time t is:

$$\max\{S_t - K, 0\}$$

> **Question 11.6**
>
> What is the intrinsic value of a put option at time t?

Thus, the intrinsic value of an option is:

- positive, if it is in-the-money
- equal to zero, if it is at-the-money
- equal to zero, if it is out-of-the-money.

Time value

The *time value* or *option value* of a derivative is defined as the excess of an option's premium over its intrinsic value. By premium here we mean the current price, not the premium originally paid. It primarily represents the value of the *choice* that the option provides to its holder. The more valuable the choice is to the holder, perhaps because of the greater uncertainty that he faces about future share price movements for example, the greater is the time value.

3 Factors affecting option prices

3.1 Introduction

A number of mathematical models are used to value options. One of the more widely used is the *Black-Scholes* model. This uses five parameters to value an option on a non-dividend-paying share. The five parameters are:

- the underlying share price, S_t
- the strike price, K
- the time to expiry, $T-t$
- the volatility of the underlying share, σ
- the risk-free interest rate, r.

In the case of a dividend-paying share, we can consider dividends to be a sixth factor.

Here the price of an option means the size of the option premium paid at the outset.

3.2 Underlying share price

The first parameter is the initial price of the underlying share at time t, S_t.

The effect of the price of the underlying share on a typical call option is shown in Figure 11.1.

Note that the dotted line in the graph represents the intrinsic value. The time value is therefore the vertical distance between the actual price, *ie* premium, and the intrinsic value. Note that for a call option on a non-dividend-paying share this is always positive.

Note that the price for a call option is always greater than the intrinsic value. This follows on from the lower bound derived in Section 5 below: namely that:

$$c_t \geq S_t - Ke^{-r(T-t)} > S_t - K$$

ie $$c_t \geq \max\left[S_t - Ke^{-r(T-t)}, 0\right] \geq \max\left[S_t - K, 0\right] = \text{intrinsic value}$$

The share price affects the option price or premium differently for call options and put options.

Question 11.7

Suppose that the price of Share X is 112 and that a put option on Share X with an exercise price of 110 is currently priced at 5. What are the intrinsic value and time value of the option?

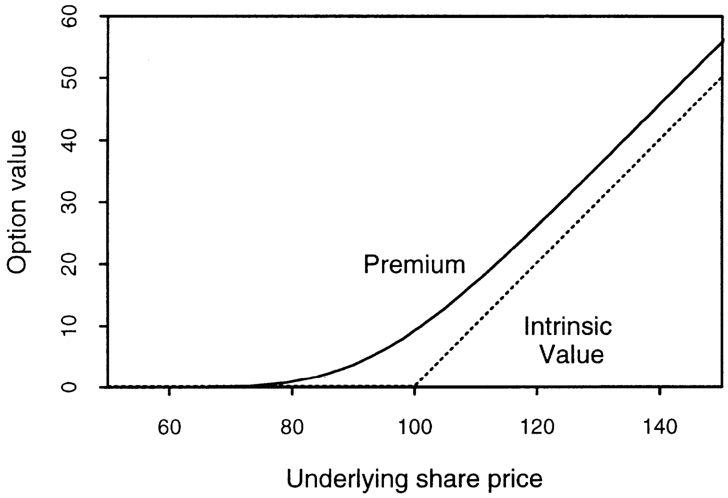


Figure 11.1: Call option premium and intrinsic value as a function of the current share price, S_t

Call option

In the case of a call option, a higher share price means a higher intrinsic value (or, where the intrinsic value is currently zero, a greater chance that the option is *in-the-money* at maturity). A higher intrinsic value means a higher premium.

Put option

For a put option, a higher share price will mean a lower intrinsic value and a lower premium.

In each case the change in the value of the option will not match precisely the change in the intrinsic value because of the later timing of the option payoff.

So, the time value is not constant with respect to share price. The graph shows that the time value is greatest when there is more uncertainty in the outcome. When the share price is well above or below the exercise price, the question of whether exercise will take place is more certain, so the time value is smaller. The uncertainty is greatest when the share price is around the exercise price.

3.3 Strike price

In the case of a call option, a higher strike price means a lower intrinsic value. A lower intrinsic value means a lower premium. For a put option, a higher strike price will mean a higher intrinsic value and a higher premium. In each case the change in the value of the option will not match precisely the change in the intrinsic value because of the later timing of the option payoff.

3.4 Time to expiry, T – t

The longer the time to expiry, the greater the chance that the underlying share price can move significantly in favour of the holder of the option before expiry. So the value of an option will increase with term to maturity. This increase is moderated slightly by the change in the time value of money.

Question 11.8

We can also note that the longer the time to expiry, the greater the chance that the underlying share price can move *against* the holder of the option before expiry. Why therefore does the value of a call option increase with term to maturity?

Whereas this is true for *all* the simple call options we consider in this course, it may not be true for a deeply in-the-money European put option. This is because the *guaranteed* gain from being able to exercise the option and get your money sooner could be worth more than the *possible* gain from the share price moving in your favour over a longer period to expiry.

3.5 Volatility of the underlying share

Within this context, volatility refers to the general level of variability in the market price of the underlying share.

The higher the volatility of the underlying share, the greater the chance that the underlying share price can move significantly in favour of the holder of the option before expiry. So the value of an option will increase with the volatility of the underlying share.

The argument here is the same as with the time to expiry. Note that the holder of an option will therefore like volatility or risk, whereas with most other assets we dislike risk and consequently place a lower value on a riskier asset.

3.6 Interest rates

An increase in the risk-free rate of interest will result in a higher value for a call option because the money saved by purchasing the option rather than the underlying share can be invested at this higher rate of interest, thus increasing the value of the option.

Buying a call option and later exercising it can be broadly compared with buying the share directly. By using the call option the buyer is deferring the payment of the bulk of the purchase price and can earn extra interest during the period until exercising the option (although the buyer will then miss out on any dividends).

For a put option, higher interest means a lower value.

This is because put options can be purchased as a way of deferring the sale of a share. Comparing this strategy with an immediate sale of the share, we see that the investor's money is tied up for longer.

The basic Black-Scholes model can be adapted to allow for a sixth factor determining the value of an option:

3.7 Income received on the underlying security

In many cases the underlying security might provide a flow of, say, dividend income. Normally such income is not passed onto the holder of an option. Then the higher the level of income received, the lower the value of a call option, because by buying the option instead of the underlying share the investor foregoes this income. The reverse is true for a put.

Q

Question 11.9

Without looking back through the above notes, list the five factors that determine the price of an American put option and, for each factor, state whether an increase in its value produces an increase or a decrease in the value of the option.

3.8 The Greeks and risk management

So far, we have discussed the factors that affect the price of a derivative contract and suggested what happens to the price when you increase or decrease each factor. In addition, we can try to quantify what the change in the price will be in response to a change in one of the factors, while holding the others constant. In other words, we can calculate the (partial) derivative of the derivative price with respect to each of the above factors. These (mathematical) derivatives are given Greek letters and collectively known as "the Greeks". In Chapter 12, we look at the Greeks in more detail.

In the same way that we can consider the price of a single derivative contract, we can also work out the effect of the various factors on a whole portfolio of shares and derivatives – we simply add up the Greeks for each constituent. These Greeks can then tell us about the exposure to risk of our portfolio. For example, what will happen to the value of our portfolio if share prices fall by 10%, or the risk-free interest rate increases by 1%, and so on. The Greeks can therefore be used to manage risk. Risk management is considered in much greater detail in Subject CA1.

4 Pricing forward contracts

4.1 Introduction

Forward contracts were first introduced in Subject CT1. There you learned what a forward contract is and also how to price them. You may find it useful to recap on this material before reading on.

The underlying asset for a forward could be one of several types, *eg* a commodity (such as gold) or a share. Here we will consider primarily shares.

A forward contract is the simplest form of derivative contract. It is also the most simple to price in the sense that the forward price can be established without reference to a model for the underlying share price.

Suppose that, besides the underlying share, we can invest in a cash account that earns interest at the continuously compounding rate of *r* per annum. In Subject CT1 this was denoted by δ.

The "forward price" is the price the holder agrees to pay at the expiry date. So it corresponds to the exercise price for an option. There is no initial premium to pay when you purchase a forward.

Recall that the forward price *K* should be set at a level such that the value of the contract at time 0 is zero (that is, no money changes hands at time 0).

Example

Investor A agrees to sell 1,500 Vodafone shares in six months to Investor B at a price of £1.40 per share. So, £1.40 is the forward price of the share, whereas the current price is £1.33.

Suppose that six months later the share price is £1.35. Investor A then has to sell 1,500 shares, with a current market value of £2,025, for £2,100. He therefore makes a profit from the futures contract of £75. Similarly, Investor B has to buy the shares for a loss of £75.

4.2 Pricing forward contracts

We now move on to the fair price for a forward contract and demonstrate two proofs of this being the case.

Proposition

The fair or economic forward price is $K = S_0 e^{rT}$.

Proof (a)

Before showing the actual proof, we will first demonstrate that when the forward is priced correctly, everything works well.

Suppose, first, that we have set the forward price at $K = S_0 e^{rT}$.

At the same time we can borrow an amount S_0 in cash (subject to interest at rate r) and buy one share. The net cost at time 0 is then zero.

The forward contract costs nothing and the share costs S_0, which we have borrowed.

At time T we will have:

- **one share worth S_T on the open market**
- **a cash debt of $S_0 e^{rT}$**
- **a contract to sell the share at the forward price K.**

Therefore we hand over the one share to the holder of the forward contract and receive K. At the same time we repay the loan: an amount $S_0 e^{rT}$.

Since $K = S_0 e^{rT}$ we have made a profit of exactly 0. There is no chance of losing money on this transaction, nor is there any chance of making a positive profit. It is a risk-free trading strategy.

So $K = S_0 e^{rT}$ appears to be the correct forward price. The actual proof is one of contradiction. We will follow the same procedure as above in the case where $K \neq S_0 e^{rT}$ and demonstrate that this would allow an arbitrage profit.

Now suppose instead that $K > S_0 e^{rT}$. We can issue one forward contract and, at the same time, borrow an amount S_0 in cash (subject to interest at rate r) and buy one share. The net cost at time 0 is zero.

At time T we will have:

- **one share worth S_T on the open market**
- **a cash debt of S_0e^{rT}**
- **a contract to sell the share at the forward price K .**

Therefore we hand over the one share to the holder of the forward contract and receive K. At the same time we repay the loan: an amount S_0e^{rT} .

Since $K > S_0e^{rT}$ we have made a guaranteed profit having made no outlay at time 0.

This is an example of arbitrage: that is, for a net outlay of zero at time 0 we have a probability of 0 of losing money and a strictly positive probability (in this case equal to 1) of making a profit greater than zero.

Q

Question 11.10

If this arbitrage opportunity existed in real life, what would happen?

Instead of issuing one contract at this price, why not issue lots of them and we will make a fortune?

In practice a flood of sellers would come in immediately, pushing down the forward price to something less than or equal to S_0e^{rT} . So the arbitrage possibility could exist briefly but it would disappear very quickly before any substantial arbitrage profits could be made.

Having shown that $K > S_0e^{rT}$ leads to arbitrage opportunities we complete the proof by contradiction, by showing that $K < S_0e^{rT}$ also leads to arbitrage opportunities.

Now suppose that $K < S_0e^{rT}$.

We follow the same principles, at time 0:

- **buy one forward contract**
- **sell one share at a price S_0**
- **invest an amount S_0 in cash.**

The net value at time 0 is zero.

At time *T*:

- **we have cash of $S_0 e^{rT}$**
- **we pay $K\left(K < S_0 e^{rT}\right)$ for one share after which our net holding of shares is zero**
- **the shareholding has zero value and we have $S_0 e^{rT} - K > 0$ cash.**

Again this is an example of arbitrage, meaning that we would not, in practice, find that $K < S_0 e^{rT}$.

Because both $K > S_0 e^{rT}$ and $K < S_0 e^{rT}$ lead to arbitrage opportunities the only possibility for the fair price is $K = S_0 e^{rT}$.

Proof (b)

We now give a second proof, one you should be familiar with from studying Subject CT1.

Let *K* be the forward price. Now compare the setting up of the following portfolios at time 0:

- **A: one long forward contract**
- **B: borrow Ke^{-rT} cash and buy one share at S_0.**

If we hold both of these portfolios up to time *T* then both have a value of $S_T - K$ at *T*.

By the principle of no arbitrage these portfolios must have the same value at all times before *T*. In particular, at time 0, portfolio B has value $S_0 - Ke^{-rT}$ which must equal the value of the forward contract.

This can only be zero (the value of the forward contract at *t* = 0) if:

$$K = S_0 e^{rT}$$

Question 11.11

A three-month forward contract exists on a zero-coupon corporate bond with a current price per £100 nominal of £42.60. The yield available on three-month government securities is 6% *pa* effective. Calculate the forward price.

5 Bounds for option prices

In this section we deduce upper and lower limits for option prices based on general reasoning. Remember that throughout this chapter we are considering options based on non-dividend-paying shares.

5.1 Lower bounds on option prices

European calls

Consider a portfolio, A, consisting of a European call on a non-dividend-paying share and a sum of money equal to $Ke^{-r(T-t)}$.

Portfolio A is a natural portfolio to consider when dealing with a call option. The call option gives the right to buy the share at time *T*, and therefore the amount of money held is just the right amount that will accumulate to the requisite amount, *ie K*.

At time *T*, portfolio *A* has a value which is equal to the value of the underlying share, provided the share price S_T is greater than *K*.

Question 11.12

Explain why this is true.

If S_T is less than *K* then the payoff from portfolio A is greater than that from the share.

In this instance, the payoff from Portfolio A is *K* – the accumulated amount of cash at the exercise date – because the option would not be exercised, since the share is only worth S_T.

Since the option plus cash produces a payoff at least as great as the share, it must have a value greater than or equal to S_t. This gives us a lower bound for c_t:

$$c_t + Ke^{-r(T-t)} \geq S_t$$

ie $$c_t \geq S_t - Ke^{-r(T-t)}$$

Example

Suppose that:

- the current price of Share X is 115
- the exercise price of a 3-month call option on X is 100
- the continuously-compounded risk-free rate of return is 12% *pa*

Then, the lower bound for the European call option price is:

$$c_t \geq 115 - 100\, e^{-0.12 \times \frac{3}{12}}$$

$$c_t \geq 17.96$$

If instead the share price were 10 higher at 125, then the lower bound would be:

$$c_t \geq 125 - 100\, e^{-0.12 \times 3/12}$$

$$c_t \geq 27.96$$

ie it would also be 10 higher.

European puts

A similar argument can be used for put options: Portfolio B contains a European put option and a share. Compare this with the alternative of cash, currently worth $Ke^{-r(T-t)}$. At time T portfolio B will be worth at least as much as the cash alternative.

Question 11.13

Why must Portfolio B be worth at least as much as the cash alternative?

Thus:

$$p_t + S_t \geq Ke^{-r(T-t)}$$

$$\Rightarrow \qquad p_t \geq Ke^{-r(T-t)} - S_t$$

Question 11.14

What is the lower bound for a 3-month European put option on Share X if the share price is 95, the exercise price 100 and the risk-free rate 12% *pa*?

Question 11.15

Suppose that you spot that the current price of a particular European put satisfies:

$$p_t < Ke^{-r(T-t)} - S_t$$

What would you do?

American calls

Recall that, unlike its European counterpart, an American option can be exercised at any date up to and including the expiry date. A surprising result, however, is that it is never optimal to exercise an American call on a non-dividend-paying share early. Hence the above relationship for European calls also holds for American calls, *ie*:

$$C_t \geq S_t - Ke^{-r(T-t)}$$

The reason for this surprising result is as follows:

If we exercise an American call option early then we receive $S_t - K$. There is nothing to stop us, instead, pretending that it's just a European call option and refusing to exercise early. So the value of the American call option must be at least that of the value of the European call option. But we already know that the value of the European option is worth more than its intrinsic value:

$$c_t \geq S_t - Ke^{-r(T-t)} > S_t - K$$

So because we always receive more by selling the option than we do by exercising it, the option to exercise early is effectively worthless.

Question 11.16

If this is true, and American call options are effectively just the same as European call options then why do people bother selling American call options in real life? Is it just a marketing ploy?

Question 11.17

State the consequence of this surprising result for the time value of a European call option on a non-dividend-paying share.

American puts

The lower bound for an American put option can be increased above that derived above for a European put option. Since early exercise is always possible we have:

$$P_t \geq K - S_t$$

So, the intrinsic value is a lower bound for the price. This condition, which is stronger than for European puts, holds because early exercise may be sensible for an American put in order to receive the exercise money earlier. By receiving the cash before the last possible expiry date, the holder of the option then benefits from receiving interest on that cash for the remaining term.

Summary

Finally note that if any of the lower bounds given above are negative we can get a tighter bound from the fact that any option has a non-negative value to the holder – ignoring the premium already paid. This means that the following bounds can be given:

$$c_t \geq \max\left\{S_t - Ke^{-r(T-t)}, 0\right\}$$
$$C_t \geq \max\left\{S_t - Ke^{-r(T-t)}, 0\right\}$$
$$p_t \geq \max\left\{Ke^{-r(T-t)} - S_t, 0\right\}$$
$$P_t \geq \max\left\{K - S_t, 0\right\}$$

Q

Question 11.18

What can you say about the time value for each of the above options?

5.2 Upper bounds on option prices

European calls

A call option gives the holder the right to buy the underlying share for a certain price. The payoff $\max\{S_T - K, 0\}$ is always less than the value of the share at time T, S_T. Therefore the value of the call option must be less than or equal to the value of the share:

$$c_t \le S_t$$

This is obvious from the fact that you could just buy the share anyway, without needing to have the call option.

European puts

For a European put option the maximum value obtainable at expiry is the strike price *K*. Therefore the current value must satisfy:

$$p_t \le Ke^{-r(T-t)}$$

ie it cannot exceed the discounted value of the sum received on exercise – which it will equal if the share price falls to zero.

These bounds make no assumptions about the behaviour of the share price. In fact, we will see later that we can improve on these inequalities if we make an assumption about the behaviour of the share price.

For certain types of stochastic model for S_t we find that we are able to write down explicit formulae for the prices of European call and put options.

For example, if we assume that the process determining the log-share price is described by Brownian motion, then we can use the Black-Scholes analysis and formulae discussed later in the course.

American calls

We can see that the call option inequality applies also to an American call option on a non-dividend-paying stock.

ie $C_t \le S_t$

American puts

On the other hand, the possibility of early exercise of an American put option presents us with much more complexity. There are no simple rules for deciding upon the time to exercise. Partly as a result of this there is no explicit formula for the price of an American put option.

For American puts the upper bound for a European put may not hold. For example, if the share price is zero, an American put is worth exactly K, not the discounted value of K. All we can be sure of is that a put will never be worth more than K.

ie $P_t \le K$

Q

Question 11.19 (Revision)

Without looking back at the previous notes, give an upper and lower bound for each type of option (American/European, call/put) on a non-dividend-paying share. In each case you should think through the logic behind the bound carefully.

6 Put-call parity

Consider the argument we used to derive the lower bounds for European call and put options on a non-dividend-paying stock. This used two portfolios:

- **A: one call plus cash of $Ke^{-r(T-t)}$**
- **B: one put plus one share.**

Recall that both portfolios included only *European* options on *non-dividend-paying* shares – this is an important condition underpinning the arguments that follow.

Both portfolios have a payoff at the time of expiry of the options of $\max\{K, S_T\}$.

We can see this as follows.

Portfolio A

First consider Portfolio A, consisting of a European call plus cash of $Ke^{-r(T-t)}$. The value of Portfolio A at the expiry date is given by:

$S_T - K + K = S_T$ if $S_T > K$ (*ie* the call option is exercised)

and $0 + K = K$ if $S_T \le K$ (*ie* the call expires worthless)

Portfolio B

Now consider Portfolio B, consisting of the underlying share plus a European put with the same expiry date and exercise price as the call. On expiry the value of Portfolio B is:

$0 + S_T = S_T$ if $S_T > K$ (*ie* the put expires worthless)

and $K - S_T + S_T = K$ if $S_T \le K$ (*ie* the put option is exercised)

Thus, the values at expiry are the same for both portfolios regardless of the share price at that time, namely $\max\{K, S_T\}$.

Since they have the same value at expiry and since the options cannot be exercised before then they should have the same value at any time $t < T$.

ie $$c_t + Ke^{-r(T-t)} = p_t + S_t$$

This relationship is known as *put-call parity*.

If the result was not true then this would give rise to the possibility of *arbitrage*: that is, for a net outlay of zero at time *t* we have a probability of 0 of losing money and a strictly positive probability (in this case 1) of making a profit greater than zero. In this case the failure of put-call parity would allow an investor to sell calls and take a cash position and buy puts and shares with a net cost of zero at time *t* and a certain profit at time *T*.

One consequence of put-call parity is that, having found the value of a European call on a non-dividend-paying share (*eg* from the Black-Scholes formula discussed later in the course) you can easily find the value of the corresponding put.

In contrast to forward pricing, put-call parity does not tell us *what* c_t and p_t are individually: only the relationship between the two. To calculate values for c_t and p_t we require a model.

In all of these sections, the pricing of derivatives is based upon the principle of *no arbitrage*.

Note that we have made very few assumptions in arriving at these results. No model has been assumed for stock prices. All we have assumed is that we will make use of buy-and-hold investment strategies. Any model we propose for pricing derivatives must, therefore, satisfy both put-call parity and the forward-pricing formula. If a model fails one of these simple tests then it is not arbitrage free.

Question 11.20

Explain why the put-call parity relationship above does not hold in the case of:

(i) American options on non-dividend-paying shares.

(ii) European options on dividend-paying shares.

In the case of dividend-paying securities, the put-call parity relationship is:

$$c_t + Ke^{-r(T-t)} = p_t + S_t e^{-q(T-t)}$$

This formula can also be found on page 47 of the Tables. Note that q is the continuously-compounded dividend rate so that we are assuming that all dividends are reinvested immediately in the same share.

Question 11.21

Company X issues 3-month European call options on its own shares with a strike price of 120p. They are currently priced at 30 pence per share. The current share price is 123p and the current force of interest is $\delta = 6\%\ pa$.

(i) If dividends are payable continuously at a rate of $q = 12\%\ pa$ then calculate the fair price for put options on the share price at the same strike price.

(ii) Explain the strategy for arbitrage profit if, instead, the price of the put options is 25p.

Further reading

Should you wish to consult an alternative reference and/or read more about options and derivatives then the book by Hull listed in the Study Guide is a good choice. It provides an alternative presentation of much of the material in Chapters 11 to 18 of this course.

7 Exam-style question

We finish this chapter with an exam-style question on put-call parity.

Question

This question is based on Subject 109 September 2000 Question 1.

(i) State what is meant by put-call parity. [2]

(ii) By constructing two portfolios with identical payoffs at the exercise date of the options, derive an expression for the put-call parity of a European option on a share that has a dividend payable prior to the exercise date. [6]

(iii) If the equality in (ii) does not hold, explain how an arbitrageur can make a riskless profit. [3]

[Total 11]

Solution

(i) ***Put-call parity***

Put-call parity is a theoretical relationship between the price of a call option and a put option on a share. The options involved have the same exercise price and exercise date.

It assumes that security markets are arbitrage-free and that there is a constant and known risk-free rate of return that can be earned on deposits. However, it makes no assumptions about the nature of the process determining share prices.

(ii) ***Put-call parity on a dividend-paying share***

Suppose that a certain amount of dividend is known to be payable at some date $t < t_1 < T$. Define d to be this amount.

Consider two portfolios as follows:

1. Portfolio A – which consists of one European call option plus cash equal in amount to the discounted value of the strike price plus the present value of the dividends to be paid at time t_1 – *ie* a cash amount of $d\,e^{-r(t_1-t)} + K\,e^{-r(T-t)}$

2. Portfolio B – which consists of one European put plus one dividend-paying share.

Then the value of Portfolio A at the exercise date T is given by:

$$S_T - K + d\,e^{r(T-t_1)} + K \;=\; S_T + d\,e^{r(T-t_1)} \quad \text{if } S_T > K$$

(*ie* the call option is exercised leaving the investor with the share plus the accumulated value of the dividend received), and:

$$0 + d\,e^{r(T-t_1)} + K \;=\; d\,e^{r(T-t_1)} + K \qquad \text{if } S_T \le K$$

(*ie* the call expires worthless and the investor is left with cash equal to the exercise price plus the accumulated value of the dividend).

Similarly, the value of Portfolio B is:

$$0 + S_T + d\,e^{r(T-t_1)} = S_T + d\,e^{r(T-t_1)} \qquad \text{if } S_T > K$$

(*ie* the put expires worthless and the investor is left with the share plus the accumulated value of the dividend received), and:

$$K - S_T + S_T + d\,e^{r(T-t_1)} = K + d\,e^{r(T-t_1)} \text{ if } S_T \le K$$

(*ie* the put option is exercised and the investor is left with cash equal to the exercise price plus the accumulated value of the dividend).

Thus, the payoffs at expiry are the same for both portfolios regardless of the share price at that time. Since they have the same value at expiry and since the options cannot be exercised before, then – in an arbitrage-free market – they should have the same value at any time $t < T$, that is:

$$c_t + d\,e^{-r(t_1-t)} + Ke^{-r(T-t)} = p_t + S_t$$

(iii) ***Riskless arbitrage profit***

Suppose that instead of the above equality, it was the case that:

$$c_t + d\,e^{-r(t_1-t)} + Ke^{-r(T-t)} < p_t + S_t$$

so that the call option is cheap and/or the put option is dear. It follows that we could make a riskless arbitrage profit by selling the put and the share (assuming that we can short sell it) and using the proceeds to buy the call. This would generate a sum of money equal to:

$$A + d\,e^{-r(t_1-t)} + Ke^{-r(T-t)} \qquad (1)$$

where:

$$A = p_t + S_t - c_t - d\,e^{-r(t_1-t)} - Ke^{-r(T-t)} > 0$$

which we could then put on deposit at the risk-free rate of return.

At the exercise date, the total value of the call plus the accumulated amount of the cash sum in (1) will be equal to:

$$S_T + A\,e^{r(T-t)} + d\,e^{r(T-t_1)} \qquad \text{if } S_T > K$$

or: $$K + A\,e^{r(T-t)} + d\,e^{r(T-t_1)} \qquad \text{if } S_T \le K$$

Whilst the accumulated value of the share (including the dividend paid) plus the put is equal to:

$$S_T + d\,e^{r(T-t_1)} \qquad \text{if } S_T > K$$

or: $$K + d\,e^{r(T-t_1)} \qquad \text{if } S_T \le K$$

Hence, whichever outcome prevails, we will make a net profit of $A\,e^{r(T-t)} > 0$.

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

Chapter 11 Summary

Derivatives

A *derivative* is a security or contract which promises to make a payment at a specified time in the future, the amount of which depends upon the behaviour of some underlying security up to and including the time of the payment.

Arbitrage

An *arbitrage opportunity* is a situation where we can make a certain profit with no risk.

The *principle of no arbitrage* states that arbitrage opportunities do not exist.

The *law of one price* says that any two portfolios that behave in exactly the same way must have the same price. For if this were not true, we could buy the "cheap" one and sell the "expensive" one to make an arbitrage (risk-free) profit.

Intrinsic and time values

The *intrinsic value* of a derivative is the value assuming expiry of the contract immediately rather than at some time in the future.

The *time value* of a derivative is the excess (if any) of the total value of an option over its intrinsic value.

Options

A *European call option* gives its holder the right, but not the obligation, to buy one share from the issuer of the contract at time T and at the strike price K.

A *European put option* gives its holder the right, but not the obligation, to sell one share to the issuer of the contract at time T and at the strike price K.

American options are similar to their European equivalents, except that they can be exercised at any time t up to expiry T.

Factors affecting option prices

- underlying share price, S_t
- strike price, K
- time to expiry, $T-t$
- volatility of the underlying share, σ
- risk-free interest rate, r
- dividend rate, q

Forwards

Because both $K > S_0 e^{rT}$ and $K < S_0 e^{rT}$ lead to an arbitrage opportunity, the only possibility for the fair price of a forward contract on a non-dividend-paying share is:

$$K = S_0 e^{rT}$$

Bounds for option prices

Option prices lie within the following ranges:

$$S_t \geq c_t \ \geq \ \max\left\{S_t - Ke^{-r(T-t)}, 0\right\}$$
$$S_t \geq C_t \ \geq \ \max\left\{S_t - Ke^{-r(T-t)}, 0\right\}$$
$$Ke^{-r(t-t)} \geq p_t \ \geq \ \max\left\{Ke^{-r(T-t)} - S_t, 0\right\}$$
$$K \geq P_t \ \geq \ \max\left\{K - S_t, 0\right\}$$

Put-call parity

$$c_t + Ke^{-r(T-t)} = p_t + S_t e^{-q(T-t)}$$

Chapter 11 Solutions

Solution 11.1

An example is to bet:

- £10 on Nag Portachio
- £5 on GG Badeye
- £5 on Horsey Mullet

which would cost you £20 in total.

If Nag Portachio or GG Badeye wins then you receive £20 and you are quits. If Horsey Mullet wins then you receive £40. So this betting strategy provides the chance of a £20 gain with no chance of a loss.

You can compare the odds for a horse race with the probabilities of each horse winning. So 3/1 compares to a probability $\frac{1}{3+1}=\frac{1}{4}$ of winning. If a bookmaker were "fair" then all these "probabilities" would add up to one. In real life, a bookmaker will make money from gamblers by having these "probabilities" add up to more than one. However, in this question they have made a mistake because the "probabilities" add up to less than one:

$$\frac{1}{1+1}+\frac{1}{1+3}+\frac{1}{1+7}=\frac{7}{8}$$

This means that there is an opportunity to make an arbitrage profit by betting on all three horses in the right proportions.

Solution 11.2

It is useful to re-express Bookmaker A's odds in terms of "unit" bets. For a 40p bet on Bruno, the payoff is £1 if he wins and £0 if he loses. For a 66.7p bet on Tyson, the payoff is £1 if he wins and £0 if he loses. Now we can recreate the "special" bet by simply adding these unit bets. With twenty 40p bets on Bruno and six 66.7p bets on Tyson, the payoff is £20 if Bruno wins and £6 if Tyson does. Bookmaker A will receive £8 + £4 = £12 for selling all of these bets. The manager of Bookmaker A can then place £11.50 on the special bet and put 50p in his pocket as arbitrage profit. Whatever the outcome of the boxing match, the payoff to the customer is matched exactly by the win from the special bet.

This approach can be carried out on any scale. The strategy is that every time bets are placed on Bruno and Tyson in the ratio 2:1, the manager puts 4.167% in his pocket as arbitrage profit and places the rest on the special bet.

In practice bookmakers will devote some time to ensuring that this sort of thing does not occur. There are some famous examples involving Sicilian bookmakers where this has happened.

Solution 11.3

If you found a different strategy in Question 11.2 then this question still forces you to think about what the "replicating portfolio" is for this special bet. You can see from the way we have answered Question 11.2 that the replicating strategy is £8 on Bruno and £4 on Tyson. This portfolio costs £12, which is different from the £11.50 price for the special bet. So two assets whose payoffs are exactly the same in all circumstances have different prices. Hence, the principle of no arbitrage does not hold.

Solution 11.4

The writer of the option makes a loss equal to $K - S_T < 0$, as he is obliged to sell the share at K when it is in fact worth $S_T > K$. As with a forward, the combined profit of the two parties is therefore always equal to zero.

Solution 11.5

The put option should be exercised if $S_T < K$ because the option holder would then be able to sell the share for K when it is worth only S_T. He would therefore make a profit of $K - S_T > 0$ from exercising the option and consequently should do so. Conversely, if $S_T > K$, then he would be silly to sell the share to the writer of the put option for K, when he could instead sell it for S_T in the open market.

Solution 11.6

The *intrinsic value* of a derivative is the value assuming expiry of the contract immediately rather than at some time in the future. Thus, the intrinsic value of a put option at time t is either:

- $K - S_t$ if the exercise price exceeds the share price – in which case it will be exercised, or
- 0 if the exercise price is less than the share price – in which case it will not be exercised.

ie the intrinsic value is $\max\{K - S_t, 0\}$.

Solution 11.7

If $S_t = 112$, $K = 110$ and $p_t = 5$, then:

- Intrinsic value = $\max\left[K - S_t, 0\right]$ = $\max\left[110 - 112\right] = 0$
- Time value = total value − intrinsic value = $5 - 0$ = 5

Solution 11.8

Consider the holder of a call option. Although a longer time to expiry does mean that there is a greater chance that the share price can fall a long way, this does not fully offset the value of the chance of greater profits if the price goes up. This is because once the share price goes below the exercise price K, the option will not be exercised and the holder simply loses the premium paid at outset.

Now, regardless of how far the share price goes down, the maximum loss that the holder can make in this instance is the premium paid at the outset, whereas the maximum profit if the price goes up is unlimited. Thus, the holder of the option faces an asymmetric risk – *ie* lots of upside profit potential and limited downside risk. Consequently, the more chance there is of the share price moving a long way, up or down, the more valuable the option.

Solution 11.9

The second column in the table indicates whether an increase in the factor concerned leads to an increase or decrease in the value of a put option.

Factor	*Value of American put option will …*
share price	decrease
exercise price	increase
time to expiry	increase
volatility of share price	increase
risk-free rate of interest	decrease

Solution 11.10

Shrewd investors would exploit the opportunity in the short-term. The market would adjust the price (typically the forward price would fall) and close out the arbitrage opportunity. For an example this simple, this whole process would be almost immediate.

Solution 11.11

$$K = 42.6e^{\frac{3}{12}\delta}$$

$$= 42.6(1+i)^{\frac{3}{12}}$$

$$= 42.6 \times 1.06^{\frac{3}{12}} = £43.23$$

Solution 11.12

With Portfolio A, the sum of money will grow with interest to be worth exactly *K* at time *T* ($Ke^{-r(T-t)}$ is the present value at time *t* of a payment of *K* at time *T*). If $S_T > K$ the call option will be exercised (leaving zero cash). The payoff is thus S_T at time *T*.

Solution 11.13

At time *T* the cash will be worth *K*.

Portfolio B (the share plus the put option) will be worth:

- *K* if $S_T < K$ (because the option will be exercised by selling the share, leaving *K*)
- S_T if $S_T > K$ (because the option will not be exercised).

Thus Portfolio B is always worth at least as much as the cash deposit at time *T*.

ie $\quad p_t + S_t \geq Ke^{-r(T-t)}$

Solution 11.14

If:

- the current price of Share X is 95
- the exercise price of a 3-month put option on X is 100
- the continuously-compounded risk-free rate of return is 12% *pa*

Then, the lower bound for the put option on Share X is:

$$p_t \ \geq \ K\, e^{-r\,(T-t)} - 95$$

$$p_t \ \geq \ 100\, e^{-0.12\times\frac{3}{12}} - 95$$

$$p_t \ \geq \ 2.04$$

Solution 11.15

You could borrow an amount $p_t + S_t$ in the cash market to purchase the share and put option. You therefore have a share, the put option and a cash loan.

You wait until the exercise time (regardless of the movements of the underlying).

Your portfolio at that time is worth

$$\begin{aligned}\max\left\{K - S_T, 0\right\} + S_T - \left(p_t + S_t\right)e^{r(T-t)} &= \max\left\{K, S_T\right\} - \left(p_t + S_t\right)e^{r(T-t)} \\ &> \max\left\{K, S_T\right\} - Ke^{-r(T-t)}e^{r(T-t)} \\ &= \max\left\{0, S_T - K\right\} \\ &\geq 0\end{aligned}$$

Overall then, for zero initial outlay, you've made a strictly positive profit, *ie* you've taken advantage of an arbitrage opportunity.

Solution 11.16

No! It is important to realise that this surprising result is only true for a non-dividend paying share (or a share with no dividends between time t and expiry). The crucial point is that if there is a dividend between time t and expiry then it may be beneficial to exercise the option early in order to receive this dividend.

Also, for options where there is not an active market with high trading volumes there is the possibility that you might not be able to find a buyer at the time you want to sell. So exercising may be your only option.

Solution 11.17

An option having a negative time value would mean that the intrinsic value is greater than the total value. Because you would never want to exercise the call option early there is no way that this is ever the case and so we know that the time value of a call option on a non-dividend-paying share must always be positive.

Previously, we said that the time value represented the value derived from the choice the holder has of whether to exercise or not. This description might suggest that the time value must always be positive. However, a more precise definition is that it is the market price of the option minus the intrinsic value. In certain situations the holder of an option might be "losing out" by not exercising, and in such cases the time value can be negative.

Solution 11.18

For the call options we have:

$$c_t \geq \max\left\{S_t - Ke^{-r(T-t)}, 0\right\}$$

and $$C_t \geq \max\left\{S_t - Ke^{-r(T-t)}, 0\right\}$$

The intrinsic value of the options is $\max\left\{S_t - K, 0\right\}$, but since $\max\left\{S_t - K, 0\right\} \leq \max\left\{S_t - Ke^{-r(T-t)}, 0\right\}$ the value of both call options is greater than the intrinsic value. The time value is therefore positive. In fact we have the tighter bound:

$$\begin{aligned}\text{time value European call} &= c_t - \max\left\{S_t - K, 0\right\} \\ &\geq \max\left\{S_t - Ke^{-r(T-t)}, 0\right\} - \max\left\{S_t - K, 0\right\}\end{aligned}$$

This is always ≥ 0 but when $K > S_t > Ke^{-r(T-t)}$, it is strictly positive.

For an American put we can say that the time value is positive:

$$\text{time value of Am put} = P_t - \max\left\{K - S_t, 0\right\} \geq \max\left\{K - S_t, 0\right\} - \max\left\{K - S_t, 0\right\} = 0$$

However, for a European put option the time value could be negative. The analysis above gives:

$$\text{time value of Euro put} = p_t - \max\left\{K - S_t, 0\right\} \geq \max\left\{Ke^{-r(T-t)} - S_t, 0\right\} - \max\left\{K - S_t, 0\right\}$$

A look at the right-hand side shows that this will be negative if $K > S_t > Ke^{-r(T-t)}$, since then $K - S_t > 0$ and $Ke^{-r(T-t)} - S_t < K - S_t$.

The reason the time value of a European put can be negative is that by holding the option, rather than selling the share, you have money tied up that is not earning interest. This doesn't occur with the American put because of the possibility of early exercise. It also doesn't happen with the call options because they work the other way round, *ie* holding the call means that you can invest money. If the share paid dividends, however, then something like this could happen with the call, since then holding the option means that you are missing out on the dividends.

Solution 11.19

	Lower bound	Upper bound
European call	$c_t \geq S_t - Ke^{-r(T-t)}$	$c_t \leq S_t$
American call	$C_t \geq S_t - Ke^{-r(T-t)}$	$C_t \leq S_t$
European put	$p_t \geq Ke^{-r(T-t)} - S_t$	$p_t \leq Ke^{-r(T-t)}$
American put	$P_t \geq K - S_t$	$P_t \leq K$

All the lower bounds are subject to a minimum of zero.

Note that it is never optimal to exercise an American call on a non-dividend-paying share before expiry. As a result an American and a European call are equivalent where the underlying share is non-dividend-paying.

Solution 11.20

(i) This is because it can be worthwhile to exercise an American put early – in which case the cash will not have accumulated fully and so the payoffs do not work out the same. This means that Portfolio B is worth more than Portfolio A.

(ii) Dividends will be received on Portfolio B, but not on Portfolio A. Again we see that Portfolio B is then worth strictly more than A.

Solution 11.21

(i) ***Put option value***

Using put-call-parity, the value of the put option should be:

$$p_t = c_t + Ke^{-r(T-t)} - S_t e^{-q(T-t)}$$

$$= 30 + 120e^{-0.06\times\frac{3}{12}} - 123e^{-0.12\times\frac{3}{12}}$$

$$= 28.8p$$

(ii) ***Arbitrage profit***

If the put options are only 25*p* then they are cheap. If things are cheap then we buy them! So looking at the put-call parity relationship, we "buy the cheap side and sell the expensive side", *ie* we buy put options and shares and sell call options and cash.

For example:

- sell 1 call option 30*p*
- buy 1 put option (25*p*)
- buy 1 share (123*p*)
- sell (borrow) cash 118*p*

This is a zero-cost portfolio and, because put-call parity does not hold, we know it will make an arbitrage profit. Let's check:

In 3 months' time, repaying the cash will cost us:

$$118e^{0.06\times 3/12} = 119.78$$

We also receive dividends d on the share.

1. If the share price is above 120 in 3 months' time then the other party will exercise their call option and we will have to give them the share. They will pay 120 for it and our profit is:

$$120 - 119.78 + d = 0.22 + d$$

(The put option is useless to us and we throw it in the bin!)

2. If the share price is below 120 in 3 months' time then we will exercise our put option and sell it for 120. Our profit is:

$$120 - 119.78 + d = 0.22 + d$$

(The call option is useless to the other party and they throw it in the bin!)

Chapter 12

The Greeks

Syllabus objectives

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

12. Demonstrate an awareness of the commonly used terminology for the first, and where appropriate second, partial derivatives (the Greeks) of an option price.

0 Introduction

Later on in the course we will consider how to price European options. We will first do this using a discrete-time model (the Binomial model) and then using a continuous time model (the Black-Scholes model). The important point to realise is that no matter how complicated the model used to value a given derivative, it cannot conflict with the simpler properties that we already know from studying Chapter 11.

For example, no matter how ingenious a model for valuing a European call option might be, if it contradicts the fact that $c_T = \max\{S_T - K, 0\}$ then it must be ruled out.

Question 12.1 (Revision)

List some other things that we already know about European options from Chapter 11.

Question 12.2

What would you say is the most important result learnt so far about European call and put options?

In this chapter we introduce the Greeks, mathematical derivatives that describe the rate at which derivative prices change as different factors are varied – often referred to as the "sensitivity" to the parameters.

For example, the Greek letter Δ (delta) is used to denote $\dfrac{\partial f(S_t)}{\partial S_t}$, *ie* the rate of change of the derivative price with respect to the share price itself (while the other parameters remain unchanged).

Question 12.3

State whether delta should be positive or negative in the case that the derivative is:

(i) a call option

(ii) a put option.

Again, the important point to realise is that no matter how complicated the model used to value a given derivative, it cannot conflict with the basic properties that we already know. So, whatever we find out about the Greeks for call and put options by studying this chapter must remain true no matter how complicated the model used to value the derivatives is.

1 Portfolio risk management

1.1 Introduction

Call options allow exposure to be gained to upside movements in the price of the underlying asset. Put options allow the downside risks to be removed. In each case, however, because of the effect of gearing both call and put options, on their own, are more risky than the underlying asset.

Question 12.4

Explain what is meant by gearing in this context.

Question 12.5

Portachio has £10,000 invested in a portfolio consisting of 1,000 shares. Benny Kiang has £10,000 invested in a portfolio of 5,000 call options on the share and the delta is 0.5. If the share price increases by 10%, what will be the value of each portfolio?

In more general terms, combinations of various derivatives and the underlying asset in a single portfolio allow us to modify our exposure to risk.

In particular, we can use derivatives to reduce the exposure of our portfolio to the risk of adverse movements in the market price of the underlying assets. If, for example, we are concerned about falls in the investment market, we might buy put options. By guaranteeing the price at which we can sell our assets, this reduces the risk associated with market falls. We will still, however, benefit from the resulting profits should the market instead go up.

Derivative contracts therefore give us more control over the market risks that we face, thereby increasing our opportunity set of possible risk and return combinations. Moreover, if we hold suitable derivatives and the underlying assets in appropriate combinations then we can sometimes eliminate almost all of the market risk facing our portfolio – though other risks such as lack of marketability or credit risk will remain. The strategy of reducing market risk in this way is known as *hedging*.

For example, in the proof of the Black-Scholes PDE, which we will meet later on in Chapter 14, **we take a mixed portfolio of a derivative and the underlying asset to create an instantaneously risk-free portfolio. This is called *delta hedging*.**

Delta hedging involves the construction of a portfolio whose overall “delta” is equal to zero. We will discuss this later.

Delta is just one of what are called *the Greeks*. The Greeks are a group of *mathematical* derivatives that can be used to help us to manage or understand the risks in our portfolio.

Let $f(t,S_t)$ be the value at time t of a derivative when the price of the underlying asset at t is S_t.

We now introduce the complete set of six Greeks.

1.2 Delta

The delta for an individual derivative is:

$$\Delta = \frac{\partial f}{\partial S_t} \equiv \frac{\partial f}{\partial S_t}(t,S_t)$$

The notation here works as follows. We are thinking of the derivative price $f(t,S_t)$ as a function of time t and share price S_t. The partial derivative with respect to S_t, estimated at time 2 when the share price equals 100 is then written as:

$$\frac{\partial f(2,100)}{\partial S_2}$$

For the underlying asset, whose value is S_t, $\Delta = 1$.

Question 12.6

Explain briefly why the delta of the underlying asset is equal to one.

When we consider delta hedging we add up the deltas for the individual assets and derivatives (taking account, of course, of the number of units held of each). If this sum is zero and if the underlying asset prices follow a diffusion then the portfolio is instantaneously risk-free.

A portfolio for which the weighted sum of the deltas of the individual assets is equal to zero is described as delta-hedged or *delta-neutral*.

Instantaneously risk-free means that if we know the value of the portfolio at time t, then we can predict its value at time $t + dt$ with complete certainty. In other words, there is no risk or uncertainty concerning the change in the value of the portfolio over the instantaneous time interval $(t, t + dt]$.

Example

Later we will meet the following "risk-free" portfolio that we will use to construct the Black-Scholes partial differential equation. It is:

- minus one derivative
- plus $\dfrac{\partial f}{\partial S_t} = \Delta$ shares.

Thus, the total delta of this portfolio is equal to:

$$(-1) \times \Delta \; + \; \Delta \times (+1) \; = \; 0$$

Consequently, the "risk-free" portfolio must be instantaneously risk-free.

Question 12.7

Show that a delta-hedged portfolio with value $V\left(t, S_t\right)$ is risk-free if the underlying process S_t is a diffusion. *Hint: apply Ito's lemma.*

If it is intended that the sum of the deltas should remain close to zero (this is what is called *delta hedging*) then normally it will be necessary to rebalance the portfolio on a regular basis. The extent of this rebalancing depends primarily on *Gamma*.

Within this context, we can distinguish between *dynamic hedging* and *static hedging*.

The process of simply constructing an initial portfolio with a total delta of zero, at time 0 say, and not rebalancing to reflect the subsequent changes in delta, is known as *static* delta hedging.

Note, however, that as the share price S_t varies with time, so does:

- the price of the derivative, $f(t, S_t)$
- $\Delta = \dfrac{\partial f(t, S_t)}{\partial S_t}$.

Hence, in order to ensure that the total portfolio delta remains equal to zero over time we need to "rebalance" the constituents of the portfolio – so as to offset the changes in delta. Strictly speaking, since delta changes continuously through time, this rebalancing process must itself be continuous.

The process of continuously rebalancing the portfolio in this way in order to maintain a constant total portfolio delta of zero is known as *dynamic* delta hedging.

1.3 Gamma

$$\Gamma = \frac{\partial^2 f}{\partial S_t^2}$$

For the underlying asset, S_t, $\Gamma = 0$.

> ***Question 12.8***
>
> Explain why the gamma of the underlying asset is equal to zero.

Gamma is the rate of change of Δ with the price of the underlying asset. It therefore measures the *curvature* or convexity of the relationship between the derivative price and the price of the underlying asset.

Suppose a portfolio is following a delta-hedging strategy. If the portfolio has a high value of Γ then it will require more frequent rebalancing or larger trades than one with a low value of gamma.

This is because a high value of Γ means that Δ is more sensitive to changes in the share price S_t. Consequently, a given change in S_t will produce a greater change in Δ, which means that a greater amount of rebalancing will be required in order to ensure that the overall portfolio delta remains unchanged, *ie* equal to zero. Conversely, 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.

It is recognised that continuous rebalancing of the portfolio is not feasible and that frequent rebalancing increases transaction costs. The need for rebalancing can, therefore, be minimised by keeping gamma close to zero.

Hence, for practical hedging purposes, a portfolio with a low Γ is preferable, as costly rebalancing will be required less frequently in order to keep the portfolio approximately delta-hedged. Note that it may also be possible to construct a *gamma-neutral* portfolio – *ie* one with an overall gamma equal to zero.

1.4 Vega

$$\nu = \frac{\partial f}{\partial \sigma}$$

For the underlying asset, S_t, $\nu = 0$.

(Unlike the other names *vega* is not a Greek letter.)

If you have previously studied Subject 109 or met the Greeks before, possibly at university, then you may have seen this derivative called kappa (a real Greek letter). Derivative traders, however, call it vega.

This is the rate of change of the price of the derivative with respect to a change in the assumed level of volatility of S_t.

Note that we refer here to the "assumed" level of volatility. As we saw, the value of an option depends on the parameter σ. However, unlike the other parameters, the value of this parameter cannot be observed directly.

The price of a derivative depends directly on the assumed volatility of the price of the underlying asset. Hence, if the volatility of the price of the underlying asset changes, then so must the price of the derivative.

The value of a portfolio with a low value of vega will be relatively insensitive to changes in volatility. Put another way: it is less important to have an accurate estimate of σ if vega is low. Since σ is not directly observable, a low value of vega is important as a risk-management tool. Furthermore, it is recognised that σ can vary over time. Since many derivative pricing models assume that σ is constant through time the resulting approximation will be better if vega is small.

1.5 Rho

$$\rho = \frac{\partial f}{\partial r}$$

Rho tells us about the sensitivity of the derivative price to changes in the risk-free rate of interest. The risk-free rate of interest can be determined with a reasonable degree of certainty but it can vary by a small amount over the (usually) short term of a derivative contract. As a result, a low value of ρ reduces risk relative to uncertainty in the risk-free rate of interest.

1.6 Lambda

$$\lambda = \frac{\partial f}{\partial q}$$

where q is the assumed, continuous dividend yield on the underlying security.

Note that in the definitions of ν, ρ and λ we are assuming that σ, r and q take constant values throughout the life of the derivative contract, but that these "constant" values could change.

> ***Question 12.9***
>
> What can we say about the lambda of a long futures position in a dividend-paying share?

1.7 Theta

$$\Theta = \frac{\partial f}{\partial t}$$

Since time is a variable which advances with certainty it does not make sense to hedge against changes in t in the same way as we do for *unexpected* changes in the price of the underlying asset.

Note that:

- theta is usually written as a capital letter Θ
- t here is the time since the start of the contract, not the remaining life, which is $T-t$.

We will discuss theta again later in the context of the Black-Scholes PDE when its intuitive interpretation will become clearer.

> ***Question 12.10***
>
> For each of the Greeks ν, Θ, ρ and λ, discuss whether its value will be positive or negative in the case of:
>
> (i) a call option
>
> (ii) a put option.

2 Exam-style question

We finish this chapter with an exam-style question on the Greeks and hedging.

Question

An investor claims to be able to value an unusual derivative on a non-dividend paying share using the pricing formula:

$$V_t = S_t^2 e^{-4S_t}$$

where S_t denotes the price of the share at time t.

(i) Derive formulae for the delta and gamma of the derivative, based on the pricing formula above. [2]

(ii) For each of the following scenarios, calculate the number of shares that must be purchased or sold along with a short holding in one derivative, in order to achieve a delta-hedged portfolio:

(a) the current share price is 1

(b) the current share price is 3. [2]

(iii) Explain which of the scenarios in (ii) is likely to involve more portfolio management in the near future if the investor is determined to maintain a delta-hedged portfolio. [2]

[Total 6]

Solution

(i) ***Delta and gamma***

Using the product rule, the delta of the derivative is:

$$\Delta = \frac{\partial V_t}{\partial S_t} = 2S_t e^{-4S_t} - 4S_t^2 e^{-4S_t}$$

$$= 2S_t e^{-4S_t}\left(1 - 2S_t\right)$$

The gamma of the derivative is:

$$\Gamma = \frac{\partial^2 V_t}{\partial S_t^2} = \frac{\partial \Delta}{\partial S_t} = \frac{\partial}{\partial S_t}\left(2S_t e^{-4S_t} - 4S_t^2 e^{-4S_t}\right)$$

Using the product rule on each term in the brackets gives:

$$\Gamma = 2e^{-4S_t} - 8S_t e^{-4S_t} - 8S_t e^{-4S_t} + 16S_t^2 e^{-4S_t}$$

$$= 2e^{-4S_t}\left(1 - 8S_t + 8S_t^2\right)$$

(ii)(a) ***Current share price is 1***

When the share price is 1, the delta of the derivative is:

$$\Delta = 2 \times 1 \times e^{-4\times1}\left(1 - 2\times1\right) = -0.03663$$

So a short holding of one derivative requires *selling* 0.03663 shares to achieve a delta-hedged portfolio.

(ii)(b) ***Current share price is 3***

When the share price is 3, the delta of the derivative is:

$$\Delta = 2 \times 3 \times e^{-4\times3}\left(1 - 2\times3\right) = -0.0001843$$

So a short holding of one derivative requires *selling* 0.0001843 shares to achieve a delta-hedged portfolio.

(iii) ***Portfolio management***

The amount of rebalancing required depends on the value of gamma. If the current share price is 1, then the gamma is:

$$\Gamma = 2e^{-4\times 1}\left(1-8\times 1+8\times 1^2\right) = 0.03663$$

If the current share price is 3, then the gamma is:

$$\Gamma = 2e^{-4\times 3}\left(1-8\times 3+8\times 3^2\right) = 0.000602$$

So, in the first scenario, we have a higher value of gamma and so the delta of the portfolio is more sensitive to changes in the share price. This means it is more likely to involve more rebalancing of the portfolio in the near future if the investor is to maintain a delta-hedged position.

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

Chapter 12 Summary

The Greeks

Delta	$\Delta = \dfrac{\partial f}{\partial S_t}$	the change of the derivative price with the share price
Gamma	$\Gamma = \dfrac{\partial^2 f}{\partial S_t^2}$	the change of delta with the share price
Theta	$\Theta = \dfrac{\partial f}{\partial t}$	the change of the derivative price with time
Vega	$\nu = \dfrac{\partial f}{\partial \sigma}$	the change of the derivative price with volatility
Rho	$\rho = \dfrac{\partial f}{\partial r}$	the change of the derivative price with the risk-free rate
Lambda	$\lambda = \dfrac{\partial f}{\partial q}$	the change of the derivative price with the dividend rate

Delta hedging

A portfolio for which the weighted sum of the deltas of the individual assets is equal to zero is described as *delta-hedged*.

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

Chapter 12 Solutions

Solution 12.1

We know the payoff functions for European options:

$$c_T = \max\{S_T - K, 0\}$$

and $$p_T = \max\{K - S_T, 0\}$$

We know the upper and lower bounds for European and American options on a non-dividend-paying share:

$$S_t \geq c_t \ \geq \ \max\left\{S_t - Ke^{-r(T-t)}, 0\right\}$$
$$S_t \geq C_t \ \geq \ \max\left\{S_t - Ke^{-r(T-t)}, 0\right\}$$
$$Ke^{-r(t-t)} \geq p_t \ \geq \ \max\left\{Ke^{-r(T-t)} - S_t, 0\right\}$$
$$K \geq P_t \ \geq \ \max\{K - S_t, 0\}$$

We know that it is never beneficial to exercise an American call option early on a non-dividend-paying share. The consequence of this is that the time value of a European call option must be positive for a non-dividend-paying share.

We know put-call parity always holds:

$$c_t + Ke^{-r(T-t)} = p_t + S_t e^{-q(T-t)}$$

Solution 12.2

The most important result is probably put-call parity. This is an incredibly important result that can be used to find the correct price of a put option from the price of a call option and vice versa:

$$c_t + Ke^{-r(T-t)} = p_t + S_t e^{-q(T-t)}$$

If a pricing model contradicts this then it is inaccurate and should not be used. We will see later with the models described here, that they are all consistent with put-call parity

Solution 12.3

(i) ***Call option***

An increase in the share price would either push a call option that is currently out-of-the-money towards being in the money or push one that is already in-the-money further into the money. Therefore delta is positive for a call option.

(ii) ***Put option***

A decrease in the share price would either push a put option that is currently out-of-the-money towards being in-the-money or push one that is already in-the-money further into the money. Therefore delta is negative for a put option.

Solution 12.4

When purchasing a call or a put option, the option premium paid is typically much smaller than the price of the underlying asset itself. Purchasing the option nevertheless enables us to obtain exposure to most of the variation in the price of the underlying asset. Consequently, the *percentage* returns obtained by purchasing the option will be much greater (or much less!) than those obtained by purchasing the underlying asset instead. This effect is referred to as gearing.

Solution 12.5

Portachio's portfolio will be worth £11,000 because each share will go up in value from £10 to £11. Benny Kiang's portfolio will be worth £12,500 because each option will go up by $0.5 \times £1 = 50\text{p}$ from £2 to £2.50.

Solution 12.6

The delta of the underlying asset is equal to the derivative of the underlying asset price with respect to itself. By definition, this must be equal to one, *ie*:

$$\Delta = \frac{\partial S_t}{\partial S_t} = 1$$

Solution 12.7

Consider the change in the value of the portfolio $dV(t, S_t)$. Ito's Lemma tells us that:

$$dV(t, S_t) = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S_t}dS_t + ½\frac{\partial^2 V}{\partial S_t^{\ 2}}(dS_t)^2$$

Now since S_t is a diffusion we have $dS_t = \mu(t, S_t)dt + \sigma(t, S_t)dZ_t$, so $(dS_t)^2 = \sigma^2 dt$ giving:

$$\begin{aligned} dV &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S_t}dS_t + ½\sigma^2\frac{\partial^2 V}{\partial S_t^{\ 2}}dt \\ &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S_t}\big(\mu(t, S_t)dt + \sigma(t, S_t)dZ_t\big) + ½\sigma^2(t, S_t)\frac{\partial^2 V}{\partial S_t^{\ 2}}dt \\ &= \left(\frac{\partial V}{\partial t} + \mu(t, S_t)\frac{\partial V}{\partial S_t} + ½\sigma^2(t, S_t)\frac{\partial^2 V}{\partial S_t^{\ 2}}\right)dt + \sigma(t, S_t)\frac{\partial V}{\partial S_t}dZ_t \end{aligned}$$

It follows that the change in the portfolio value V will be deterministic if the stochastic dZ_t term vanishes, *ie* if $\Delta = \dfrac{\partial V}{\partial S_t} = 0$.

The portfolio is said only to be *instantaneously* risk-free because as the share price changes with time, so will the value of delta, and hence the delta-hedged position will no longer hold.

Solution 12.8

Gamma is the second derivative, *ie* the derivative of delta. We saw in the previous question that the delta of the underlying asset is equal to one and hence constant. The derivative of a constant is zero and hence so must be the gamma of the underlying asset, *ie*:

$$\Gamma = \frac{\partial \Delta}{\partial S_t} = \frac{\partial}{\partial S_t}(1) = 0$$

Solution 12.9

If we buy a future on a dividend-paying share, instead of the share itself, then we do not receive the dividends paid by the share over the lifetime of the future. The greater the value of these dividends, the more income we miss out on by holding the future and the lower is the value of the future. Consequently, it must be the case that lambda is negative for a future.

Solution 12.10

Vega

If the underlying security becomes more volatile then there is a greater chance of the price moving in favour of the option holder. Although there is also an increased chance of it moving against the holder, the downside loss is capped. Therefore vega should be positive for both a call option and a put option.

Theta

The intuitive argument is very similar to before. The greater the time to expiry, the more chance that the share price will move in the holder's favour, with the downside loss again being capped. Thus, because time t works against time to expiry $T-t$, Theta is usually negative for both a call and a put option.

There are circumstances where theta may be positive but these are considered beyond the scope of Subject CT8 and are left to Subject ST6.

Rho

You can think of holding a call option as having cash in the bank waiting to buy the share. If interest rates rise then the holder of a call option will benefit in the mean time. The holder of a put option may already own a share and is waiting to sell it for cash. So if interest rates rise then they will lose out on that interest in the mean time. So rho is positive for a call option and negative for a put.

Lambda

Again, you can think of holding a call option as having cash in the bank waiting to buy the share. If the dividend rate rises then the holder of a call option will lose out in the mean time. The holder of a put option may already own a share and is waiting to sell it for cash. So if dividend rates rise then they will benefit from the extra dividends in the mean time. So lambda is negative for a call option and positive for a put.

Chapter 13

The binomial model

Syllabus objectives

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

4. Show how to use binomial trees and lattices in valuing options and solve simple examples.

5. Derive the risk-neutral pricing measure for a binomial lattice and describe the risk-neutral pricing approach to the pricing of equity options.

6. Explain the difference between the real-world measure and the risk-neutral measure. Explain why the risk-neutral pricing approach is seen as a computational tool (rather than a realistic representation of price dynamics in the real world).

7. State the alternative names for the risk-neutral and state price deflator approaches to pricing.

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 this chapter, we start to develop simple models that can be used to value derivatives. In particular we use *binomial trees* or *lattices* to find the value at time 0 of a derivative contract that provides a payoff at a future date based on the value of a non-dividend-paying share at that future date.

The analysis throughout this and the subsequent chapters applies the no-arbitrage principle in order to value derivative contracts. The basic idea is that if we can construct a portfolio that replicates the payoff from the derivative under every possible circumstance, then that portfolio must have the same value as the derivative. So, by valuing the replicating portfolio we can value the derivative.

The models discussed in this chapter represent the underlying share price as a stochastic process in discrete time and with a discrete state space - in fact a geometric random walk. In subsequent chapters, we will discuss the continuous-time and continuous-state space analogue, geometric Brownian motion (or the lognormal model), which can be interpreted as the limiting case of the binomial model as the size of the time steps tends to zero.

1 Background

Here we will consider a model for stock prices in discrete time. This is the first point where we introduce a model. The model will seem very simple and naïve at first sight, but:

(a) it introduces the key concepts of financial economic pricing and

(b) it leads us to the celebrated Black-Scholes model as a limiting case.

1.1 Assumptions

In the binomial model it is assumed that:

- **there are no trading costs** or taxes
- **there are no minimum or maximum units of trading**
- **stock and bonds can only be bought and sold at discrete times 1, 2, ...**
- the principle of no arbitrage applies.

As such the model appears to be quite unrealistic. However, it does provide us with good insight into the theory behind more realistic models. Furthermore it provides us with an effective *computational* tool for derivatives pricing.

As well as being the principle used to determine the derivative price, the principle of no arbitrage leads to a constraint on the parameters used in the binomial model.

1.2 Definitions

The share price process

We will use S_t to represent the price of a non-dividend-paying stock at discrete time intervals $t\,(t = 0,1,2,...)$. For $t > 0, S_t$ is random.

You have already met this notation at the beginning of Chapter 10. Here, in the binomial model, we are working in the specific case where t is a positive integer.

Note that in this instance "stock" specifically means a share or equity as opposed to a bond. For the time being we ignore the possibility of dividends, which would otherwise unnecessarily complicate matters.

The stock price S_t is assumed to be a random or stochastic process. Over any discrete time interval from $t-1$ to t, we assume that S_t either goes up or goes down. We also assume that we cannot predict beforehand which it will be and so future values of S_t cannot be predicted with certainty. We will, however, be able to attach probabilities to each possibility and we also assume that the sizes of the jumps up or down are known.

The cash process

Besides the stock we can also invest in a bond or a cash account which has value B_t at time *t* per unit invested at time 0. This account is assumed to be risk-free and we will assume that it earns interest at the constant risk-free continuously-compounding rate of *r* per annum. So $B_t = e^{rt}$.

We usually assume that $r > 0$, so that:

- $B_t > B_0 = 1$ and $B_t > B_{t-1}$ for $t > 0$
- B_t increases in an entirely predictable manner as we move through time, *ie* at the continuously-compounding rate of r per time period.

At all points in time there are no constraints (positive or negative) on how much we hold in stock or cash.

Question 13.1

If we invested one unit in the cash account at time 0, what will its value be at time 1?

2 The one-period model

2.1 Basic structure

The aim in this chapter is to find the value at an arbitrary time 0 – *ie now* – of a derivative that provides a payoff at some future date based on the value of the stock at that future date. As a starting point, we consider a one-period binomial model. In this model, we start at time $t = 0$, when the stock price is equal to S_0. Over the one-period time interval to $t = 1$, the stock price will do one of two things. It will either:

1. jump upwards or
2. jump downwards.

Here we are trying to find the value of a derivative that pays out an amount that depends directly on the value of the stock price at time 1, *ie* on S_1.

We have two possibilities for the price at time 1:

$$S_1 = \begin{cases} S_0 u & \text{if the price goes up} \\ S_0 d & \text{if the price goes down} \end{cases}$$

Here u is a fixed number bigger than 1 and d is a fixed number less than 1.

This is represented in Figure 13.1 (part of the Core Reading).

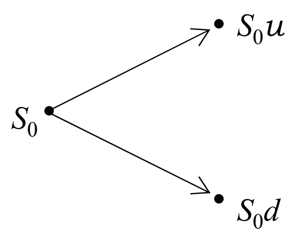


Figure 13.1: One-period binomial model for stock prices

Figure 13.1 shows a one-period binomial model for the stock price. *One-period* because we consider only the time interval from time 0 to time 1, *binomial* because there are only two possible ways in which the stock price can move, *ie* up to S_0u or down to S_0d.

Such a model is referred to as a *binomial tree* and each of the paths from S_0 to S_0u and from S_0 to S_0d as a *branch*. The points at each end of the branches are sometimes known as *nodes*.

We can now see the implication of there being no arbitrage in this model. **In order to avoid arbitrage we must have $d < e^r < u$. Suppose this is not the case:**

For example, if $e^r < d < u$, then we could borrow £1 of cash and buy £1 of stock. At time 0 this would have a net cost of £0. At time 1 our portfolio would be worth:

- $d - e^r$ **or**
- $u - e^r$

both of which are greater than 0. This is an example of arbitrage.

Or more generally, if one asset (*eg* the share) is certain to earn more over the one step than the other (*eg* the cash), then we simply borrow the other asset to invest in it as many times over as we like, thus making an arbitrage profit. The only way to avoid this scenario is if the share performs either "better" or "worse" than the cash, *ie* $d < e^r < u$.

2.2 *Determination of the derivative price at time 0*

Suppose that we have a derivative which pays c_u if the price of the underlying stock goes up and c_d if the price of the underlying stock goes down.

So, the value of the payment made by the derivative, which we can denote by the random variable *C*, depends on the underlying stock price.

At what price should this derivative trade at time 0?

At time 0 suppose we hold ϕ units of stock and ψ units of cash. The value of this portfolio at time 0 is V_0.

Question 13.2

What is V_0 equal to?

The portfolio is sometimes represented as an ordered pair (ϕ, ψ). ϕ and ψ are the Greek letters "phi" and "psi".

At time 1 the same portfolio has the value:

$$V_1 = \begin{cases} \phi S_0 u + \psi e^r & \text{if the stock price went up} \\ \phi S_0 d + \psi e^r & \text{if the stock price went down} \end{cases}$$

eg if the stock price went up, then the value of one unit of the stock has increased to S_0u and that of an initial unit in the cash account to e^r.

So (ϕ,ψ) could be any values and this portfolio could consist of an amount of shares and an amount of cash in whatever amounts we choose. We are going to choose ϕ and ψ in order that the portfolio replicates the payoff of the derivative, no matter what the outcome of the share price process.

Let us choose ϕ and ψ so that $V_1 = c_u$ if the stock price goes up and $V_1 = c_d$ if the stock price goes down. Then:

$$\phi S_0 u + \psi e^r = c_u$$

and:

$$\phi S_0 d + \psi e^r = c_d$$

This choice of ϕ and ψ therefore ensures that the value of the portfolio (ϕ,ψ) at time 1 is equal to the derivative payoff whether the stock price goes up or down. Hence, by the no-arbitrage principle, the value of this portfolio at time 0, V_0, must also be the value of the derivative contract at that time. So, if we can solve these simultaneous equations to find ϕ and ψ, then we can determine V_0, which must be equal to the value of the derivative at time 0.

So, we have two linear equations in two unknowns, ϕ and ψ. The easiest way to solve these equations is to subtract them. This enables us to find ϕ. We can then substitute into either equation to find ψ. **We solve this system of equations and find that:**

$$\phi = \frac{c_u - c_d}{S_0(u-d)}$$

$$\psi = e^{-r}\left(\frac{c_d u - c_u d}{u-d}\right)$$

$$\Rightarrow \qquad V_0 = \phi S_0 + \psi$$

We now notice that all the terms in the numerators for ϕ and ψ involve either c_u or c_d. Therefore it is possible to rearrange V_0 so that is a sum of a multiple of c_u and a multiple of c_d.

$$\Rightarrow \qquad V_0 = \frac{c_u - c_d}{u-d} + \frac{e^{-r}(c_d u - c_u d)}{u-d}$$

$$= \frac{1}{u-d}\left[c_u(1-de^{-r}) + c_d(ue^{-r}-1)\right]$$

$$= \frac{e^{-r}}{u-d}\left[(e^r - d)c_u + (u-e^r)c_d\right]$$

$$= e^{-r}\left[\frac{e^r-d}{u-d}c_u + \frac{u-e^r}{u-d}c_d\right]$$

ie $\quad \boldsymbol{V_0 = e^{-r}\left(qc_u + (1-q)c_d\right)}$

where $\boldsymbol{q = \dfrac{e^r - d}{u-d}}$ **and** $\boldsymbol{1-q = \dfrac{u-e^r}{u-d}}$**.**

Note that the no-arbitrage condition $\boldsymbol{d < e^r < u}$ ensures that $\boldsymbol{0 < q < 1}$.

Recall that the formula for V_0 must, by the no-arbitrage principle, be equal to the price of the derivative at time 0. Because $0 < q < 1$, let's just "pretend" for now that q is a probability. If we do this then we can see that the right hand side simplifies to:

$$V_0 = e^{-r}E[C_1]$$

ie, the discounted value of the expected derivative payoff at time 1. But remember we are just "pretending" that q is a probability when in actual fact it isn't a real-world probability. Because we are only "pretending" we can add a Q to the notation to remind ourselves:

$$V_0 = e^{-r}E_Q[C_1]$$

We will see that, although it doesn't reflect reality, this "pretending" idea is a good way of thinking about it.

If we denote the payoff of the derivative at $t=1$ by the random variable C_1, we can write:

$$V_0 = e^{-r} E_Q(C_1)$$

where Q is a probability measure which gives probability q to an upward move in prices and $1-q$ to a downward move. We can see that q depends only upon u, d and r and not upon the potential derivative payoffs c_u and c_d. Note, by convention $E_Q\left[Y|F_0\right]$ is sometimes written as $E_Q[Y]$.

You can think of a probability measure as a set of probabilities (more about this later). Note also that Q doesn't depend in any way on the real-world probabilities.

This is a surprising result. We might have thought at the start that the price of the derivative should be some sort of discounted value of the expected payoff. The above result is telling us that this is the case, but the probabilities used to calculate the expectation are not the real-world probabilities of the two possible derivative values! Instead we calculate the expectation using the "probabilities" q and $1-q$.

Question 13.3

Write down an expression for C_1.

2.3 Probability measures

General definition

Within the context of a binomial tree, a *probability measure* is simply a function that assigns a real number in the interval [0,1] to each of the branches in our binomial tree such that at any node they sum to 1. We can thus interpret each number as a probability.

Note carefully that the probabilities assigned by a probability measure do *not* need to correspond to the real-world probabilities that the stock price *actually* moves up or down between times 0 and 1. In fact, any function that assigns a value $q \in [0,1]$ to the up-branch and $1-q$ to the down-branch is a possible probability measure. Q is therefore the particular measure, amongst many possible measures, that happens to assign the probability $q = \dfrac{e^r - d}{u-d}$ to the up-branch between times 0 and 1, and $1-q = \dfrac{u-e^r}{u-d}$ to the down-branch.

$E_Q(C_1)$ therefore represents the expectation of the derivative payoff C_1 with respect to the probability measure Q – *ie* the expectation of C_1 based on an up-probability of q and a down-probability of $1-q$.

We could equally evaluate the expectation of C_1 based on a different probability measure, P say, which assigns a different probability $0 \leq p \leq 1$ to the up-branch. In that case, $E_P(C_1)$ would be the expectation of C_1 with respect to P.

ie $\quad E_P(C_1) = p\,c_u + (1-p)\,c_d$

We could in fact use any probability measure we like to calculate an expectation – including the real-world probabilities – different measures producing different expectations. The probability measure Q is, however, especially useful as it enables us to calculate the derivative price.

In fact Q is called the risk-neutral probability measure and you should answer the following question to see why.

Question 13.4

Evaluate $E_Q[S_1]$ and hence suggest a reason why Q is called the risk-neutral probability measure.

Replicating portfolio

The portfolio (ϕ,ψ) is called a *replicating portfolio* because it replicates, precisely, the payoff at time 1 on the derivative without any risk.

In other words, $V_1 = c_u$ if the stock price goes up and $V_1 = c_d$ if the stock price goes down. So, there is *never* any risk or possibility that $V_1 \neq C_1$. Hence, in a world that is free of arbitrage, the values of the derivative and the replicating portfolio (ϕ,ψ) at time 0 must be equal – *ie* they are both equal to $V_0 = e^{-r}E_Q(C_1)$.

It is also a simple example of a hedging strategy: that is, an investment strategy which reduces the amount of risk carried by the issuer of the contract. In this respect not all hedging strategies are replicating strategies.

Recall that a hedging strategy is one that reduces the extent of, or in this case eliminates, any variation in the market value of a portfolio.

So, if we hold the portfolio (ϕ,ψ) and sell the derivative C, then the total value of the resulting portfolio at time 1 will be zero, however the stock price moves over the interval to time 1. An immediate consequence of this is that in an arbitrage-free world the value of the combined portfolio is zero at time 0. Hence, the value of the combined portfolio does not change from zero as we move from $t = 0$ to $t = 1$ and so we are said to have a perfectly-hedged position. This will also be true if we instead sell the portfolio (ϕ,ψ) and hold the derivative C.

It is also possible to have an intermediate position where the risk is reduced but not eliminated. This would be the case for example if we sold the derivative C and held the portfolio $(\tfrac{1}{2}\phi,\tfrac{1}{2}\psi)$. This portfolio is *not* a replicating strategy.

The real-world probability measure, P

Up until now we have not mentioned the *real-world* probabilities of up and down moves in prices. Let these be p and $1-p$ where $0<p<1$, defining a probability measure P.

So P is a set of probabilities that assigns the *actual* or *real-world* probability p of an upward jump in the stock price to an up-branch of the binomial tree and $1-p$ to a down-branch.

Other than by total coincidence, p will not be equal to q – as p is the actual probability of the stock price moving upwards, whereas q is simply a number defined as:

$$q=\frac{e^r-d}{u-d}$$

So, q depends upon u, d and r, but not p.

When we wish to emphasise that q is *not* a real-world probability, it is often referred to as a *synthetic probability*. The important thing to note is that the real-world probability p (if indeed this can be determined) is irrelevant to our calculation of the derivative price, which is based solely on the synthetic probability q.

2.4 *The risk-neutral probability measure*

Let us consider the expected stock price at time 1. Under P this is:

$$E_P(S_1)=S_0\left(pu+(1-p)d\right)$$

This is of course just the expectation of the stock price at time 1 with respect to the real-world probability measure P.

and under *Q* it is:

$$E_Q(S_1) = S_0\left(qu + (1-q)d\right) = S_0\left(\frac{u(e^r - d)}{u-d} + \frac{d(u-e^r)}{u-d}\right) = S_0 e^r$$

Under *Q* we see that the expected return on the *risky* stock is the same as that on a risk-free investment in cash – as the last equation shows that the expected stock price $E_Q(S_1)$ is simply equal to the accumulation of the initial stock price S_0 at the risk-free rate of return r. (So, the expected rate of return on the stock must be equal to the risk-free rate.)

In other words under the probability measure *Q* investors are neutral with regard to risk: they require no additional return for taking on more risk.

This is because both risk-free cash and the risky stock are priced so as to yield the same expected return with respect to measure Q. So if we use the measure Q this is equivalent to assuming that investors do not require an additional risk premium to compensate them for the additional risk that they incur when investing in the risky stock, *ie* they are risk-neutral.

Note carefully that we are *not* saying that investors in the real world are risk-neutral – or equivalently that investors are risk-neutral under the real-world probability measure P. We are simply saying that they can be assumed to be risk-neutral under the non-real-world probability measure Q, which is something entirely different.

This is why *Q* is sometimes referred to as a *risk-neutral probability measure.* It is the probability measure with respect to which any asset, whether risky or risk-free, offers the same expected return to investors, namely, the risk-free rate of return.

Under the real-world measure *P* the expected return on the stock will not normally be equal to the return on risk-free cash. Under normal circumstances investors demand higher expected returns in return for accepting the risk in the stock price. So we would normally find that $p > q$. However, this makes no difference to our analysis.

Question 13.5

Show why we would normally find that $p > q$.

2.5 Numerical example

The above computations will become clearer if we consider a simple numerical example.

Let's consider a one-period binomial model of a stock whose current price is 40. Suppose that:

- over the single period under consideration, the stock price can either move up to 60 or down to 30
- the actual probability of an up-movement is equal to ½
- the continuously-compounded risk-free rate of return is 5% per time period
- we wish to find the current value of a one-period European call option, V_0, that has an exercise price of 45.

The binomial tree in respect of the stock price is therefore as follows:

$$\begin{tikzcd}[column sep=large] & S_1 = 60 \\ S_0 = 40 \arrow[ur] \arrow[dr] & \\ & S_1 = 30 \end{tikzcd}$$

We can likewise consider a corresponding binomial tree in respect of C_1, the payoff provided by the call option at time 1 – *ie* the profit paid at exercise:

$$\begin{tikzcd}[column sep=large] & C_1 = c_u = 60 - 45 = 15 \\ C_0 = ? \arrow[ur] \arrow[dr] & \\ & C_1 = c_d = 0 \end{tikzcd}$$

Recall that V_0, the value of the replicating portfolio (ϕ, ψ) at time 0, must be equal to C_0, the derivative value at time 0.

Question 13.6

Why does $c_d = 0$?

In order to find V_0, and hence C_0, we can calculate the risk-neutral probabilities, q and $1-q$, and then use the result that:

$$V_0 = e^{-r} E_Q(C_1)$$

ie $$V_0 = e^{-r}\left[q\, c_u + (1-q)\, c_d\right]$$

Recall from above that the risk-neutral probability q is equal to:

$$q = \frac{e^r - d}{u - d}$$

In this instance:

$$d = \frac{30}{40} = 0.75$$

and $$u = \frac{60}{40} = 1.5$$

So $$q = \frac{e^{0.05} - 0.75}{1.5 - 0.75} = 0.40169$$

and $$1 - q = 0.59831$$

Hence:

$$V_0 = e^{-0.05}\left[0.40169 \times 15 + 0.59831 \times 0\right] = 5.732$$

ie $$C_0 = 5.732$$

Question 13.7

Find the constituents of the replicating portfolio (ϕ, ψ) and show that it costs 5.732 to set up this portfolio.

Three final points to note are that:

1. We did not use the actual or real-world probability of $p = ½$ to find V_0, *ie* V_0 is independent of p.

2. $p = ½ > 0.40169 = q$, as noted above.

3. V_0 differs from the discounted value of the expected derivative payoff based on the actual probability of $p = ½$, which is equal to:

$$V_0' = e^{-0.05}\left[0.5\times 15 + 0.5\times 0\right] = 7.134$$

Note that $V_0' > V_0$. This is because $p > q$, and so the real-world probability measure places greater weight upon the good outcome when the share price increases to 60.

3 Two-period binomial tree

3.1 Basic structure

The one-period binomial model is a good starting point for our analysis of derivative pricing. The stochastic processes underlying stock price movements will, however, typically be more complex than can be represented by the one-period model. We therefore need to extend the previous model into a multi-period context. The obvious next step is thus to develop a two-period binomial model. It turns out that the ideas discussed above carry over quite naturally to the two-period model.

To keep things simple, we assume that the periods we are dealing with are years.

Warning: The notation used in this section makes the calculations look more complicated than they really are. When you carry out numerical calculations using a diagram, it's actually quite a straightforward process.

We now look at a two-period binomial model. The most general model is as follows (see Figure 13.2):

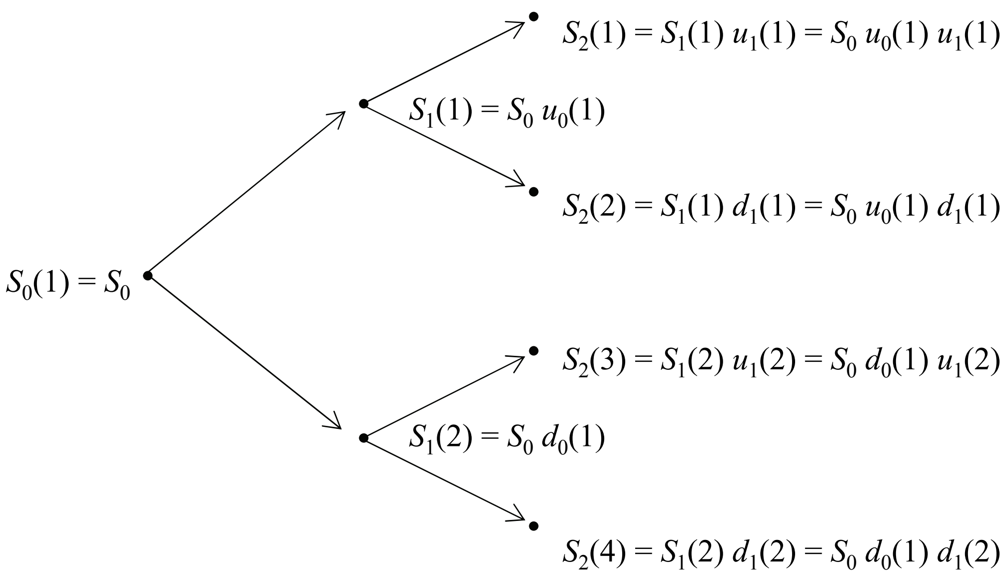


Figure 13.2: Two-period binomial model. Where the price at a particular node is denoted $S_t(j)$ this means the price in state (t, j).

The subscripts here denote time and the arguments in brackets denote the vertical position, counting from the top of the tree.

Year 1:

$$\boldsymbol{S_0} \rightarrow \begin{cases} \boldsymbol{S_0\, u_0(1)} & \textbf{if up, to state (1,1)} \\ \boldsymbol{S_0\, d_0(1)} & \textbf{if down, to state (1,2)} \end{cases}$$

The first year is essentially the same as our previous one-period model. The stock price therefore starts at S_0 at time 0 and either goes up or down over the year from $t=0$ to $t=1$. We do, however, need to be more careful with our notation once we move into a multi-period model.

Accordingly, the stock price is assumed to start from an initial state or node that is referred to as State (0,1) – *ie State 1 at time 0.* There is, of course, only one possible state at time 0, defined by the stock price S_0. The stock price then changes over the year from time 0 to time 1. It either moves along the up-branch to State (1,1), the first state (or up-node) at time 1, or down to State (1,2) the second state (or down-node) at time 1.

The stock price at State (1,1) following an up-movement is therefore $S_0\, u_0(1)$, where $u_t(j)$ denotes the proportionate increase in the stock price in an up-movement over the time interval from time t to time $t+1$, starting from State (t, j). Similarly, should the stock price moves down to State (1,2), then it will be equal to $S_0\, d_0(1)$.

Year 2:

From state $(\mathbf{1,1})$**,** *ie* following a price *increase* in the first time interval**:**

$$\boldsymbol{S_0 u_0}(\mathbf{1}) \rightarrow \begin{cases} \boldsymbol{S_0\, u_0(1)\, u_1(1)} & \textbf{if up, to state (2,1)} \\ \boldsymbol{S_0\, u_0(1)\, d_1(1)} & \textbf{if down, to state (2,2)} \end{cases}$$

So, for example, if the stock price falls in the second period, then it moves from State (1,1) to State (2,2) via the down-branch $d_1(1)$.

From state $(\mathbf{1,2})$**,** *ie* following a price *decrease* in the first time interval**:**

$$\boldsymbol{S_0 d_0}(\mathbf{1}) \rightarrow \begin{cases} \boldsymbol{S_0\, d_0(1)\, u_1(2)} & \textbf{if up, to state (2,3)} \\ \boldsymbol{S_0\, d_0(1)\, d_1(2)} & \textbf{if down, to state (2,4)} \end{cases}$$

Recall that $u_1(2)$ denotes the up-branch price ratio from time 1 to time 2, starting at State (1,2).

3.2 *Determination of the derivative price at time 0*

Suppose that the derivative gives a payoff at time 2 of $c_2(j)$ if the price at time 2 of the stock is in state $(2, j)$.

How do we calculate the price of the derivative at time 0?

The approach used turns out to be the natural extension of that used in the one-period model. Here we first find the intermediate price at time 1 and then use this price to find the price at time 0.

We do this by working backwards from time 2. So we calculate the value of the contract at time 1 for each of the possible states at time 1 – *ie* State (1,1) following an upward movement in the first period and State (1,2) following a price fall.

Let $V_1(j)$ be the value of the contract if we are in state *j* at time 1. Then, by analogy with the one-period model:

$$V_1(1) = e^{-r}\left(q_1(1)c_2(1) + \left(1 - q_1(1)\right)c_2(2)\right)$$

$$V_1(2) = e^{-r}\left(q_1(2)c_2(3) + \left(1 - q_1(2)\right)c_2(4)\right)$$

where: $q_1(1) = \dfrac{e^r - d_1(1)}{u_1(1) - d_1(1)}$ **and:** $q_1(2) = \dfrac{e^r - d_1(2)}{u_1(2) - d_1(2)}$

So, the value of the derivative at time 1, is equal to the expectation at time 1, of the derivative payoff at time 2, calculated with respect to the risk-neutral probability measure Q (discounted at the risk-free rate of return).

No-arbitrage conditions imply that $d_t(j) < e^r < u_t(j)$ (and hence $0 < q_t(j) < 1$) for all *t* and *j*.

Question 13.8

Why do no-arbitrage conditions imply that $d_t(j) < e^r < u_t(j)$?

The price at time 0 is found by treating the values $V_1(1)$ and $V_1(2)$ in the same way as derivative payoffs at time 1.

So:

$$V_0(1) = e^{-r}\left[q_0(1)V_1(1) + \left(1 - q_0(1)\right)V_1(2)\right]$$

So, the value of the derivative at time 0, is equal to the expectation at time 0, of the derivative payoff at time 1, calculated with respect to the risk-neutral probability measure Q and discounted at the risk-free rate of return.

Combining the two steps above we can then conclude that the value of the derivative at time 0 is equal to the expectation at time 0 of the derivative payoff *at time 2*, calculated with respect to the risk-neutral probability measure Q and discounted at the risk-free rate of return over two time periods. We return to this below.

Let C_2 be the random derivative payoff at time 2 (that is, it takes one of the values $c_2(j)$ for $j = 1,2,3,4$).

Let V_t be the random value of the contract at time *t*.

Let F_t be the history of the process up to and including time *t* (that is, the sigma-algebra generated by the sample paths up to and including time *t*) and let *F* be the sigma-algebra generated by all sample paths (up to the final time considered by the model).

"Sigma-algebra" is a technical term that is not important here. Recall that F_t, which we called the *filtration* in Chapter 8, is basically the information known about the process S_t by time *t*. *F* is often used to denote the information known by the final time of the model, in this case the payoff date of the derivative at time 2.

Amongst the information that F_t gives us we have:

- our current position in the binomial tree, $S_t(j)$
- the current stock price S_t.

So, $V_t(j)$, the value of the derivative contract if we are in State *j* at time *t*, must depend on F_t.

Let *Q* be the probability measure generated by the probabilities $q_0(1), q_1(1)$ and $q_1(2)$ – *ie* the risk-neutral probability measure**.**

Then:

$$V_1(1) = e^{-r} E_Q\left[C_2 \mid up\ in\ year\ 1\right]$$

$$V_1(2) = e^{-r} E_Q\left[C_2 \mid down\ in\ year\ 1\right]$$

or $$V_1 = e^{-r} E_Q\left[C_2 \mid F_1\right]$$

> ***Question 13.9***
>
> Explain briefly in words what the final expression above means.

Likewise:

$$V_0 = e^{-r} E_Q\left[V_1 \mid F_0\right]$$

$$= e^{-r} E_Q\left[e^{-r} E_Q\left\{C_2 \mid F_1\right\} \mid F_0\right]$$

$$= e^{-2r} E_Q\left[C_2 \mid F_0\right]$$

So, the no-arbitrage value of the derivative at time 0, is equal to:

- the time 0 *expectation* of the derivative payoff paid at time 2
- calculated with respect to the *risk-neutral probability measure Q*
- and the *information set* (*sigma-algebra*) generated by the history of the stock price up to and including time 0
- and *discounted* at the continuously-compounded risk-free rate of return.

Note that the intermediate expression above is derived by applying the so-called *tower property* of conditional expectations. This states that for any random variable X, probability measure P and sigma-algebra F_i:

$$E_P[E_P\{X|F_j\}|F_i] = E_P[X|F_i]$$

for $i \le j$. This is a generalisation of the formula $E[E[A|B]] = E[A]$.

3.3 Numerical example

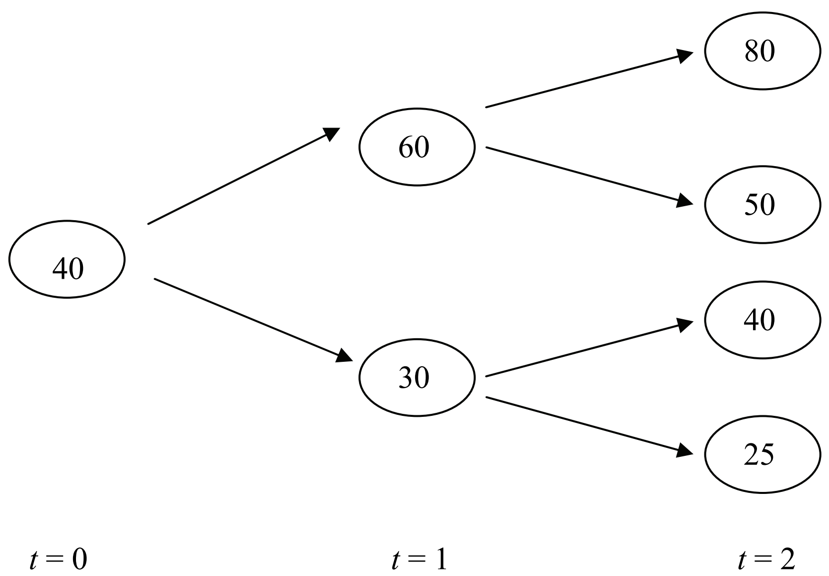
Again a numerical example should help to clarify matters. This example extends the previous one-period example to a two-period scenario.

Let's now consider a two-period binomial model of a stock whose current price is 40 (as before). Suppose that:

- Over the first period, the stock price can either move up to 60 or down to 30 (as before).
- Following an up-movement in the first period, the stock price can either move up to 80 or down to 50.
- Following a down-movement in the first period, the stock price can either move up to 40 or down to 25.
- The real-world probability of an up-movement is always equal to ½ (as before).
- The continuously-compounded risk-free rate of return is 5% per time period.
- We wish to find V_0, the current value of a two-period European call option, that has an exercise price of 45.

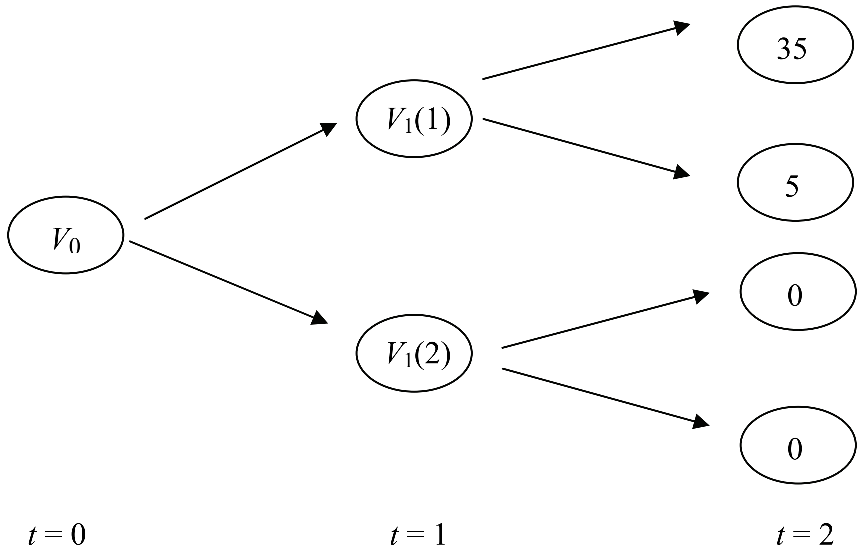
The binomial tree in respect of the *stock price* is therefore as follows:

Share prices



We can also draw a corresponding binomial tree showing the derivative value at each state or node:

Derivative values



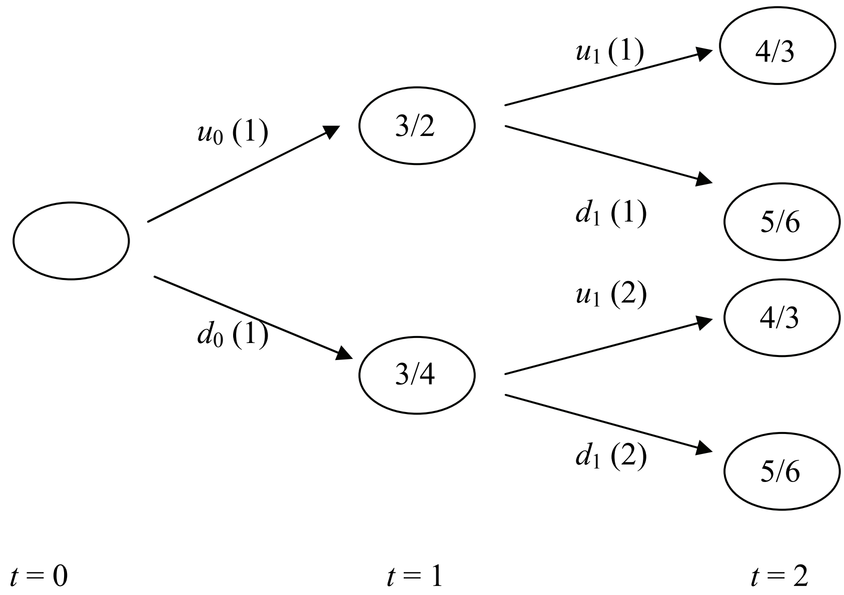
Note that the derivative value at each final node is always equal to the payoff at that node – which is certain once that particular node has been reached.

Question 13.10

Draw a further binomial tree corresponding to the above trees and annotate it with the states $S_t(j)$ and the risk-neutral probabilities. (You do not need to work out any numbers yet – just show the symbols.)

Finally, we can also annotate the branches of the binomial tree with the price movement factors, u and d – and show the relevant values in the nodes:

Share price movements



To find V_0, the value of the derivative at time 0, we need to work backwards through the tree. In doing so, we can essentially consider the binomial tree as three distinct one-period trees – two in the second period and one in the first. For each subtree we apply the one-period approach of the previous section and then combine our results to find the two-period derivative price. We therefore proceed in three steps as follows:

1. We first determine $V_1(1)$, the derivative's value at time 1, assuming that we are then at State (1,1). This is found in exactly the way as in the one-period model, but using the second-period risk-neutral probability $q_1(1)$.

2. We repeat the first step to find $V_1(2)$, the derivative's values at time 1, assuming that we are then at State (1,2), using the second-period risk-neutral probability $q_1(2)$.

3. We then repeat the procedure to find V_0, the derivative's value at time 0, as the expected and discounted value of V_1 using the first-period risk-neutral probability $q_0(1)$.

Step 1: Starting at State (1,1)

The price movement factors are $u_1(1) = 4/3$ and $d_1(1) = 5/6$. So, the risk-neutral probability is:

$$q_1(1) = \frac{e^r - d_1(1)}{u_1(1) - d_1(1)} = \frac{e^{0.05} - 5/6}{4/3 - 5/6} = 0.43588$$

Starting from State (1,1), the possible derivative payoffs at time 2 are $c_2(1) = 35$ and $c_2(2) = 5$.

Hence:

$$V_1(1) = e^{-r}\left(q_1(1)c_2(1) + (1 - q_1(1))c_2(2)\right)$$

ie $$V_1(1) = e^{-0.05}\left(0.43588 \times 35 + (1 - 0.43588) \times 5\right) = 17.195$$

Step 2: Starting at State (1,2)

The price movement factors are $u_1(2) = 4/3$ and $d_1(2) = 5/6$.

Starting from State (1,2), the possible derivative payoffs at time 2 are $c_2(1) = 0$ and $c_2(2) = 0$.

So, the risk-neutral probability is:

$$q_1(2) = \frac{e^r - d_1(2)}{u_1(2) - d_1(2)} = \frac{e^{0.05} - 5/6}{4/3 - 5/6} = 0.43588$$

Note that this answer is the same as in the first step.

> ***Question 13.11***
>
> Why?

Finally, we have that:

$$V_1(2) = e^{-r}\left(q_1(2)c_2(3) + (1 - q_1(2))c_2(4)\right)$$

ie $$V_1(2) = e^{-0.05}\left(0.43588 \times 0 + (1 - 0.43588) \times 0\right) = 0$$

This is equal to zero as both of the final possible stock prices are below the exercise price and so the derivative payoff at time 2 must be zero – given that the stock price fell in the first period.

So, the calculations in Step 2 were pointless! We gave them for completeness but in the exam you should always be aware of a chance to simplify the calculations like this.

Step 3: Starting at State (0,1)

Q

Question 13.12

Repeat the previous steps for the first time period to show that:

(i) the risk-neutral probability is $q_0(1) = 0.40169$

(ii) the derivative price at time 0 is $V_0 = 6.570$.

4 n-period binomial tree

The final step in the development of the binomial model approach to derivative pricing is to extend the model to a more general, multi-period or *n*-period context. Once again, the ideas discussed above carry over quite naturally into the general multi-period model.

4.1 Basic structure

We can generalise all the preceding results to *n* periods.

At time *t* there are 2^t possible states $(t,1),(t,2),...,(t,2^t)$.

Recall that in the two-period model:

- at time 0, there is $2^0 = 1$ state
- at time 1, there are $2^1 = 2$ states
- at time 2, there are $2^2 = 4$ states.

In state (t,j) the price of the underlying stock is $S_t(j)$ – exactly as before**.**

From this state the price can:

- **go up to $S_{t+1}(2j-1) = S_t(j)u_t(j)$ and state $(t+1,2j-1)$**
- **or down to $S_{t+1}(2j) = S_t(j)d_t(j)$ and state $(t+1,2j)$.**

The $2j$'s just come in because of the way the states are numbered in each column. For example, if we are currently in State 3 (say) then we'll move to State 5 in the next column after an up-movement or State 6 after a down-movement.

Aside

This model, as well as those in the preceding sections, allows share prices to move in a way which is clearly much simpler than reality. However, models are always approximations to reality and the quality of a model can be gauged by how closely answers provided by the model resemble reality.

All models are, of course, by definition simplified characterisations of reality. They attempt to capture the important features of a particular situation in order to help us understand that situation. The strengths, weaknesses and assumptions of the model are as important as the results the model produces, and their explicit consideration often provides additional insight into the situation being modelled.

In this case the binomial model is recognised as an effective model (provided the time to maturity is broken up into a suitable number of sub-periods) for pricing and valuing derivative contracts. In this respect we might describe the binomial model as a good *computational tool*.

4.2 Risk-neutral probability measure

The risk-neutral probabilities

If the risk-free rate of interest, *r*, is constant with $d_t(j) < e^r < u_t(j)$ then this induces the probabilities:

$$q_t(j) = \frac{e^r - d_t(j)}{u_t(j) - d_t(j)}$$

(as in the one-period and two-period models) of an up move from state (t, j).

The probability measure Q

Putting the sample paths and the *q*–probabilities together gives us the probability measure *Q*.

We have noted before that $E_Q(S_1) = S_0 e^r$ giving rise to the use of the name *risk-neutral measure* for *Q*. This follows from the fact that the time 0 expectation of the time 1 stock price with respect to measure *Q*, is equal to the accumulated value of the initial stock price at the risk-free rate.

Similarly in the *n*-period model we have:

$$E_Q\left[S_{t+1} | F_t\right] = S_t e^r$$

Q

Question 13.13

Prove that $E_Q[S_{t+1} | F_t] = S_t e^r$ for $t = 0,1,2,...$

Now applying the tower law, *ie* $E_Q[E_Q\{X|F_t\}|F_s] = E_Q[X|F_s]$ for $s \le t$:

$$E_Q[S_{t+1}|F_0] = E_Q\left[E_Q\left[S_{t+1} \mid F_t\right] \mid F_0\right] = E_Q\left[S_t e^r \mid F_0\right] = e^r E_Q\left[S_t \mid F_0\right]$$

It follows that:

$$\boldsymbol{E_Q\left[S_{t+1}\right] = E_Q\left[S_{t+1} \middle| F_0\right] = S_0 e^{(t+1)r}}$$

by induction. So the use of the expression *risk-neutral measure* for *Q* is still valid – as the expectation with respect to Q of the stock price in $t+1$ periods' time is equal to the initial stock price accumulated at the risk-free rate of return over those $t+1$ periods.

4.3 Finding the derivative price at time 0

We can write $C_n(j)$ for the payoff under a derivative maturing at time *n* in state (n, j) and $V_t(j)$ for the price of the derivative at time *t* in state (t, j). The corresponding random variables are denoted by C_n and V_t respectively. As was shown in the two-period model we calculate prices by starting at the maturity date and working backwards.

So:

$$\boldsymbol{V_n(j) = C_n(j)}$$

and for $t < n$:

$$\boldsymbol{V_t = e^{-r} E_Q\left[V_{t+1} \middle| F_t\right]}$$

So, moving backwards through the $n-t$ stages of the tree from time *n* to time *t* we get:

$$\boldsymbol{= e^{-r(n-t)} E_Q\left[C_n \middle| F_t\right]}$$

As we work our way backwards through the binomial tree by analogy with the one-period model we can also construct a replicating strategy (ϕ_{t+1},ψ_{t+1}) where:

$\phi_{t+1}(j)$ **= number of units of stock when in state (t,j) at time t**

$$= \frac{V_{t+1}(2j-1)-V_{t+1}(2j)}{S_t(j)\,(u_t(j)-d_t(j))}$$

and $\psi_{t+1}(j)$ **= amount held in cash when in state (t,j) at time t**

$$= e^{-r}\left(\frac{V_{t+1}(2j)\,u_t(j)-V_{t+1}(2j-1)\,d_t(j)}{u_t(j)-d_t(j)}\right)$$

$\phi_{t+1}(j)$ and $\psi_{t+1}(j)$ denote the appropriate holdings of shares and cash respectively for the replicating portfolio over the time interval from t to t+1.

Note here that the holding of $\phi_{t+1}(j)$ shares means that the difference between the value of the holding in shares when prices go up and when prices go down precisely matches the difference between the two possible values of the option.

ie $\phi_{t+1}(j)\times\left[S_t(j)\left(u_t(j)-d_t(j)\right)\right] = V_{t+1}(2j-1)-V_{t+1}(2j)$

This means that all risk has been removed.

Question 13.14

Which of the following statements would be true if you wanted to have a replicating portfolio at all times in a general n-period binomial tree?

1. Once you set up the initial replicating portfolio you would leave it unchanged throughout.

2. There is a particular replicating strategy required for each time interval and you would have to adjust your portfolio at each step.

3. A different replicating strategy is required for every node in the tree and you would have to adjust your portfolio at each step depending on which node you are at.

5 Recombining binomial trees

5.1 Ease of computation

The model in the previous subsection is very flexible given that it allows for different levels of volatility when in different states.

Question 13.15

How does the previous model allow for different levels of volatility in different states?

However, the usefulness of the model is severely limited by the number of states which exist even for relatively low numbers of time periods up to maturity (that is, 2^n states), since computation times even for simple derivative securities are at best proportional to the number of states.

One solution to this problem is to assume instead that the volatility is the same in all states, so that the price ratios for the up-steps and down-steps are the same size, irrespective of where they appear in the binomial tree. This may not be an unreasonable assumption given that the steps are expressed in proportionate or percentage rather than absolute terms.

Note that this now means that 2 up-steps and 1 down-step (say) over 3 time periods will take you to the same share price whatever order the steps occur in.

Suppose that we assume the sizes of the up-steps and down-steps are the same in all states.

ie $\quad u_t(j) = u$ **and** $d_t(j) = d$

$\Rightarrow \quad q_t(j) = q$

for all t, j with $d < e^r < u$, and $0 < q < 1$.

Note that the constancy of the risk-neutral probability q, follows directly from that of the step sizes, given that q is a direct function of them:

$$q_t(j) = \frac{e^r - d_t(j)}{u_t(j) - d_t(j)}$$

ie $\quad q_t(j) = \dfrac{e^r - d}{u - d} = q$, constant

Then we have:

$$S_t = S_0 u^{N_t} d^{t-N_t}$$

where N_t is the number of up-steps between time 0 and time *t*. This means that we have $n+1$ possible states at time *n* instead of 2^n. So, in a ten-period model, for example, the number of possible states at time 10 is reduced from 1024 to just 11.

Consequently, computing times are substantially reduced, provided the payoff on the derivative is not path-dependent: that is, it depends upon the number of up-steps and down-steps but not their order. For such non-path-dependent derivatives we have $C_n = f(S_n)$ for some function *f*: for example, for a European call option we have $f(x) = \max\{x-K, 0\}$, where *K* is the strike price.

Question 13.16

What is $f(x)$ for a European put option?

Question 13.17

Would the following derivatives satisfy the non-path-dependent assumption mentioned here?

- Derivative A pays you in one year's time a cash amount equal to the highest value the share price reached during the year.

- Derivative B pays you in one year's time the average of the share price at the start and end of the year.

This special form for the *n*-period model allows us to call it a *recombining* binomial tree or a *binomial lattice* (see Figure 13.3).

The term *recombinant* is also used in this context.

A further implication of this model is that, unlike the non-recombining model discussed in the previous sections, there will usually be more than one route from the initial node to any particular final node.

5.2 Binomial distribution

Under this model the *q*-probabilities are, as stated above, equal, and all steps are made independent of one another. So the number of up steps up to time t, N_t, has a binomial distribution with parameters *t* and *q*. Furthermore, for 0 < *t* < *n*:

- **N_t is independent of $N_n - N_t$** – *ie* the number of up-steps (and hence down-steps) in non-overlapping time intervals is independent
- **and $N_n - N_t$ has a binomial distribution with parameters $n-t$ and *q*.**

The price at time *t* of the derivative is:

$$V_t = e^{-r(n-t)} \sum_{k=0}^{n-t} f\left(S_t u^k d^{n-t-k}\right) \frac{(n-t)!}{k!(n-t-k)!} q^k (1-q)^{n-t-k}$$

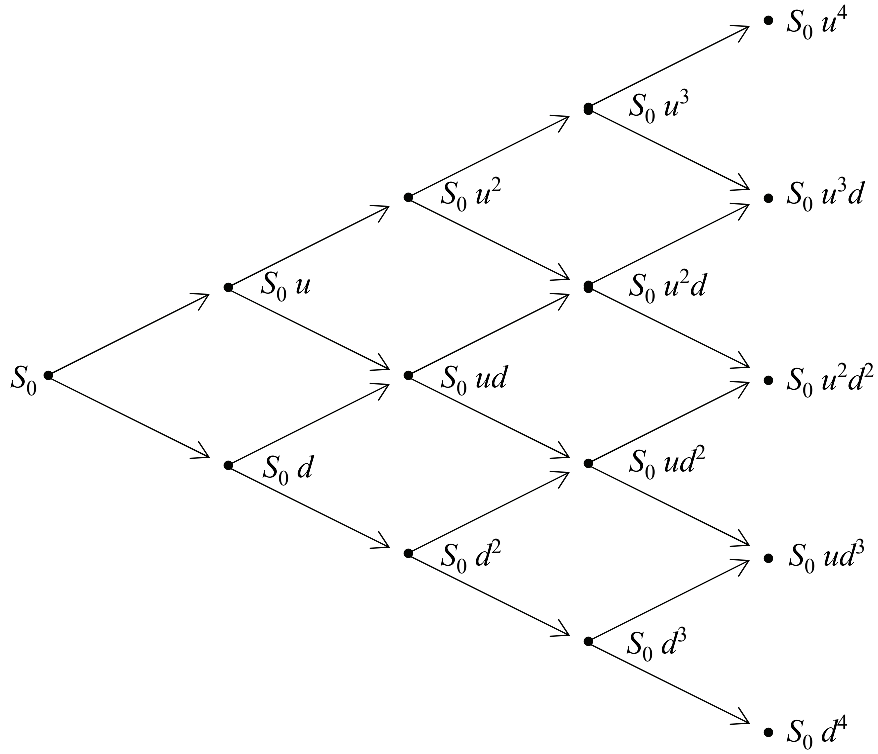


Figure 13.3: Recombining binomial tree or binomial lattice

This figure is part of the Core Reading.

Question 13.18

Derive the above expression for V_t.

Question 13.19

Consider a binomial lattice model for a 2-month call option with an exercise price of 200. Suppose that the share price either goes up by 4% or down by 3% each month, that the risk-free continuously-compounded rate is ½% per month and that the current share price is also 200.

Use the formula above to estimate the value of the option.

6 Calibrating binomial models

It is often convenient when calibrating the binomial model to have the mean and variance implied by the binomial model correspond to the mean and variance of a lognormal distribution. The reasoning will become clearer when considering continuous-time versions in later chapters.

For recombining binomial models, an additional condition that leads to a unique solution is:

$$u = \frac{1}{d}$$

ie, an up-step and a down-step would mean the share price, after two steps, is the same as it is at time 0.

Recall from Chapter 9, Section 1.3, the solution to the SDE for geometric Brownian motion. Here, we showed that if $dS_t = \alpha S_t \, dt + \sigma S_t \, dB_t$, then S_t / S_0 has a lognormal distribution with parameters $\left(\alpha - \frac{1}{2}\sigma^2\right) t$ and $\sigma^2 t$.

Question 13.20

If S_t / S_0 has a lognormal distribution with parameters $\left(\alpha - \frac{1}{2}\sigma^2\right) t$ and $\sigma^2 t$, then give formulae for the mean and variance of S_t / S_0.

If we parameterise the lognormal distribution (under the risk-neutral law) so that:

$$\ln\{S(t)/S(t_0)\} \sim N\left[\left(r - \frac{\sigma^2}{2}\right)(t - t_0), \sigma^2 (t - t_0)\right]$$

then the conditions that must be met are:

$$E\left[S(t+\delta t)/S(t)\right] = \exp(r\delta t) \qquad \textbf{(1)}$$

and $$\text{var}\left[\ln\{S(t+\delta t)/S(t)\}\right] = \sigma^2 \delta t \qquad \textbf{(2)}$$

where:

- **δt is the time interval of each step in the binomial model**
- **$S(t)$ denotes the price of the asset at time t.**

Noting that:

$$\boldsymbol{E\left[S(t+\delta t)/S(t)\right]=qu+(1-q)d}$$

it follows from equation (1) that:

$$qu+(1-q)d=\exp(r\delta t)$$

Rearranging to make q the subject, we get:

$$\boldsymbol{q=\frac{\exp(r\delta t)-d}{u-d}}$$

Using the equation (2) and the assumption that $\boldsymbol{u=\frac{1}{d}}$:

$$\boldsymbol{\begin{aligned}\operatorname{var}\left[\ln\{S(t+\delta t)/S(t)\}\right]&=q(\ln u)^2+(1-q)(-\ln u)^2-E\left[\ln\{S(t+\delta t)/S(t)\}\right]^2\\&=(\ln u)^2-E\left[\ln\{S(t+\delta t)/S(t)\}\right]^2\end{aligned}}$$

The last term involves terms of higher order than $\boldsymbol{\delta t}$.

ie $\quad E\left[\ln\{S(t+\delta t)/S(t)\}\right]^2=f\left\{(\delta t)^2\right\}$

which tends to 0 as $\delta t\to 0$.

So, if we ignore the $\boldsymbol{E\left[\ln\{S(t+\delta t)/S(t)\}\right]^2}$ terms and equate the expression to $\boldsymbol{\sigma^2\delta t}$, we get:

$$\operatorname{var}\left[\ln\{S(t+\delta t)/S(t)\}\right]=(\ln u)^2=\sigma^2\delta t$$

Then, we solve to get:

$$\boldsymbol{u=\exp\left(\sigma\sqrt{\delta t}\right)}$$

and hence also:

$$\boldsymbol{d=\exp\left(-\sigma\sqrt{\delta t}\right)}$$

When a (continuously payable) dividend is paid on the underlying asset, it is convenient and conventional to adjust the up-steps and down-steps to be:

$$\boldsymbol{u} = \mathbf{exp}\left(\boldsymbol{\sigma}\sqrt{\boldsymbol{\delta t}} + \boldsymbol{v\delta t}\right)$$

$$\boldsymbol{d} = \mathbf{exp}\left(-\boldsymbol{\sigma}\sqrt{\boldsymbol{\delta t}} + \boldsymbol{v\delta t}\right)$$

where v is the continuously payable dividend rate.

These formulae can be found on page 45 of the *Tables* (using the letter q instead of v). In some CT8 exam questions, explicit values for u and d are not given. In this case, you should assume that $ud = 1$ and use these formulae to calculate u and d.

Question 13.21

A non-dividend paying share has volatility $\sigma = 20\%\,pa$. Calculate the values of u and d for the share price movements over one month.

7 The state price deflator approach

In this section we will present a different, but equivalent, approach to pricing.

7.1 1-period case

Recall the one-period binomial model where:

$$V_1 = \begin{cases} c_u & \text{if } S_1 = S_0 u \\ c_d & \text{if } S_1 = S_0 d \end{cases}$$

Then:

$$\begin{aligned} V_0 &= e^{-r} E_Q[V_1] \\ &= e^{-r}\left[q c_u + (1-q) c_d \right] \end{aligned}$$

So far we have been pricing derivatives on a risk-neutral basis. Here the fair price for the derivative is the discounted value of:

"probability" share price goes up × derivative value if share price goes up

+

"probability" share price goes down × derivative value if share price goes down

We can re-express this value in terms of the real-world probability p :

$$\begin{aligned} V_0 &= e^{-r}\left[p \frac{q}{p} c_u + (1-p) \frac{(1-q)}{(1-p)} c_d \right] \\ &= E_P[A_1 V_1] \end{aligned}$$

Where A_1 is a random variable with:

$$A_1 = \begin{cases} e^{-r} \dfrac{q}{p} & \text{if } S_1 = S_0 u \\ e^{-r} \dfrac{1-q}{1-p} & \text{if } S_1 = S_0 d \end{cases}$$

The expression for value of the derivative takes the same form as before. It is the discounted value of:

"probability" share price goes up × derivative value if share price goes up

+

"probability" share price goes down × derivative value if share price goes down

However, in this case we have real-world probabilities and a different discount factor. The discount factor A_1 depends on whether the share price goes up or goes down. This means it is random and so we call it a *stochastic discount factor*.

A_1 is called a *state price deflator*. It also has a variety of other names:

- **deflator**
- **state price density**
- **pricing kernel**
- **stochastic discount factor** (as mentioned above)**.**

Question 13.22

Explain why the state price deflator will take a higher value if the share price goes down.

Question 13.23

Let the value of a share at time 0 be $S_0 = 100$ and the continuously-compounded risk-free rate be 3% per time period. In one time period's time the share price will either have gone up to 120 or down to 85. The probability that the share price goes up is 0.6. Calculate the possible values of the state price deflator A_1.

Note that:

(a) if $V_1 = 1$ then:

$$V_0 = E_P[A_1 \times 1] = e^{-r}$$

This is not surprising. If a derivative pays 1 in one period's time regardless of the outcome of the share price then, by no-arbitrage, the fair price to pay for this derivative is just the discounted value of 1.

(b) if $V_1 = S_1$ then:

$$V_0 = E_P[A_1 \times S_1] = S_0$$

Again, this is not surprising. If a derivative pays the value of the share in one period's time regardless of what this turns out to be then, by no-arbitrage, the fair price to pay for this derivative is the value of the share now.

7.2 *n-period case, binomial lattice*

We now extend the theory to the *n*-period case.

Recall that under the risk-neutral approach to pricing we have:

$$V_n = f\left(S_n\right)$$

Now:

$$S_n = S_0 u^i d^{n-i}$$

where i is the number of up steps.

Over n steps the share price will go up i times and go down $n-i$ times. Each time it goes up, its price is multiplied by u and each time it goes down its price is multiplied by d.

So let us define:

$$V_n\left(i\right) = f\left(S_0 u^i d^{n-i}\right)$$

Then we have:

$$V_0 = e^{-rn} E_Q[V_n]$$

$$= e^{-rn} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} q^k (1-q)^{n-k} f(S_0 u^k d^{n-k})$$

Although the algebra is heavier, this is the same form as before and is an expression in terms of the risk-neutral probability q. In the same way as the 1-period case, we now re-express this in terms of the real-world probability p.

$$V_0 = e^{-rn} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \left(\frac{q}{p}\right)^k \left(\frac{1-q}{1-p}\right)^{n-k} V_n(k)$$

$$= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} A_n(k)\, V_n(k)$$

$$= E_P[A_n V_n]$$

where $A_n = e^{-rn} \left(\frac{q}{p}\right)^{N_n} \left(\frac{1-q}{1-p}\right)^{n-N_n}$

and N_n is the number of up steps up to time n.

So, again, the discount factor A_n is random and so we call it a stochastic discount factor.

A_n is again called the state price deflator. An important property of A_n is that for all n = 1, 2, … we have:

$$A_n = A_{n-1} \times e^{-r} \left(\frac{q}{p}\right)^{I_n} \left(\frac{1-q}{1-p}\right)^{1-I_n}$$

where $I_n = \begin{cases} 1 & \text{if } S_n = S_{n-1} u \\ 0 & \text{if } S_n = S_{n-1} d \end{cases}$

It follows that:

- $S_n = S_{n-1} u^{I_n} d^{1-I_n}$ **and**
- $N_n = \sum_{k=0}^{n} I_k$

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_0. Theoretically, they are the same; they only differ in the way that they present the calculation of a derivative price.

One presents it using normal discount factors and the risk-neutral probability q and the other presents it using stochastic (random) discount factors and the real-world probability p.

Finally, note that:

$$E_P[A_n] = e^{-rn}$$

If a derivative pays 1 in n periods' time regardless of the outcome of the share price then, by no-arbitrage, the fair price to pay for this derivative is just the discounted value of 1.

We also have:

$$E_P[A_n S_n] = S_0$$

If a derivative pays the value of the share in n periods' time regardless of what this turns out to be then, by no-arbitrage, the fair price to pay for this derivative is the value of the share now.

The state price deflator approach can be adapted to price a derivative at any time t and it is straightforward to show that:

$$V_t = \frac{E_P[A_T V_T]}{A_t}$$

In words, this formula above is saying that the fair price for the derivative is the discounted value of the expected value of the derivative at time T. But because we are taking that expectation with respect to the real-world probability P, the discount factor we use is the state price deflator A.

$\dfrac{A_T}{A_t}$ is analogous to the deterministic discount factor $\dfrac{v^T}{v^t} = v^{T-t}$, *ie* the "present value" at time t of a payment at time T.

8 Exam-style questions

We finish this chapter with two exam-style questions.

Question 1

The market price of a security can be modelled by assuming that it will either increase by 12% or decrease by 15% each month, independently of the price movement in other months. No dividends are payable during the next two months. The continuously-compounded monthly risk-free rate of interest is 1%. The current market price of the security is 127.

(i) Use the binomial model to calculate the value of a two-month European put option on the security with a strike price of 125. [3]

(ii) Calculate the value of a two-month American put option on the same security with the same strike price. [3]

(iii) Calculate the value of a two-month American call option on the same security with the same strike price. [2]

[Total 8]

Solution 1

(i) ***Calculate the value of the European put option***

We are given $u = 1.12$, $d = 0.85$ and $r = 0.01$ in the question. The risk-neutral probability under the binomial model is:

$$q = \frac{e^r - d}{u - d} = \frac{e^{0.01} - 0.85}{1.12 - 0.85} = 0.5928$$

The possible values for the security price after two months are:

$$S_{uu} = 1.12^2 \times 127 = 159.3088$$

$$S_{ud} = S_{du} = 1.12 \times 0.85 \times 127 = 120.904$$

$$S_{dd} = 0.85^2 \times 127 = 91.7575$$

Therefore the possible payoffs of the put option after two months are:

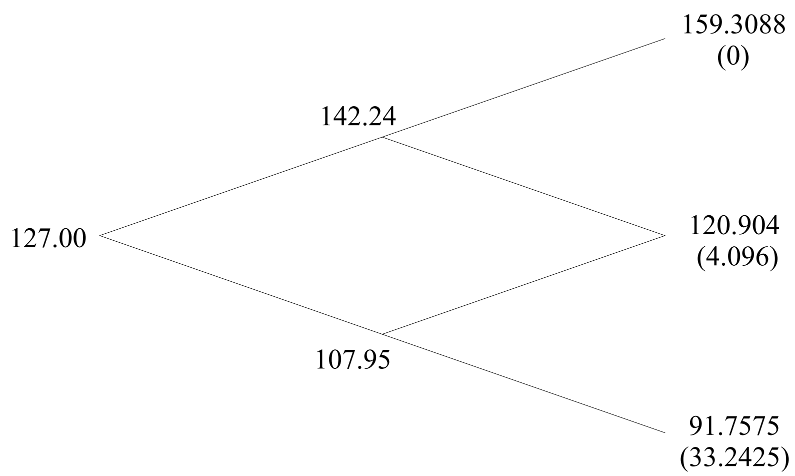
$$p_{uu} = 0$$

$$p_{ud} = 125 - 120.904 = 4.096$$

$$p_{dd} = 125 - 91.7575 = 33.2425$$

Because the up and down steps are the same over each month, this binomial tree is recombining.

The binomial tree of security prices looks like this, with the final put payoffs in italics.



We can discount the expected value of the payoff under the risk-neutral probability to find the initial value:

$$e^{-2r}E_Q[P_2]=e^{-2r}\left(q^2p_{uu}+2q(1-q)p_{ud}+(1-q)^2p_{dd}\right)$$

$$=e^{-0.02}\left(0+2\times 0.5928\times 0.4072\times 4.096+0.4072^2\times 33.2425\right)$$

$$=7.342$$

(ii) ***Calculate the value of the American put option***

The only difference from part (i) is that with an American option, it may be beneficial to exercise the option early, after one month.

If the security price rises over the first month then there would be no point in exercising the put option because the payoff would be zero. However, if the security price falls during the first month, the payoff from exercising the option early is:

$$125-127\times 0.85=17.05$$

Using the risk-neutral probability we can calculate the value of holding on to the put option:

$$V_1(2)=e^{-0.01}\{q\times 4.096+(1-q)\times 33.2425\}=15.8062$$

Since this is less than 17.05, if the security price falls, it would be advantageous to exercise after one month. The extra value at time 0 of having this option after one month is:

$$e^{-0.01}(1-q)\times(17.05-15.8062)=0.5015$$

Finally the value of the American option is the value of the European option plus the extra value of the option to exercise after one month:

$$V_0=7.342+0.5015=7.843$$

(iii) ***Calculate the value of the American call option***

Because you would never exercise an American call option early on a non-dividend-paying security, the value of the American call option is the same as that of a European call option.

Using put-call parity we can derive this value from the value of the European put option in part (i):

$$c_0 = p_0 + S_0 - Ke^{-r(T-t)} = 7.342 + 127 - 125e^{-0.01\times 2} = 11.817$$

Alternative methods are valid, eg:

$$c_0 = e^{-0.02}q^2 \times (159.3088 - 125) = 11.817$$

Question 2

A company share price is to be modelled using a 5-step recombining binomial tree, with each step in the tree representing one day. Each day, it is assumed that the share price:

- increases by 2%, or
- decreases by 1%

You may assume that the force of interest is $\delta = 5.5\%$ pa and that there are 365 days in a year. No dividends are to be paid over the next five days.

(i) Calculate the risk-neutral probability of an up-step on any given day. [2]

(ii) Calculate the fair price of a 5-day at-the-money call option on £10,000 worth of shares in this company. [5]

A special option is available where the payoff after 5 days is:

$$\max\left\{S_5^* - K, 0\right\}$$

where S_5^* is the arithmetic average share price recorded at the end of each of the 5 days and K is the strike price.

(iii) Calculate the fair price of a 5-day special option (strike price $K = 1.06 S_0$) on £10,000 worth of shares in this company. [4]

(iv) Explain whether an at-the-money special option is likely to have a higher value of vega than a standard "vanilla" call option. [3]

[Total 14]

Solution 2

(i) ***Risk-neutral probability***

In the risk-neutral world, the expected return on the share is equal to the return on the risk-free asset:

$$E_Q\left[S_1\right] = qS_0 \times 1.02 + (1-q)S_0 \times 0.99 = S_0 e^{0.055/365}$$

Rearranging for q, we get:

$$q = \frac{e^{0.055/365} - 0.99}{1.02 - 0.99} = 0.338357$$

Alternatively, you could use the formula for the up-step probability given on page 45 of the Tables.

(ii) ***Call option value***

For simplicity, we can assume that we are valuing one call option on a share worth £10,000. We can use the up and down steps to calculate the six possible final share prices in this binomial tree:

$$S_5 = £10,000 \times 1.02^5 \times 0.99^0 = £11,040.81 \quad \text{for 5 up jumps}$$
$$S_5 = £10,000 \times 1.02^4 \times 0.99^1 = £10,716.08 \quad \text{for 4 up jumps}$$
$$S_5 = £10,000 \times 1.02^3 \times 0.99^2 = £10,400.90 \quad \text{for 3 up jumps}$$
$$S_5 = £10,000 \times 1.02^2 \times 0.99^3 = £10,094.99 \quad \text{for 2 up jumps}$$
$$S_5 = £10,000 \times 1.02^1 \times 0.99^4 = £9,798.08 \quad \text{for 1 up jump}$$
$$S_5 = £10,000 \times 1.02^0 \times 0.99^5 = £9,509.90 \quad \text{for 0 up jumps}$$

So, the payoffs from a call option, with strike price $K = £10,000$ are:

£1,040.81	for 5 up jumps
£716.08	for 4 up jumps
£400.90	for 3 up jumps
£94.99	for 2 up jumps
£0	for 1 up jump
£0	for 0 up jumps

So, the fair price of the call option is:

$$V_0 = e^{-5\times 0.055/365}\begin{Bmatrix} 1,041.81q^5 + 716.08\times 5q^4(1-q) \\ +400.90\times 10q^3(1-q)^2 + 94.99\times 10q^2(1-q)^3 \end{Bmatrix}$$

Substituting in the value of q, this is:

$$\begin{aligned} V_0 &= 0.999247\{4.62 + 31.05 + 67.98 + 31.50\} \\ &= £135.05 \end{aligned}$$

(iii) ***Special option value***

Again, for simplicity, we can assume that we are valuing one special option on a share worth £10,000. By investigation we find that the average share price only exceeds £10,600 if the share price goes up on all 5 days. If we try 4 up jumps and 1 down jump, we get:

$$\frac{1.02 + 1.02^2 + 1.02^3 + 1.02^4 + 1.02^4\times 0.99}{5} = 1.055 < 1.06$$

The average share price over the 5 days, if it goes up every day, is:

$$£10,000\times\frac{1.02 + 1.02^2 + 1.02^3 + 1.02^4 + 1.02^5}{5} = £10,616.24$$

The payoff of the special option in this case is:

$$\max\{10,616.24 - 10,600, 0\} = £16.24$$

So, the fair price of the special option is:

$$\begin{aligned} V_0 &= e^{-5\times 0.055/365}\times 16.24q^5 \\ &= £0.07 \end{aligned}$$

(iv) ***Value of vega***

Vega is the rate of change of an option value f with respect to the volatility σ of the underlying asset, *ie* the shares in the company:

$$\nu = \frac{\partial f}{\partial \sigma}$$

The special option's payoff is dependent on the average share price, which is a much smoother process than the current share price. So, because the option's final payoff is less dependent on the current share price, the option's value will vary less with changes in the volatility of the share price.

So, the special option will have a lower value of vega.

This observation is reinforced by noting that the payoff from the vanilla call option in part (ii) with strike £10,600 lies in the range $(0, 440.81)$ *according to this model, whereas the payoff for the special option lies in the much narrower range* $(0,16.24)$.

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

Chapter 13 Summary

Binomial model assumptions

- Assets may be bought and sold at integer times $t = 0, 1, 2, 3, \ldots$
- Assets may be held in any amount.
- There are no taxes or transaction costs.
- There are no arbitrage opportunities $\Leftrightarrow d < e^r < u$.

One-period model: replicating portfolio and risk-neutral valuation

In the one-period model:

$$\begin{aligned} V_0 &= \phi S_0 + \psi \\ &= e^{-r}\left(q\,c_u + (1-q)\,c_d\right) \\ &= e^{-r} E_Q(C_1) \end{aligned}$$

where:

$$\phi = \frac{c_u - c_d}{S_0(u-d)} \qquad \psi = e^{-r}\left(\frac{c_d u - c_u d}{u-d}\right)$$

$$q = \frac{e^r - d}{u-d} \qquad 1-q = \frac{u-e^r}{u-d}$$

The portfolio consisting of ϕ shares and ψ cash is a *replicating portfolio*.

Q is the *risk-neutral probability measure*, which gives the risk-neutral probability q to an upward move in prices and $1-q$ to a downward move.

The risk-neutral probabilities ensure that the underlying security yields an expected return equal to the risk-free rate.

Finding the derivative price using the risk-neutral probabilities is referred to as *risk-neutral valuation*.

Two-period model: risk-neutral valuation

Here the value of the derivative is:

$$V_0 = e^{-2r} E_Q[C_2 | F_0]$$

So, the no-arbitrage value of the derivative at time 0, is equal to:

- the time 0 expectation of the payoff paid at time 2
- calculated with respect to the risk-neutral probability measure Q and
- the information set (sigma-algebra) F_0 generated by the history of the stock price up to and including time 0 and
- discounted at the continuously-compounded risk-free rate of return.

n-period model

The results of the one-period and two-period binomial models generalise to the multi-period or n-period context. In this case:

- the risk-neutral probability from State (t, j) is:

$$q_t(j) = \frac{e^r - d_t(j)}{u_t(j) - d_t(j)}$$

- the expectation of the stock price in n periods' time calculated with respect to the risk-neutral measure Q, is equal to the current stock price, accumulated at the continuously-compounded risk-free rate of return over those m periods, *ie* $E_Q[S_{t+n} | F_t] = S_t e^{rn}$

- the derivative price in State (t, j) is $V_t = e^{-r(n-t)} E_Q[C_n | F_t]$

- the number of units of stock when in State (t, j) at time t is:

$$\phi_{t+1}(j) = \frac{V_{t+1}(2j-1) - V_{t+1}(2j)}{S_t(j)\big(u_t(j) - d_t(j)\big)}$$

- the amount held in cash when in State (t, j) at time t is:

$$\psi_{t+1}(j) = e^{-r}\left(\frac{V_{t+1}(2j)\,u_t(j) - V_{t+1}(2j-1)\,d_t(j)}{u_t(j) - d_t(j)}\right)$$

Recombining binomial trees

A recombining binomial tree (or binomial lattice) is one in which values of u and d, and consequently the risk-neutral probabilities, are the same in all states.

With such models:

- the volume of computation required is greatly reduced
- N_t, the number of up-steps up to time t, has a binomial distribution with parameters t and q
- $S_t = S_0 u^{N_t} d^{t-N_t}$

Calibrating binomial models

If we assume that $ud = 1$, then a binomial tree model with steps of length δt can be calibrated to have the same mean and variance as a continuous-time model. In this instance, we need:

$$q = \frac{\exp(r\delta t) - d}{u - d}$$

$$u = \exp\left(\sigma\sqrt{\delta t} + \nu\delta t\right)$$

$$d = \exp\left(-\sigma\sqrt{\delta t} + \nu\delta t\right)$$

The state price deflator approach in the one-period binomial tree

The state price deflator is:

$$A_1 = \begin{cases} e^{-r}\dfrac{q}{p} & \text{if } S_1 = S_0 u \\ e^{-r}\dfrac{1-q}{1-p} & \text{if } S_1 = S_0 d \end{cases}$$

Then the fair price for the derivative is:

$$V_0 = E_P[A_1 V_1]$$

The state price deflator approach in the n-period binomial tree

The state price deflator is:

$$A_n = e^{-rn}\left(\frac{q}{p}\right)^{N_n}\left(\frac{1-q}{1-p}\right)^{n-N_n}$$

where N_n is the number of up-steps up to time n. The fair price for the derivative is then:

$$V_0 = E_P[A_n V_n]$$

A *state price security* is a contract that agrees to pay one unit of currency if a particular state occurs at a particular time in the future and zero in all other states.

Chapter 13 Solutions

Solution 13.1

The cash account at time 1 is equal to $B_1 = e^{r\times 1} = e^r$

Solution 13.2

$$V_0 = \phi \times S_0 + \psi \times 1$$

ie $$V_0 = \phi S_0 + \psi$$

Solution 13.3

C_1 is the derivative payoff at time 1, which is equal to:

$$C_1 = \begin{cases} c_u & \text{if the stock price went up (to } S_0u) \\ c_d & \text{if the stock price went down (to } S_0d) \end{cases}$$

Solution 13.4

$$E_Q[S_1] = qS_0u + (1-q)S_0d$$

$$= \frac{e^r - d}{u-d} S_0u + \frac{u - e^r}{u-d} S_0d$$

$$= \frac{e^r u - ud + ud - de^r}{u-d} S_0$$

$$= \frac{e^r(u-d)}{u-d} S_0 = S_0 e^r$$

So, under the probability measure Q, we see that the expected rate of return on the share is simply the risk-free rate.

In a world in which all investors were risk-neutral, all assets would offer the same expected rate of return (as investors would care only about expected return and would be indifferent to risk.) The probability measure Q is the set of probabilities such that all risky assets do give the same expected return, namely the risk-free rate. It is for this reason that Q is sometimes referred to as the *risk-neutral probability measure*.

Solution 13.5

Under normal circumstances investors demand higher expected returns in return for accepting the risk inherent in the random stock price. So, the expected return calculated with respect to the real-world probability measure P should be higher than that based on the risk-neutral probability measure Q, so as to compensate investors for the additional risk inherent in the stock. In order to generate this additional expected return, the expected stock price at time 1 must be higher under P than under Q, *ie*:

$$E_P(S_1) > E_Q(S_1)$$

$$\Leftrightarrow \quad S_0\left(pu+(1-p)d\right) > S_0\left(qu+(1-q)d\right)$$

$$\Leftrightarrow \quad pu-pd > qu-qd$$

$$\Leftrightarrow \quad \left(p-q\right)(u-d) > 0$$

Given that $u-d>0$ by assumption, we would therefore normally find that $p>q$.

Solution 13.6

$c_d = 0$ because if the stock price goes down over the single period under consideration, then it ends up at 30. As this is less than the exercise price of 45, the call option will not be exercised and will expire worthless. Hence, the option payoff is then equal to 0.

Solution 13.7

Recall from above that:

$$\phi = \frac{c_u - c_d}{S_0(u-d)}$$

and $$\psi = e^{-r}\left(\frac{c_d u - c_u d}{u-d}\right)$$

So, plugging in the relevant values gives:

$$\phi = \frac{15-0}{40(1.5-0.75)} = 0.5$$

$$\psi = e^{-0.05}\left(\frac{0\times 1.5-15\times 0.75}{1.5-0.75}\right) = -14.268$$

Hence, the replicating portfolio consists of:

- a positive holding of half a share
- a negative holding of 14.268 of cash – *ie* we borrow cash.

Alternatively, you could have solved the simultaneous equations:

$$60\phi+\psi e^{0.05}=15 \text{ and } 30\phi+\psi e^{0.05}=0$$

Finally, the cost of the replicating portfolio is equal to:

$$V_0 = \phi S_0+\psi$$

and plugging in the relevant values from above gives:

$$V_0 = 0.5\times 40-14.268 = 5.732$$

as required.

Solution 13.8

This is the same as the restriction described in Section 2.1. Suppose that this is not the case, that we are at time 1 in State $(1, j)$ and that $e^r < d_1(j)$ (and $u_1(j)$), say. Then we can buy a share and borrow the amount of cash needed to pay for this. At time 1 this would have a net cost of £0. At time 2 our portfolio will be worth either $d_1(j)-e^r$ or $u_1(j)-e^r$ both of which are greater than 0 – according to our assumption. So, we have a violation of the no-arbitrage condition.

Solution 13.9

The expression is:

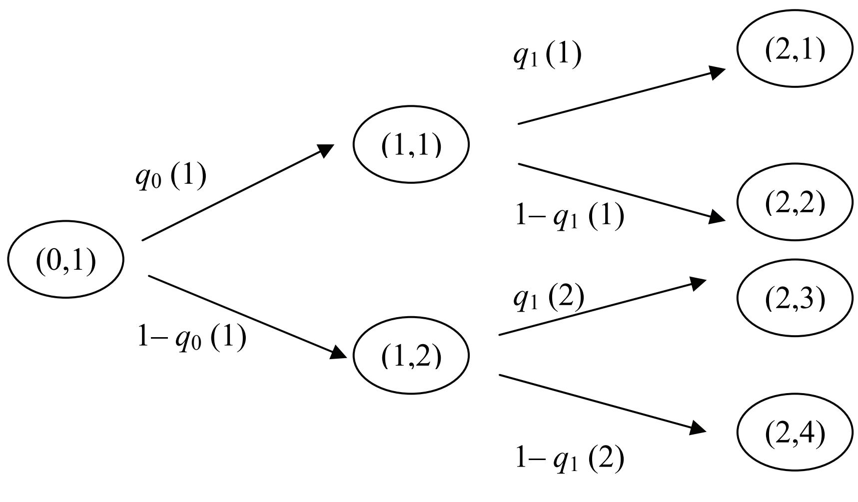
$$V_1 = e^{-r} E_Q[C_2 | F_1]$$

It says that:

- the value of the derivative at time 1 is equal to the expectation of the derivative payoff at time 2
- calculated with respect to both the risk-neutral probability measure Q and the information set F_1, generated by the history of the stock price movements up to and including time 1
- discounted at the continuously-compounded risk-free rate of return r.

Solution 13.10

States and risk-neutral probabilities



Solution 13.11

It is the same because the u and d price movement factors happen to be the same. This will not always be the case. Note also that the risk-free rate of return has not changed either. In the general case, this could be different for each period.

Solution 13.12

(i) ***Risk-neutral probability***

For the first time period, the price movement factors are $u_0(1) = 3/2$ and $d_0(1) = 3/4$.

So, the risk-neutral probability is:

$$q_0(1) = \frac{e^r - d_0(1)}{u_0(1) - d_0(1)} = \frac{e^{0.05} - 3/4}{3/2 - 3/4} = 0.40169$$

which is, of course, the same as in the one-period model in the previous section.

(ii) ***Derivative price at time 0***

Starting from State (0,1), the possible derivative values at time 1 are $V_1(1) = 17.195$ and $V_1(2) = 0$.

Hence:

$$V_0 = e^{-r}\left(q_0(1)V_1(1) + (1 - q_0(1))V_1(2)\right)$$

ie $$V_0 = e^{-0.05}\left(0.40169 \times 17.195 + (1 - 0.40169) \times 0\right) = 6.570$$

Solution 13.13

We know that $q = \dfrac{e^r - d}{u - d}$. It follows that:

$$E_Q\left[S_{t+1} \mid F_t\right] = qS_t u + (1 - q)S_t d = S_t\left[\frac{e^r - d}{u - d}(u - d) + d\right] = S_t e^r$$

Solution 13.14

Statement 3 would be true. You would usually have to rebalance your replicating portfolio in a different way at each node you visit.

Solution 13.15

The previous model allows for different levels of volatility in different states by allowing for different up and down price factors in different states, *ie* $u_t(j)$ and $d_t(j)$ vary with t and j.

Solution 13.16

For a European put option, $f(x) = \max\{K - x, 0\}$ where K is the strike price.

Solution 13.17

Derivative A is path-dependent since the highest value would be different if, for example, we had the sequences $uudd$ and $dduu$.

Derivative B is not path-dependent, since an average calculated based only on the initial and final values does not depend on the particular path taken in between.

Solution 13.18

The general expression for the value of the derivative contract at time t is (irrespective of whether the tree is recombining or not):

$$\begin{aligned} V_t &= e^{-r(n-t)}\, E_Q[C_n | F_t] \\ &= e^{-r(n-t)}\, E_Q[f(S_n) | F_t] \\ &= e^{-r(n-t)} \sum_s f(s) \times P_Q[S_n = s | F_t] \end{aligned}$$

where the summation is over all the possible values that S_n can take and the probabilities are the risk-neutral probabilities. If the tree *is* recombining then:

$$V_t = e^{-r(n-t)} \sum_{k=0}^{n-t} f(S_t u^k d^{n-t-k}) \times P_Q[S_n | F_t]$$

or $$V_t = e^{-r(n-t)} \sum_{k=0}^{n-t} f\left(S_t u^k d^{n-t-k}\right) \frac{(n-t)!}{k!(n-t-k)!} q^k (1-q)^{n-t-k}$$

Solution 13.19

Here:

$$u = 1.04$$

$$d = 0.97$$

$$r = 0.005$$

$$n-t = 2$$

So, the risk-neutral probability is equal to:

$$q = \frac{e^r - d}{u-d} = \frac{e^{0.005} - 0.97}{1.04 - 0.97} = 0.500179$$

Hence:

Share price at exercise	Call option payoff	Probability of payoff
$200 \times 1.04^2 = 216.32$	16.32	$q^2 = 0.250179$
$200 \times 1.04 \times 0.97 = 201.76$	1.76	$2q(1-q) = 0.500000$
$200 \times 0.97^2 = 188.18$	0	$(1-q)^2 = 0.249821$

Finally, the current call option price is therefore:

$$V_0 = e^{-0.005 \times 2}\left[16.32 \times 0.250179 + 1.76 \times 0.500000 + 0 \times 0.249821\right]$$

$$= 4.914$$

Solution 13.20

The properties of the lognormal distribution give us the expectation and variance of S_t / S_0:

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

Solution 13.21

Applying the formula for *u*, we get:

$$u = \exp\left(\sigma\sqrt{\delta t} + \nu\delta t\right)$$

$$= \exp\left(0.2\sqrt{1/12} + 0/12\right)$$

$$= 1.0594$$

When dividends are zero, we have:

$$d = \frac{1}{u} = 0.9439$$

Solution 13.22

If the share price goes up then the state price deflator A_1 is:

$$e^{-r}\frac{q}{p}$$

whereas if the share price goes down A_1 is:

$$e^{-r}\frac{1-q}{1-p}$$

So if $q < p$ then A_1 will take a higher value if the share price goes down.

Recall that we derived q on a risk-neutral basis and that under Q, statistically speaking, we expect the share to "act like cash". In real life we don't expect a share to act like cash because it is a riskier investment. Because, in real life, investors want to be compensated for extra risk with higher expected return, it must be the case that, under P, the expected return on the share is higher.

Therefore we have that:

$$q < p$$

and the state price deflator A_1 will take a higher value if the share price goes down.

Solution 13.23

We can calculate the risk-neutral probability q in the same way as earlier in the chapter:

$$u = \frac{120}{100} = 1.2$$

$$d = \frac{85}{100} = 0.85$$

$$q = \frac{e^r - d}{u - d} = \frac{e^{0.03} - 0.85}{1.2 - 0.85} = 0.51558$$

We can then use the formula for the state price deflator A_1 and the real-world probability given ($p = 0.6$) to calculate the value of A_1:

$$A_1 = \begin{cases} e^{-r}\dfrac{q}{p} & \textit{if } S_1 = S_0 u \\ e^{-r}\dfrac{1-q}{1-p} & \textit{if } S_1 = S_0 d \end{cases}$$

$$= \begin{cases} e^{-0.03}\dfrac{0.51558}{0.6} & \textit{if } S_1 = 120 \\ e^{-0.03}\dfrac{0.48442}{0.4} & \textit{if } S_1 = 85 \end{cases}$$

$$= \begin{cases} 0.8339 & \textit{if } S_1 = 120 \\ 1.1752 & \textit{if } S_1 = 85 \end{cases}$$

Chapter 14

The Black-Scholes option pricing formula

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)*

9. Show how to use the Black-Scholes model in valuing options and solve simple examples.

10. Discuss the validity of the assumptions underlying the Black-Scholes model.

0 Introduction

In the previous chapter we developed a simple binomial model to price derivatives in discrete time with a discrete state space. The assumed share price process was a geometric random walk. We now extend our analysis to continuous time and a continuous state space, in an attempt to more realistically model the processes driving stock prices. Since a random walk becomes Brownian motion in the continuous-time limit, we assume our share price follows a geometric Brownian motion or lognormal model, as described in Chapters 9 and 10.

We will discuss two methods of pricing derivatives in the continuous-time case – these are basically alternative proofs of the same result, the *Black-Scholes* pricing formula. Both methods can be described as no-arbitrage approaches. In this chapter we focus primarily on the *risk-free* construction or PDE (partial differential equation) approach. The proof allows us to set up a partial differential equation (PDE) for the value of the derivative. It can then be shown that the Garman-Kohlhagen formula is a solution of this equation, and also satisfies the boundary conditions for a call or put option. In Part 4 we turn our attention to the development of a derivative pricing approach based on the use of *replicating strategies,* as in the discrete-time case.

Section 2.4 sets up the PDE by applying Ito's lemma, which you will recall from Chapter 9. The boundary conditions are also given. In principle, it is straightforward to differentiate the Garman-Kohlhagen formula and show that it does satisfy the PDE and boundary conditions. However, to do so we need to calculate "the Greeks". As you are aware from Chapter 12, these Greeks are not only of interest for solving the PDE, but also have other applications. We will return to our intuitive interpretation of delta, gamma and theta within the context of the Black-Scholes PDE.

The Black-Scholes analysis of option prices is underpinned by a number of key assumptions. We discuss these first in Section 1 and consider how realistic they are in practice. Even though the assumptions do not all hold in practice, this does not prevent the Black-Scholes model providing a good approximation to reality. The approach offers valuable insight into option pricing and is widely used in practice.

1 The assumptions underlying the Black-Scholes model

1.1 The assumptions

The assumptions underlying the Black-Scholes model are as follows:

1. **The price of the underlying share follows a geometric Brownian motion.**

   ie the share price changes *continuously* through time according to the stochastic differential equation:

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

   Recall that this is the same as the lognormal model discussed in Chapter 10 and therefore has the same properties.

2. **There are no risk-free arbitrage opportunities.**

3. **The risk-free rate of interest is constant, the same for all maturities and the same for borrowing or lending.**

   In fact, this simplifying assumption is not critical and can be relaxed.

4. **Unlimited short selling (that is, negative holdings) is allowed.**

   So, we are allowed to sell unlimited amounts of securities that we do not own – *eg* market makers can do this through repo arrangements (which are discussed briefly in Subject CA1). This is necessary because, in order to hedge a derivative whose price is positively correlated with that of the underlying asset – *eg* a call option, which will have a positive delta – we need to hold a negative quantity of the underlying asset.

5. **There are no taxes or transaction costs.**

   This is important since we will need to continuously rebalance some risk-free portfolios.

6. **The underlying asset can be traded continuously and in infinitesimally small numbers of units.**

   Infinite divisibility of securities is necessary to ensure that perfect hedges can be achieved. Continuous trading requires that security markets are open 24 hours a day (and 365¼ days a year!).

> ***Question 14.1***
>
> Why does the Black-Scholes analysis require the assumption of continuous trading?

The key general implication of the underlying assumptions is that the market in the underlying share is complete: that is, all derivative securities have payoffs which can be replicated. This consequence is at odds with the real world and implies problems with the underlying assumptions

1.2 How realistic are the assumptions?

> ***Question 14.2 (Revision)***
>
> What were the six main defects of the lognormal model of security prices discussed in Chapter 10?

It is clear that each of these assumptions is unrealistic to some degree: for example,

- **Share prices can jump. This invalidates assumption 1 since geometric Brownian motion has continuous sample paths.**

  An important consequence of discontinuous share prices is that it is not possible to rebalance the risk-free portfolio at each moment so as to eliminate movements in the value of the portfolio. Hence, the portfolio is not entirely risk-free.

  However, hedging strategies can still be constructed which substantially reduce the level of risk.

- **The risk-free rate of interest does vary and in an unpredictable way.**

  We might, for example, assume that the risk-free rate is either the base rate set by the central bank or the yield on Treasury bills, both of which can vary over time.

  However, over the short term of a typical derivative the assumption of a constant risk-free rate of interest is not far from reality. (More specifically the model can be adapted in a simple way to allow for a stochastic risk-free rate, provided this is a predictable process.)

  In addition, different rates may apply for borrowing and lending.

- **Unlimited short selling may not be allowed except perhaps at penal rates of interest. These problems can be mitigated by holding mixtures of derivatives which reduce the need for short selling. This is part of a suitable risk management strategy as discussed in Section 2 below.**

  You probably have to pay a higher interest rate on your overdraft than you receive on your savings account. So too do financial institutions!

- **Shares can normally only be dealt in integer multiples of one unit, not continuously and dealings attract transaction costs: invalidating assumptions 4, 5 and 6. Again we are still able to construct suitable hedging strategies which substantially reduce risk.**

  Transactions costs do arise in practice, their impact depending upon their size. Several extensions to the standard Black-Scholes model have been developed to allow for the effect of transactions costs on option prices.

- **Distributions of share returns tend to have fatter tails than suggested by the log-normal model, invalidating assumption 1.**

  The assumption that share prices follow a geometric Brownian motion of the form $dS_t = S_t\left(\mu\, dt + \sigma\, dZ_t\right)$ implies that the future share price S_T, $T > t$, is lognormally distributed. Actual share prices, however, experience large up and down movements more commonly than suggested by a lognormal distribution. A particular consequence of this is that large jumps make it more difficult to maintain a delta-neutral portfolio.

Despite all of the potential flaws in the model assumptions, analyses of market derivative prices indicate that the Black-Scholes model does give a very good approximation to the market.

It is worth stressing here that all models are only approximations to reality. It is always possible to take a model and show that its underlying assumptions do not hold in practice.

This does not mean that a model has no use. A model is useful if, for a specified problem, it provides answers which are a good approximation to reality or if it provides insight into underlying processes.

A model is a stylised representation of a more complex situation and as such aims to characterise the most important features of that situation in a way that enables the situation to be analysed. It thereby provides useful insight into that situation. More complex models often provide greater insight, but at the cost of greater complexity and perhaps reduced tractability.

In this respect the Black-Scholes model is a good model since it gives us prices which are close to what we observe in the market (despite the fact that we can criticise quite easily the individual assumptions) and because it provides insight into the usefulness of dynamic hedging.

In addition, it is widely used by derivative traders, who can make adjustments to allow for its known deficiencies. The model can be calibrated to reproduce observable market prices and it is valuable in understanding the sensitivity of option prices to the various factors discussed in Chapter 11.

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:

$$dX_t = \mu\left(t, X_t\right) dt + \sigma\left(t, X_t\right) dZ_t$$

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 $g(t,z) = S_0 \exp[(\mu - \frac{1}{2}\sigma^2)t + \sigma z]$ so that $S_t = g(t, Z_t)$. Now apply Ito's formula to $g(t, Z_t)$:

$$\begin{aligned} dg(t,Z_t) &= \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial z} dZ_t + \frac{1}{2}\frac{\partial^2 g}{\partial z^2}(dZ_t)^2 \\ &= \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial z} dZ_t + \frac{1}{2}\frac{\partial^2 g}{\partial z^2} dt \\ &= S_t\left((\mu - \frac{1}{2}\sigma^2)dt + \sigma dZ_t + \frac{1}{2}\sigma^2 dt\right) \\ &= S_t(\mu dt + \sigma dZ_t). \end{aligned}$$

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**
- **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 for the change in the derivative price as a function of the change in the share price. Here $df(S_t,t)$ means the change in the value of the derivative over a very small time period.

An expression for $df(t,S_t)$

By Ito's formula we have:

$$\boldsymbol{df(t,S_t) = \frac{\partial f}{\partial t} \times dt + \frac{\partial f}{\partial S_t} \times dS_t + \frac{1}{2}\frac{\partial^2 f}{\partial {S_t}^2} \times (dS_t)^2}$$

Recalling that for an exponential Brownian motion with drift:

$$dS_t = \left[\mu\, dt + \sigma\, dZ_t\right] S_t$$

and hence:

$$\begin{aligned}(dS_t)^2 &= \left[\mu^2 (dt)^2 + \sigma^2 (dZ_t)^2 + 2\mu\sigma\, dt\, dZ_t\right] S_t^2 \\ &= \left[\mu^2 (dt)^2 + \sigma^2 (dt) + 2\mu\sigma\, dt\, dZ_t\right] S_t^2 \\ &= \sigma^2 S_t^2 dt\end{aligned}$$

We then have from the equation above:

$$\begin{aligned}df(t,S_t) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S_t} dS_t + \frac{1}{2}\frac{\partial^2 f}{\partial S_t^2}(dS_t)^2 \\ &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S_t}\left[\left(\mu dt + \sigma dZ_t\right)S_t\right] + \frac{1}{2}\frac{\partial^2 f}{\partial S_t^2}\left[\sigma^2 S_t^2 dt\right] \\ &= \boldsymbol{\left(\frac{\partial f}{\partial t} + \mu S_t \frac{\partial f}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 f}{\partial {S_t}^2}\right) dt + \sigma S_t \frac{\partial f}{\partial S_t} dZ_t}\end{aligned}$$

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 $\frac{\partial f}{\partial S_t}$ to mean $\frac{\partial f}{\partial S_t}(t,S_t)$ *etc.* The correct way to apply Ito's formula 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 $\frac{\partial f}{\partial S_t}(t,S_t)$ shares.**

Let $V(t,S_t)$ be the value of this portfolio, *ie*:

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

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*:

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

Q

Question 14.5

Derive the previous equation.

We are assuming 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 that although the above expression representing the change in the portfolio involves S_t, which is random, we are thinking about the change in the infinitesimal time interval $(t,t+dt)$. Hence S_t can be treated as a constant.

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

Note that although the above expression representing the change in the portfolio involves S_t, which is random, we are thinking about the change *at time t*, hence S_t is a known constant.

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 $S_t > 0$:

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

Question 14.6

Derive the right-hand side of this equation.

Recall that $V(t,S_t) = -f\left(t,S_t\right)+\dfrac{\partial f}{\partial S_t}S_t$ is the value of the portfolio.

We now have two different expressions for $-df(t,S_t)+\dfrac{\partial f}{\partial S_t}dS_t$. If we equate these, we get:

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

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

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 $f(T, S_T) = \max\left\{S_T - K, 0\right\}$

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:

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

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.

3 The Black-Scholes model for dividend-paying shares

3.1 Introduction

In previous sections we assumed that the underlying asset produced no income, so that the price of the underlying asset would give us the total return directly.

So far we have considered only non-dividend-paying shares and, in doing so, we have treated the share as a non-dividend-paying *pure asset*, *ie* an asset that provides an investment return only via capital growth. In practice, of course, most shares do pay a stream of dividends and so we must modify our previous analysis to allow for this possibility.

We now, essentially, repeat Section 2 in the case that the underlying security has income payable continuously at a rate q. This leads to the *Garman-Kohlhagen formula* for the value of a call option on a dividend-paying share in a continuous-time framework.

3.2 The underlying SDE with dividends

Suppose instead that dividends are payable continuously at the constant rate of q per annum per share: that is, the dividend payable over the interval $(t, t+dt]$ is:

$$qS_t dt$$

Note that the dividend amount is proportional to the value of the share at that time. Note also that this q has nothing to do with the risk-neutral probability measure discussed earlier.

Suppose that S_t is subject to the same stochastic differential equation as before:

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

although μ may be different from the rate of growth on the non-dividend-paying asset described before.

Recall that the total return on a share is the sum of the growth rate and the dividend rate.

Question 14.10

State Ito's Lemma for a function $f(X_t)$ of a diffusion process X_t where:

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

where Z_t is a standard Brownian motion.

Modification due to dividends

Once we introduce dividends into our model, the problem we face is that the share price process S_t no longer represents the whole value of the asset. More specifically, if we buy the share for S_0 at time 0, then by time T the total value of what we have bought is equal to:

- S_T, the share price at time T
- *plus* the total of the accumulated dividends received to date.

Moreover, we have defined the dividend payable over the interval $(t, t + dt]$ to be $qS_t dt$ – *ie* it depends on S_t. So, the total of the accumulated dividends received to date must depend on the share price at every instant over the interval 0 to T.

We therefore need to construct a new process, $\tilde{S}_t$ say, that is closely-related to S_t and that does represent the entire value of the asset we purchase by buying the share – *ie* the total of the past dividend income *and* the capital growth. This is achieved by considering the following investment.

Suppose that we start with one unit of the share worth S_0 at time 0. Subsequently, any dividends that we receive are assumed to be reinvested instantaneously, by purchasing additional units of the same share. Note that as the dividends are assumed to be received continuously, the reinvestment process will itself be continuous. Over any time interval $[t, t+dt)$ the dividend payout on one share is equal to $qS_t dt$, which can be reinvested at time t to purchase an additional qdt units of the share.

Question 14.11

How many units of the share will we hold at time t?

It is the value of this portfolio, allowing for the continuous reinvestment of dividends, that we define as the process of interest and that we denote by $\tilde{S}_t$. Many unit-linked savings funds offer a facility where investors can choose either to receive dividends by cheque or to reinvest them in the fund. In the first case the value of the fund corresponds to S_t, in the second case $\tilde{S}_t$.

Let $\tilde{S}_t$ be the value of an investment of $\tilde{S}_0 = S_0$ at time 0 in the underlying asset assuming that all dividends are reinvested in the same asset at the time of payment of the dividend. Sometimes $\tilde{S}_t / \tilde{S}_0$ is described as the *total return* on the asset from time 0 to time t.

It is important to note that $\tilde{S}_t$ is the tradable asset and not S_t in the following sense. If we pay S_0 at time 0 for the asset then we are buying the right to future dividends as well as future growth of the capital. In other words the value of the asset at time t should account for the accumulated value of the dividends as well as the value of the capital at time t.

It is straightforward to see that the stochastic differential equation for $\tilde{S}_t$ is:

$$d\tilde{S}_t = \tilde{S}_t \left((\mu + q)dt + \sigma dZ_t \right)$$

As the dividends are received continuously at the rate of q *per annum* and are reinvested immediately, so the growth rate or drift of the *total value* of the share asset at time t must be the drift in the share price alone, μ, *plus* the instantaneous income yield, q.

Solving this we find that :

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

Question 14.12

Show that $d\tilde{S}_t = \tilde{S}_t \left\{ \left(\mu + q \right) dt + \sigma \, dZ_t \right\}$ is the corresponding SDE to the above equation.

Effectively, the tradable asset $\tilde{S}_t$ is just the share price process accumulated at the fixed rate q, where S_t follows the geometric Brownian motion as we assumed earlier.

Question 14.13

How are $\tilde{S}_t$ and S_t related?

Note that if the model is assumed to start at time zero then S_0 and $\tilde{S}_0$ are the same.

3.3 The Garman-Kohlhagen formula

Let us consider a European call option on the underlying asset S_t with strike price K and time of maturity T. The payoff on this option will be $\max\{S_t - K, 0\}$ as before. However, in valuing the option we must take account of the fact that dividends are payable on the underlying asset which do not feed through to the holder of the option. Let us denote the value of this option at time t by $f(t, S_t)$.

Proposition 14.2 (The Garman-Kohlhagen formula for a call option on a dividend-paying share)

For such a call option:

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

where:

$$d_1 = \frac{\log\frac{S_t}{K} + \left(r - q + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = d_1 - \sigma\sqrt{T-t}$$

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

This formula is given on page 47 of the Tables.

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

As before we will give here two proofs of this result, one here using the partial differential equation (PDE) approach and the other in Chapter 16 using the martingale approach.

3.4 *The PDE approach*

An expression for df(t,S_t)

By Ito's lemma we have:

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

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

The risk-free portfolio

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

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

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

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

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 including the dividend payment: that is:

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

So, the difference here from before is that the change in the value of the portfolio also includes the dividend income $qS_t dt$.

Now note that the expression for $-df(t,S_t) + \frac{\partial f}{\partial S_t}(dS_t + qS_t dt)$ involves dt but not dZ_t so that the investment gain over the next instant 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. Therefore we must have:

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

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

$$\Rightarrow \quad \frac{\partial f}{\partial t}+(r-q)S_t\frac{\partial f}{\partial S_t}+\frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 f}{\partial S_t^2}=rf$$

Notice that the only difference from before is that dividends are now included. The r factor on the LHS is replaced by $r-q$.

The value of the derivative is found by specifying appropriate boundary conditions and solving the PDE, in exactly the same way as we did with no dividends.

The boundary conditions are:

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

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

Question 14.14

Verify that the Garman-Kohlhagen formula for a European call option on a dividend-paying share satisfies the above boundary condition.

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.

3.5 Summary

This leaves us to verify that the Garman-Kohlhagen formula does satisfy the PDE. This will involve deriving the Greeks based on the Garman-Kohlhagen formula. The algebra involved in this is in the Subject ST6 Core Reading. However, it has also been tested in the CT8 exam. You should make sure you are familiar with the methods used in April 2008 Q11 (i) and September 2008 Q9 (ii).

Finally, note that the Black-Scholes formula is simply the Garman-Kohlhagen formula with $q = 0$ and so by verifying that the Garman-Kohlhagen formula satisfies the above PDE, we will implicitly verify that the Black-Scholes formula satisfies the same PDE with $q = 0$.

Delta

In particular, it is worth noting the following results relating to delta, Δ, as this may help with exam questions. The following results are derived by differentiating the Black-Scholes and Garman-Kohlhagen formulae with respect to S_t:

- For a European call option on a non-dividend-paying share, $\Delta = \Phi(d_1)$.
- For a European put option on a non-dividend-paying share, $\Delta = -\Phi(-d_1)$.
- For a European call option on a dividend-paying share, with a continuously-compounded dividend yield, q, $\Delta = e^{-q(T-t)}\Phi(d_1)$.
- For a European put option on a dividend-paying share, with a continuously-compounded dividend yield, q, $\Delta = -e^{-q(T-t)}\Phi(-d_1)$.

4 Implied volatility

Although the concept of implied volatility does not appear in the Core Reading, it is something the Examiners might expect you to be able to work out based on what does appear in the Core Reading. This short appendix therefore explains the basic idea.

Both the standard Black-Scholes formula and the Garman-Kohlhagen formula for a dividend-paying share require knowledge of σ, the volatility of the price of the underlying asset. Unlike the other parameters in these formulae, the volatility cannot be observed directly in the market. It must therefore be estimated if either formula is to be used to price a derivative in practice.

One way of doing this is to use an observed option price from the market and derive the volatility that is consistent with this price. This is possible because we can observe the values of the current price of the underlying asset, the risk-free rate of interest and the dividend yield on the underlying asset. This information is added to the strike price and the maturity date leaving the volatility as the only unknown quantity in the Black-Scholes formula. The resulting estimate is known as the *implied volatility*.

By differentiating the Black-Scholes formula for a European call option (in its Garman-Kohlhagen form) with respect to σ, we find that:

$$\nu = se^{-qu}\,\phi(d_1)\sqrt{u} > 0$$

Thus, the price of a call option is strictly increasing as the volatility increases.

Suppose we know that the price of an option is 6.87. It is not possible for us to write down a nice formula for σ in terms of the other parameters. However, it is straightforward to find the value of σ working backwards by trial and error. Since the option price is a strictly increasing function of σ, the solution to this problem is unique.

Suppose we find that with σ = 0.17, Black-Scholes implies that the corresponding option price would be 6.841 and with σ = 0.18 the price would be 7.006. Using linear interpolation we can then estimate the volatility as σ = 0.17175. This estimate is called the *implied volatility*.

5 Exam-style questions

We finish this chapter with two tricky exam-style questions. The concepts tested in these questions are also frequently tested in the CT8 exam. You should spend some time getting to grips with them.

Question 1

A building society issues a one-year bond that entitles the holder to the return on a weighted-average share index (ABC500) up to a maximum level of 30% growth over the year. The bond has a guaranteed minimum level of return so that investors will receive at least $x\%$ of their initial investment back. Investors cannot redeem their bonds prior to the end of the year.

(i) Explain how the building society can use a combination of call and put options to prevent making a loss on these bonds. [4]

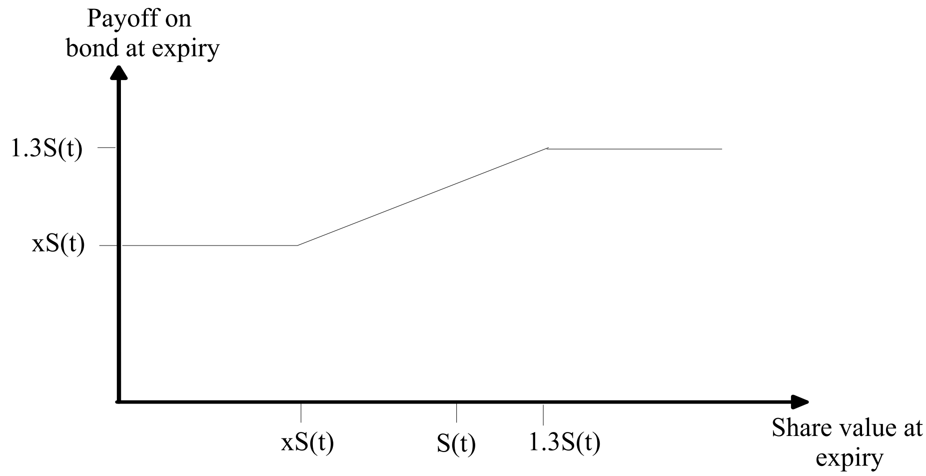
(ii) The volatility of the ABC500 index is $30\%\,pa$ and the continuously-compounded risk-free rate of return is $4\%\,pa$. Assuming no dividends, use the Black-Scholes pricing formulae to determine the value of x (to the nearest 1%) that the building society should choose to make neither a profit nor a loss. [6]

[Total 10]

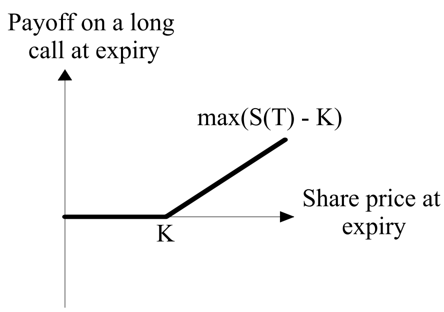
Solution 1

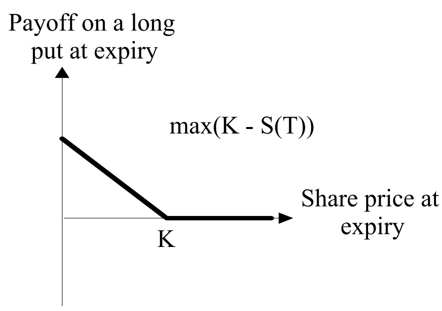
(i) ***Preventing a loss***

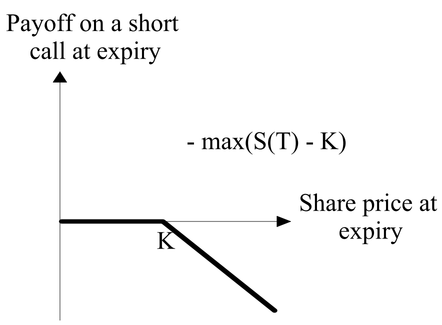
A helpful way of tackling tricky Black-Scholes questions such as these is to draw a graph comparing the payoff on the bond at expiry with the value of the underlying asset at expiry. In the diagram below, S(t) is the value of the initial investment.

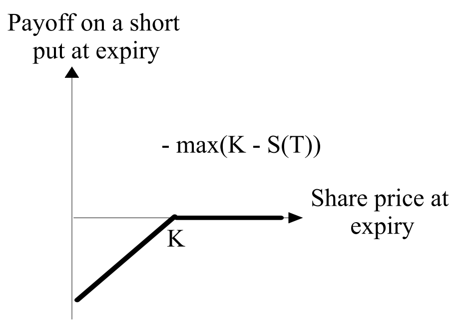


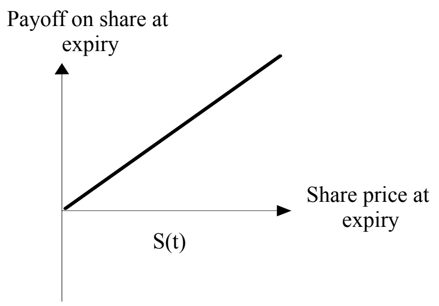
We can compare the shape of this graph against graphs for the payoffs on call and put options, and on the underlying shares.











The graphs demonstrate that a combination of the underlying shares, a long put with exercise price xS_t and a short call with exercise price $1.3S_t$ will replicate the graph and hence the payoff on the bond.

If an investor buys a bond the building society can invest the money in the ABC500 so that it is not exposed to movements in the ABC500 index. However, the building society is guaranteeing that investors will receive at least $x\%$ of their initial investment back. The building society can hedge this loss by buying a put option on the index with a strike price of $x\%$ of the current share price. This put option will cost money – let's say p.

The building society is also limiting the investors' return to 130% of their initial investment. This they can do by selling call options with a strike price of 130% of the current share price. This call option will be priced at c, say. If $c \geq p$ then the building society will not make a loss.

(ii) ***No profit or loss***

If $c = p$ then the building society will not make a profit or a loss. So the problem requires us to work out the price of the call option c and then work out the value of x such that $c = p$.

We will be using the Black-Scholes formula to price the options and, because the numbers are all relative, we can assume that the initial index price is 100, say.

Using the formulae on page 47 of the *Tables*, we can calculate the price of a call option:

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

$$= \frac{\ln\left(\dfrac{100}{130}\right) + \left(0.04 + \dfrac{0.3^2}{2}\right)\times 1}{0.3\sqrt{1}} = -0.5912$$

$$d_2 = d_1 - \sigma\sqrt{T-t} = -0.5912 - 0.3\sqrt{1} = -0.8912$$

$$c = S_t\Phi(d_1) - K\Phi(d_2)e^{-r(T-t)}$$

$$= 100\times\Phi(-0.5912) - 130\times\Phi(-0.8912)e^{-0.04}$$

$$= 100\times 0.277 - 130e^{-0.04}\times 0.186$$

$$= 4.44$$

We now need to work out the value of x so that $p = 4.44$. We will try $K = 90$ to begin with:

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

$$= \frac{\ln\left(\dfrac{100}{90}\right) + \left(0.04 + \dfrac{0.3^2}{2}\right)\times 1}{0.3\sqrt{1}} = 0.6345$$

$$d_2 = d_1 - \sigma\sqrt{T-t} = 0.6345 - 0.3\sqrt{1} = 0.3345$$

$$p = K\Phi(-d_2)e^{-r(T-t)} - S_t\Phi(-d_1)$$

$$= 90\times\Phi(-0.3345)e^{-0.04} - 100\times\Phi(-0.6345)$$

$$= 90e^{-0.04}\times 0.369 - 100\times 0.263$$

$$= 5.62$$

This is higher than the required value, and so we try a lower value for the strike price, say $K = 80$:

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

$$= \frac{\ln\left(\frac{100}{80}\right) + \left(0.04 + \frac{0.3^2}{2}\right)\times 1}{0.3\sqrt{1}} = 1.0271$$

$$d_2 = d_1 - \sigma\sqrt{T-t} = 1.0271 - 0.3\sqrt{1} = 0.7271$$

$$p = K\Phi(-d_2)e^{-r(T-t)} - S_t\Phi(-d_1)$$

$$= 80\times\Phi(-0.7271)e^{-0.04} - 100\times\Phi(-1.0271)$$

$$= 80e^{-0.04}\times 0.234 - 100\times 0.152$$

$$= 2.74$$

We require a put option with a premium of $p = 4.44$. So we linearly interpolate to find the value of K that will give us this:

$$K = 80 + (90-80)\times\frac{4.44-2.74}{5.62-2.74} = 85.90$$

So $x = 86\%$ approximately. So, the building society can use $x = 86\%$ and the discussed hedging portfolio to avoid making a loss.

In fact, the exact figure is 86.43%, which is still 86% to the nearest 1%!

Question 2

A company's directors have decided to provide senior managers with a performance bonus scheme. The bonus scheme entitles the managers to a cash payment of £10,000 should the company share price have increased by more than 20% at the end of the next 6 months. In addition, the managers will be entitled to 5,000 free shares each, should the share price have increased by more than 10% at the end of the next 6 months.

You are given the following data:

Current share price	£7.81
Risk-free rate	5% pa (continuously-compounded)
Share price volatility	25% pa

No dividends to be paid over the next 6 months.

(i) By considering the terms of the Black-Scholes call option pricing formula, calculate the value of the bonus scheme to one manager. [6]

(ii) Explain the main disadvantages of this bonus scheme as an incentive for managers to perform. [2]

(iii) Some shareholders are concerned that this scheme might cause an undesirable distortion to the managers' behaviour. Suggest some modifications to the scheme that will ensure that the managers' aims coincide with the long-term objectives of the shareholders. [3]

[Total 11]

Solution 2

(i) ***Value of the bonus scheme***

First, consider the share options. The manager will receive 5,000 shares if the share price in 6 months' time is greater than:

$$£7.81 \times 1.1 = £8.59$$

This is like having 5,000 call options on the share with a strike price of £8.59, except that no payment is actually required. We can use the Black-Scholes formula for one call option, which is on page 47 of the Tables (the Garman-Kohlhagen formula with $q = 0$):

$$C_{[K=8.59]} = 7.81\Phi(d_1) - 8.59e^{-0.05\times 6/12}\Phi(d_2)$$

But the manager will not need to pay the £8.59 so the second term is not required. So, the value of each share option is:

$$7.81\Phi(d_1) = 7.81\Phi\left[\frac{\ln(S/K) + \left(r + \sigma^2/2\right)(T-t)}{\sigma\sqrt{T-t}}\right]$$

$$= 7.81\Phi\left[\frac{\ln(1/1.1) + \left(0.05 + 0.25^2/2\right)\times 0.5}{0.25\sqrt{0.5}}\right]$$

$$= 7.81\Phi(-0.30935)$$

$$= 7.81[1 - 0.6215]$$

$$= £2.96$$

So, the share options for each manager are worth $5,000 \times £2.96 = £14,800$.

Next, consider the £10,000 cash bonus. The managers will receive this if the share price is greater than:

$$£7.81 \times 1.2 = £9.37$$

The value of a call option here would be:

$$C_{[K=9.37]} = 7.81\Phi(d_1) - 9.37e^{-0.05\times 6/12}\Phi(d_2)$$

The second term here corresponds to the value of the strike price £9.37 that would be paid if the share price is greater than £9.37 in 6 months' time.

The cash bonus is made in the same circumstances as this. However, the amount is £10,000 rather than £9.37.

So the value of the cash bonus is:

$$\begin{aligned}10{,}000e^{-0.05\times\frac{6}{12}}\Phi\left(d_2\right) &= 10{,}000\times 0.97531\Phi\left[\frac{\ln\left(S/K\right)+\left(r-\sigma^2/2\right)\left(T-t\right)}{\sigma\sqrt{T-t}}\right]\\ &= 9{,}753.1\Phi\left[\frac{\ln\left(7.81/9.37\right)+\left(0.05-0.25^2/2\right)\times 0.5}{0.25\sqrt{0.5}}\right]\\ &= 9{,}753.1\Phi\left(-0.97833\right)\\ &= 9{,}753.1\left(1-0.8360\right)\\ &= £1{,}600\end{aligned}$$

Finally, the total value of the bonus scheme is:

£14,800+£1,600

=£16,400

(ii) ***Incentive for managers to perform?***

Both the £10,000 cash bonus and the share options do not give managers any incentive to help the company share price once the 6-month period is over.

Managers may be able to sell their free shares, add the proceeds to the £10,000 and may have little interest in how the company subsequently performs.

In addition, because the managers do not receive any bonus at all whether the share price increases by 9%, or decreases by 50% say.

As such, they may be tempted to undertake a riskier investment strategy that is not in the best interests of the shareholders.

Any increase above 20% is not further rewarded.

(iii) ***Improvement to the scheme***

The share options will provide managers more of a long-term incentive if they are restricted from selling the shares for a fixed time period after they are awarded, 3 years say.

In addition, a condition may be imposed that the shares will only be awarded provided the manager continues to work for the company for a fixed time period, 3 years say.

These will ensure the managers have an interest in how the company performs after the 6-month period is over.

Instead of a cash bonus, the managers could be given the equivalent amount in more bonus shares, again with the restrictions mentioned above.

The number of free shares issued could be made to depend more gradually on the company's share price performance, eg 100 free shares for each percentage point performance above a specified benchmark level.

This may stop any manager being tempted to employ an all-or-nothing approach in their business/investment strategy.

It may be possible to pay the managers' salaries almost entirely in shares so that their interests are the same as that of the shareholders.

6 End of Part 3

What next?

1. Briefly **review** the key areas of Part 3 and/or re-read the **summaries** at the end of Chapters 11 to 14.
2. Attempt some of the questions in Part 3 of the **Question and Answer Bank**. If you don't have time to do them all, you could save the remainder for use as part of your revision.
3. Attempt **Assignment X3**.

Chapter 14 Summary

Assumptions of Black-Scholes

- the underlying share price follows geometric Brownian motion
- the market is arbitrage-free
- the risk-free rate r is constant and the same for all borrowing and lending
- assets may be bought and sold at any time $t > 0$
- assets may be held in any amount
- there are no taxes or transaction costs

An implication of these is that the market in the underlying share is complete. The validity of each assumption can be questioned.

Black Scholes PDE

$$rf(S_t) = \Theta + \Delta(r-q)S_t + \frac{\sigma^2}{2}\Gamma S_t^{\,2}$$

For any derivative to be fairly priced, it must:

- satisfy the boundary conditions, *ie* have the correct payoff at expiry and
- satisfy the above PDE.

Garman-Kohlhagen formulae for options on a dividend-paying share

European call

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

European put

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

where:

$$d_1 = \frac{\ln\left(S_t/K\right) + (r - q + \tfrac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\ , \qquad d_2 = d_1 - \sigma\sqrt{T-t}$$

The Black-Scholes formula for a non-dividend-paying share is the same but with $q = 0$.

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

Chapter 14 Solutions

Solution 14.1

Continuous trading is necessary so that the hedging portfolio can be rebalanced *continuously*.

Solution 14.2

The continuous-time lognormal model may be inappropriate for modelling investment returns because:

- The volatility σ may not be constant over time. Estimates of volatility from past data are critically dependent on the time period chosen for the data and also how often the price history is sampled

- The long-term drift μ may not be constant over time. In particular, interest rates will impact the drift.

- There is evidence in real markets of mean-reverting behaviour, which is inconsistent with the independent increments assumption.

- There is evidence in real markets of momentum effects, which is inconsistent with the independent increments assumption.

- The distribution of security returns $\log\left(S_u/S_t\right)$ has a higher peak in reality than that implied by the normal distribution. This is because there are more days of little or no movement in the share price.

- The distribution of security returns $\log\left(S_u/S_t\right)$ has fatter tails in reality than that implied by the normal distribution. This is because there are frequent big "jumps" in security prices.

Solution 14.3

Recall Ito's lemma from Chapter 9. $f(t, X_t)$ is also a diffusion process and its SDE is:

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \mu(t, X_t)\frac{\partial f}{\partial X_t} + ½\sigma^2(t, X_t)\frac{\partial^2 f}{\partial X_t^2} \right) dt + \sigma(t, X_t)\frac{\partial f}{\partial X_t} dZ_t$$

This can be checked by using a Taylor's series expansion in the two variables t and X_t to write:

$$df(t, X_t) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X_t}dX_t + ½\frac{\partial^2 f}{\partial X_t^2}(dX_t)^2$$

Remember that because X_t is stochastic, we need to retain the second-order term, whereas we don't need the $½\frac{\partial^2 f}{\partial t^2}(dt)^2$ term because $(dt)^2 = o(dt) = 0$.

Then recall from the multiplication table for increments on page 19 of Chapter 9 that all second-order increments are zero except $(dZ_t)^2 = dt$. It follows that $(dX_t)^2 = \sigma^2(t, X_t)dt$. So, substituting in for dX_t and $(dX_t)^2$ gives:

$$\begin{aligned} df(t, X_t) &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X_t}\left(\mu(t, X_t)dt + \sigma(t, X_t)dZ_t\right) + ½\frac{\partial^2 f}{\partial X_t^2}\sigma^2(t, X_t)dt \\ &= \left(\frac{\partial f}{\partial t} + \mu(t, X_t)\frac{\partial f}{\partial X_t} + ½\sigma^2(t, X_t)\frac{\partial^2 f}{\partial X_t^2} \right) dt + \sigma(t, X_t)\frac{\partial f}{\partial X_t} dZ_t \end{aligned}$$

Solution 14.4

Using put-call parity and the result from Proposition 14.1, we have:

$$\begin{aligned} p_t &= Ke^{-r(T-t)} - S_t + c_t = Ke^{-r(T-t)} - S_t + S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2) \\ &= Ke^{-r(T-t)}(1-\Phi(d_2)) - S_t(1-\Phi(d_1)) \\ &= Ke^{-r(T-t)}\Phi(-d_2) - S_t\Phi(-d_1) \end{aligned}$$

Notice that the bottom line answer is also given on page 47 of the Tables. Again $q = 0$.

Solution 14.5

Using Ito's Lemma:

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

So we have:

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

Finally, using the result that $(dS_t)^2 = \sigma^2 S_t^2 dt$, this becomes:

$$\begin{aligned}
&= -\frac{\partial f}{\partial t}dt - \frac{1}{2}\frac{\partial^2 f}{\partial S_t{}^2}\sigma^2 S_t^2 dt \\
&= \left(-\frac{\partial f}{\partial t} - \frac{1}{2}\frac{\partial^2 f}{\partial S_t{}^2}\sigma^2 S_t^2\right)dt
\end{aligned}$$

as required.

Solution 14.6

We can derive the right-hand side of the equation $(rV\,dt)$ as follows. The value of the portfolio must accumulate to $e^{rdt}V_t$ during the short time interval *dt*. If we expand the exponential as a series, and ignore second-order and higher-order terms (because *dt* is infinitesimal), then we see that the accumulated value is $V_t(1+rdt) = V_t + rV_t dt$. The change in the value must therefore be $rV_t dt$.

This is analogous to the equation for the cash bond, which was $dB_t = rB_t dt$.

Solution 14.7

We must verify that it:

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

Solution 14.8

Amongst the assumptions of Black-Scholes are that the risk-free rate and volatility are constant. We have not introduced dividends yet but when we do these will be at a constant rate q.

Solution 14.9

(i) ***Boundary condition***

At expiry:

$$f(S_T, T) = S_T - K$$

which is what we would expect. The investor will pay K and receive a share worth S_T.

(ii) ***Black-Scholes PDE***

We first differentiate $f(S_t, t)$ with respect to S_t and t to find the Greeks:

$$\Delta = \frac{\partial f}{\partial S_t} = 1 \qquad \Gamma = \frac{\partial \Delta}{\partial S_t} = 0 \qquad \Theta = \frac{\partial f}{\partial t} = -rKe^{-r(T-t)}$$

Using these Greeks we see that:

$$\Theta + \Delta r S_t + \frac{1}{2}\sigma^2 \Gamma S_t^2$$

$$= -rKe^{-r(T-t)} + 1 \times rS_t + 0$$

$$= r\left(S_t - Ke^{-r(T-t)}\right)$$

$$= rf(S_t, t)$$

and so the PDE is satisfied.

Solution 14.10

Ito's lemma is the chain rule for stochastic calculus. It states that if f is a deterministic and twice continuously differentiable function, then $Y_t = f(X_t)$ is also a diffusion process:

$$df\left(X_t\right) = \left(\mu\left(t, X_t\right)\frac{\partial f}{\partial X} + ½\sigma^2\left(t, X_t\right)\frac{\partial^2 f}{\partial X^2}\right)dt + \sigma\left(t, X_t\right)\frac{\partial f}{\partial X}dZ_t$$

This can be shown by:

- first using Taylor's theorem in the limit to write:

$$dY_t = \frac{\partial f}{\partial X_t}dX_t + ½\frac{\partial^2 f}{\partial X_t^2}(dX_t)^2 \qquad \text{(note that } \frac{\partial f}{\partial t} = 0 \text{ here)}$$

- then apply the table of products that says all second order increments are zero except $\left(dZ_t\right)^2 = dt$. It follows that $\left(dX_t\right)^2 = \sigma^2\left(t, X_t\right)dt$. The result follows by substituting in for dX_t and $\left(dX_t\right)^2$.

Solution 14.11

Given that we started with one unit of share at time 0, if we purchase additional shares continuously at the rate of q – *ie* we purchase qdt over the interval $(t, t+dt]$ for *each share that we already hold* – then by time t our total holding will be e^{qt}.

Solution 14.12

If:

$$\tilde{S}_t = \tilde{S}_0 \exp\left[\left(\mu + q - ½\sigma^2\right)t + \sigma Z_t\right]$$

then the differential of $\tilde{S}_t$ can be found by Ito's lemma.

Let $\tilde{S}_t = f\left(X_t\right)$ where $X_t = \left(\mu - ½\sigma^2\right)t + \sigma Z_t$ and $f\left(x\right) = \tilde{S}_0 e^x$.

It follows that:

$$\begin{aligned}
dS_t = df\left(X_t\right) &= f'\left(X_t\right)dX_t + ½ f''\left(X_t\right)\left(dX_t\right)^2 \\
&= \tilde{S}_t\left(\left(\mu - ½\sigma^2\right)dt + \sigma dZ_t + ½\sigma^2 dt\right) \\
&= \tilde{S}_t\left[\mu dt + \sigma dZ_t\right]
\end{aligned}$$

Note that if we consider *t* an explicit variable and think of $\tilde{S}_t = g\left(t, Z_t\right)$ then we need an extra $\dfrac{\partial \tilde{S}_t}{\partial t}$ term:

$$\begin{aligned}
d\tilde{S}_t(t,Z_t) &= \frac{\partial \tilde{S}_t}{\partial t}dt + \frac{\partial \tilde{S}_t}{\partial Z_t}dZ_t + \frac{1}{2}\frac{\partial^2 \tilde{S}_t}{\partial Z_t^2}(dZ_t)^2 \\
&= \left[\mu + q - ½\sigma^2\right]\tilde{S}_t\, dt + \sigma\, \tilde{S}_t\, dZ_t + ½\sigma^2 \tilde{S}_t (dZ_t)^2 \\
&= \left[\mu + q - ½\sigma^2\right]\tilde{S}_t\, dt + \sigma\, \tilde{S}_t\, dZ_t + ½\sigma^2 \tilde{S}_t\, dt \\
&= \tilde{S}_t\left\{\left(\mu + q\right)dt + \sigma\, dZ_t\right\}
\end{aligned}$$

as required.

Solution 14.13

As the tradable asset $\tilde{S}_t$ is just the share price process accumulated at the fixed rate *q*, so it must be the case that:

$$\tilde{S}_t = e^{qt} S_t\,.$$

We can see from the last Core Reading equation that:

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

Solution 14.14

We need to check that:

$$\lim_{t \to T} f(S_t, t) = \max\left[S_T - K, 0\right]$$

Because $d_2 = d_1 - \sigma\sqrt{T-t}$, the limit of d_1 and d_2 are equal:

$$\lim_{t \to T} d_1 = \lim_{t \to T} d_2 = \begin{cases} \dfrac{+ve+0}{0} = +\infty & \text{if } S_T > K \\ \dfrac{-ve+0}{0} = -\infty & \text{if } S_T < K \end{cases}$$

If $S_T = K$, and writing d_1 as:

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

we see that:

$$\lim_{t \to T} d_1 = 0 + 0 = 0$$

Finally, noting that $\Phi(\infty) = 1$, $\Phi(-\infty) = 0$ and $\Phi(0) = 0.5$, we have:

$$\lim_{t \to T} f(S_t, t) = \begin{cases} S_T - K & \text{if } S_T > K \\ 0 & \text{if } S_T < K \\ 0.5 \times \left(S_T - K\right) = 0 & \text{if } S_T = K \end{cases}$$

$$= \max\left[S_T - K, 0\right]$$

Chapter 15

The 5-step method in discrete 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:*

 - *Explain what is meant by a complete market.*
 - *Explain what is meant by risk-neutral pricing and the equivalent martingale measure.*
 - *Demonstrate how to price and hedge a simple derivative contract using the martingale approach. (part)*

0 Introduction

In this chapter we build up the theory required for an alternative proof of the Garman-Kohlhagen formula. The proof itself will be demonstrated in Chapter 16. It centres on the development of a derivative pricing approach based on the use of *replicating strategies,* as in the discrete time case. It is referred to here as the 5-step method. Before presenting the 5-step method in the continuous case, in this chapter we will demonstrate the structure of the proof in the simpler discrete case, *ie* for the binomial model.

The approach in the binomial case involves constructing a replicating portfolio. In other words, at a given time $t-1$ we need to be able to set up a portfolio of shares and cash that will evolve by time t to be worth the same as the derivative at that time, no matter how the share price moves. In the binomial case there are only two possibilities for the share price movement over each time step. This will allow us to calculate the correct amounts of shares ϕ_t and cash ψ_t simply by solving two simultaneous equations.

However, there is a problem in continuous time with such a replicating strategy. In the continuous case the number of possible share price movements is infinite, in fact they form a continuum, and so this simple method might seem doomed to failure.

The solution lies in recognising that the concepts underlying the replication method can be formulated in terms of *martingales.* A powerful theorem known as the *martingale representation theorem* then comes to the rescue. An application of this theorem confirms the existence of the ϕ and ψ we need for replication. This forms the basis of the *martingale approach* to pricing or *risk-neutral valuation*.

The formula we end up with is the general derivative pricing formula:

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

We will see in Chapter 16 that this same formula can be used to derive the Black-Scholes and Garman-Kohlhagen formulae for the call and put options, which we met previously.

Some of the material in this chapter and Chapter 16 is quite technical. Should you wish to see an alternative presentation of this material, then it may be worth looking at the book by Baxter & Rennie listed in the Study Guide.

1 Preliminary concepts

It is helpful to recognise that the basic set-up for the continuous case is analogous to the one used in the discrete time framework outlined in Chapter 13. Consequently, we will on occasion refer back to the binomial model. First of all, however, we introduce some preliminary concepts.

Suppose that we are given some probability triple (Ω, F, P) where P is the *real-world* probability measure (sometimes also referred to as the *physical* or *natural* or *objective* measure). Here Ω is the sample space (set of all possible outcomes) as defined on page 16 of Chapter 8.

Question 15.1

What is represented by F?

1.1 Background

The share price process and the filtration

We will make use in this and subsequent sections of a stochastic process S_t for prices, where S_t is measurable with respect to F_t (that is, given F_t we know the value of S_u for all $u \le t$). Let F_t be the sigma algebra generated by S_u (and B_u if this is stochastic) for $0 \le u \le t$: that is, F_t gives the history of the process up to time t.

So S_t will be used to denote the value of the share at time t. Because share prices are random, this is not known before time t, and so we model it with a stochastic process. As mentioned in Chapter 8, where it was called a filtration, you should think of F_t as meaning that we are at time t, and we know the history of the share value up until time t.

The real-world probability measure P

Let A be some event contained in F (for example, suppose that A is the event that S_1 is greater than or equal to 100). Then $P(A)$ is the *actual* probability that the event A will occur.

On a more intuitive level with m independent realisations of the future instead of one we would find that the event A occurs on approximately a proportion $P(A)$ occasions (with the approximation getting better as m gets larger and larger).

So at time 0 we may think that the probability that the share price will be greater than or equal to 100 at time t is 0.2. Another way of saying this is that, if we could run the future up until time t lots of times, we would expect the share price to be greater than or equal to 100 about one fifth of the time.

In practice we can only ever *estimate* the real-world probabilities, using a mathematical model of a situation, since it is impossible for us to "replay" the real world.

Cash

Suppose also that we have a risk-free cash bond which has a value at time t of B_t. We will again refer to this as either the *cash bond* or simply as *cash.*

Sometimes the risk-free rate of interest will itself follow a stochastic process, but in the sections which follow we will assume that the risk-free rate of interest is constant: that is, B_t is deterministic and equal to $B_0 e^{rt}$ for some constant r.

Assuming that $r > 0$, we have the following properties:

- $B_0 = 1$ is the value of cash at $t = 0$
- $B_t > B_0 = 1$ and $B_t > B_{t-1}$ for $t > 0$
- B_t increases in an entirely predictable manner as we move through time – *ie* at the continuously-compounding rate of r per time period.

You must be careful not to confuse this with when B_t is used as the notation for Brownian motion. It should be clear from the context in which B_t is being used but we will try to add comments in sections where the distinction is not so obvious.

1.2 *Tradeable assets*

In the basic form of the Black-Scholes model that we will see in the next section, we will assume that S_t represents the price of a non-dividend-paying share. This means that if we invest S_0 at time 0, S_t represents the *total return* on the investment up to time t assuming that we hold onto the share until that time.

This, as well as the cash bond B_t, is an example of a *tradeable asset.* Such an asset is one where its price at time t is equal to the total return on that investment up to time t with no dividend income payable or inputs of cash required.

1.3 Self-financing strategies

We now consider the properties that are necessary for a replicating portfolio.

The portfolio

Suppose that at time t we hold the portfolio (ϕ_t, ψ_t) where:

- **ϕ_t represents the number of units of S_t held at time t**
- **ψ_t is the number of units of the cash bond held at time t.**

So a portfolio is an ordered pair of processes ϕ_t and ψ_t that describes the number of units of each security held at time *t*. Note that:

- we do not constrain the values of ϕ_t and ψ_t to be non-negative – *ie* short selling of securities is permitted.
- the choice of the portfolio (ϕ_t, ψ_t) at time *t* represents a *dynamic* strategy, as the values of ϕ_t and ψ_t can change continuously through time.

We assume that S_t is a tradeable asset as described above.

Previsible processes

The only significant requirement on (ϕ_t, ψ_t) is that they are *previsible*: that is, that they are F_{t^-}-measurable (so ϕ_t and ψ_t are known based upon information up to *but not including* time t).

So, ϕ_t and ψ_t are dependent only on the history of stock prices up to *but not including* time *t*. This is sometimes referred to as the history of stock prices up to time t^- and the corresponding information set is written as F_{t^-}. In the binomial case, $F_{t^-} = F_{t-1}$.

The reason we require ϕ_t and ψ_t to be previsible is that we want to be able to replicate the portfolio in advance. Consider the binomial case again. At time $t-1$ we need to be able to set up a portfolio that replicates the value of the derivative at time t, no matter what happens to the share price. It is no use being able to do this in retrospect! This would not be useful.

In continuous time we would need to continuously change the holdings ϕ_t and ψ_t. This is obviously not possible in practice, but we should be able to approximate the theoretical ideal by rebalancing the portfolio on a regular basis.

Changes in the value of the portfolio

Let $V(t)$ be the value at time t of this portfolio: that is:

$$V(t) = \phi_t S_t + \psi_t B_t$$

Now consider the instantaneous pure investment gain in the value of this portfolio over the period from time t up to $t + dt$: that is, assuming that there is no inflow or outflow of cash during the period $[t, t + dt)$.

Any increase or decrease in the value of the portfolio is therefore due entirely to capital gains or losses generated within the fund. We are assuming that the stock does not pay dividends. There is no net investment/new money injected into or withdrawn from the fund.

This instantaneous pure investment gain is equal to:

$$\phi_t dS_t + \psi_t dB_t$$

Question 15.2

Why is the pure investment gain equal to $\phi_t dS_t + \psi_t dB_t$?

Let's see what the change in the value would be if we did have inflows or outflows, *ie* if we were to buy $d\phi_t$ new shares and pay in $d\psi_t$ units of the cash bond.

The instantaneous change in the value of the portfolio, allowing for cash inflows and outflows, is given by:

$$dV_t = (\phi_t + d\phi_t)(S_t + dS_t) + (\psi_t + d\psi_t)(B_t + dB_t) - (\phi_t S_t + \psi_t B_t)$$

which simplifies to:

$$dV(t) \equiv V(t + dt) - V(t) = \phi_t dS_t + S_t d\phi_t + d\phi_t dS_t + B_t d\psi_t + \psi_t dB_t + d\psi_t dB_t$$

If $d\phi_t = d\psi_t = 0$, this simplifies to:

$$dV_t = \phi_t dS_t + \psi_t dB_t$$

Self-financing

The portfolio strategy is described as *self-financing* if $dV(t)$ is equal to $\phi_t dS_t + \psi_t dB_t$: that is, at $t + dt$ there is no inflow or outflow of money necessary to make the value of the portfolio back up to $V(t + dt)$.

ie $\left(\phi_t, \psi_t\right)$ is self-financing $\Leftrightarrow$ $dV_t = \phi_t dS_t + \psi_t dB_t$

So, the required change in the value of a self-financing portfolio is equal to the pure investment gain. The significance of being self-financing will become clearer once we talk about replicating strategies.

1.4 Replicating strategies and complete markets

Let X be any derivative payment contingent upon F_U where U is the time of payment of X: that is, X is F_U-measurable and, therefore, depends upon the path of S_t up to time U. The time of payment U may be fixed or it may be a random stopping time.

"X is F_U-measurable" means that if you know the history of stock prices up to and including time U, then you know the value of X. In this particular instance, F_U tells us the value of S_U, the stock price at time U, and hence the value of X, the derivative payment contingent upon S_U.

Question 15.3

For which types of derivative is the time of payment U fixed? For which type is it not fixed?

Replicating strategy

A *replicating strategy* is a self-financing strategy (ϕ_t, ψ_t), defined for $0 \le t < U$, such that:

$$V(U) = \phi_U S_U + \psi_U B_U = X$$

So, for an initial investment of $V(0)$ at time 0, if we follow the self-financing portfolio strategy (ϕ_t, ψ_t) we will be able to reproduce the derivative payment without risk.

"Following the self-financing portfolio strategy (ϕ_t, ψ_t)" means that we rebalance our portfolio continuously so that we have ϕ_t units of stock and ψ_t units of cash at time *t*, and we do this without any topping up or siphoning off.

Hence, the self-financing portfolio strategy (ϕ_t, ψ_t) replicates the derivative payment *X* at the time of payment *U*, irrespective of what S_U, the stock price at time *U*, and hence the value of *X*, actually turns out to be. Moreover, if we choose the initial portfolio (ϕ_0, ψ_0) at time 0, then we can replicate the derivative payment at time *U* without the need for any further cash injections into the portfolio.

This is because the change in the value of the portfolio (ϕ_t, ψ_t) over any time interval will be such that its value at the end of that time interval will exactly equal the cost of the portfolio that must then be purchased to maintain the replicating strategy over the next time interval.

Complete market

The market is *complete* if for any such contingent claims X there is a replicating strategy (ϕ_t, ψ_t).

This is important because it means that, in a complete market, we will always be able to price a contingent claim *X*, such as a derivative, based on the arbitrage-free approach and using a replicating strategy (ϕ_t, ψ_t).

We have already seen one example of a complete market: the binomial model. In that model we saw that we could replicate any derivative payment contingent on the history of the underlying asset price.

To replicate a call option or put option we needed shares and cash, which were both available assets. We will see shortly that to replicate these derivatives in continuous time we can use these same assets.

Another example of a complete market is the continuous-time log-normal model for share prices.

Here the model for the share prices is:

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

where Z_t is a standard Brownian motion.

1.5 Equivalent measures

Two measures P and Q which apply to the same sigma-algebra F are said to be *equivalent* if for any event E in F:

$$P(E) > 0 \text{ if and only if } Q(E) > 0$$

where $P(E)$ and $Q(E)$ are the probabilities of E under P and Q respectively.

We have already seen equivalent measures when we looked at the binomial model in Chapter 13.

You can think of this intuitively as meaning that any event E that is possible/impossible under probability measure P is also possible/impossible under probability measure Q. However, remember that when infinite sequences of events are involved, any particular sequence you specify will normally have probability zero.

Here we will be using P to denote the real-world probabilities and Q for the risk-neutral probabilities.

From the above definition of equivalence the only constraint on the real-world measure P is that at any point in the binomial tree the probability of an up-move lies strictly between 0 and 1. The only constraint on Q is the same but this can be equated to the requirement that the risk-free return must lie strictly between the return on a down-move and the return on an up-move.

Question 15.4 (Revision)

Why is this?

This gives us considerable flexibility in the range of possible equivalent measures.

We could, if we wanted to, come up with a whole range of other probability measures, but we will mostly be interested in the real-world and the risk-neutral ones.

1.6 Sample paths problem

In the binomial model, up to a given finite time horizon each sample path has a probability greater than zero so equivalence is straightforward to prove. This is because there are only 2 paths that the share price process can take at each individual node. The share price goes up with probability $p \neq 0$ or down with probability $1-p \neq 0$.

In many continuous-time models (for example, Brownian motion) all individual sample paths have zero probability.

> ***Question 15.5***
>
> Why do all the sample paths have zero probability?

This makes equivalence more difficult to establish.

Remember that, for equivalence, the probabilities for any event (*ie* set of sample paths) have to either both equal zero or be strictly positive together.

Continuous-time example

Suppose that Z_t is a standard Brownian motion under P:

$$Z_t \sim N(0,t) \text{ for all } t$$

Let $X_t = \gamma t + \sigma Z_t$ be a Brownian motion with drift under P.

Is there a measure Q under which X_t is a standard Brownian motion and which is equivalent to P?

The answer is *yes* if $\sigma = 1$ but *no* if $\sigma \neq 1$ (but we do not give a proof here). So we can change the drift of the Brownian motion but not the volatility.

This example can be expressed more formally in the following theorem.

1.7 *The Cameron-Martin-Girsanov theorem*

Suppose that Z_t is a standard Brownian motion under P. Furthermore suppose that γ_t is a previsible process. Then there exists a measure Q equivalent to P and where:

$$\tilde{Z}_t = Z_t + \int_0^t \gamma_s ds$$

is a standard Brownian motion under Q.

Conversely, if Z_t is a standard Brownian motion under P and if Q is equivalent to P then there exists a previsible process γ_t such that:

$$\tilde{Z}_t = Z_t + \int_0^t \gamma_s ds$$

is a Brownian motion under Q.

Note that the converse of the Cameron-Martin-Girsanov Theorem tells us that we can change the drift but not the volatility of the Brownian motion.

Geometric Brownian motion

As we saw in Chapter 8, the process $S_t = S_0 \exp[(\mu - \frac{1}{2}\sigma^2)t + \sigma Z_t]$ where Z_t is a standard Brownian motion under P is sometimes called *geometric Brownian motion*.

Let us consider the discounted asset price $e^{-rt}S_t$. We have:

$$E_P[e^{-rt}S_t] = S_0 e^{(\mu - r)t}$$

So $e^{-rt}S_t$ is not a martingale under P (unless $\mu = r$).

Question 15.6

Show that $E_P\left[e^{-rt}S_t\right] = S_0 e^{(\mu-r)t}$

Now take $\gamma_t = \gamma = \dfrac{\mu - r}{\sigma}$, and define:

$$\tilde{Z}_t = Z_t + \int_0^t \gamma_s ds = Z_t + \frac{(\mu - r)}{\sigma} t$$

Then:

$$\begin{aligned} S_t &= S_0 \exp\left[(\mu - \tfrac{1}{2}\sigma^2)t + \sigma Z_t \right] \\ &= S_0 \exp\left[(r - \tfrac{1}{2}\sigma^2)t + \sigma Z_t + \mu t - rt \right] \\ &= S_0 \exp\left[(r - \tfrac{1}{2}\sigma^2)t + \sigma \left(Z_t + \frac{\mu - r}{\sigma} t \right) \right] \\ &= S_0 \exp\left[(r - \tfrac{1}{2}\sigma^2)t + \sigma \tilde{Z}_t \right] \end{aligned}$$

Also by the Cameron-Martin-Girsanov theorem there exists a measure Q equivalent to P and where $\tilde{Z}_t$ is a standard Brownian motion under Q.

Furthermore, we find, for $u < t$, that:

$$\begin{aligned} E_Q[e^{-rt} S_t \mid F_u] &= e^{-rt} S_u E_Q[\exp\{(r - \tfrac{1}{2}\sigma^2)(t-u) + \sigma(\tilde{Z}_t - \tilde{Z}_u)\}] \\ &= e^{-rt} S_u e^{(r - \frac{1}{2}\sigma^2)(t-u) + \frac{1}{2}\sigma^2(t-u)} \\ &= e^{-ru} S_u \end{aligned}$$

So $e^{-rt} S_t$ is a martingale under Q.

2 The martingale representation theorem

Our ability to replicate the derivative payoff relies on setting up a self-financing, previsible portfolio. In the binomial model this was straightforward, but as mentioned above, in the continuous case it isn't. Here we need to use the *martingale representation theorem.*

2.1 Theorem

Suppose that X_t is a martingale with respect to a measure P :

ie **for any $t < s$, $E_P[X_s \mid F_t] = X_t$**

Suppose also that Y_t is another martingale with respect to P. The *martingale representation theorem* states that there exists a unique previsible process ϕ_t such that (in continuous time):

$$Y_t = Y_0 + \int_0^t \phi_s dX_s$$

or $$dY_t = \phi_t dX_t$$

if and only if there is no other measure equivalent to P under which X_t is a martingale.

Recall that **a measure P' is equivalent to P (written $P' \sim P$) if $P'(A) = 0 \Leftrightarrow P(A) = 0$.**

The theorem also holds when X_t is a vector of martingales.

Note that the integral in equation for Y_t above is a stochastic integral, as introduced in Chapter 9. The subsequent equation is the same relationship expressed as a stochastic differential equation.

It will turn out that in the continuous case, ϕ_t is the amount of stock to be held in the replicating portfolio. As mentioned before, it is therefore essential that ϕ_t be a previsible process. This is what the theorem guarantees.

The proof of the martingale representation theorem is beyond the syllabus, but we will show why such a result might hold by analogy with the discrete case. In the discrete case, we have finite differences, $\Delta X_t, \Delta Y_t$ so that the theorem essentially becomes:

- the process $\phi_t = \dfrac{\Delta Y_t}{\Delta X_t}$ is previsible.

Note that ϕ_t would not be well-defined here if the volatility of X_{t-1}, *ie* ΔX_{t-1}, was zero.

2.2 *Proof of the martingale representation theorem in discrete time*

We can illustrate the proof of the martingale representation theorem in the context of a discrete time random walk. Note that this is slightly different to the binomial model because the process in the binomial model was assumed to be a geometric random walk, *ie* we multiplied by a random factor each time, rather than *adding* a random term.

If we apply this to the discrete time binomial model we might have:

$$\boldsymbol{X_{t+1} = \begin{cases} X_t + u(t, X_t) & \text{with probability } q \\ X_t + d(t, X_t) & \text{with probability } 1-q \end{cases}}$$

Note that $X_{t+1} = X_t + \Delta X_{t+1}$ where ΔX_{t+1} is a random variable which takes the value $u(t, X_t)$ with probability q and the value $d(t, X_t)$ with probability $1 - q$. Here u is a positive quantity and d is a negative quantity.

If X_t is a martingale with respect to the implied measure Q (*ie* the probabilities q and $1-q$) then:

$$q\,u(t, X_t) + (1-q)\,d(t, X_t) = 0$$

$$\Rightarrow \qquad q = \frac{-d}{u-d}$$

Notice that this uniquely specifies *Q*.

Question 15.7

Find an expression for $d(t, X_t)$ in terms of q and $u(t, X_t)$.

Now if Y_t is also a martingale with respect to Q (*ie* based on the same probabilities q and $1-q$) then, first, Y_t must also follow a binomial model with:

$$\boldsymbol{Y_{t+1}=\begin{cases} Y_t+\tilde{u}(t,Y_t) & \text{with probability } q \\ Y_t+\tilde{d}(t,Y_t) & \text{with probability } 1-q \end{cases}}$$

since both X_t and Y_t must be measurable with respect to the same sigma-algebra F_t for all t.

It is important to realise that the fact that the processes are measurable with respect to the same information set, implies that $\Delta X_{t+1}=u\left(t,X_t\right)$ if and only if $\Delta Y_{t+1}=\tilde{u}\left(t,X_t\right)$.

Second, since Y_t is a martingale with respect to Q we must have:

$$E_Q\left[Y_{t+1}|F_t\right] = Y_t$$

ie $\boldsymbol{q\tilde{u}(t,Y_t)+(1-q)\tilde{d}(t,Y_t)=0}$

or $\boldsymbol{\tilde{d}(t,Y_t)=\dfrac{-q\tilde{u}(t,Y_t)}{1-q}}$

so that the one-step martingale property is satisfied.

Now consider $\phi_{t+1}=\dfrac{\Delta Y_{t+1}}{\Delta X_{t+1}}$. Since the denominator and numerator are both random white noise terms, it would appear that ϕ_{t+1} will also be random, and not known until time $t+1$. However, consider the outcomes for ϕ_{t+1}.

One possibility is:

$$\phi_{t+1}=\frac{\Delta Y_{t+1}}{\Delta X_{t+1}}=\frac{\tilde{d}\left(t,Y_t\right)}{d\left(t,X_t\right)}$$

On the other hand, we might have:

$$\phi_{t+1}=\frac{\Delta Y_{t+1}}{\Delta X_{t+1}}=\frac{\tilde{u}\left(t,Y_t\right)}{u\left(t,X_t\right)}$$

Now we know that:

$$d(t,X_t) = \frac{-q\,u(t,X_t)}{1-q} \qquad \text{and} \qquad \tilde{d}(t,Y_t) = \frac{-q\,\tilde{u}(t,Y_t)}{1-q}$$

So dividing these gives:

$$\frac{\tilde{u}(t,Y_t)}{u(t,X_t)} = \frac{\tilde{d}(t,Y_t)}{d(t,X_t)}$$

So, at time t there is actually only *one* possible outcome. Hence ϕ_{t+1} is known at time t and therefore previsible.

Then if $\phi_{t+1} = \tilde{u}(t,Y_t)/u(t,X_t)$ (so ϕ_t is previsible: that is F_{t-} or F_{t-1} measurable) we have:

$$Y_t = Y_0 + \sum_{s=1}^{t} \phi_s \Delta X_s$$

where $\Delta X_s = X_s - X_{s-1}$

or $\Delta Y_t = \phi_t \Delta X_t$

What we have done here is to calculate the total change in Y between time 0 and time t, which is found by adding up the steps:

$$Y_t = Y_0 + \Delta Y_1 + \Delta Y_2 + \ldots + \Delta Y_t$$

ie $$Y_t - Y_0 = \Delta Y_1 + \Delta Y_2 + \ldots + \Delta Y_t$$

If Y experiences an up-movement at time s then $\Delta Y_s = \tilde{u}(s-1, Y_{s-1})$, which from the definition of ϕ_s is equal to $\phi_s\, u(s-1, X_{s-1})$. Since $u(s-1, X_{s-1})$ is the size of the up-movement in X at time s, this corresponds to $\phi_s \Delta X_s$ when there is an up-movement. (Down-movements are connected in a similar way.)

But $\Delta Y_t = \phi_t \Delta X_t$. So we can write:

$$Y_t = Y_0 + \phi_1 \Delta X_1 + \phi_2 \Delta X_2 + \ldots + \phi_t \Delta X_t$$

ie $$Y_t = Y_0 + \sum_{s=1}^{t} \phi_s \Delta X_s$$

So, one Q-martingale, Y_t say, can be represented in terms of a different but equivalent Q-martingale, X_t.

More specifically, the change in Y_t over any interval $[t-1,\ t]$ is equal to the corresponding change in X_t, ΔX_t, scaled up by the ratio of the relative sizes of the up-steps in the respective binomial trees.

So, the actual value of Y_t can be represented in terms of its initial value Y_0 and the ΔX_s's.

Question 15.8

Explain in words how ϕ_t is related to the two martingale processes X_t and Y_t.

2.3 Diffusion models

Our proof in continuous time will be based on a diffusion model for the share price.

Diffusion models are characterised by their stochastic differential equation:

$$dX_t = \mu_X(t, X_t)dt + \sigma_X(t, X_t)dZ_t$$

where Z_t is a standard Brownian motion with respect to a measure P.

Recall that μ represents the drift and σ the volatility.

If X_t is a martingale with respect to P then we must have:

$$\mu_X(t, X_t) = 0 \text{ for all } t,\ X_t$$

So, the process X_t has no drift.

If Y_t is also a martingale then we have:

$$dY_t = \sigma_Y(t, Y_t)dZ_t = \phi_t \sigma_X(t, X_t)dZ_t = \phi_t dX_t$$

where $\phi_t = \dfrac{\sigma_Y(t, Y_t)}{\sigma_X(t, X_t)}$ provided $\sigma_X(t, X_t) > 0$ with probability 1.

3 Another look at the binomial model

3.1 Introduction

We are now going to use the definitions and theorems we've built up to put together a proof of the formula $V_t = e^{-r(n-t)} E_Q\left[C_n | F_t\right]$ the fair price for a derivative. First, we should give a worded description of how we are going to use various results. We start by working backwards from what we are trying to show, in order to give an intuitive idea of why the proof involves the steps that it does.

- 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.

You should refer back to this section to help you understand the purpose of any of the steps in the 5-step method, whether the method is in the discrete-time case (as here), or in Chapter 16 for the continuous-time case.

3.2 The 5-step method

Before we look at the continuous-time model let us apply some of these results to the binomial model from Chapter 13. The 5-step proof that you will meet in Chapter 16 has exactly the same structure as the one given in this section.

We work through a series of steps which can be used to solve the problems of pricing and hedging of derivatives.

The idea of this section is to reformulate the risk-neutral pricing method of the binomial model in a way that can be applied to the continuous-time case. In so doing, we show that the fair price for a derivative contract is the discounted value of the expected payoff, where the expectation is taken with respect to the risk-neutral probability measure. This probability measure is defined as the assignment of probabilities that make the discounted share price process a martingale.

The proof relies on constructing a self-financing, replicating portfolio, made out of shares and cash. The martingale representation theorem can be used to guarantee that this portfolio is previsible.

When working through the 5 steps you should bear in mind that we are reformulating the replicating proof from scratch. For example, we don't yet know what q is. We will help illustrate the proof by reference to the numerical example of a 2-period binomial tree discussed in Section 3.3 of Chapter 13.

Step 1

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

In fact, we found such a measure for the binomial model in Chapter 13.

Question 15.9

Find the measure Q (according to this definition) when the share price process is a geometric random walk (*ie* it increases by a factor of u or decreases by a factor of d at each step), as in the binomial model.

Note that this definition is equivalent to finding a measure with respect to which the expected share price evolves at the risk-free rate, *ie* it is a risk-neutral probability measure.

Question 15.10

Recall the numerical example of a 2-period binomial tree discussed in Section 3.3 of Chapter 13. In that example we used the tree to find (amongst other things) the value of the call option at time 1 $\left(V_1(1)=17.195\right)$ when the share price was 60. In the example:

- the risk-free rate was 5% per period
- the risk-neutral probability for the up-step was $q = 0.43588$
- the share price either went up to 80 or down to 50 at time 2 (the exercise date)
- the call option had a strike price of 45.

Suppose that the share price at time 1, S_1, is equal to 60. Find the corresponding value of D_1, the discounted asset process, and show that the martingale property is satisfied.

Step 2 (proposition)

Define $V_t = e^{-r(n-t)}E_Q[C_n \mid F_t]$ where the random variable C_n is the derivative payoff at time n. It is proposed that this is the fair price of the derivative and we will prove this over the next few steps.

This is what the 5-step method is trying to show, *ie* that the fair price for the derivative is the discounted value of the expected payoff, where the expectation is taken with respect to Q found in Step 1. This formula is exactly the same as the one that we saw in Chapter 13.

Question 15.11

Recall again the numerical example in Section 3.3 of Chapter 13. Show that the value of the call option at time 1 when the share price is 60 can be calculated using the above formula.

The quantity V_t is the value at time t of the replicating portfolio we will be using.

Step 3

Let $E_t = B_n^{-1} E_Q[C_n \mid F_t] = e^{-rt} V_t$.

This is a definition of a new process E_t. It is the discounted value of the replicating portfolio, which will turn out to be the same as the discounted value of the derivative process.

B_t^{-1} is sometimes referred to as the discount process. It is equal to the inverse of the value of an initial unit of cash at time t. Hence:

$$B_t^{-1} = e^{-rt}$$

In particular $B_n^{-1} = e^{-rn}$ and so we have:

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

using the definition of V_t from Step 2.

Question 15.12

Recall again the numerical example in Section 3.3 of Chapter 13. Calculate the value of E_1 when the share price is 60 at time 1.

Under Q, E_t is a martingale.

ie **for $s > 0$ $E_Q[E_{t+s} \mid F_t] = E_Q[B_n^{-1} E_Q\{C_n \mid F_{t+s}\} \mid F_t] = B_n^{-1} E_Q[C_n \mid F_t] = E_t$**

by the *tower property* of conditional expectation.

Step 4

Steps 1 and 3 have given us two martingales with respect to the equivalent martingale measure *Q*, namely D_t and E_t.

Since the measure *Q* is the unique martingale measure, by the martingale representation theorem there exists a previsible process ϕ_t (that is, ϕ_t is F_{t-1}-measurable) such that:

$$\begin{aligned}\Delta E_t &\equiv E_t - E_{t-1} \\ &= \phi_t (D_t - D_{t-1}) \\ &\equiv \phi_t \Delta D_t\end{aligned}$$

The next few equations use the binomial tree notation introduced in Section 4.3 of Chapter 13, where the number in brackets shows the position, counting from the top of the tree. It is assumed that the share price is in state j at time $t-1$.

Let us see if we can establish what ϕ_t is. Now:

$$\Delta E_t = \begin{cases} e^{-rt}V_t(2j-1) - e^{-r(t-1)}V_{t-1}(j) & \text{if up} \\ e^{-rt}V_t(2j) \quad - e^{-r(t-1)}V_{t-1}(j) & \text{if down} \end{cases}$$

and $$\Delta D_t = \begin{cases} e^{-rt}S_{t-1}(j)u_{t-1}(j) - e^{-r(t-1)}S_{t-1}(j) & \text{if up} \\ e^{-rt}S_{t-1}(j)d_{t-1}(j) - e^{-r(t-1)}S_{t-1}(j) & \text{if down} \end{cases}$$

In fact, we found the relationship between the derivative values in Chapter 13, which we derived using a replicating portfolio.

Recall that $V_{t-1}(j) = e^{-r}\left(q_{t-1}(j)V_t(2j-1) + (1-q_{t-1}(j))V_t(2j)\right)$.

We can substitute this expression for $V_{t-1}(j)$ and simplify.

Then we can see that:

$$\Delta E_t = \begin{cases} e^{-rt}\left(V_t(2j-1) - V_t(2j)\right)\left(1-q_{t-1}(j)\right) \\ e^{-rt}\left(V_t(2j-1) - V_t(2j)\right)\left(-q_{t-1}(j)\right) \end{cases}$$

We also had a formula for the risk-neutral probability q.

Furthermore, since $q_{t-1}(j) = \dfrac{e^r - d_{t-1}(j)}{u_{t-1}(j) - d_{t-1}(j)}$, we can also see that:

$$\Delta D_t = \begin{cases} e^{-rt} S_{t-1}(j)\left(u_{t-1}(j) - d_{t-1}(j)\right)\left(1 - q_{t-1}(j)\right) \\ e^{-rt} S_{t-1}(j)\left(u_{t-1}(j) - d_{t-1}(j)\right)\left(-q_{t-1}(j)\right) \end{cases}$$

Therefore $\Delta E_t = \phi_t \Delta D_t$ where:

$$\begin{aligned} \phi_t(j) = \frac{\Delta E_t(j)}{\Delta D_t(j)} &= \frac{E_t(2j-1) - E_{t-1}(j)}{D_t(2j-1) - D_{t-1}(j)} \\ &= \frac{E_t(2j-1) - \left(q_{t-1}(j) E_t(2j-1) + \left(1 - q_{t-1}(j)\right) E_t(2j)\right)}{D_t(2j-1) - \left(q_{t-1}(j) D_t(2j-1) + \left(1 - q_{t-1}(j)\right) D_t(2j)\right)} \\ &= \frac{E_t(2j-1) - E_t(2j)}{D_t(2j-1) - D_t(2j)} \\ &= \frac{V_t(2j-1) - V_t(2j)}{S_{t-1}(j)\left[u_{t-1}(j) - d_{t-1}(j)\right]} \end{aligned}$$

ie $\quad \phi_t = \dfrac{V_t(2j-1) - V_t(2j)}{S_{t-1}(j)\left(u_{t-1}(j) - d_{t-1}(j)\right)}$

Note that we've used $\phi_t(j)$ to emphasise the fact that ϕ_t does depend on j in general.

So we've shown that the martingale representation theorem does work in this case.

As we expected this is what we found in Chapter 13.

Question 15.13

Recall again the numerical example in Section 3.3 of Chapter 13. Calculate the value of ϕ_2 when the share price is 60 at time 1.

So, we define $\phi_t = \dfrac{\Delta E_t}{\Delta D_t} = \dfrac{E_t - E_{t-1}}{D_t - D_{t-1}}$ and the martingale representation theorem guarantees us that this is previsible. ϕ_t will be the number of shares in our replicating portfolio.

Step 5

Let $\psi_t = E_{t-1} - \phi_t D_{t-1}$. Note that ψ_t is previsible by construction, the other previsible process needed to apply the martingale representation theorem. ψ_t will be the number of units of the cash bond in our replicating portfolio.

In this final step we need to show that we can construct a previsible, self-financing, replicating portfolio.

Between times $t-1$ and t^- suppose that we hold the portfolio consisting of:

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

So we hold the portfolio $(\phi_t = \Delta E_t / \Delta D_t, \psi_t = E_{t-1} - \phi_t D_{t-1})$.

The value of this portfolio at time $t-1$ is:

$$\phi_t S_{t-1} + \psi_t B_{t-1} = e^{r(t-1)}\left(\phi_t D_{t-1} + \psi_t\right) = e^{r(t-1)} E_{t-1} = V_{t-1}$$

At time t the value of the portfolio will be:

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

Writing D_t as $D_{t-1} + \Delta D_t$ gives:

$$= e^{rt}\left(\phi_t D_{t-1} + \psi_t + \phi_t \Delta D_t\right)$$

Then using $\psi_t = E_{t-1} - \phi_t D_{t-1}$ and $\Delta E_t = \phi_t \Delta D_t$ gives:

$$= e^{rt}\left(E_{t-1} + \Delta E_t\right)$$

$$= e^{rt} E_t$$

$$= V_t$$

Therefore the portfolio is self-financing. Furthermore, $V_n = C_n$, the derivative payoff at time n.

The formula in Step 2 above gives, $V_n = e^{-r(n-n)} E_Q\left[C_n | F_n\right] = C_n$.

Therefore the hedging strategy (ϕ_t, ψ_t) is replicating, implying that:

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

is the fair price at time t for this contract.

This completes the proof, *ie* that the formula proposed in Step 2 is correct, which therefore proves the result we derived using the binomial tree approach in Chapter 13.

Question 15.14

Describe in words what the above formula says about the fair price for the derivative at time t.

Question 15.15

Recall again the numerical example in Section 3.3 of Chapter 13. Calculate the value of ψ_2 when the share price is 60 at time 1 and use it to calculate the corresponding option price.

3.3 Summary

Let us summarise here what we have done because we will repeat the steps when we look at the continuous-time model in the next chapter.

- **We have established the equivalent martingale measure Q.**
- **We have proposed a fair price, V_t, for a derivative and its discounted value $E_t = e^{-rt}V_t$.**
- **We have used the martingale representation theorem to construct a hedging strategy (ϕ_t, ψ_t).**
- **We have shown that this hedging strategy replicates the derivative payoff at time n.**
- **Therefore V_t is the fair value of the derivative at time t.**

Here "hedging" is used as a synonym for "replicating".

Question 15.16

Have we assumed in this proof that the markets are arbitrage-free?

4 Exam-style question

We finish this chapter with an exam-style question.

Question

(i) A risky asset has value A_t at time *t*. A probability measure R is defined so that $A_t v^t$ is an R-martingale, where v is calculated using the risk-free interest rate. Explain why R can be described as a risk-neutral probability measure. [2]

(ii) Let $X_t = v^T E_R\left[C \middle| F_t\right]$, where C is a discrete random variable occurring at time $T > t$. Prove that X_t is an R-martingale. [2]

Hint: if X is a discrete random variable and $\underline{Y}$ is a vector of random variables, then $E\left[E\left[X \middle| \underline{Y}\right]\right] = E\left[X\right]$.

(iii) Stating any results that you use, deduce that θ_t is previsible, where $dX_t = \theta_t dD_t$, where $D_t = A_t v^t$. [2]

[Total 6]

Solution

(i) ***Risk-neutral R***

If $A_t v^t$ is an R-martingale then we have:

$$E_R\left[A_t v^t \middle| F_s\right] = A_s v^s$$

$$\Leftrightarrow \qquad E_R\left[A_t \middle| F_s\right] = A_s v^{s-t} = A_s(1+i)^{t-s}$$

So, using R as our probability measure we see that, statistically speaking, we expect A_t to "act like cash", *ie* to increase at the risk-free rate. The riskiness within it has been "neutralised" and therefore R is the risk-neutral probability measure.

(ii) ***Prove that X_t is a martingale***

For a martingale we need:

$$E_R\left[X_t \middle| F_s\right] = X_s$$

The LHS is:

$$E_R\left[X_t \middle| F_s\right] = E_R\left[v^T E_R\left[C \middle| F_t\right] \middle| F_s\right]$$

By analogy with the hint in the question, we have:

$$E_R\left[E_R\left[C \middle| F_t\right] \middle| F_s\right] = E_R\left[C \middle| F_s\right]$$

ie the tower law. So, we can simplify the RHS of the previous equation to give:

$$v^T E_R\left[E_R\left[C \middle| F_t\right] \middle| F_s\right] = v^T E_R\left[C \middle| F_s\right] = X_s$$

(iii) ***Previsible*** $\boldsymbol{\theta_t}$

We make use of the martingale representation theorem in the following form:

Let D_t and X_t be R-martingales and let D_t have non-zero volatility. Then there exists a unique process θ_t such that $dX_t = \theta_t dD_t$. Furthermore, θ_t is previsible.

Since θ_t satisfies this equation and the conditions for the theorem to apply, we conclude that θ_t is previsible.

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

Chapter 15 Summary

Preliminary concepts

ϕ_t is *previsible* if it is known based on information up to but not including time *t*.

A portfolio (ϕ_t, ψ_t) is *self-financing* if ϕ_t, ψ_t are previsible and $dV_t = \phi_t dS_t + \psi_t dB_t$,

ie the change in the value of the portfolio is equal to the pure instantaneous investment gain.

Consider a derivative with random variable payoff X at time T. A self-financing portfolio V_t is a *replicating* portfolio for *X* if $V_T = X$.

So, for an initial investment of V_0 at time 0, if we follow the self-financing portfolio strategy we will be able to reproduce the derivative payment exactly and without risk.

An investment market is *complete* if for every derivative in that market, there exists a replicating strategy for that derivative.

X_t *is a* Q-martingale *if* $E_Q[X_u \mid F_t] = X_t$ *whenever* $t < u$.

The Cameron-Martin-Girsanov theorem

Suppose that Z_t is a standard Brownian motion under P. Furthermore suppose that γ_t is a previsible process. Then there exists a measure Q equivalent to P and where:

$$\tilde{Z}_t = Z_t + \int_0^t \gamma_s ds$$

is a standard Brownian motion under Q.

Conversely, if Z_t is a standard Brownian motion under P and if Q is equivalent to P then there exists a previsible process γ_t such that:

$$\tilde{Z}_t = Z_t + \int_0^t \gamma_s ds$$

is a Brownian motion under Q.

Martingale representation theorem (discrete time version)

Let X_t be a P-martingale and let Y_t be a second P-martingale. Then there exists a unique previsible process ϕ_t such that:

$$Y_t = Y_0 + \sum_{s=1}^{t} \phi_s \Delta X_s \quad \text{where} \quad \Delta X_s = X_s - X_{s-1}$$

or $$\Delta Y_t = \phi_t \Delta X_t$$

if and only if there is no other measure equivalent to P under which X_t is a martingale.

Martingale representation theorem (continuous time version)

Let X_t be a P-martingale and let Y_t be a second P-martingale. Then there exists a unique previsible process ϕ_t such that:

$$Y_t = Y_0 + \int_0^t \phi_s dX_s$$

or $$dY_t = \phi_t dX_t$$

if and only if there is no other measure equivalent to P under which X_t is a martingale.

The 5-step approach

The binomial model result can be proved using the martingale approach, which consists of five steps:

Step 1

Find the equivalent martingale measure Q under which $D_t = e^{-rt} S_t$ is a martingale.

Step 2

Let $V_t = e^{-r(n-t)} E_Q\left[C_n | F_t\right]$ where C_n is the derivative payoff at time n. This is proposed as the fair price of the derivative at time t.

Step 3

Let $E_t = e^{-rn} E_Q\left[C_n | F_t\right] = e^{-rt} V_t$. This is a martingale under Q.

Step 4

By the martingale representation theorem, there exists a previsible process ϕ_t such that $\Delta E_t = \phi_t \Delta D_t$.

Step 5

Let $\psi_t = E_{t-1} - \phi_t D_{t-1}$ and at time t hold the portfolio:

- ϕ_t units of the tradable asset S_t
- ψ_t units of the cash account.

At time $t-1$ the value of this portfolio is equal to V_{t-1}. At time t the value of this portfolio is equal to V_t. Also V_T = X. Therefore, the hedging strategy (ϕ_t, ψ_t) is replicating and so V_t is the fair price at time t.

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

Chapter 15 Solutions

Solution 15.1

F represents the set of all information concerning the stock price process S_t that could (eventually) be known.

Solution 15.2

Recall that the pure investment gain is $\phi_t dS_t + \psi_t dB_t$.

In this equation, dB_t represents the infinitesimal change in the value of one unit of cash during the time interval $(t, t + dt)$. dS_t represents the corresponding change in the stock price, which is a random variable.

The absence of any new net investment into or out of the portfolio means that there are no new purchases or sales of either the stock or the cash bond. Hence, ϕ_t and ψ_t remain unchanged over this short time interval. Consequently, any changes to the value of the portfolio over the time interval $(t, t + dt)$ must arise solely through changes in the stock price dS_t and /or the increase in the value of cash dB_t.

Solution 15.3

The time of payment U is fixed for:

- a European option, which can be exercised only at the fixed exercise or expiry date
- a future or a forward, both of which have a fixed delivery date.

The time of payment is not fixed for an American option, where the holder can choose to exercise the option at any time on or before the exercise date.

Solution 15.4

As was the case in Chapter 13, this restriction is required in order to avoid any violation of the no-arbitrage condition – as we are pricing the derivative using an arbitrage-free approach. If arbitrage opportunities were present, then the value of our replicating portfolio might differ from that of the derivative claim that we are trying to price.

Recall that in order to avoid arbitrage we must have $d < e^r < u$, where u and d are the up-step and down-step factors respectively in our binomial tree. Recall from Chapter 13 that:

$$q = \frac{e^r - d}{u - d}$$

If $0 < q < 1$, then:

$$0 < q = \frac{e^r - d}{u - d} < 1$$

Multiplying this through by $u - d$ gives:

$$0 < e^r - d < u - d$$

ie $\quad d < e^r < u$

Solution 15.5

In continuous time, the probability that a random variable takes an exact value will be zero. *ie* $P[X = x] = 0$. We must work with areas to get non-zero probabilities, $P[x - 0.5 < X < x + 0.5] \neq 0$. You can think of a sample path as the realisation of a "collection of random outcomes up to a certain point" and this logic applies equally to paths.

Solution 15.6

We know that:

$$\log S_t = \log S_0 + (\mu - ½\sigma^2)t + \sigma Z_t$$

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

This means that S_t has a lognormal distribution with the same parameters. Using the formula in the Tables, the mean of S_t is:

$$E[S_t] = \exp\left\{\log S_0 + (\mu - ½\sigma^2)t + ½\sigma^2 t\right\} = S_0 e^{\mu t}$$

Since r is a constant, we then have:

$$E[e^{-rt} S_t] = e^{-rt} S_0 e^{\mu t} = S_0 e^{(\mu - r)t}$$

Solution 15.7

We have just seen that:

$$q = \frac{-d(t, X_t)}{u(t, X_t) - d(t, X_t)}$$

Rearranging this gives:

$$d(t, X_t) = \frac{-q\, u(t, X_t)}{1 - q}$$

Solution 15.8

ϕ_t is equal to the ratio of the changes in the two martingale processes, and it reflects the relative *volatilities* of the two processes.

Solution 15.9

Recall that the martingale condition requires that:

$$E_Q[D_{t+1}|F_t] = D_t$$

ie $\quad E_Q[e^{-r(t+1)}S_{t+1}|F_t] = e^{-rt}S_t$

$\Leftrightarrow \quad q\,u\,S_t + (1-q)\,d\,S_t = S_t\,e^r$

$\Leftrightarrow \quad q\,u + (1-q)\,d = e^r$

$\Leftrightarrow \quad q = \dfrac{e^r - d}{u-d}$

Solution 15.10

Using the definition of D_1 we have:

$$D_1 = e^{-r\times 1}S_1 = e^{-0.05\times 1}\times 60 = 57.074$$

We can also find the two possible values of D_t at time 2. If the share price goes up to 80, then:

$$D_2 = e^{-r\times 2}S_2 = e^{-0.05\times 2}80 = 72.387$$

If the share price goes down to 50, then:

$$D_2 = e^{-r\times 2}S_2 = e^{-0.05\times 2}50 = 45.242$$

Thus, using the probability measure with $q = 0.43588$:

$$E\left[D_2 \mid F_1\right] = 72.387\times 0.43588 + 45.242(1-0.43588) = 57.074$$

ie the martingale property is satisfied.

Solution 15.11

Applying the formula in Step 2 of the proof gives:

$$V_1 = e^{-r(2-1)} E_Q\left[C_2|F_1\right] = e^{-0.05\times 1}\left[35\times 0.43588 + 5\times(1-0.43588)\right] = 17.195$$

This agrees with the value that we found in Chapter 13.

Solution 15.12

Applying the formula just given, we get:

$$E_1 = B_2^{-1} E_Q\left[C_2|F_1\right] = e^{-0.05\times 2}\left[35\times 0.43588 + 5\times(1-0.43588)\right] = 16.356$$

Alternatively, we can calculate it by discounting the value of V_1, which we found in the previous question:

$$E_1 = e^{-r}V_1 = e^{-0.05}\times 17.195 = 16.356$$

Solution 15.13

Applying the formula in the Core Reading gives:

$$\phi_2 = \frac{35-5}{80-50} = 1$$

NB A value of 1 is obtained here because the option is in-the-money whether the share price goes up or down.

Solution 15.14

The expression is:

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

It says that:

- the value of the derivative at time t is equal to the expectation of the derivative payoff at time n
- taken with respect to both the risk-neutral probability measure Q and the information set F_t, generated by the history of the stock price movements up to and including time t
- discounted at the continuously-compounded risk-free rate of return r.

Solution 15.15

Here:

$$\psi_2 = E_1 - \phi_2 D_1 = 16.356 - 1 \times 57.074 = -40.718$$

So the value of the call option at time 1 when the share price is 60 is given by:

$$\phi_1 S_1 + \psi_1 B_1 = 1 \times 60 - 40.718 \times e^{0.05} = 17.195$$

This value again agrees with the answer found in Section 3.3 of Chapter 13.

Solution 15.16

Yes. Otherwise, it would not necessarily follow that the replicating portfolio has the same value as the derivative.

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 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.

Q

Question 16.2 (Revision)

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

Q

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 (page 10) 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:

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

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

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*.

Q

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

Suppose that $X = \max\{S_T - K, 0\}$. Then:

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

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

Given F_t (so we are at time t and we know the history up until time t), **we have:**

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

where $\tilde{Z}_T - \tilde{Z}_t \sim N(0, T-t)$ under Q. Therefore S_T is log-normally distributed.

In fact:

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

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

Let us write:

$$\boldsymbol{e^{-r(T-t)} S_T \mid F_t = \exp(\alpha + \beta U)}$$

where:

- $\boldsymbol{U \sim N(0,1)}$ **under** $\boldsymbol{Q}$
- $\boldsymbol{\alpha = \log S_t - \tfrac{1}{2}\sigma^2 (T-t)}$ **and** $\boldsymbol{\beta = \sigma\sqrt{T-t}}$ (1)

Now we use the derivative pricing formula and let $Ke^{-r(T-t)} = e^{\alpha + \beta u}$.

Then:

$$\begin{aligned} \boldsymbol{V_t} &= \boldsymbol{e^{-r(T-t)} E_Q[\max\{S_T - K, 0\} \mid F_t]} \\ &= \boldsymbol{E_Q\left[\max\left\{e^{\alpha+\beta U} - e^{\alpha+\beta u}, 0\right\}\right]} \end{aligned}$$

where:

$$\boldsymbol{u = \frac{\log\left(Ke^{-r(T-t)}\right) - \alpha}{\beta}} \tag{2}$$

$$\boldsymbol{= \frac{\log\dfrac{K}{S_t} - (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}}$$

Question 16.8

Check that this is the correct formula for u.

If we take out the factor $e^{\alpha+\beta u}$, this becomes:

$$V_t = e^{\alpha+\beta u} E_Q\left[\max\left\{e^{\beta(U-u)} - 1, 0\right\}\right]$$

Using the definition of the expectation for a continuous random variable, we have:

$$V_t = e^{\alpha+\beta u} \int_{-\infty}^{\infty} \max\left\{e^{\beta(v-u)} - 1, 0\right\} \phi(v) dv$$

Since the maximum will be zero when $v < u$, this is:

$$\boldsymbol{V_t = e^{\alpha+\beta u} \int_u^\infty \left(e^{\beta(v-u)} - 1\right) \phi(v) dv}$$

where $\phi(v)$ is the probability density function of the standard normal distribution

ie $\qquad \boldsymbol{\phi(v) = (2\pi)^{-1/2} \exp\left(-\tfrac{1}{2} v^2\right)}$

So:

$$\begin{aligned}
V_t &= e^{\alpha} \int_u^\infty \frac{1}{\sqrt{2\pi}} e^{\beta v - \frac{1}{2} v^2} dv - e^{\alpha+\beta u} \left(1 - \Phi(u)\right) \\
&= e^{\alpha+\frac{1}{2}\beta^2} \int_u^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(v-\beta)^2} dv - e^{\alpha+\beta u} \Phi(-u) \\
&= e^{\alpha+\frac{1}{2}\beta^2} P\left[N(\beta,1) > u\right] - e^{\alpha+\beta u} \Phi(-u) \\
&= e^{\alpha+\frac{1}{2}\beta^2} \left(1 - \Phi(u-\beta)\right) - e^{\alpha+\beta u} \Phi(-u) \\
&= e^{\alpha+\frac{1}{2}\beta^2} \Phi(\beta - u) - e^{\alpha+\beta u} \Phi(-u) \qquad (3)
\end{aligned}$$

But, using (1) above we have:

$$\alpha + \frac{1}{2}\beta^2 = \log S_t - \frac{1}{2}\sigma^2(T-t) + \frac{1}{2}\sigma^2(T-t) = \log S_t$$

We have also seen that we can rearrange equation (2) to get:

$$\alpha + \beta u = \log\left(Ke^{-r(T-t)}\right)$$

Finally, using (1) and (2) we get:

$$\begin{aligned}\beta - u &= \sigma\sqrt{T-t} - \frac{\log\frac{K}{S_t} - (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ &= \frac{\log\frac{S_t}{K} + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ &= d_1\end{aligned}$$

and:

$$-u = (\beta - u) - \beta = d_1 - \sigma\sqrt{T-t} = d_2$$

Therefore, substituting all these expressions into equation (3), we get:

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

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

1.5 Example: replication of the payoff on a European call option

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 16.9***
>
> Derive these three relationships.

By Ito's formula:

$$\begin{aligned}
dV_t &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial s}dS_t + \frac{1}{2}\frac{\partial^2 V}{\partial s^2}(dS_t)^2 \\
&= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial s^2}\right)dt + \frac{\partial V}{\partial s}B_t(rD_t dt + dD_t) \\
&= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial s^2} + rS_t \frac{\partial V}{\partial s}\right)dt + \frac{\partial V}{\partial s}B_t dD_t
\end{aligned}$$

Then, using the third equation above:

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

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:

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

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:

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

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

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

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

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

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

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

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 16.10

What is the replicating strategy for a European call option?

Question 16.11

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 16.12

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 16.13

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.

Question 16.14

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 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:

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

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:

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

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:

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

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

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:

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

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

By analogy with the non-dividend-paying stock:

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

where:

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

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

Also:

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

and:

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

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

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 16.15

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

Question 16.16

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:

$$
\begin{aligned}
df\left(S_t\right) = d(\log S_t) &= \frac{1}{S_t} dS_t + \frac{1}{2}\left(-\frac{1}{S_t^2}\right)(dS_t)^2 \\
&= \frac{1}{S_t}(rS_t\, dt + \sigma S_t\, dB_t) - \frac{1}{2S_t^2}(rS_t\, dt + \sigma S_t\, dB_t)^2 \\
&= r\, dt + \sigma\, dB_t - \tfrac{1}{2}\sigma^2\, dt \\
&= (r - \tfrac{1}{2}\sigma^2)dt + \sigma dB_t
\end{aligned}
$$

Changing the t's to s's and integrating this equation between limits of $s = 0$ and $s = t$, we get:

$$\left[\log S_s\right]_{s=0}^{s=t} = (r - \tfrac{1}{2}\sigma^2)\int_0^t ds + \sigma \int_0^t dB_s$$

$$\Rightarrow \qquad \log S_t - \log S_0 = (r - \tfrac{1}{2}\sigma^2)\, t + \sigma B_t$$

$$\Rightarrow \qquad S_t = S_0\, e^{(r - ½\sigma^2)t + \sigma B_t}$$

(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:

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

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

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

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

Put option

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

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

$$dS_t = \left[\mu\, dt + \sigma\, dZ_t\right] S_t$$

$$\begin{aligned}
\Rightarrow \quad (dS_t)^2 &= \left[\mu^2 (dt)^2 + \sigma^2 (dZ_t)^2 + 2\mu\sigma\, dt\, dZ_t\right] S_t^2 \\
&= \left[\mu^2 (dt)^2 + \sigma^2 (dt) + 2\mu\sigma\, dt\, dZ_t\right] S_t^2 \\
&= \sigma^2 S_t^2 dt
\end{aligned}$$

Using these results and Ito's lemma, we get:

$$\begin{aligned} df(S_t,t) &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S_t}dS_t + \frac{1}{2}\frac{\partial^2 f}{\partial S_t^2}(dS_t)^2 \\ &= \Theta dt + \Delta dS_t + \frac{1}{2}\Gamma\sigma^2 S_t^2 dt \end{aligned}$$

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:

$$\begin{aligned} dV_t &= -df(S_t,t) + \Delta\left(dS_t + qS_t dt\right) \\ &= -\left\{\Theta dt + \Delta dS_t + \frac{1}{2}\Gamma\sigma^2 S_t^2 dt\right\} + \Delta\left(dS_t + qS_t dt\right) \\ &= -\left\{\Theta - q\Delta S_t + \frac{1}{2}\Gamma\sigma^2 S_t^2\right\} dt \end{aligned}$$

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 of 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 $Z_t - Z_s \sim N(0, t-s)$

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 $E[X_t | F_s] = X_s$

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

Start from the equation defining u :

$$Ke^{-r(T-t)} = e^{\alpha+\beta u}$$

Take logs:

$$\log\left(Ke^{-r(T-t)}\right) = \alpha + \beta u$$

Make u the subject:

$$u = \frac{\log\left(Ke^{-r(T-t)}\right) - \alpha}{\beta} = \frac{\log K - r(T-t) - \alpha}{\beta}$$

Now substitute the definitions of α and β:

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

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

Solution 16.9

First equation

Since $S_t = S_0 e^{(r-\frac{1}{2}\sigma^2)t + \sigma\tilde{Z}_t}$, we have:

$$D_t = e^{-rt} S_t = e^{-rt} S_0 e^{(r-\frac{1}{2}\sigma^2)t+\sigma\tilde{Z}_t} = S_0 e^{-\frac{1}{2}\sigma^2 t + \sigma\tilde{Z}_t}$$

Let $X_t = -\tfrac{1}{2}\sigma^2 t + \sigma\tilde{Z}_t$, so that $dX_t = -\tfrac{1}{2}\sigma^2 dt + \sigma d\tilde{Z}_t$ and $D_t = e^{X_t}$.

Now apply Ito's formula:

$$dD_t = \left\{-\tfrac{1}{2}\sigma^2 e^{X_t} + \tfrac{1}{2}\sigma^2 e^{X_t}\right\} dt + \sigma e^{X_t} d\tilde{Z}_t = \sigma D_t d\tilde{Z}_t$$

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:

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t$$

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 16.10

Hold:

- $\phi_t = \dfrac{\partial V}{\partial S} = \Phi(d_1)$ 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 16.11

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 $$\begin{aligned} V_t - \phi_t S_t &= 475\Phi(d_1) - 500e^{-0.05\times 2/12}\Phi(d_1 - 0.2\sqrt{2/12}) \\ &\quad -0.314 \times 475 = -142 \text{ cash} \end{aligned}$$

Solution 16.12

Because it doesn't require an understanding of stochastic calculus, the PDE approach is quicker and easier to describe, and more easily understood.

Solution 16.13

It's a martingale. See Question 16.4.

Solution 16.14

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 16.15

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 16.16

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.

Chapter 17

The term structure of interest rates

Syllabus objectives

(x) Demonstrate a knowledge and understanding of models of the term structure of interest rates.

1. *Describe the desirable characteristics of a model for the term structure of interest rates.*

2. *Describe, as a computational tool, the risk-neutral approach to the pricing of zero-coupon bonds and interest-rate derivatives for a general one-factor diffusion model for the risk-free rate of interest.*

3. *Describe, as a computational tool, the approach using state price deflators to the pricing of zero-coupon bonds and interest-rate derivatives for a general one-factor diffusion model for the risk-free rate of interest.*

4. *Demonstrate an awareness of the Vasicek, Cox-Ingersoll-Ross and Hull & White models for the term structure of interest rates.*

5. *Discuss the limitations of these one-factor models and show an awareness of how these issues can be addressed.*

0 Introduction

In this chapter we will look at stochastic models for the term structure of interest rates. In particular, we will focus on models that are framed in continuous time and are arbitrage-free.

Continuous-time models are generally used for their greater tractability compared to discrete-time ones. An arbitrage-free model of the term structure is one that generates arbitrage-free bond prices.

There are three main types of models used to describe interest rates mathematically:

1. The *Heath-Jarrow-Morton* approach uses an Ito process to model the forward rate for an investment with a fixed maturity. We will not consider this approach here.

2. *Short-rate models* use an Ito process to model the short rate. We will look at three specific models of this type: the Vasicek model, the Cox-Ingersoll-Ross model and the Hull & White model.

3. *State price deflator models* assume that bond prices can be modelled in terms of a deflator function, which is assumed to follow an Ito process.

As you can see, Ito processes are a key feature of these models. So you might find it helpful to review the topics of Brownian motion, Ito's formula and stochastic differential equations from Chapter 9.

1 Notation and preliminaries

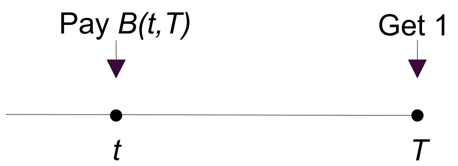
1.1 Notation

Modelling interest rates is more complicated than modelling share prices because interest rates depend not only on the current time (which we will denote by t), but also on the term of the investment. For example, an investor with a 10-year bond will normally earn a different rate of interest than an investor with a 5-year bond.

We will make use of the following notation:

- **$B(t,T)$ = zero-coupon bond price**

  = price at t for £1 payable at T



- **$r(t)$ = instantaneous risk-free rate of interest at t**

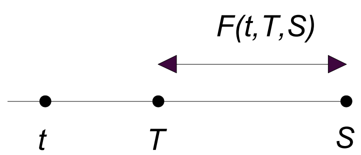
  So $r(t)$ is the force of interest that applies in the market at time t. It is often referred to as the *short rate*.

- **$C(t)$ = unit price for investment at the risk-free rate**

  So $C(t)$ is the accumulated value at time t of 1 unit invested at time 0.

- **$F(t,T,S)$ = forward rate at t for delivery between T and S**

  At time t we can consider the market prices of two investments, one maturing at time T and one at a later time S. These two prices will imply a certain constant force of interest applicable between time T and time S. This is $F(t,T,S)$. In an arbitrage-free market, it represents the force of interest at which we can agree at time t to borrow or lend over the period from T to S.



- **$f(t,T)$ = instantaneous forward-rate curve**

  So $f(t,T)$ is the force of interest at future time T implied by the current market prices at time t.

- **$R(t,T)$ = spot-rate (zero-coupon yield) curve**

  The spot rate $r(t,T)$ is the constant force of interest applicable over the period from time t to time T that is implied by the market prices at time t.

Question 17.1

Suppose that the current time corresponds to $t=5$ and that the force of interest has been a constant 4% *pa* over the last 5 years. Suppose also that the force of interest implied by current market prices is a constant 4% *pa* for the next 2 years and a constant 6% *pa* thereafter.

If $T=10$ and $S=15$, write down or calculate each of the six quantities $B(t,T)$, $r(t)$, $C(t)$, $F(t,T,S)$, $f(t,T)$ and $R(t,T)$ using the notation above.

In real life, interest rates will have a more complex pattern, or *term structure*.

Figure 17.1 below shows the yield curve for the UK Government bond market on 31 December 2003. The "term" plotted on the x-axis corresponds to $T-t$.

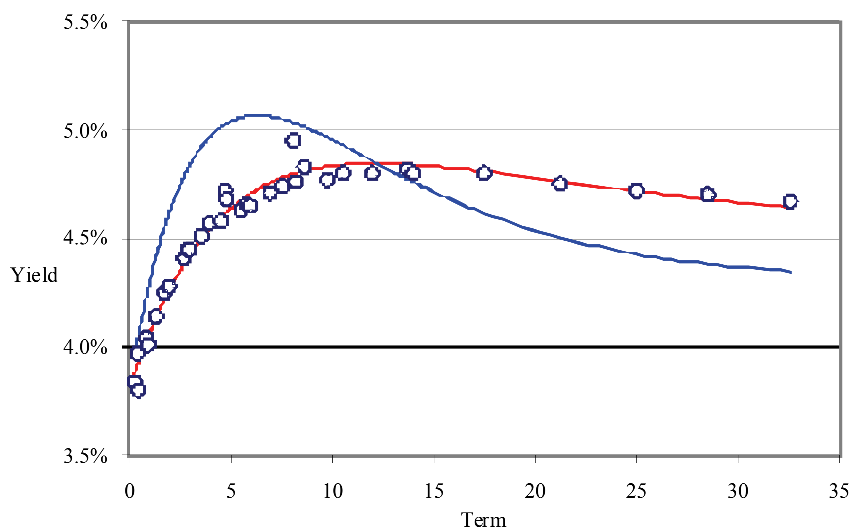


Figure 17.1: UK yield curves 31 December 2003

The circles show the remaining term and the gross redemption yield for each of the available bonds. These approximate to the spot rates $R(t,T)$ for zero-coupon bonds with corresponding terms. A mathematical curve has been fitted to these points. At this particular time the yield curve had a humped shape.

The other curve shown is the forward rate curve $f(t,T)$.

1.2 Relationships between interest rates and bond prices

The six quantities defined above can all be used to describe the term structure of interest rates in a particular market. It is important to know the interrelationships between these quantities.

Zero-coupon bond prices

Zero-coupon bond prices are related to the spot-rate and forward-rate curves in the following way:

$$R(t,T) = -\frac{1}{T-t}\log B(t,T) \quad \text{for } t<T$$

or $$B(t,T) = \exp\left[-R(t,T)(T-t)\right]$$

These two equations involving the spot rate $R(t,T)$ are equivalent. The second one is analogous to the relationship $v^n = e^{-\delta n}$ from Subject 102 or CT1. The forward rate is related to the zero-coupon bond price as follows:

$$F(t,T,S) = \frac{1}{S-T}\log\frac{B(t,T)}{B(t,S)} \quad \text{for } t<T<S$$

Q

Question 17.2

Derive this relationship.

The instantaneous forward rate $f(t,T)$ is the forward rate over an instant of time. So it is the limit of the forward rate $F(t,T,S)$ as S approaches T.

$$f(t,T) = \lim_{S\to T} F(t,T,S) = -\frac{\partial}{\partial T}\log B(t,T)$$

or $\boldsymbol{B(t,T) = \exp\left[-\int_t^T f(t,u)du\right]}$

Question 17.3

How does $f(t,T) = -\dfrac{\partial}{\partial T}\log B(t,T)$ follow?

We can also deduce the following relationship, which shows that the spot rate is an average of the forward rates:

$$R(t,T) = -\frac{1}{T-t}\log B(t,T) = \frac{1}{T-t}\int_t^T f(t,u)du$$

The final formula indicates a similarity between the forward-rate curve, $\boldsymbol{f(t,T)}$, in this stochastic setting and the force of interest in a deterministic setting.

In Subject CT1 the right-hand side of the equation for $B(t,T)$ would have been written as $\exp\left[-\int_t^T \delta(u)du\right]$, which is the factor that would be used for discounting the value of a payment from time T to time t.

Cash account

We will usually refer to $\boldsymbol{C(t)}$ as the cash account. $\boldsymbol{C(t)}$ represents the value at time $\boldsymbol{t}$ of £1 invested at time 0 at the risk-free rate of interest. Thus the stochastic differential equation for $\boldsymbol{C(t)}$ is:

$$\boldsymbol{dC(t) = r(t)C(t)dt}$$

That is, the investment gain made on a unit of the cash account from $\boldsymbol{t}$ to $\boldsymbol{t+dt}$ is equal to the interest earned at the risk-free rate over that short period.

This equation may not appear to be stochastic since it doesn't contain a Brownian increment term dW_t. But don't forget that the current interest rate $r(t)$ can vary randomly as time progresses.

The solution to the SDE is:

$$\boldsymbol{C(t) = \exp\left[\int_0^t r(u)du\right]}$$

Question 17.4

Derive this solution from the SDE:

$$dC(t) = r(t)C(t)dt \, .$$

Models that describe the dynamics of $B(t,T)$, $r(t)$, $f(t,T)$ and $R(t,T)$ over time are called, amongst other things, *term structure models* or *models for the term structure of interest rates* or *interest rate models*. We will see later how these processes must be related if we are to have an arbitrage-free model.

Question 17.5

Under one particular term structure model:

$$f(t,T) = 0.03e^{-0.1(T-t)} + 0.06(1 - e^{-0.1(T-t)}) \, .$$

Sketch a graph of $f(t,T)$ as a function of T, and derive expressions for $B(t,T)$ and $R(t,T)$.

2 Desirable characteristics of a term structure model

Question 17.6

What do you think term structure models are used for?

We will now discuss characteristics of a term structure model that are regarded as desirable features.

- **The model should be arbitrage free.**

  In very limited circumstances this is not essential, but in the majority of modern actuarial applications, this is essential. Most obviously, anything involving dynamic hedging would immediately identify and exploit any arbitrage opportunities.

  The markets for Government bonds and interest rate derivatives are generally assumed to be pretty much arbitrage-free in practice.

- **Interest rates should be positive.** Negative *nominal* interest rates are almost non-existent in practice, though negative *real* interest rates have arisen from time to time.

  Banks will always have to offer investors a positive return to prevent them from withdrawing paper cash and putting it "under the bed". This might be impractical for a large life office or pension fund but, nevertheless, it holds in practice.

  Some term structure models do allow interest rates to go negative. One such example is the Vasicek model we will see later in this chapter.

  Whether or not this is a problem depends on the probability of negative interest rates within the timescale of the problem in hand and their likely magnitude if they can go negative.

- $r(t)$ **and other interest rates should exhibit some form of mean-reverting behaviour.** Again this is because the empirical evidence suggests that interest rates do tend to mean revert in practice.

  This might not be particularly strong mean reversion but it is essential for many actuarial applications where the time horizon of a problem might be very long.

- **How easy is it to calculate the prices of bonds and certain derivative contracts?**

  This is a computational issue. It is no good in a modelling exercise to have a wonderful model if it is impossible to perform pricing or hedging calculations within a reasonable amount of time. This is because we need to act quickly to identify any potential arbitrage opportunities or to rebalance a hedged position.

  Thus we aim for models that either give rise to simple formulae for bond and option prices or that make it straightforward to compute prices using numerical techniques.

- **Does the model produce realistic dynamics?**

  For example, can it reproduce features that are similar to what we have seen in the past with reasonable probability? Does it give rise to a full range of plausible yield curves – *ie* upward-sloping, downward-sloping and humped?

- **Does the model, with appropriate parameter estimates, fit historical interest rate data adequately?**

- **Can the model be calibrated easily to current market data?**

  If so, is this calibration perfect or just a good approximation? This is an important point when we are attempting to establish the fair value of liabilities. If the model cannot fit observed yield curves accurately then it has no chance of providing us with a reliable fair value for a set of liabilities.

- **Is the model flexible enough to cope properly with a range of derivative contracts?**

Question 17.7

When carrying out actuarial valuations of pension funds, pension consultants usually assume a constant interest rate over all terms. This model of interest rates is *not* arbitrage-free.

How can this be reconciled with the first point listed above?

3 The risk-neutral approach to pricing

3.1 Introduction

We will now consider one-factor, Markov, diffusion models for the short rate of interest, $r(t)$.

A *one-factor* model is one in which there is assumed to be only one source of randomness affecting the short rate. This randomness is usually modelled using Brownian motion.

The general stochastic differential equation (SDE) for $r(t)$ is:

$$dr(t) = a\left(t, r(t)\right)dt + b\left(t, r(t)\right)dW(t)$$

where:

- $a(t, r)$ **is the drift**
- $b(t, r)$ **is the volatility**
- $W(t)$ **is a standard Brownian motion under the real-world measure P.**

In actuarial work we are used to assuming a *fixed* rate of interest in calculations. However, in this chapter we are considering *stochastic* models where future interest rates behave randomly. This means that we need to specify which probability measure we are using.

If we are to have a model that is arbitrage-free then we need to consider the prices of tradable assets, with the most natural of these being the zero-coupon bond prices $B(t,T)$.

Modelling the short rate $r(t)$ does not tell us directly about the prices of the assets traded in the market. To see whether arbitrage opportunities exist or not, we need to examine these prices.

We can use an argument similar to the original derivation of the Black-Scholes model to demonstrate that:

$$B(t,T) = E_Q\left[\exp\left(-\int_t^T r(u)du\right) \middle| r(t)\right]$$

where Q is called the risk-neutral measure.

Since we are considering Markov models, any information about the value of the short rate before time t is irrelevant, and so we can replace $r(t)$ with its filtration F_t. We can then write this equation in the equivalent form:

$$B(t,T) = E_Q\left[\exp\left(-\int_t^T r(u)du \right) \times 1 \,\middle|\, F_t \right]$$

Since the payoff from the bond at maturity is $B(T,T) = 1$, we can see that this equation is analogous to the valuation formula $V_t = E_Q\left[e^{-r(T-t)} X \middle| F_t \right]$, which we met in Chapter 15 for derivatives based on shares.

If we had a derivative whose payoff X was dependent on the future value of the bond, then its value at time t would be:

$$V_t = E_Q\left[\exp\left(-\int_t^T r(u)du \right) X \,\middle|\, F_t \right]$$

This formula is interesting because it tells us that, if we assume that the market is arbitrage-free, we can work out the value of the bond at the current time t based on the current short rate $r(t)$ and the possible future values of $r(u)$ for $t < u \le T$. We don't need to know the entire term structure, *ie* the values of $f(t,u)$ for $t < u \le T$, at the current time.

At this stage we need to pause. When we considered the Black-Scholes model for equity derivative pricing the choice of Q was unambiguous, because we started off by writing down the dynamics of a tradable asset.

The tradable asset was the share price S_t, which we modelled directly using the SDE $dS_t = S_t\left(\mu\, dt + \sigma dW_t\right)$.

Here the situation is different. We have started off with a process for $r(t)$ that is not a tradable asset. However, since this is a one-factor model, as soon as we introduce one tradable asset, $B(t,T)$, we will be able to determine what the *market price of risk* is.

We will see that the market price of risk is a key quantity that features prominently in the algebraic calculations.

3.2 *The market price of risk*

Take a specific bond with maturity at T_1. Suppose its SDE under the real-world measure P is:

$$dB(t,T_1) = B(t,T_1)\left[m(t,T_1)dt + S(t,T_1)dW(t)\right]$$

where, besides $S(t,T_1)$, $m(t,T_1)$ might also be stochastic.

Remember that $W(t)$ is standard Brownian motion, and therefore has mean zero. So $m(t,T_1)$ represents the expected rate at which the value of this bond is growing.

The *market price of risk* is defined as:

$$\gamma(t,T_1) = \frac{m(t,T_1) - r(t)}{S(t,T_1)}$$

$\gamma(t,T_1)$ represents the excess expected return over the risk-free rate per unit of volatility in return for an investor taking on this volatility.

We may also be interested in the risk premium, which is just the excess expected return over the risk-free rate. This can be obtained by just multiplying both sides of the previous equation by the volatility coefficient $S(t,T_1)$.

The *risk premium* on the bond is thus:

$$\gamma(t,T_1)S(t,T_1) = m(t,T_1) - r(t)$$

It is not difficult to show (but we won't do it here) **that, if we have a one-factor model, then we must have, at any given time t, the same market price of risk $\gamma(t)$ for bonds of all maturities: that is, $\gamma(t,T) = \gamma(t,T_1) = \gamma(t)$ for all $T > t$.**

With this identity we find that for all $t < T$:

$$\begin{aligned} dB(t,T) &= B(t,T)\left[m(t,T)dt + S(t,T)dW(t)\right] \\ &= B(t,T)\left[\left(r(t) + \gamma(t)S(t,T)\right)dt + S(t,T)dW(t)\right] \\ &= B(t,T)\left[r(t)dt + S(t,T)\left(dW(t) + \gamma(t)dt\right)\right] \\ &= B(t,T)\left[r(t)dt + S(t,T)d\tilde{W}(t)\right] \end{aligned}$$

where $d\tilde{W}(t) = dW(t) + \gamma(t)dt$ is the standard Brownian motion under Q.

Example

Explain the steps in the previous calculation.

Solution

The first line here is just the model we assumed at the start of this section, but now we are considering a bond with a generic maturity date T, since $\gamma(t)$ does not depend on the specific maturity date.

To obtain the second line, we rearrange the definition of the market price of risk:

$$\gamma(t,T) = \frac{m(t,T) - r(t)}{S(t,T)}$$

to make $m(t,T)$ the subject, then substitute this expression into the first line.

The third line is obtained by grouping together the terms containing an $S(t,T)$ factor.

The fourth line just uses the definition of the new process $\tilde{W}(t)$.

What we have done here is to apply a change of measure (from P to Q) that changes the $m(t,T)$ in the drift coefficient to $r(t)$.

Note that when we make this transformation from the SDE under *P* to *Q*, the drift under *Q* (namely $B(t,T)r(t)$) **of all tradable assets must always be equal to the price of the security times the risk-free rate of interest.**

In other words, two bonds with the same value will experience the same drift, even if they have different terms.

The market price of risk was also considered in Chapter 7.

3.3 The risk-neutral measure as a computational tool

For a one-factor model we have seen above the broad principle that transforms us from P to Q. In order to say more about the basic price processes we must look at the effect of this transformation on $r(t)$. Thus:

$$\begin{aligned} dr(t) &= a(t,r(t))dt + b(t,r(t))dW(t) \quad \text{under } P \\ &= a(t,r(t))dt + b(t,r(t))\left(d\tilde{W}(t) - \gamma(t)dt\right) \\ &= \left(a(t,r(t)) - \gamma(t)b(t,r(t))\right)dt + b(t,r(t))d\tilde{W}(t) \\ &= \tilde{a}(t,r(t))dt + b(t,r(t))d\tilde{W}(t) \end{aligned}$$

where $\tilde{a}(t,r(t)) = a(t,r(t)) - \gamma(t)b(t,r(t))$.

> ***Question 17.8***
>
> Explain the steps in this calculation.

The final two lines give us the dynamics (*ie* the SDE) **of $r(t)$ under the artificial measure Q. We then use this to determine**

$$B(t,T) = E_Q\left[\exp\left(-\int_t^T r(u)du\right) \middle| r(t)\right]$$

for specific models.

It is important to remember that Q is an artificial computational tool. It is determined by combining:

(a) the model for $r(t)$ under the real-world measure P and

(b) the market price of risk established from knowledge of the dynamics of one bond.

3.4 Practical implementation

When modellers use this approach to pricing, from the practical point of view they normally start by specifying the dynamics of $r(t)$ under Q in order to calculate bond prices. Second, they specify the market price of risk as a component of the model, and this allows us to determine the dynamics of $r(t)$ under P.

4 The state price deflator approach to pricing

Other modellers prefer to take a different approach to model building using *state price deflators* (also sometimes called *deflators* or *pricing kernels*).

With this approach we first specify a strictly positive diffusion process $A(t)$ with SDE under P :

$$dA(t) = A(t)\left(\mu_A(t)dt + \sigma_A(t)dW(t)\right)$$

where $\mu_A(t)$ and $\sigma_A(t)$ will be stochastic.

Note that, with this approach, the dynamics of the state price deflator $A(t)$ are defined in terms of the real-world probability measure.

Building a term structure model is then deceptively simple. We define zero-coupon bond prices according to the formula

$$B(t,T) = \frac{E_P\left[A(T) \mid F_t\right]}{A(t)}$$

Question 17.9

Check that this formula works when $t = T$.

Example

Suppose that $A(t)$ is a deterministic process with $\mu_A(t) = -\delta$ and $\sigma_A(t) = 0$, where δ is a constant. Show that $A(t) = ke^{-\delta t}$ and evaluate $B(t,T)$.

Solution

Here the SDE for $A(t)$ becomes $dA(t) = -A(t)\delta\, dt$, which is an ordinary differential equation. This can be written in the form:

$$\frac{1}{A(t)}\frac{dA(t)}{dt} = -\delta \quad \text{or} \quad \frac{d}{dt}\log A(t) = -\delta$$

Integrating gives:

$$\log A(t) = -\delta t + \text{constant}$$

So: $A(t) = ke^{-\delta t}$, where k is a constant.

$B(t,T)$ can now be calculated from the formula $B(t,T) = \dfrac{E_P\left[A(T) \mid F_t\right]}{A(t)}$, *ie*:

$$B(t,T) = \frac{E_P\left[ke^{-\delta T} \mid F_t\right]}{ke^{-\delta t}}$$

Since $A(t)$ is not random here, we can simplify this to get:

$$B(t,T) = \frac{ke^{-\delta T}}{ke^{-\delta t}} = e^{-\delta(T-t)}$$

This example illustrates that the process $A(t)$ acts like a discount factor that applies to payments at time t. It is a generalisation of the familiar v^t factor.

It can be shown (but we won't do it here) **that this method gives rise to an arbitrage-free model for bond prices. The formula for $B(t,T)$ is a very simple looking formula, but any potential difficulty comes in working out $E_P\left[A(T) \mid F_t\right]$ as we need "interesting" models for $A(t)$ in order to get interesting and useful models for $r(t)$ and $B(t,T)$.**

> **Question 17.10**
>
> Suppose that $\mu_A(t) = -\delta$ and $\sigma_A(t) = \sigma$, where δ and σ are constants.
>
> Show that $A(t) = A(0)e^{(-\delta - \frac{1}{2}\sigma^2)t + \sigma W_t}$ and $B(t,T) = e^{-\delta(T-t)}$.

Using this approach the risk-free rate of interest can be shown (but again we won't do it here) **to be simply:**

$$r(t) = -\mu_A(t)$$

It follows that our model will have positive interest rates if and only if $\mu_A(t)$ is negative for all t with probability 1. This means that $A(t)$ must be a strictly positive supermartingale (that is, $E_P\left[A(T) \mid F_t\right] < A(t)$ for all t, T).

You should be familiar with martingales, which have the property that $E\left[X_t \mid F_s\right] = X_s$ whenever $t > s$, *ie* the process has zero drift. A *supermartingale* has the property that $E\left[X_t \mid F_s\right] \le X_s$, *ie* the drift is negative or zero.

> **Question 17.11**
>
> What do you think a submartingale is? What can you say about a process that is both a supermartingale and a submartingale?

> **Question 17.12**
>
> How do we know that $A(t)$ is a supermartingale?

This looks like it is quite different from the risk-neutral approach to pricing. However, the two approaches are, in fact, exactly equivalent: any model developed in one framework has an equivalent form under the other framework. In general this can be done using the *Radon-Nikodym derivative*, dQ/dP.

For some models it is equally easy to take the risk-neutral or the state price deflator approach. In other cases one approach (the one used to develop the model in the first place) looks relatively straightforward while the other approach might appear quite intractable.

5 Models for the term structure of interest rates

5.1 Introduction

The models we've looked at so far have focused on how the short rate of interest varies over calendar time. In this section we will focus on how interest rates vary as a function of the term $T-t$.

Several of the more straightforward models are based on the short-rate $r(t)$ in the risk-neutral framework, of which the Vasicek and Cox-Ingersoll-Ross (CIR) models are the best known. These two models are *time homogeneous*: that is, the future dynamics of $r(t)$ only depend upon the current value of $r(t)$ rather than what the present time t actually is.

5.2 The Vasicek model

The Vasicek model has the dynamics, under Q:

$$dr(t) = \alpha\left(\mu - r(t)\right)dt + \sigma d\tilde{W}(t)$$

where $\tilde{W}(t)$ is a standard Brownian motion under Q.

The parameter α takes a positive value.

The graph below show a simulation of this process based on the parameter values $\alpha = 0.1$, $\mu = 0.06$ and $\sigma = 0.02$.

Figure 17.2: Example simulation of short rate from the Vasicek model

Example

The SDE for the Vasicek model is the same as for the Ornstein-Uhlenbeck process, which we met in Chapter 9. Deduce a formula for $r(t)$ for the Vasicek model.

Solution

As we saw in Chapter 9, the solution to this equation can be expressed as:

$$r(t) = r(0)e^{-\alpha t} + \mu\left(1 - e^{-\alpha t}\right) + \sigma\int_0^t e^{-\alpha(t-u)}d\tilde{W}_u$$

Example

What are the key properties of this model that can be deduced from this SDE?

Solution

This SDE tells us about the changes in the short rate over each instant.

The drift coefficient can be written as $-\alpha\,[r(t) - \mu]$. So the drift depends on the current interest rate. Since $0 < \alpha < 1$, this coefficient always directs the movements towards μ. So the process is mean-reverting towards the constant mean value μ.

Since the random component of the movements is based on Brownian motion, the movements in the short rate are normally distributed. The volatility parameter σ specifies how big the random movements are.

We then have the following formula for bond prices:

$$B(t,T) = e^{a(\tau) - b(\tau)r(t)}$$

where

- $\tau = T - t$
- $b(\tau) = \dfrac{1 - e^{-\alpha\tau}}{\alpha}$ **and** $a(\tau) = (b(\tau) - \tau)\left(\mu - \dfrac{\sigma^2}{2\alpha^2}\right) - \dfrac{\sigma^2}{4\alpha}b(\tau)^2$.

We won't derive these formulae but, as you can see, they show how the prices vary according to the term τ.

Example

Show that the instantaneous forward rate for the Vasicek model can be expressed as:

$$f(t,T) = r(t)e^{-\alpha\tau} + \left(\mu - \frac{\sigma^2}{2\alpha^2}\right)(1-e^{-\alpha\tau}) + \frac{\sigma^2}{2\alpha^2}(1-e^{-\alpha\tau})e^{-\alpha\tau}$$

where $\tau = T - t$.

Solution

We have a formula for $B(t,T)$ for this model. So the instantaneous forward rate can be derived from this using the relationship:

$$f(t,T) = -\frac{\partial}{\partial T}\log B(t,T)$$

This gives:

$$f(t,T) = -\frac{\partial}{\partial T}[a(\tau) - b(\tau)r(t)] = -a'(\tau) + b'(\tau)r(t)$$

From the definitions of $b(\tau)$ and $a(\tau)$, we find that:

$$b'(\tau) = \frac{d}{d\tau}\left(\frac{1-e^{-\alpha\tau}}{\alpha}\right) = e^{-\alpha\tau}$$

and

$$\begin{aligned}
a'(\tau) &= \frac{d}{d\tau}\left[(b(\tau)-\tau)\left(\mu - \frac{\sigma^2}{2\alpha^2}\right) - \frac{\sigma^2}{4\alpha}b(\tau)^2\right] \\
&= (b'(\tau)-1)\left(\mu - \frac{\sigma^2}{2\alpha^2}\right) - \frac{\sigma^2}{4\alpha}\times 2b(\tau)b'(\tau) \\
&= (e^{-\alpha\tau}-1)\left(\mu - \frac{\sigma^2}{2\alpha^2}\right) - \frac{\sigma^2}{2\alpha}\left(\frac{1-e^{-\alpha\tau}}{\alpha}\right)e^{-\alpha\tau} \\
&= -(1-e^{-\alpha\tau})\left(\mu - \frac{\sigma^2}{2\alpha^2}\right) - \frac{\sigma^2}{2\alpha^2}(1-e^{-\alpha\tau})e^{-\alpha\tau}
\end{aligned}$$

Substituting these expressions into the general formula for $f(t,T)$ gives the required answer.

Question 17.13

Write down an expression in terms of the model parameters for the long rate, *ie* the instantaneous forward rate corresponding to $T-t=\infty$, according to the Vasicek model.

The curves shown on the graph of gilt yields in Section 1 were fitted using a Vasicek model with parameter values $\alpha=0.131$, $\mu=0.083$ and $\sigma=0.037$.

Question 17.14

"The particular model used for the graph implies that interest rates are mean-reverting to the value $\mu=0.083$."

True or false?

Example

Given that the current time is t, what can we say about the statistical properties of $r(T)$ based on the Vasicek model?

Solution

When the current time is t and we are looking ahead to time T, the solution to the Ornstein-Uhlenbeck equation is:

$$r(T)=r(t)e^{-\alpha(T-t)}+\mu(1-e^{-\alpha(T-t)})+\sigma\int_t^T e^{-\alpha(T-u)}d\tilde{W}_u$$

Under the risk-neutral probability measure, the Ito integral here has mean zero and variance:

$$\sigma^2\int_t^T e^{-2\alpha(T-u)}du=\frac{\sigma^2}{2\alpha}\left(1-e^{-2\alpha(T-t)}\right)$$

So:

$$r(T)\big|r(t)\sim N\left[r(t)e^{-\alpha(T-t)}+\mu\left(1-e^{-\alpha(T-t)}\right),\frac{\sigma^2}{2\alpha}\left(1-e^{-2\alpha(T-t)}\right)\right] \text{ under } Q$$

With this model we can also exploit the bivariate normality of both $r(T)$ and $\int_t^T r(u)du$ to develop simple formulae for the prices of European options on both zero-coupon and coupon-bearing bonds.

Example

How would we derive a formula for the value of a call option on a bond that gave the holder the option to buy the bond at a specified time s (where $t < s < T$) at an agreed price K?

Solution

We would use a risk-neutral pricing formula like the ones we saw in Section 3. Here the payoff at time s is $\max[B(s,T)-K,0]$. So the pricing formula for the value of the option at the current time t would be:

$$V_t = E_Q\left[\exp\left(-\int_t^s r(u)du\right)\times\max\left[B(s,T)-K,0\right]\middle| F_t\right]$$

Option price formulae for zero-coupon bonds closely resemble the Black-Scholes formula for equity option prices.

The Vasicek model suffers from a number of drawbacks. The most obvious of these is that interest rates can go negative.

Here is another simulation of the Vasicek process with the same parameter values as before. In this simulation the short rate has become negative at certain times.

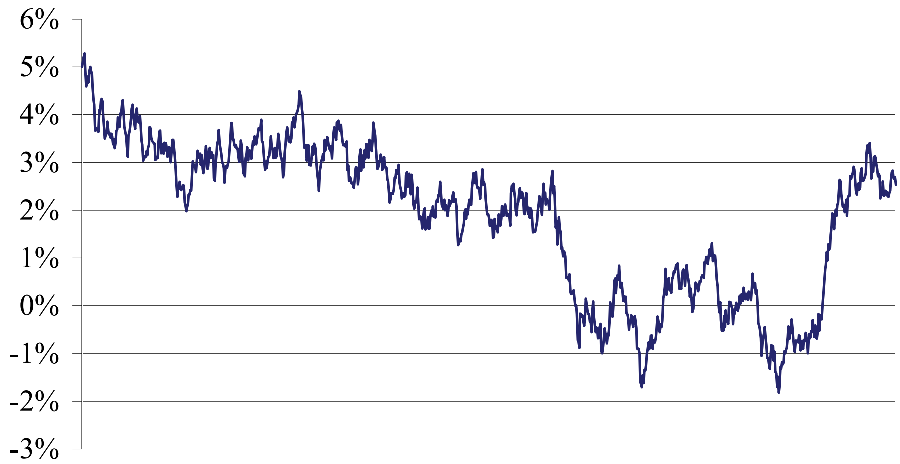


Figure 17.3: Vasicek simulation with negative interest rates

Sometimes this is not a particular problem: for example, if the probability of negative rates is small possibly because the time horizon is short. In other cases (especially longer-term actuarial applications) the probability and severity of negative interest rates can be significant.

5.3 The Cox-Ingersoll-Ross (CIR) model

Basic process

The Cox-Ingersoll-Ross model ensures that all interest rates remain positive, thereby countering one of the main drawbacks of the Vasicek model. The SDE for $r(t)$ under Q is:

$$dr(t) = \alpha\left(\mu - r(t)\right)dt + \sigma\sqrt{r(t)}d\tilde{W}(t)$$

The graph below show a simulation of this process based on the parameter values $\alpha = 0.1$, $\mu = 0.06$ and $\sigma = 0.1$.

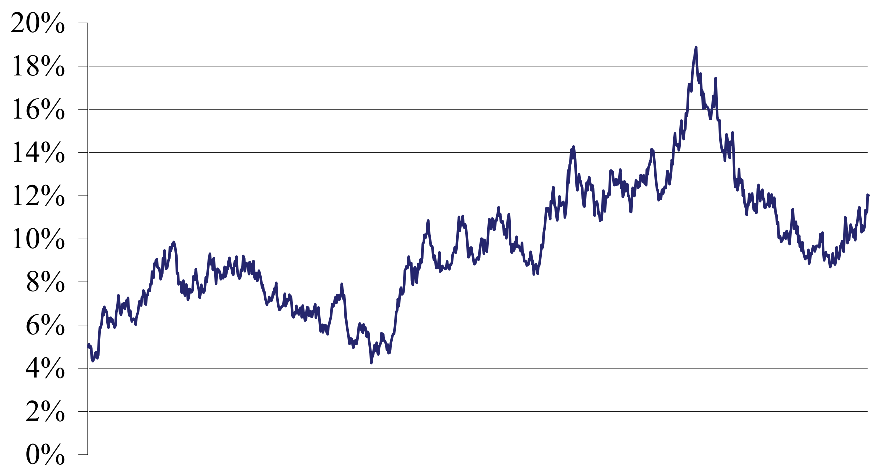


Figure 17.4: Simulation from Cox-Ingersoll-Ross model

We can see that the form of the drift of $r(t)$ is the same as for the Vasicek model. The critical difference between the two models occurs in the volatility, which is increasing in line with the square root of $r(t)$. Since this diminishes to zero as $r(t)$ approaches zero, and provided σ^2 is not too large ($r(t)$ will never hit zero provided $\sigma^2 \le 2\alpha\mu$), we can guarantee that $r(t)$ will not hit zero. Consequently all other interest rates will also remain strictly positive.

The key point here is that the random term is multiplied by a power of $r(t)$, which means that as $r(t)$ approaches zero, the random increments get smaller and smaller. The power of ½ (*ie* the square root) has been chosen because this is the value that "just" prevents the process from reaching zero.

Other features of the CIR model

We can make the following observations:

Prices of zero-coupon bonds are given by:

$$B(t,T) = e^{a(\tau)-b(\tau)r(t)}$$

This equation is the same as for the Vasicek model. As before, $\tau = T - t$.

But now:

- $$b(\tau) = \frac{2(e^{\theta\tau}-1)}{(\theta+\alpha)(e^{\theta\tau}-1)+2\theta}$$
- $$a(\tau) = \frac{2\alpha\mu}{\sigma^2}\log\left(\frac{2\theta e^{(\theta+\alpha)\tau/2}}{(\theta+\alpha)(e^{\theta\tau}-1)+2\theta}\right)$$
- $$\theta = \sqrt{\alpha^2+2\sigma^2}\ .$$

Pricing formulae for European call and put options on zero-coupon bonds look similar to those for the Vasicek model and to the Black-Scholes formulae for equity options. However, where the latter models use the cumulative distribution function of the Normal distribution, the CIR formulae use the cumulative distribution function of the non-central chi-squared distribution.

If $X_1, X_2, \ldots, X_n$ are independent random variables, each with a $N(0,1)$ distribution, then $Y = X_1^2 + X_2^2 + \cdots + X_n^2$ has a chi-square distribution with n degrees of freedom.

If $X_1, X_2, \ldots, X_n$ are independent random variables and $X_i \sim N(d_i, 1)$, then $Y_d = X_1^2 + X_2^2 + \cdots + X_n^2$ is said to have a *non-central* chi-square distribution with n degrees of freedom and non-centrality parameter $d = \sum_{i=1}^{n} d_i^2$.

So the non-central chi-square distribution can be thought of as a lopsided version of the ordinary chi-square distribution.

Question 17.15

What is the mean of the non-central chi-square distribution with n degrees of freedom and non-centrality parameter d ?

From the point of view of implementation, the CIR model is slightly more tricky than the Vasicek model.

5.4 The Hull & White (HW) model

We will now look at a simple extension of the Vasicek model. Recall the SDEs for both the Vasicek and CIR models gave us time-homogeneous models. This means that bond prices at time t depend only on $r(t)$ and on the term to maturity. This results in a lack of flexibility when it comes to pricing related contracts. For example, on any given date theoretical bond prices will probably not match exactly observed market prices. We can re-estimate $r(t)$ to improve the match and even re-estimate the constant parameters α, μ and σ but we will still, normally, be unable to get a precise match.

A simple way to get theoretical prices to match observed market prices is to introduce some elements of *time-inhomogeneity* into the model. The Hull & White (HW) model does this by extending the Vasicek model in a simple way. We define the SDE for $r(t)$ under Q as follows

$$dr(t) = \alpha\left(\mu(t) - r(t)\right)dt + \sigma d\tilde{W}(t)$$

where $\mu(t)$ is a deterministic function of t. $\mu(t)$ has the natural interpretation of being the local mean-reversion level for $r(t)$.

This is the same equation as for the Vasicek model, except that μ is no longer a constant, but can vary over time.

The Hull & White model can be easily extended to include a time-varying but deterministic $\sigma(t)$. This allows us to calibrate the model to traded option prices as well as zero-coupon bond prices.

Since $\mu(t)$ is deterministic the Hull & White model is just as tractable as the Vasicek model. It has only lost the simplicity of time homogeneity.

An advantage of the HW model is that it allows us to price interest-rate linked contracts more accurately. This is important for a variety of reasons.

First, in insurance, the fair value of fixed liabilities (with no option characteristics) must accurately reflect the current observed term-structure of interest rates. The use of the Vasicek (or CIR) model, even after fitting the model to the current term structure, might introduce some bias into the fair value.

Second, bond and interest rate derivatives traders want to be able to quote prices that are in line with prices being quoted by other traders. This is facilitated by the use of models like the HW model and other, more sophisticated, no-arbitrage models. The HW model suffers from the same drawback as the Vasicek model, namely that interest rates might become negative.

6 Limitations of one-factor models

6.1 Limitations

Question 17.16

What is a one-factor terms structure model?

One-factor models have certain limitations that it is important to be familiar with.

First, if we look at historical interest rate data we can see that changes in the prices of bonds with different terms to maturity are not perfectly correlated, as one would expect to see if a one-factor model was correct. Sometimes we even see, for example, that short-dated bonds fall in price while long-dated bonds go up. Recent research has suggested that around three factors, rather than one, are required to capture most of the randomness in bonds of different durations.

Second, if we look at the long run of historical data we find that there have been sustained periods of both high and low interest rates with periods of both high and low volatility. Again these are features that are difficult to capture without introducing more random factors into a model. This issue is especially important for two types of problem in insurance:

1. **the pricing and hedging of long-dated insurance contracts with interest-rate guarantees**

2. **asset-liability modelling and long-term risk-management.**

Third, we need more complex models to deal effectively with derivative contracts that are more complex than, say, standard European call options. For example, any contract that makes reference to more than one interest rate should allow these rates to be less than perfectly correlated.

One-factor models do, nevertheless have their place as tools for: valuation of simple liabilities with no option characteristics; or short-term, straightforward derivatives contracts.

For other problems it is appropriate to make use of models that have more than one source of randomness: so-called *multifactor* models.

6.2 Multifactor models

A simple example of a multifactor model is the 2-factor Vasicek model. This models two processes: $r(t)$, as before, and $m(t)$, the local mean-reversion level for $r(t)$. Thus:

$$dr(t) = \alpha_r \left(m(t) - r(t)\right) dt + \sigma_{r1} d\tilde{W}_1(t) + \sigma_{r2} d\tilde{W}_2(t)$$

$$dm(t) = \alpha_m \left(\mu - m(t)\right) dt + \sigma_{m1} d\tilde{W}_1(t)$$

where $\tilde{W}_1(t)$ and $\tilde{W}_2(t)$ are independent, standard Brownian motions under the risk-neutral measure Q. This looks superficially like the Hull & White model, but the HW model has a deterministic mean-reversion level, whereas here $m(t)$ is stochastic.

In general, zero-coupon bond prices can always be calculated using the risk-neutral approach using the formula:

$$B(t,T) = E_Q\left[e^{-\int_t^T r(u)du} \,\middle|\, F_t \right]$$

For the 2-factor Vasicek model we can write this as:

$$B(t,T) = E_Q\left[e^{-\int_t^T r(u)du} \,\middle|\, r(t), m(t) \right]$$

since the model is Markov and all relevant information available at time t about the future is given by $r(t)$ and $m(t)$.

There are a great many multifactor models. Often these have been designed in order to tackle specific problems. Students should always remember that models that were designed for and are good for one specific task may not be suitable for quite a different task. For example, models that are widely used by investment banks for pricing and hedging short-term derivatives may not be suitable for use in long-term asset-liability modelling studies.

Question 17.17

What are the three main limitations of one-factor models?

7 Exam-style question

We finish this chapter with an exam-style question.

Question

Compare the Vasicek and Hull & White models of interest rates. [5]

Solution

Compare means describe the similarities and the differences.

The equations defining the two models are similar:

Vasicek: $dr(t) = \alpha[\mu - r(t)]dt + \sigma d\tilde{W}(t)$

H&W: $dr(t) = \alpha[\mu(t) - r(t)]dt + \sigma d\tilde{W}(t)$

Both are:

- continuous-time Markov models of the short rate of interest
- Ito processes defined by a stochastic differential equation
- one-factor models (*ie* they incorporate only one source of randomness)
- usually defined in terms of a standard Brownian motion under the risk-neutral probability measure.

Both models imply that the short rate is mean-reverting.

Both models imply that the future short rate has a normal distribution.

Both models allow negative values for the short rate.

The key difference is that the Vasicek model is time-homogeneous (with constant μ), whereas the Hull & White model is not (with time-dependent μ).

The Hull & White model has to be calibrated (*ie* a function has to be selected for $\mu(t)$) to match the current pattern of bond prices.

Both models are mathematically tractable, although the Hull & White model is slightly more complicated algebraically.

Chapter 17 Summary

Notation

- t current time
- $B(t,T)$ zero-coupon bond price
- $r(t)$ instantaneous risk-free rate
- $C(t)$ accumulated value of 1 unit at the risk-free rate
- $F(t,T,S)$ forward rate for the period (T,S)
- $f(t,T)$ instantaneous forward rate
- $R(t,T)$ spot rate

Relationships

- $R(t,T) = -\dfrac{1}{T-t}\log B(t,T) = \dfrac{1}{T-t}\displaystyle\int_t^T f(t,u)du$
- $B(t,T) = \exp[-R(t,T)(T-t)]$
- $F(t,T,S) = \dfrac{1}{S-T}\log\dfrac{B(t,T)}{B(t,S)}$
- $f(t,T) = \lim\limits_{S\to T} F(t,T,S) = -\dfrac{\partial}{\partial T}\log B(t,T)$
- $B(t,T) = \exp\left[-\displaystyle\int_t^T f(t,u)du\right]$
- $dC(t) = r(t)C(t)dt$
- $C(t) = \exp\left[\displaystyle\int_0^t r(u)du\right]$

Desirable characteristics of a term structure model

- The model should be arbitrage-free.
- Interest rates should be positive.
- Interest rates should be mean-reverting.
- Bonds and derivative contracts should be easy to price.
- It should produce realistic interest rate dynamics.
- It should fit historical interest rate data adequately.
- It should be easy to calibrate to current market data.
- It should be flexible enough to cope with a range of derivatives.

General one-factor model for the short rate

Under P

$$dr(t) = a(t, r(t))\, dt + b(t, r(t))\, dW(t)$$

$$dB(t,T) = B(t,T)\,[m(t,T)\, dt + S(t,T)\, dW(t)]$$

Risk premium

$$\gamma(t)\, S(t,T) = m(t,T) - r(t)$$

Market price of risk

$$\gamma(t) = \frac{m(t,T) - r(t)}{S(t,T)}$$

Change of measure

$$d\tilde{W}(t) = dW(t) + \gamma(t)\, dt$$

Under Q

$$dr(t) = \tilde{a}(t, r(t))\, dt + b(t, r(t))\, d\tilde{W}(t),$$

$$dB(t,T) = B(t,T)\,[r(t)\, dt + S(t,T)\, d\tilde{W}(t)]$$

where:

$$\tilde{a}(t, r(t)) = a(t, r(t)) - \gamma b(t, r(t))$$

Risk-neutral valuation formula for a zero-coupon bond:

$$B(t,T) = E_Q\left[\exp\left(-\int_t^T r(u)du\right) \middle| r(t)\right]$$

Risk-neutral valuation formula for a derivative with payoff X_T:

$$V_t = E_Q\left[\exp\left(-\int_t^T r(u)du\right) X_T \middle| F_t\right]$$

State price deflators

$$dA(t) = A(t)\left[\mu_A(t)\, dt + \sigma_A(t)\, dW(t)\right] \text{ under } P$$

where:

$$r(t) = -\mu_A(t)$$

Valuation formula for a zero-coupon bond:

$$B(t,T) = \frac{E_P\left[A(T) \mid F_t\right]}{A(t)}$$

Examples of one-factor models

Vasicek model

$$dr(t) = \alpha\left[\mu - r(t)\right]dt + \sigma d\tilde{W}(t) \quad \text{under } Q$$

Cox-Ingersoll-Ross model

$$dr(t) = \alpha\left[\mu - r(t)\right]dt + \sigma\sqrt{r(t)}d\tilde{W}(t) \quad \text{under } Q$$

Hull & White model

$$dr(t) = \alpha\left[\mu(t) - r(t)\right]dt + \sigma d\tilde{W}(t) \quad \text{under } Q$$

where $\mu(t)$ is a deterministic function.

Two-factor Vasicek model

Model:

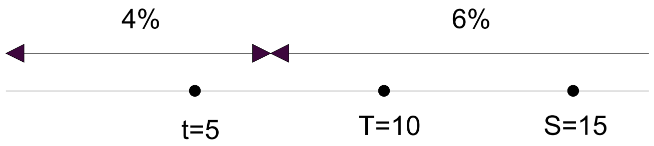
$$dr(t) = \alpha_r\left[m(t) - r(t)\right]dt + \sigma_{r1}d\tilde{W}_1(t) + \sigma_{r2}d\tilde{W}_2(t) \quad \text{under } Q$$
$$dm(t) = \alpha_m\left[\mu - m(t)\right]dt + \sigma_{m1}d\tilde{W}_1(t)$$

Valuation formula for a zero-coupon bond:

$$B(t,T) = E_Q\left[\exp\left(-\int_t^T r(u)du\right)\middle| F_t\right] = E_Q\left[\exp\left(-\int_t^T r(u)du\right)\middle| r(t), m(t)\right]$$

Chapter 17 Solutions

Solution 17.1



- $B(5,10) = e^{-(2\times 0.04 + 3\times 0.06)} = e^{-0.26} = 0.771$
- $r(5) = 0.04$
- $C(5) = e^{5\times 0.04} = e^{0.2} = 1.2214$
- $F(5,10,15) = 0.06$
- $f(5,10) = 0.06$
- $R(5,10) = \dfrac{1}{5}(2\times 0.04 + 3\times 0.06) = 0.052$

Solution 17.2

One way to see this is to say that the price for a bond purchased at time t and maturing at time S should equal the price for a similar bond maturing at the earlier time T but discounted for the extra period $S-T$ at the forward rate applicable over that period. In other words:

$$B(t,S) = B(t,T)\exp\left[-F(t,T,S)(S-T)\right]$$

Rearranging this equation gives the required result:

$$F(t,T,S) = \frac{1}{S-T}\log\frac{B(t,T)}{B(t,S)}$$

Solution 17.3

The instantaneous rate $f(t,T)$ is defined as $\lim\limits_{S \to T} F(t,T,S)$. Since S is a later time than T, we can write this as:

$$f(t,T) = \lim_{h \to 0} F(t,T,T+h)$$

Using the equation $F(t,T,S) = \dfrac{1}{S-T} \log \dfrac{B(t,T)}{B(t,S)}$, this is:

$$\begin{aligned} f(t,T) &= \lim_{h \to 0} \frac{1}{h} \log \frac{B(t,T)}{B(t,T+h)} \\ &= \lim_{h \to 0} \frac{\log B(t,T) - \log B(t,T+h)}{h} \\ &= -\lim_{h \to 0} \frac{\log B(t,T+h) - \log B(t,T)}{h} \\ &= -\frac{\partial}{\partial T} \log B(t,T) \end{aligned}$$

Solution 17.4

This is an ordinary differential equation, which we can write in the form:

$$\frac{1}{C(t)} \frac{dC(t)}{dt} = r(t)$$

or $$\frac{d}{dt} \log C(t) = r(t)$$

If we integrate this over the range $(0,t)$, using u for the dummy variable, we get:

$$\left[\log C(t)\right]_0^t = \int_0^t r(u)du$$

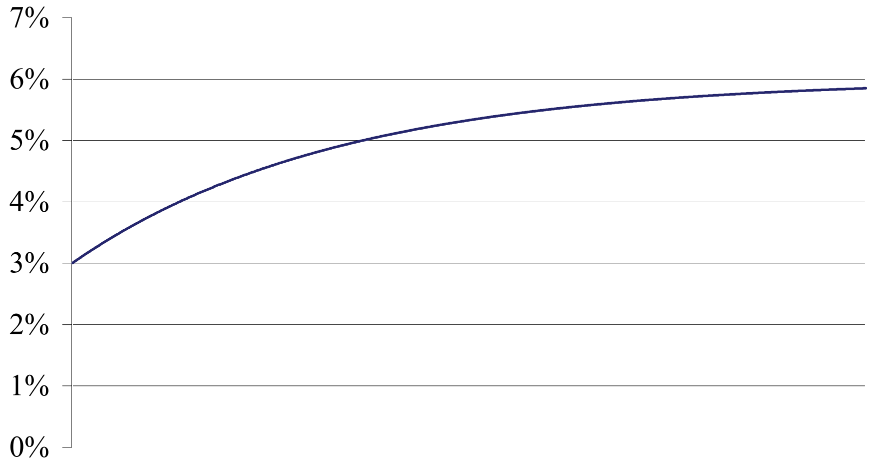
ie $$\log C(t) - \log C(0) = \int_0^t r(u)du$$

Using the initial condition $C(0) = 1$ and exponentiating gives the required result:

$$C(t) = \exp\left[\int_0^t r(u)du\right]$$

Solution 17.5

The factor $e^{-0.1(T-t)}$ equals 1 when $T=t$, but then decreases exponentially to zero as $T \to \infty$. So $f(t,T)$ is a weighted average of 0.03 and 0.06, and its graph will increase from 0.03 (*ie* a force of interest of 3%) to 0.06.



We can find $B(t,T)$ from the relationship $B(t,T) = \exp\left[-\int_t^T f(t,u)du\right]$:

$$\begin{aligned}
B(t,T) &= \exp\left[-\int_t^T f(t,u)du\right] \\
&= \exp\left[-\int_t^T \left\{0.03e^{-0.1(u-t)} + 0.06(1-e^{-0.1(u-t)})\right\}du\right] \\
&= \exp\left[-\int_t^T \left\{0.06 - 0.03e^{-0.1(u-t)}\right\}du\right] \\
&= \exp\left(-\left[0.06u + 0.3e^{-0.1(u-t)}\right]_t^T\right) \\
&= \exp\left(-0.06(T-t) - 0.3e^{-0.1(T-t)} + 0.3\right)
\end{aligned}$$

We can then find $R(t,T)$ from the relationship $R(t,T) = \dfrac{-1}{T-t}\log B(t,T)$:

$$\begin{aligned}
R(t,T) &= \frac{-1}{T-t}\log B(t,T) \\
&= \frac{-1}{T-t}\left(-0.06(T-t) - 0.3e^{-0.1(T-t)} + 0.3\right) \\
&= 0.06 - 0.3\left[\frac{1-e^{-0.1(T-t)}}{T-t}\right]
\end{aligned}$$

Solution 17.6

The main uses of term structure (interest rate) models are:

- by bond traders looking to identify and exploit price inconsistencies
- for calculating the price of interest rate derivatives
- by investors with a portfolio involving bonds or loans who want to set up a hedged position
- for asset-liability modelling.

Solution 17.7

Pension fund trustees are long-term investors who do not trade actively in the bond markets. The relationship between the prices of bonds with different terms is not crucial. If they misjudge future interest rates slightly, this will be corrected at the next valuation and will not result in a large financial loss. So it is usually sufficiently accurate, as an approximation, to assume a single interest rate applicable to the typical term of the fund's liabilities.

On the other hand, bond traders, for example, trade actively in the bond markets, trying to make a profit by exploiting price inconsistencies. They are competing directly with other bond traders and so it is essential that the models they use do not have in-built price inconsistencies.

Solution 17.8

The first line is just the definition of the model for the short rate, which was introduced at the start of Section 3.

The second line involves a change to the risk-neutral measure. This is done by applying the relationship $d\tilde{W}(t) = dW(t) + \gamma(t)dt$ from Section 3.2.

The third line is obtained by rearranging the terms to separate the dt and the $d\tilde{W}(t)$ terms.

The last line uses the new definition given for $\tilde{a}(t, r(t))$ to write the drift coefficient as a single function.

Solution 17.9

When $t = T$ we have:

$$B(T,T) = \frac{E_P\left[A(T) \mid F_T\right]}{A(T)}$$

The expectation here is conditional on the history of the process $A(t)$ up to time T. Consequently, there is no additional randomness in $A(T)$, which can be considered to be a known value. So we have:

$$B(T,T) = \frac{A(T)}{A(T)} = 1$$

This is exactly what we would expect, since the bond matures at time T and has a value of 1 at that time.

Solution 17.10

The SDE for $A(t)$ now becomes:

$$dA(t) = A(t)\left[-\delta dt + \sigma dW_t\right]$$

This has the same form as the SDE for the lognormal model for share prices. So its solution is:

$$A(t) = A(0)e^{(-\delta - ½\sigma^2)t + \sigma W_t}$$

$B(t,T)$ can now be calculated from the formula $B(t,T) = \dfrac{E_P\left[A(T) \mid F_t\right]}{A(t)}$:

$$\begin{aligned} B(t,T) &= \frac{E_P\left[A(0)e^{(-\delta - ½\sigma^2)T + \sigma W_T} \mid F_t\right]}{A(0)e^{(-\delta - ½\sigma^2)t + \sigma W_t}} \\ &= \frac{e^{(-\delta - ½\sigma^2)(T-t)}}{e^{\sigma W_t}} E_P\left[e^{\sigma W_T} \mid F_t\right] \\ &= e^{(-\delta - ½\sigma^2)(T-t)} E_P\left[e^{\sigma(W_T - W_t)} \mid F_t\right] \end{aligned}$$

This expectation is conditional on the history up to time t. But $W_T - W_t$ relates to the "future" time interval (t,T), and so is statistically independent. So we can write:

$$B(t,T) = e^{(-\delta-½\sigma^2)(T-t)} E_P\left[e^{\sigma(W_T-W_t)}\right]$$

Since $W_T - W_t \sim N(0,T-t)$, the expectation is now just the MGF of a normal distribution, so that:

$$\begin{aligned} B(t,T) &= e^{(-\delta-½\sigma^2)(T-t)} e^{½\sigma^2(T-t)} \\ &= e^{-\delta(T-t)} \end{aligned}$$

Solution 17.11

A submartingale has the "opposite" property, namely $E\left[X_t \mid F_s\right] \geq X_s$, *ie* the drift is non-negative.

If a process is both a supermartingale and a submartingale, then both the inequalities

$$E\left[X_t \mid F_s\right] \leq X_s \quad \text{and} \quad E\left[X_t \mid F_s\right] \geq X_s$$

apply. This means that we must have $E\left[X_t \mid F_s\right] = X_s$. In other words, the process X_t is a martingale.

Solution 17.12

$A(t)$ takes positive values and $\mu_A(t)$ is negative. Therefore the drift coefficient $A(t)\mu_A(t)$ in the SDE for $A(t)$ is negative. So the expected value of $dA(t)$ is negative, *ie* the process has negative drift, which means that it is a supermartingale.

Solution 17.13

Letting $T \to \infty$ (and hence $\tau \to \infty$) in the equation for $f(t,T)$ gives:

$$f(t,\infty) = r(t) \times 0 + \left(\mu - \frac{\sigma^2}{2\alpha^2}\right)(1-0) + \frac{\sigma^2}{2\alpha^2}(1-0) \times 0 = \mu - \frac{\sigma^2}{2\alpha^2}$$

Solution 17.14

The dynamics of $r(t)$ for this particular Vasicek model are:

$$dr(t) = -0.131\,[r(t) - 0.083]dt + 0.037d\tilde{W}(t)$$

under the *risk-neutral* probability measure Q. Under this measure $\tilde{W}(t)$ is standard Brownian motion and therefore has zero drift and the process mean-reverts to the value 0.083.

However, under the *real-world* probability measure P, $\tilde{W}(t)$ would have non-zero drift and the process will mean-revert to a different value. In fact, the long-term rate in the real world can be found from the formula derived in the previous question, namely:

$$\mu - \frac{\sigma^2}{2\alpha^2} = 0.0431 \quad ie \;\; 4.31\%$$

Solution 17.15

Since $X_i \sim N(d_i, 1)$, we find that:

$$E[X_i^2] = Var(X_i) + [E(X_i)]^2 = 1 + d_i^2$$

It follows that:

$$E[Y_d] = E[X_1^2 + X_2^2 + \cdots + X_n^2] = \sum_{i=1}^{n}(1 + d_i^2) = n + d$$

Solution 17.16

Recall that a one-factor model is a model in which interest rates are assumed to be influenced by a single source of randomness.

Solution 17.17

The three main limitations of one-factor models are:

1. They are consistent with perfect correlation of bond prices, which is inconsistent with the empirical evidence.

2. They assume a constant volatility of interest rates, which is again inconsistent with the empirical evidence.

3. They may be unsuitable for pricing complex options, *eg* those whose payoffs depend on more than one interest rate.

Chapter 18

Credit risk

Syllabus objectives

(xi) Demonstrate a knowledge and understanding of simple models for credit risk.

1. *Define the terms credit event and recovery rate.*

2. *Describe the different approaches to modelling credit risk:*
   - *structural models*
   - *reduced form models*
   - *intensity-based models.*

3. *Demonstrate a knowledge and understanding of the Merton model.*

4. *Demonstrate a knowledge and understanding of a two-state model for credit ratings with a constant transition intensity.*

5. *Describe how the two-state model can be generalised to the Jarrow-Lando-Turnbull model for credit ratings.*

6. *Describe how the two-state model can be generalised to incorporate a stochastic transition intensity.*

0 Introduction

This chapter is based, with permission, on Chapter 11 in "*Interest Rate Models: An Introduction*" by Andrew Cairns (Princeton University Press, 2004).

This chapter addresses credit risk – the risk that a person or an organisation will fail to make a payment they have promised. This is quite a new area (*eg* the Jarrow-Lando-Turnbull (JLT) model was only published in 1997) that has grown in importance with the recent introduction of credit derivatives (which we cover in Subjects ST5, ST6 and SA5, not in this course).

We start with some definitions to set the scene, then go on to look at the Merton model, which is a simple model that relates the price of a bond subject to default to the price of an option. We then look at two models that are applications of the theory of continuous-time jump processes, which you may (or may not!) remember from Subject CT4. The first model is a two-state model. The second is the JLT model, which is a more general multi-state model.

1 Credit events and recovery rates

In earlier chapters it has been assumed that bonds are default-free.

Question 18.1

Explain what it means for a bond to be default-free.

This is not an unreasonable assumption when considering most government bonds. However, for corporate bonds where the bond is issued by a corporate entity there is a possibility that they will default at some point during the term of the contact.

These corporate entities consist mainly of large companies and banks. For example, in May 2014, Barclays had a 5¾% bond redeemable in September 2026 and Tesco had a 5½% bond redeemable in December 2019. These companies entered into a contract to make interest payments on set dates to the bondholders and to repay the face value of the bond on the redemption date. Failure to do this would result in the bonds being in default.

The outcome of a default may be that the contracted payment stream is:

- **rescheduled**
- **cancelled by the payment of an amount which is less than the default-free value of the original contract**
- **continued but at a reduced rate**
- **totally wiped out.**

An example of cancellation would be where the bondholders agree to accept a reduced one-off cash payment of 75% of the face value instead of the contractual payments.

A *credit event* is an event that will trigger the default of a bond and includes the following:

- **failure to pay either capital or a coupon**
- **loss event** (*ie* where the company says that it is not going to make a payment)
- **bankruptcy**
- **rating downgrade of the bond by a rating agency such as Standard and Poor's or Moody's.**

In the event of a default, the fraction of the defaulted amount that can be recovered through bankruptcy proceedings or some other form of settlement is known as the *recovery rate*.

2 Modelling credit risk

2.1 Structural models

Structural models are explicit models for a corporate entity issuing both equity and debt. They aim to link default events explicitly to the fortunes of the issuing corporate entity.

A default event is the same as a credit event.

The models are simple and cannot be realistically used to price credit risk. However, studying them does give an insight into the nature of default and the interaction between bondholders and shareholders. An example of a structural model is the Merton model.

These models are called "structural" because they focus on the financial structure – the split between debt and equity – of the company issuing the bond. We will discuss the Merton model in the next section.

2.2 Reduced-form models

Reduced-form models are statistical models, which use observed market statistics rather than specific data relating to the issuing corporate entity. The market statistics most commonly used are the credit ratings issued by credit rating agencies such as Standard and Poor's and Moody's. The credit rating agencies will have used detailed data specific to the issuing corporate entity when setting their rating. They will also regularly review the data to ensure that the rating remains appropriate and will re-rate the bond either up or down as necessary.

Reduced-form models use market statistics along with data on the default-free market to model the movement of the credit rating of the bonds issued by a corporate entity over time. The output of such models is a distribution of the time to default.

These models are called "reduced-form" because they ignore any information relating to the particular company issuing the bond. Instead, they model the different levels of creditworthiness and how companies move from one status to another. They are discussed further in Subjects ST6 and SA5.

2.3 *Intensity-based models*

An intensity-based model is a particular type of continuous-time reduced form model. It typically models the "jumps" between different states (usually credit ratings) using transition intensities.

This approach uses a continuous-time multiple state model (jump process) where the states correspond to the creditworthiness of the company. The "intensities" (denoted by λ) are the rates driving the company to switch from one state to another as time passes. Intensity-based models are an example of a reduced-form model.

3 The Merton model

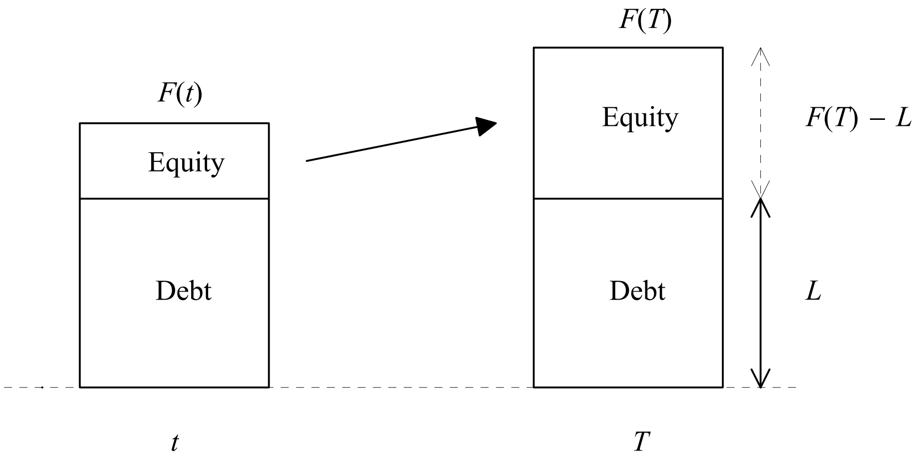
The Merton model is a simple example of a structural model. We introduce it here in Subject CT8 and it is discussed further in Subjects ST5, ST6 and SA5.

Merton's model assumes that a corporate entity has issued both equity and debt such that its total value at time t is $F(t)$. This value varies over time as a result of actions by the corporate entity, which does not pay dividends on its equity or coupons on its bonds.

For example, the value of a company will change along with investors' perceptions of the future prospects of that company.

Part of the corporate entity's value is zero-coupon debt with a promised repayment amount of L at a future time T. At time T the remainder of the value of the corporate entity will be distributed to the shareholders and the corporate entity will be wound up.

Recall from Subject CT2 that debt holders rank above shareholders in the wind-up of a company. So, provided the company has sufficient funds to pay the debt, the shareholders will receive $F(T)-L$.



The corporate entity will default if the total value of its assets, $F(T)$ is less than the promised debt repayment at time T.

ie $\quad F(T) < L$

In this situation, the bondholders will receive $F(T)$ instead of L and the shareholders will receive nothing.

Combining these two cases, we see that the shareholders will receive a payoff of $\max[F(T)-L,0]$ at time T.

This can be regarded as treating the shareholders of the corporate entity as having a European call option on the assets of the company with maturity *T* and a strike price equal to the value of the debt. The Merton model can be used to estimate either the risk-neutral probability that the company will default or the credit spread on the debt.

Because of the risk of default, a bond issued by a company will have a lower market price than a similar bond issued by a government. So the yield on the company bond will be higher than the yield on the government bond. The credit spread refers to the difference between these two rates.

Example

Consider a company initially (at time t) worth £10 million with no debt:

Debt	£0
Equity	£10m
Total asset value	£10m

Suppose that, at time t, the company borrows £5m by issuing bonds. This will increase the value of the company's assets to £15m (since it will now be sitting on £5m in cash). But its *net asset value* will still be £10m (because it now has a liability of £5m for the bonds).

Debt	£5m	
Equity	£10m	
Total asset value	£15m	(Net asset value = £10m)

If on the redemption date, at time T, the company's total value $F(T)$ is less than £5m (*ie* the company's equity is negative), then the shareholders will receive nothing because the bondholders must be paid first.

So, the total payoff for the two categories of investors will be:

Shareholders: $\max\left[F(T)-5,0\right]$

Bondholders: $\min\left[F(T),5\right]$

Question 18.2

Suggest how this model could be used to calculate the value of the risky corporate bonds at time t and also the risk-neutral probability of default.

4 Two-state models for credit ratings

We will now consider a two-state intensity-based model, which is the simplest continuous-time reduced-form model.

A simple model can be set up, in continuous time, with two states:

1. **N = not previously defaulted**
2. **D = defaulted.**

Under this model it is assumed that the default-free interest rate term structure is deterministic with $r(t) = r$ for all t.

In other words, the risk-free rate r is assumed to be constant.

If the transition intensity, under the real-world measure P, from N to D at time t is denoted by $\lambda(t)$, this model can be represented as:

$$\boxed{\textbf{No default, N}} \xrightarrow{\lambda(t)} \boxed{\textbf{Default, D}}$$

Recall from Subject CT4 that this is a two-state continuous-time Markov model. It has the same structure as the two-state mortality model, with "No default" corresponding to "Alive", "Default" corresponding to "Dead" and $\lambda(t)$ corresponding to the force of mortality at time t.

> ***Question 18.3***
>
> Is the model described here time-homogeneous or time-inhomogeneous?

If $X(t)$ is the state at time t, the transition intensity $\lambda(t)$ can be interpreted in terms of the transition probabilities over the infinitesimal time period $[t, t+dt]$.

ie $$P\left[X(t+dt) = N \middle| X(t) = N\right] = 1 - \lambda(t)dt + o(dt)$$

and $$P\left[X(t+dt) = D \middle| X(t) = N\right] = \lambda(t)dt + o(dt)$$

These equations tell us the probabilities that a company that has not defaulted by time t doesn't or does default during the next instant dt. Recall that $o(dt)$ is a "smaller order" quantity that disappears as the length of the time interval dt tends to zero.

Let τ be a stopping time defined as:

$$\tau = \inf\left\{t \middle| X(t) = D\right\}$$

The timeline stretching from now into the future contains an infinite number of times t. The mathematical symbol "inf" (short for "infimum") is a generalised form of "min" (minimum) that can be applied to infinite sets. So τ represents the earliest time that the company is in State D. If the company never moves to State D, then $\tau = \infty$.

Let $N(t)$ be a counting process defined as:

$$N(t) = \begin{cases} 0 & \text{if } \tau > t \\ 1 & \text{if } \tau \le t \end{cases}$$

In fact, this is a very simple "counting process" – it can only count up to 1! $N(t)$ is either 0 or 1 (an "on/off" indicator) depending on whether or not the company has defaulted yet.

τ can be interpreted as the time of default and $N(t)$ can be interpreted as the number of defaults up to and including time t.

It is assumed that, if the corporate entity defaults, all bond payments will be reduced by a known, deterministic factor:

$$1 - \delta$$

where δ is the recovery rate.

So, for example, if $\delta = 0.9$, we are assuming that, once the company has gone into default, all future interest payments and the redemption payment will be reduced by 10% (since $1 - \delta = 0.1$). So bondholders will receive 90% (the recovery rate) of the full amounts.

So, for a zero-coupon bond that is due to pay 1 at time T, the actual payment at time T will be:

$$\text{payment} = \begin{cases} 1 & \text{if } \tau > T \\ \delta & \text{if } \tau \le T \end{cases}$$

Question 18.4

Write down a single formula (*ie* one without if's) for the payment at time T.

As well as the transition intensity $\lambda(t)$ under P, there is an alternative probability measure Q with a different intensity $\tilde{\lambda}(t)$.

The notation here is consistent with earlier chapters. P is the real-world probability measure and Q is the risk-neutral probability measure. If we work in terms of Q, this will allow us to value the bond using the risk-neutral pricing formula.

Recall from Chapter 15 that, for P and Q to be equivalent, it is necessary that, for each $t > 0$:

$$\lambda(t) > 0 \text{ if and only if } \tilde{\lambda}(t) > 0$$

Question 18.5

Explain this result in words.

Question 18.6

Write down a formula for the probability (under Q) that a company that is in State N at time t will remain in State N until time T.

Hint: Use the survival probability formula ${}_t p_x = \exp\left(-\int_0^t \mu_{x+s} ds\right)$ *for the corresponding two-state mortality model from Subject CT4.*

Result 16.1

Let $B(t,T)$ be the price at time t of a zero-coupon bond that matures at time T. Then there exists a risk-neutral measure Q equivalent to P under which:

$$\begin{aligned} B(t,T) &= e^{-r(T-t)} E_Q\left[\text{payoff at time } T \middle| F_t\right] \\ &= e^{-r(T-t)} E_Q\left[1-(1-\delta)N(T) \middle| F_t\right] \end{aligned}$$

Since we are assuming that δ is constant, we can simplify this further to get:

$$B(t,T) = e^{-r(T-t)}\left\{1-(1-\delta)E_Q[N(T) \mid F_t]\right\}$$

The company cannot be in default when it first issues the bond at time t, *ie* $N(t)=0$. So, conditioning on F_t is equivalent to calculating the expectation given that $N(t)=0$. The expectation $E_Q[N(T)\mid F_t]$ in the previous equation then becomes:

$$\begin{aligned}E_Q[N(T)\mid N(t)=0] &= 1\times P[N(T)=1\mid N(t)=0]+0\times P[N(T)=0\mid N(t)=0]\\ &= P[N(T)=1\mid N(t)=0]\\ &= 1-P[N(T)=0\mid N(t)=0]\end{aligned}$$

We deduced informally a formula for this probability in the previous self-assessment question (by analogy with the two-state mortality model). So we know that:

$$E_Q[N(T)\mid N(t)=0]=1-\exp\left(-\int_t^T \tilde{\lambda}(s)ds\right)$$

In the general case where $\tilde{\lambda}(s)$ is allowed to vary randomly over time, we can use the following result (which is not proved in Subject CT8):

$$\boldsymbol{E_Q\left[N(T)\middle|N(t)=0\right]=E_Q\left[1-\exp\left(-\int_t^T\tilde{\lambda}(s)ds\right)\right]}$$

Note that, in the general case, the integral of $\tilde{\lambda}(s)$ will be a random quantity, so an extra expectation is required in this equation. We can now substitute this expectation into the formula given above to find the bond price $B(t,T)$.

However, we will assume here that $\tilde{\lambda}(s)$ is deterministic, which implies that:

$$\boldsymbol{B(t,T)=e^{-r(T-t)}\left[1-(1-\delta)\left\{1-\exp\left(-\int_t^T\tilde{\lambda}(s)ds\right)\right\}\right]}$$

Multiplying both sides by $e^{r(T-t)}$, then multiplying out the right-hand side, gives:

$$\begin{aligned}e^{r(T-t)}B(t,T) &= 1-\left\{1-\exp\left(-\int_t^T\tilde{\lambda}(s)ds\right)\right\}+\delta\left\{1-\exp\left(-\int_t^T\tilde{\lambda}(s)ds\right)\right\}\\ &= \exp\left(-\int_t^T\tilde{\lambda}(s)ds\right)+\delta-\delta\exp\left(-\int_t^T\tilde{\lambda}(s)ds\right)\\ &= (1-\delta)\exp\left(-\int_t^T\tilde{\lambda}(s)ds\right)+\delta\end{aligned}$$

Moving the constant δ to the left-hand side, then taking logs, we get:

$$e^{r(T-t)}B(t,T)-\delta=(1-\delta)\exp\left(-\int_t^T \tilde{\lambda}(s)ds\right)$$

$$\log\left[e^{r(T-t)}B(t,T)-\delta\right]=\log(1-\delta)-\int_t^T \tilde{\lambda}(s)ds$$

If we now differentiate both sides with respect to T, we get:

$$\frac{\partial}{\partial T}\log\left[e^{r(T-t)}B(t,T)-\delta\right]=-\tilde{\lambda}(T)$$

Remember that differentiating an integral (with respect to the upper limit) takes you back to the original function.

Finally, renaming the variable T as s, swapping the two sides and flipping the signs gives:

$$\boldsymbol{\tilde{\lambda}(s)=-\frac{\partial}{\partial s}\log\left[e^{r(s-t)}B(t,s)-\delta\right]}$$

This gives us a formula for estimating $\tilde{\lambda}(s)$, the risk-neutral transition rate at a future time s, based on observed bond prices at the current time t. (This calculation is similar to estimating a person's future force of mortality based on a set of life assurance premium quotations we've obtained for pure endowment policies maturing at different ages.)

More generally, knowing both the default-free and bond term structures and making an assumption about the recovery rate enables the implied risk-neutral transition intensities to be determined.

Dividing companies into just two states – "No default" (N) and "Default" (D) – is a rather crude, black and white approach. In reality, the "N" state is a heterogeneous category and we can improve this model by introducing some shades of grey. This involves subdividing the "N" state.

In the next section we will extend the two-state model to cover multiple credit ratings.

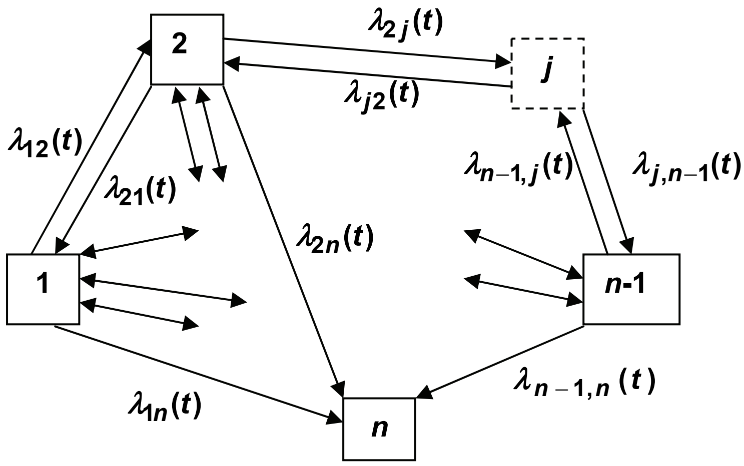
5 The Jarrow-Lando-Turnbull (JLT) model

There are several established credit rating agencies, such as Standard & Poor's and Moody's, who publish credit ratings for all the major companies. On Standard & Poor's scale, companies that have already defaulted are Grade D. Companies that have not defaulted are given one of the seven grades: AAA, AA, A, BBB, BB, B, CCC (which can be fine-tuned further with +'s and –'s). For example, for the two companies whose bonds were mentioned earlier, HBOS has a high credit rating of AA, whereas Boots has a lower rating of BBB+. (These may have changed since the time of writing.)

A more general and more realistic model, with multiple credit ratings rather than the simplistic default/no default model used above, was developed by Jarrow, Lando and Turnbull. In this model there are $n-1$ credit ratings plus default.

So, for example, we could model the Standard & Poor's rating system if we used $n=8$. This would give $n-1=7$ credit ratings, from AAA (= State 1) down to CCC (= State 7), for companies that are not already in default. The nth state (= state 8) would be for companies that are already in default (which we assume they stay in for ever).

We define the transition intensity, under the real-world measure P, from State i to State j at time t to be $\lambda_{ij}(t)$. If the transition intensities $\lambda_{ij}(t)$ are assumed to be deterministic, then this model for default risk can be represented by the following diagram:



In this n-state model, transfer is possible between all states except for the default State n, which is absorbing.

If $X(t)$ is the state or credit rating at time t, then, for $i = 1, 2, \ldots, n-1$, we have:

$$P\left[X(t+dt) = j \middle| X(t) = i\right] = \begin{cases} \lambda_{ij}(t)\,dt + o(dt) & \text{for } j \neq i \\ 1 - \sum\limits_{j \neq i} \lambda_{ij}(t)\,dt + o(dt) & \text{for } j = i \end{cases}$$

$$= \begin{cases} \lambda_{ij}(t)\,dt + o(dt) & \text{for } j \neq i \\ 1 + \lambda_{ii}(t)\,dt + o(dt) & \text{for } j = i \end{cases}$$

where:

- $\lim\limits_{dt \to 0} o(dt) = 0$
- $\lambda_{ii}(t) = -\sum\limits_{j \neq i} \lambda_{ij}(t)$

Also, as State n (default) is absorbing, $\lambda_{nj}(t) = 0$ for all j and for all t.

The λ's here correspond to the entries in the *generator matrix*, which you may remember from the chapter on Markov jump processes in Subject CT4.

The value of λ when the two subscripts are equal (*ie* λ_{ii}) does not correspond to an actual transition – if you start and end in State i, you haven't actually made a jump! However, defining λ_{ii} to equal minus the sum of all the actual transition rates out of State i allows us to write the set of equations for this process in a compact form.

Question 18.7

For a particular 4-state version of the JLT model:

$$\lambda_{12} = \lambda_{21} = 0.1,\ \lambda_{23} = 0.2,\ \lambda_{24} = 0.1,\ \lambda_{32} = 0.1,\ \lambda_{34} = 0.4,\ \lambda_{13} = \lambda_{14} = \lambda_{31} = 0$$

Construct the complete generator matrix.

In Subject CT4 we saw how to use the generator matrix to write down a set of differential equations (the Kolmogorov differential equations), which could then be solved to find the probabilities of making specified transitions over specified periods.

In fact it is possible to calculate the transition probabilities directly from the generator matrix, using a method involving matrix calculations, which we will describe now. (You'll be pleased to know that no differential equations are involved!)

We introduce the following notation:

- $\Lambda(t)$ **is an** $n \times n$ **intensity matrix,** $\Lambda(t) = \left(\lambda_{ij}(t)\right)_{i,j=1}^{n}$
- $p_{ij}(s,t) = P\left[X(t) = j \middle| X(s) = i\right]$ **for** $t > s$
- $\Pi(s,t) = \left(p_{ij}(s,t)\right)_{i,j=1}^{n}$ **is the matrix of transition probabilities**

> ***Question 18.8***
>
> State in words what $\Lambda(t)$, $p_{ij}(s,t)$ and $\Pi(s,t)$ represent.

Then it can be shown that:

$$\Pi(s,t) = \exp\left[\int_s^t \Lambda(u)\,du\right]$$

where, for an $n \times n$ **matrix** M**, we define:**

$$e^M \equiv I + \sum_{k=1}^{\infty} \frac{1}{k!} M^k$$

The proof of this is not given in Subject CT8.

The e^x function on your calculator can be generalised to an exponential function for matrices, written as $\exp(M)$ or e^M. You can see that its definition corresponds to the familiar series $e^x = 1 + x + \frac{1}{2}x^2 + \cdots$.

In the matrix version, $I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$ is the identity matrix, which has 1's down the main diagonal and zeros everywhere else. The matrix exponential function shares most of the familiar properties of the scalar version, such as $\exp(\mathbf{0}) = I$, where $\mathbf{0}$ is a matrix consisting entirely of zeros, and $\exp(A)\exp(A) = \exp(2A)$.

Question 18.9

In a particular two-state intensity-based model with constant transition intensities, the integrated generator matrix $M = \int_s^t \Lambda(u)du$ has the form $M = \begin{pmatrix} -a & a \\ b & -b \end{pmatrix}$, where $a = 0.2(t-s)$ and $b = 0.1(t-s)$.

(i) Show that, in this case, $M^2 = -(a+b)M$.

(ii) Hence deduce an explicit formula for e^M in terms of a and b.

(iii) Hence deduce a set of formulae for the transition probabilities $p_{ij}(s,t)$.

As before, if we want to use this model to calculate bond prices, we need to work with the *risk-neutral* probabilities.

As well as the transition probabilities and intensities under P, there is also a set of deterministic transition intensities $\tilde{\lambda}_{ij}(t)$ for $i, j = 1,\ldots,n$ under a risk-neutral measure Q with:

- $\tilde{\lambda}_{ij}(t) > 0 \Leftrightarrow \lambda_{ij}(t) > 0$ **for all** $j \neq i$
- $\tilde{\lambda}_{ii}(t) = -\sum_{j \neq i} \tilde{\lambda}_{ij}(t)$
- $\tilde{\lambda}_{nj}(t) = 0$ **for all** j **and for all** t**.**

The first condition above is necessary for P and Q to be equivalent measures.

These equivalent intensities give the equivalent matrix of transition probabilities:

$$\tilde{\Pi}(s,t) = \exp\left[\int_s^t \tilde{\Lambda}(u)du\right]$$

The price of a zero-coupon bond maturing at time T, which pays 1 if default has not yet occurred and δ if default has occurred and for which the credit rating of the underlying corporate entity is i, is:

$$B(t,T,X(t)) = P(t,T)\left\{1-(1-\delta)P_Q\left[X(T)=n \,|\, F_t\right]\right\}$$

assuming that interest rates and default transitions are independent. (The proof of this formula is not given in Subject CT8.)

In this equation, $P(t,T)$ is the market price at time t of a *default-free* zero-coupon bond (*eg* a government bond) that pays 1 unit at time T, so it acts as a discount factor. Using this factor in place of $e^{-r(T-t)}$ means that we can apply the same formula in situations where the risk-free interest rate is not constant over time.

If we know that the company is currently in State i, we can write this formula in the more direct form:

$$B(t,T,X(t)=i)=P(t,T)\left\{1-(1-\delta)P_Q[X(T)=n\,|\,X(t)=i]\right\}$$

So, if we can evaluate $P_Q[X(T)=n\,|\,X(t)=i]$, which is the risk-neutral probability $\tilde{p}_{in}(t,T)$, we can use the model to calculate the fair price of the bond, allowing for the possibility of default.

This formula is a generalisation of the pricing formula we had for the two-state model, which we could have written in the form:

$$B(t,T)=e^{-r(T-t)}\left\{1-(1-\delta)P_Q[N(T)=1\,|\,F_t]\right\}$$

This model allows us to calculate a price for a bond subject to default risk, and can be extended to deal with coupon-bearing bonds.

This calculated value will provide a useful benchmark when deciding whether to purchase a new issue whose market value is not yet known.

6 The two-state model with a stochastic transition intensity

The Jarrow, Lando and Turnbull approach assumes that the transition intensities between states are deterministic.

Although we have allowed the values of the transition intensities $\lambda(t)$ between any two states to vary over time, we have assumed that the functions involved are known with certainty at the outset. In real life, however, economic conditions can change unpredictably. If, for example, a recession struck, we would expect the $\lambda(t)$'s corresponding to jumps to a higher-numbered state (*ie* a state closer to the default State n) to increase significantly, as companies struggled to remain profitable.

An alternative approach would be to assume that the transition intensity between states, $\lambda(t)$, is stochastic and dependent on a separate state variable process, $X(t)$.

Note that this process $X(t)$ is different to the process $X(t)$ you met earlier in the chapter. This process would model a macro-economic indicator that we think is a good measure of the overall well-being of companies in the economy at time t.

Using this approach in a situation where there are just two states:

1. ***N* = not previously defaulted**
2. ***D* = previously defaulted,**

the time of default τ will be the first jump time in a Cox process with intensity $\lambda(t)$.

The Cox process referred to here (not to be confused with the Cox-Ingersoll-Ross model in the previous chapter) is a Poisson process in which the Poisson parameter $\lambda(t)$ is a function of time. We are only interested in the time till the first event (*ie* default) occurs.

By using a stochastic approach, $\lambda(t)$ can be allowed to vary with company fortunes and other economic factors. For example, a rise in interest rates may make default more likely and so $X(t)$ could include appropriate allowance for changes in interest rates. This approach can be used to develop models for credit risk that combine the structural modelling and intensity-based approaches.

7 Exam-style questions

We finish this chapter with two exam-style questions.

Question 1

To fund an expansion in its operations, a company has just issued 5-year zero-coupon bonds with a total face value of £10 million, taking its total asset value up to £15 million.

(i) Explain how the value of the bonds can be expressed in terms of a European put option. [3]

(ii) Hence calculate the fair price of a holding of the company bonds with a face value of £100 using the Black-Scholes model, given that the price of a 5-year zero-coupon government bond is £77.88. Assume that the annualised volatility of the company's assets over the 5-year period is 25%. [4]

(iii) Explain what is meant by a credit spread and calculate its value for the company bonds. [3]

[Total 10]

Solution 1

(i) ***Expressing the bond value as an option***

This is the Merton model.

After the bond issue the total current value of the company will be $F(0)=15$ (working in £million). In 5 years' time the company will have an unknown value of $F(5)$.

The bondholders have first call on the company's assets at that time. So they will receive 10 if $F(5)\geq 10$ and $F(5)$ if $F(5)<10$.

So the redemption payment from the bonds is $\min[F(5),10]$, which can be written as $10-\max[10-F(5),0]$. The function $\max[10-F(5),0]$ is the payoff for a European put option on $F(t)$ maturing at time 5 with strike price 10.

(ii) ***Calculate the fair price of the bonds***

The parameters for valuing the put option with payoff $\max[10-F(5),0]$ using the Black-Scholes formula are:

$$F(0)=15\ (=S_0),\quad K=10,\quad T-t=5,\quad \sigma=0.25\quad (\text{and } q=0)$$

The risk-free interest rate is found from the equation:

$$100e^{-5r}=77.88 \quad \Rightarrow r=0.050$$

Using page 47 of the *Tables*, we find that:

$$d_1=1.45204,\quad d_2=0.89302$$

$$\begin{aligned} p &= 10e^{-0.25}\Phi(-0.89302)-15\Phi(-1.45204) \\ &= 7.788\times 0.18592-15\times 0.07325=0.349 \end{aligned}$$

So the total value of the bonds now (remembering that the 10 in the payoff function is a future payment and needs to be discounted) is:

$$Ke^{-r(T-t)}-p=7.788-0.349=7.439 \quad \textit{ie } £7.439 \text{ million}$$

So the fair price is £74.39 per £100 face value.

(iii) ***Credit spread***

The price of the company bonds (£74.39) is less than the price of the government bonds (£77.88) because of the risk of default, *ie* the company may not make the redemption payment in full on the due date.

As a result, the yield r_B (assuming full payment) will be slightly higher. It can be found from the equation:

$$100e^{-5r_B} = 74.39 \quad \Rightarrow r = 0.0592$$

The difference between the bond yield of 5.92% and the default-free rate of 5% is called the *credit spread*. So here the credit spread is 0.92% (per annum, continuously compounded).

Question 2

Company X has just issued some 5-year zero-coupon bonds. A continuous-time two-state model is to be used to model the status of the company and to calculate the fair price of the bonds. It is believed that the risk-neutral transition rate for failure of the company is $\lambda(t) = 0.002t$, where t is the time in years since the issue of the bonds. The 5-year risk-free spot yield is 5.25% expressed as an annual effective rate.

(i) Calculate the risk-neutral probability that the company will have failed by the end of 5 years. [2]

(ii) In the event of failure of the company, the bonds will make a reduced payment at the maturity date. The recovery rate for a payment due at time t is:

$$\delta(t) = 1 - 0.05t$$

Calculate the fair price to pay for £100 nominal of a Company X bond, taking into account the possibility of company failure. [3]

(iii) An analyst is concerned that the estimate of $\lambda(t)$ may be too simplistic. Explain the possible reasons for his concern and how the model could be developed to deal with this. [3]

[Total 8]

Solution 2

(i) ***Probability of company failure***

The risk-neutral probability of company failure, by time n, can be expressed in terms of the transition rate $\lambda(t)$:

$$p(n)=1-\exp\left(-\int_0^n \lambda(t)dt\right)$$

With the transition rate given, we have:

$$p(n)=1-\exp\left(-\int_0^n 0.002t\,dt\right)$$

$$=1-\exp\left[-0.001t^2\right]_0^n$$

$$=1-e^{-0.001n^2}$$

For $n=5$ this is:

$$p(5)=1-e^{-0.001\times 5^2}=0.02469$$

(ii) ***Fair price of a Company X bond***

The recovery rate at time 5 is:

$$\delta(5)=1-0.05(5)=0.75$$

The risk-neutral expected payment at maturity can then be found using the probability calculated in part (i):

$$0.02469\times 0.75+(1-0.02469)\times 1=0.99383$$

We can discount this using the 5-year spot rate to get the fair price of the bond:

$$0.99383\times(1.0525)^{-5}=0.7695$$

ie £76.95 per £100 nominal.

(iii) ***Stochastic transition rates***

The analyst may be concerned because the fair price of the bond is critically dependent on an accurate assessment of the default transition intensity. If this assessment is incorrect then the bond could be mispriced.

An alternative approach would be to assume that the transition intensity $\lambda(t)$ is stochastic and dependent on a separate state variable process, $X(t)$ say.

The time of default will be the first jump time in a Cox process with intensity $\lambda(t)$.

By using a stochastic approach, $\lambda(t)$ can be allowed to vary with company fortunes and other economic factors.

For example, a rise in interest rates may make default more likely and so $X(t)$ could include appropriate allowance for changes in interest rates. This approach can be used to develop models for credit risk that combine the structural modelling and intensity-based approaches.

8 End of Part 4

What next?

1. Briefly **review** the key areas of Part 4 and/or re-read the **summaries** at the end of Chapters 15 to 18.
2. Attempt some of the questions in Part 4 of the **Question and Answer Bank**. If you don't have time to do them all, you could save the remainder for use as part of your revision.
3. Attempt **Assignment X4**.

Time to consider – "rehearsal" products

Mock Exam A / AMP and Marking – There are three separate mock exam papers that you can attempt and get marked. A recent student survey found that students who do a mock exam of some form have significantly higher pass rates. Students have said:

> "Excellent comments; very helpful. Especially pointing out my weaknesses in both exam techniques and areas of study. Thank you for your valuable advice."

You can find lots more information on our website at www.ActEd.co.uk.

Buy online at www.ActEd.co.uk/estore

And finally ...

Good luck!

Chapter 18 Summary

Credit events and recovery rates

A *credit event* is an event that will trigger the default of a bond. In the event of a default, the fraction δ of the defaulted amount that can be recovered through bankruptcy proceedings or some other form of settlement is known as the *recovery rate*.

Modelling credit risk

Structural models are models for a company issuing both shares and bonds, which aim to link default events explicitly to the fortunes of the issuing company.

Reduced-form models are statistical models that use observed market statistics such as credit ratings, as opposed to specific data relating to the company.

Intensity-based models are a particular type of continuous-time reduced-form models. They typically model the "jumps" between different states (usually credit ratings) using transition intensities.

The Merton model

Merton's model is a structural model. It assumes that a company has issued both equity and debt such that its total value at time t is $F(t)$. The total value of the bonds issued and the shareholders' interests equals $F(t)$.

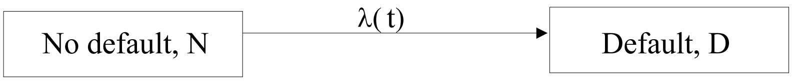
The shareholders of the company can be regarded as having a European call option on the assets of the company with maturity T and a strike price equal to the face value (L) of the debt.

The value of the bonds is equal to the discounted value of L minus the value of a European put option on the assets of the company with maturity T and strike price L. This put represents the shareholders' option to default on the bonds.

The Merton model is tractable and gives us some insight into the nature of default and the interaction between bondholders and shareholders. It can be used to estimate either the risk-neutral probability that the company will default or the credit spread on the debt.

The two-state model for credit ratings

- State N = not previously defaulted
- State D = previously defaulted



If $\lambda(t)$ is the transition intensity from N to D, then the transition probabilities are:

$$P\left[X(t+dt)=N \middle| X(t)=N\right]=1-\lambda(t)\,dt+o(dt)$$
$$P\left[X(t+dt)=D \middle| X(t)=N\right]=\lambda(t)\,dt+o(dt)$$

Let $B(t,T)$ be the price at time t of a risky zero-coupon bond that matures at time T. Then:

$$\begin{aligned} B(t,T) &= e^{-r(T-t)}E_Q\left[\text{payoff at time } T \middle| F_t\right] \\ &= e^{-r(T-t)}E_Q\left[1-(1-\delta)N(T) \middle| F_t\right] \\ &= e^{-r(T-t)}\left[1-(1-\delta)\left\{1-\exp\left(-\int_t^T \tilde{\lambda}(s)ds\right)\right\}\right] \end{aligned}$$

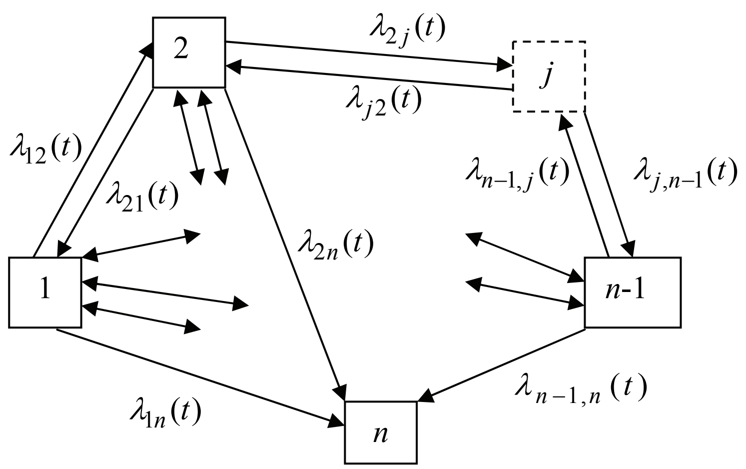
where:

- δ is the recovery rate
- $N(T)$ is the number of defaults by time T
- $\tilde{\lambda}(s)$ is the risk-neutral transition rate or intensity.

The risk-neutral transition rate at a future time s can be estimated from the observed bond prices as:

$$\tilde{\lambda}(s)=-\frac{\partial}{\partial s}\log\left[e^{r(s-t)}B(t,s)-\delta\right]$$

The Jarrow-Lando-Turnbull (JLT) model



In this n-state model, transfer is possible between all states except for State n (default), which is absorbing. If $X(t)$ is the state or credit rating at time t, then, for $i = 1, 2, \ldots, n-1$, the transition probabilities over the short time interval $(t, t+dt)$ are:

$$P\left[X(t+dt) = j \,\middle|\, X(t) = i\right] = \begin{cases} \lambda_{ij}(t)\,dt + o(dt) & \text{for } j \neq i \\ 1 + \lambda_{ii}(t)\,dt + o(dt) & \text{for } j = i \end{cases}$$

If $p_{ij}(s,t) = P\left[X(t) = j \,\middle|\, X(s) = i\right]$ for $t > s$, and the matrix of transition probabilities is:

$$\Pi(s,t) = \left(p_{ij}(s,t)\right)_{i,j=1}^{n}$$

then:

$$\Pi(s,t) = \exp\left[\int_s^t \Lambda(u)\,du\right]$$

where:

$\Lambda(t) = \left(\lambda_{ij}(t)\right)_{i,j=1}^{n}$ is the matrix of transition intensities.

If a zero-coupon bond maturing at time T pays:

- 1 if default has not yet occurred
- δ if default has occurred

and the credit rating of the underlying corporate entity is i, the fair price of the bond is:

$$B\left(t,T,X\left(t\right)=i\right)=P\left(t,T\right)\left\{1-\left(1-\delta\right)P_Q\left[X\left(T\right)=n\,\middle|\,X(t)=i\right]\right\}$$

Two-state models with stochastic transition intensity

The transition intensity $\lambda(t)$ can be allowed to vary stochastically to reflect other economic factors, such as the level of interest rates. This is done by introducing a separate state variable process $X(t)$.

This approach can be used to develop models for credit risk that combine the structural modelling and intensity-based approaches.

Chapter 18 Solutions

Solution 18.1

A bond is default-free if the stream of payments due from the bond will *definitely* be paid *in full* and *on time*.

Solution 18.2

Since we are viewing the equity as a call option on the total value of the company, we could use an option pricing method, such as the Black-Scholes option pricing formula, to calculate how much the equity at time T is worth now (*ie* at time t). The value of the bonds could then be found by subtracting this from the current value of the company $F(t)$, *ie*:

$$B(t) = F(t) - E(t).$$

where $B(t)$ and $E(t)$ are the current value of the company's risky corporate bonds and the equity respectively.

In fact, it can be shown using an option pricing formula analogous to Black-Scholes that $E(t)$ is equal to:

$$E(t) = F(t)\Phi\left(d_1\right) - Le^{-r(T-t)}\Phi\left(d_2\right) \tag{1}$$

where:

$$d_1 = \frac{\ln\left(\dfrac{F(t)}{L}\right) + \left(r + ½\sigma_F^2\right)(T-t)}{\sigma_F\sqrt{T-t}}$$

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

We also know that the total value of the company's equity is equal to its market capitalisation – the number of shares issued times the share price. What we don't know, are the value of the company $F(t)$ and its volatility σ_F.

However, it can be shown that under certain assumptions (about the nature of the stochastic processes governing $F(t)$ and $E(t)$):

$$\sigma_E E(t) = \sigma_F F(t) \frac{\partial E(t)}{\partial F(t)} = \sigma_F F(t) \Phi(d_1) \qquad (2)$$

where σ_E, the volatility of the equity value, can be observed from market data.

So, Equations (1) and (2) then represent two equations in two unknowns, $F(t)$ and σ_F, which can be solved numerically.

The current value of the company's risky corporate bond can then be found as:

$$B(t) = F(t) - E(t)$$

In addition, and as with any call option, $\Phi(d_2)$ represents the risk-neutral probability that the option will be exercised, which here means that the debt will be repaid.

So:

$$1 - \Phi(d_2) = \Phi(-d_2)$$

must represent the risk-neutral probability that the shareholders' option to repay the debt is not exercised, *ie* it represents the *risk-neutral probability of default*.

Solution 18.3

In this model, the transition intensity $\lambda(t)$ is allowed to vary over time. So this is a time-*in*homogeneous model. If we simplified the model further by assuming that $\lambda(t) = \lambda$ (a constant), then it would be time-homogeneous.

Solution 18.4

We can use the indicator variable $N(t)$ to do this. The correct formula for the payment is $1-(1-\delta)N(T)$.

If default has not occurred by time T, then $N(T)=0$ and the formula gives 1, *ie* the full payment will be made. If default *has* occurred, then $N(T)=1$ and the formula gives $1-(1-\delta)=\delta$, *ie* the reduced or "recovered" payment of δ will be made.

We will need this formula for Result 18.1.

Solution 18.5

For the probability measures P and Q to be equivalent, the set of times at which transitions are possible must be the same under each measure.

So, for example, if a transition is possible only between time 5 and time 10 under P, the same must be true under Q. Otherwise, they are not equivalent.

Solution 18.6

The probability we are interested in can be written as $P[X(T)="N" \mid X(t)="N"]$ or $P[N(T)=0 \mid N(t)=0]$.

The formula given in the hint tells us the probability that a person who is alive at age x will still be alive t years later.

So the corresponding formula here is:

$$P[N(T)=0 \mid N(t)=0]=\exp\left(-\int_t^T \tilde{\lambda}(s)ds\right)$$

Solution 18.7

Entering the values we've been given, we get:

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} ? & 0.1 & 0 & 0 \\ 0.1 & ? & 0.2 & 0.1 \\ 0 & 0.1 & ? & 0.4 \\ ? & ? & ? & ? \end{bmatrix} \end{array}$$

We know that State 4 (the default state) is absorbing. So the first three entries along the bottom row will be zero.

The entries down the diagonal are equal to minus the sum of the other entries for that row. For example, $\lambda_{22} = -(0.1+0.2+0.1) = -0.4$.

So the final generator matrix looks like this:

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} -0.1 & 0.1 & 0 & 0 \\ 0.1 & -0.4 & 0.2 & 0.1 \\ 0 & 0.1 & -0.5 & 0.4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Solution 18.8

$\Lambda(t)$ is the generator matrix at time t. The positive entries in this matrix represent the actual transition rates (intensities) from one state to another at that time. The negative entries on the main diagonal are notional values equal to minus the sum of the other entries in the same row (which means that the entries in each row add up to zero).

$p_{ij}(s,t)$ is the probability that a company that is in State i at a particular time s will be in State j at a specified future time t. If s is the current time and t is the time a bond payment is due, this probability tells us how likely it is that the company (currently in State i) will be in each state – and hence how likely it is to default. So these are what we want to use the model to work out.

$\Pi(s,t)$ is a matrix showing the complete set of $n \times n$ transition probabilities for the period from time s to time t. It gives us the same information as working out all the individual probabilities.

Solution 18.9

(i) If $M = \begin{pmatrix} -a & a \\ b & -b \end{pmatrix}$, we find that:

$$M^2 = \begin{pmatrix} -a & a \\ b & -b \end{pmatrix}\begin{pmatrix} -a & a \\ b & -b \end{pmatrix}$$

$$= \begin{pmatrix} (-a)(-a)+ab & (-a)a + a(-b) \\ b(-a)+(-b)b & ba+(-b)(-b) \end{pmatrix}$$

$$= \begin{pmatrix} a^2+ab & -a^2-ab \\ -ba-b^2 & ba+b^2 \end{pmatrix}$$

$$= \begin{pmatrix} a(a+b) & -a(a+b) \\ -b(a+b) & b(a+b) \end{pmatrix}$$

$$= -(a+b)\begin{pmatrix} -a & a \\ b & -b \end{pmatrix}$$

$$= -(a+b)M$$

Extending this to higher powers, we see that:

$$M^3 = M^2M = -(a+b)MM = -(a+b)M^2 = (a+b)^2 M$$

and $M^k = [-(a+b)]^{k-1}M$

(ii) So, in this case, the exponential function is:

$$\exp(M) = I + \sum_{k=1}^{\infty}\frac{1}{k!}M^k$$

$$= I + \sum_{k=1}^{\infty}\frac{1}{k!}[-(a+b)]^{k-1}M$$

$$= I - \frac{1}{a+b}\left(\sum_{k=1}^{\infty}\frac{1}{k!}[-(a+b)]^{k}\right)M$$

$$= I - \left(\frac{e^{-(a+b)}-1}{a+b}\right)M = I + \left(\frac{1-e^{-(a+b)}}{a+b}\right)M$$

(iii) We can then substitute the values of a and b to find the matrix of transition probabilities:

$$\Pi(s,t) = I + \left(\frac{1-e^{-(a+b)}}{a+b}\right)M$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \left(\frac{1-e^{-(a+b)}}{a+b}\right)\begin{pmatrix} -a & a \\ b & -b \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \left(\frac{1-e^{-0.3(t-s)}}{0.3}\right)\begin{pmatrix} -0.2 & 0.2 \\ 0.1 & -0.1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{3}+\frac{2}{3}e^{-0.3(t-s)} & \frac{2}{3}-\frac{2}{3}e^{-0.3(t-s)} \\ \frac{1}{3}-\frac{1}{3}e^{-0.3(t-s)} & \frac{2}{3}+\frac{1}{3}e^{-0.3(t-s)} \end{pmatrix}$$

So the transition probabilities are:

$$p_{11}(s,t) = \tfrac{1}{3}+\tfrac{2}{3}e^{-0.3(t-s)}, \ p_{12}(s,t) = \tfrac{2}{3}-\tfrac{2}{3}e^{-0.3(t-s)}$$

$$p_{21}(s,t) = \tfrac{1}{3}-\tfrac{1}{3}e^{-0.3(t-s)}, \ p_{22}(s,t) = \tfrac{2}{3}+\tfrac{1}{3}e^{-0.3(t-s)}$$

Part 1 – Questions

Introduction

The Question and Answer Bank is divided into 5 parts. The first 4 parts include a combination of:

- bookwork questions
- developmental questions and
- exam-style questions.

We have indicated the type of question throughout. All bookwork questions could be asked in the exam so you should start immediately learning this bookwork well. Developmental questions are not as likely to be asked in the exam, especially in terms of the style of the question. That is not to say that they are not useful! These questions are to help you understand the course better in order that you are in a much stronger position to attempt exam-style questions. Ultimately, the more questions you attempt, the greater your understanding of the course and the greater the chances of being able to answer whatever is asked in the exam.

For all questions, you should attempt the questions by yourself, before looking at the solutions provided. This distinction represents the difference between active studying and passive studying. Given that the examiners will be aiming to set questions to make you think (and in doing so they will be devising questions you have not seen before), it is much better if you practice the skills that they will be testing.

It may also be useful to you if you group a number of the exam-style questions together to attempt under exam time conditions. Ideally three hours should be set aside, but anything from one hour (*ie* 35 marks) upwards will help your time management. Once you have exhausted Parts 1 to 4 of the Question and Answer Bank, you can move on to Part 5, which contains a set of exam-style questions covering the whole course.

Question 1.1 (Bookwork)

(i) Describe what is meant by an "efficient market". [2]

(ii) Describe the three different forms of the Efficient Markets Hypothesis. [3]

(iii) Discuss the implications of the Efficient Markets Hypothesis. [5]
[Total 10]

Question 1.2 (Bookwork)

(i) In the context of semi-strong market efficiency, explain what is meant by informational efficiency. Describe briefly the main difficulties in testing for informational efficiency. [3]

(ii) Discuss in detail the empirical evidence concerning informational efficiency. [11]
[Total 14]

Question 1.3 (Exam-style)

At the quarterly meeting of the Auger Close Investment Club, four members are making proposals for new equity investment for the club.

Albert wants to buy shares in Armadillo Adventures, claiming that they have performed poorly in recent weeks and are due an upturn.

Brian wants to invest in Biscuits-R-Us. They have recruited a new head of marketing, who has had success at other companies. Brian feels that this new appointment will have a positive effect on the firm.

Colin selects shares at random. This quarter he is recommending the club buy into Cash 4 Kidneys PLC.

Dennis wants the club to buy shares in Diamond Dentists ("DD"). His brother works for a major health insurer and has insider information that DD's shares will rise sharply in the near future, when it is announced that his company has appointed DD as its "dentist of choice".

For each club member, describe how their share selection strategy would work in strongly efficient, semi-strongly efficient, weakly efficient and inefficient markets. [7]

Question 1.4 (Exam-style)

(i) Explain what is meant by an "excessively volatile" market. [2]

(ii) Describe how you would test if a market is "excessively volatile". [7]

(iii) Explain the practical and conceptual difficulties in using a test of an excessively volatile market to establish whether or not a market is efficient. [4]

[Total 13]

Question 1.5 (Developmental)

Colin's preferences can be modelled by the utility function such that:

$$U'(w) = 3 - 2w, \ (w > 0).$$

(i) Determine the range of values over which this utility function can be satisfactorily applied. [1]

(ii) Explain how Colin's holdings of risky assets will change as his wealth decreases. [3]

(iii) Which of the following investments will he choose to maximise his expected utility?

Investment A		*Investment B*		*Investment C*	
outcome	*probability*	*outcome*	*probability*	*outcome*	*probability*
0.1	0.3	0	0.3	0.2	0.45
0.3	0.4	0.2	0.2	0.3	0.1
0.5	0.3	0.9	0.5	0.4	0.45

[3]

[Total 7]

Question 1.6 (Developmental)

By considering the relationship $R(w) = w.A(w)$, explain which of the following statements is true.

1. If an investor's preferences display decreasing relative risk aversion then they must also display decreasing absolute risk aversion.

2. If an investor's preferences display decreasing absolute risk aversion then they must also display decreasing relative risk aversion.

[4]

Question 1.7 (Developmental)

A consumer has the choice of the following four consumption bundles:

(a) 20 units of X and 10 units of Y

(b) 15 units of X and 15 units of Y

(c) 10 units of X and 20 units of Y

(d) 12 units of X and 12 units of Y

Explain, as fully as possible, how a consumer would rank these four bundles if the consumer is indifferent between Bundle (a) and Bundle (c). You should assume that X and Y are not perfect substitutes for each other. [6]

Question 1.8 (Exam-style)

Explain the four axioms that are required to derive the expected utility theorem. [6]

Question 1.9 (Exam-style)

Jenny has a quadratic utility function of the form $U(w) = w - 10^{-5}w^2$. She has been offered a job with Company X, in which her salary would depend upon the success or otherwise of the company. If it is successful, which will be the case with probability ¾, then her salary will be $40,000, whereas if it is unsuccessful she will receive $30,000.

(i) Assuming that Jenny has no other wealth, state the salary range over which $U(w)$ is an appropriate representation of her individual preferences. [2]

(ii) Calculate the expected salary and the expected utility offered by the job. [2]

(iii) Suppose she was also to be offered a fixed salary by Company Z. Determine the minimum level of fixed salary that she would accept to work for Company Z in preference to Company X. [3]

(iv) Suppose that the owners of Company X are both risk-neutral and very keen that Jenny should join them and not Company Z. Determine whether the firm should agree to pay her a fixed wage, and, if so, how much. Comment briefly on your answer. [1]

[Total 8]

Question 1.10 (Exam-style)

Suppose that Lance and Allan each have a log utility function and an initial wealth of 100 and 200 respectively. Both are offered a gamble such that they will receive a sum equal to 30% of their wealth should they win, whereas they will lose 10% of their wealth should they lose. The probability of winning is ¼.

(i) State whether or not the gamble is fair. [1]

(ii) Calculate Lance's certainty equivalent for the gamble alone and comment briefly on your answer. [2]

(iii) Repeat part (ii) in respect of Allan and compare your answer with that in part (ii). [2]

(iv) Confirm that your comments in part (iii) apply irrespective of the individual's wealth. [2]

[Total 7]

Question 1.11 (Exam-style)

Consider the two risky assets, A and B, with cumulative probability distribution functions:

$$F_A(w) = w$$

$$F_B(w) = w^{1/2}$$

In both cases, $0 \le w \le 1$.

(i) Show that A is preferred to B on the basis of first-order stochastic dominance. [3]

(ii) Verify explicitly that A also dominates B on the basis of second-order stochastic dominance. [3]

[Total 6]

Question 1.12 (Exam-style)

(i) Within the context of behavioural finance, explain fully what is meant by overconfidence. [4]

The board of directors of an actively managed investment trust are concerned that the decisions of the trust's investment manager may subject to overconfidence bias, which could adversely affect the performance of the trust.

(ii) Discuss possible actions that the board could take in order to try and limit the impact of the investment manager's overconfidence bias. [6]

[Total 10]

Question 1.13 (Exam-style)

(i) Define framing and outline how it relates to some of the other behavioural finance themes. [5]

(ii) The marketing team of a unit trust management company are updating the promotional literature for their actively managed UK Equity trust. The trust produced a return of 5% over the last calendar year, compared to 4% for the FTSE 100 index. Give examples of how the marketing team could present the past performance of the fund favourably to prospective investors. [4]

(iii) A financial advisor is due to present a range of possible investment strategies to a high net worth individual. Explain the decisions she will need to make with regard to the presentation of the possible strategies. [4]
[Total 13]

Question 1.14 (Developmental)

Adam, Brett and Charlie are all offered the choice of investing their entire portfolio in either a risk-free asset or a risky asset. The risk-free asset offers a return of 0% *pa*, whereas the returns on the risky asset are uniformly distributed over the range –5% to +10% *pa*. Assuming that each individual makes his investment choice in order to minimise his expected shortfall, and that they have benchmark returns of –2%, 0% and +2% *pa* respectively, who will choose which investment? Comment briefly on your answer. [9]

Question 1.15 (Developmental)

(i) Define "shortfall probability" for a continuous random variable. [1]

(ii) An investor holds an asset that produces a random rate of return, R, over the course of a year. Calculate the shortfall probability using a benchmark rate of return of 1%, assuming:

(a) R follows a lognormal distribution with $\mu = 5\%$ and $\sigma^2 = (5\%)^2$

(b) R follows an exponential distribution with a mean return of 5%. [3]

(iii) Explain with the aid of simple numerical examples the two main limitations of the shortfall probability as a basis for making investment decisions. [4]
[Total 8]

Question 1.16 (Exam-style)

Consider a zero-coupon corporate bond that promises to pay a return of 10% next period. Suppose that there is a 10% chance that the issuing company will default on the bond payment, in which case there is an equal chance of receiving a return of either 5% or 0%.

(i) Calculate values for the following measures of investment risk:

(a) downside semi-variance

(b) shortfall probability based on the risk-free rate of return of 6%

(c) the expected shortfall below the risk-free return conditional on a shortfall occurring. [5]

(ii) Discuss the usefulness of downside semi-variance as a measure of investment risk for an investor. [3]

[Total 8]

Question 1.17 (Exam-style)

An investor is contemplating an investment with a return of £R, where:

$$R = 250{,}000 - 100{,}000N$$

and N is a Normal $[1,1]$ random variable.

Calculate each of the following measures of risk:

(a) variance of return

(b) downside semi-variance of return

(c) shortfall probability, where the shortfall level is £50,000

(d) Value at Risk at the 5% level

(e) Tail Value at Risk at the 5% level, conditional on the VaR being exceeded.

[13]

Hint: For part (e), you may wish to use the formula for the truncated first moment of a normal distribution given on page 18 of the Tables.

Question 1.18 (Bookwork)

Explain what is meant by the following terms, in the context of mean-variance portfolio theory:

(i) efficient frontier [1]

(ii) indifference curves [1]

(iii) optimal portfolio. [2]
[Total 4]

Question 1.19 (Developmental)

An investor can invest in only two risky assets A and B. Asset A has an expected rate of return of 10% and a standard deviation of return of 20%. Asset B has an expected rate of return of 15% and a standard deviation of return of 30%. The correlation coefficient between the returns of Asset A and the returns of Asset B is 0.6.

(i) What is the expected rate of return if 20% of an investor's wealth is invested in Asset A and the remainder is invested in Asset B? [1]

(ii) What is the standard deviation of return on the portfolio if 20% of an investor's wealth is invested in Asset A and the remainder is invested in Asset B? [1½]

(iii) Explain why an investor who invests 20% of his wealth in Asset A and the remainder in Asset B is risk-averse. [1½]
[Total 4]

Question 1.20 (Bookwork)

Consider a portfolio, P, which consists of N assets held in equal proportions. Let R_P represent the return on the portfolio, and let R_i represent the return on asset i. The covariance of the return on asset i with that on asset j is C_{ij}.

(i) State the total number of data items needed to calculate $E(R_P)$ and $\text{var}(R_P)$. [2]

(ii) Write down an expression for $\text{var}(R_P)$. [2]

(iii) Using your expression from part (ii), show that the specific risk of the portfolio (*ie* the risk associated with the individual assets) tends to zero in a well-diversified portfolio. [3]
[Total 7]

Question 1.21 (Exam-style)

(i) Describe in detail the assumptions underlying the use of mean-variance portfolio theory. [3]

Consider a two-security world in which the returns yielded by Assets 1 and 2 are perfectly positively correlated, though they have different expected returns.

(ii) Using the method of Lagrangian multipliers or otherwise, derive the equation of the efficient frontier in expected return-standard deviation space. [6]

(iii) Use your answer in (i) to:

(a) determine the gradient of the efficient frontier

(b) show that the efficient frontier is a straight line in expected return-standard deviation space that passes through the points representing Assets 1 and 2. [5]
[Total 14]

Question 1.22 (Exam-style)

Consider a world in which there are only 2 securities, 1 and 2, such that:

$$E_1 = 5\%, \quad V_1 = (10\%)^2$$

$$E_2 = 10\%, \quad V_2 = (20\%)^2$$

Let ρ denote the correlation coefficient between the returns yielded by the two securities.

(i) Derive the equation of the opportunity set in *E–V* space. [5]

(ii) Derive expressions for the portfolio expected return *E* and the portfolio proportion x_1 invested in Security 1 at the point of global minimum variance and hence comment briefly on how *E* and x_1 vary with ρ. [5]

[Total 10]

Question 1.23 (Exam-style)

Show that in the single-index model of asset returns:

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

$$V_i = \beta_i^2 V_M + V_{ei}$$

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

where V_{ei} is the variance of e_i. [8]

Question 1.24 (Developmental)

Consider the data in the table below, which relates to Securities 1, 2 and 3.

	Security		
	1	2	3
α_i	0.0	2.0	–2.2
β_i	1.1	0.6	2.0
$V_{\varepsilon i}$	2.2	1.3	1.2

You are given that:

- the expected return and standard deviation of the market return are 10 and 5 respectively
- the returns of each security can be modelled using an appropriate single-index model.

(i) Calculate:

(a) the expected return and standard deviation of return for each security

(b) the covariance of returns between each pair of securities. [4]

(ii) Consider a portfolio which consists of Securities 1, 2 and 3 in equal proportions. Calculate:

(a) the variance of the portfolio

(b) the systematic risk of the portfolio

(c) the specific risk of the portfolio. [3]

[Total 7]

Question 1.25 (Bookwork)

Distinguish between the three main classes of multifactor model. [6]

Part 1 – Solutions

Solution 1.1

(i) ***Definition of efficient market***

An efficient market is one in which every security's price equals its investment value at all times. [1]

In an efficient market information is fully reflected in the price. [½]

This means that share prices adjust instantaneously and without bias to new information. [½]
[Total 2]

(ii) ***Three forms of Efficient Markets Hypothesis***

The strong form requires that prices reflect all information that is currently known – whether or not it is publicly available. [1]

The semi-strong form requires that prices reflect all information that is publicly available. [1]

The weak form requires that prices fully reflect all information contained in the past history of prices. [1]
[Total 3]

(iii) ***Implications of the Efficient Markets Hypothesis***

The past history of prices is a subset of publicly available information, so a market must be weak form efficient if it is semi-strong form efficient. Similarly, if it is strong form efficient it must also be semi-strong and weak form efficient. [1]

The Efficient Markets Hypothesis does not imply that beating the market is impossible, since investors could out-perform the market by chance, or by accepting above average levels of market or systematic risk. [1]

However, it does imply that it is not possible consistently to achieve superior (*ie* risk-adjusted) investment performance net of costs without access to superior information.
[½]

Weak form efficiency implies that it is impossible to achieve excess risk-adjusted investment returns purely by using trading rules based upon the past history of prices and trading volumes. It therefore suggests that technical analysis cannot be justified. [½]

If only weak form efficiency applies, excess risk-adjusted returns are still possible by good fundamental analysis of public information. [½]

The semi-strong form means that prices adjust instantaneously and without bias to newly published information. This implies that it is not possible to trade profitably on information gained from public sources. So neither fundamental analysis (without insider information) nor technical analysis will yield excess risk-adjusted returns. [½]

Fundamental analysis may still, however, aid the investor in selecting the investments that are most suitable for meeting its investment needs and objectives. [½]

If the strong form is correct then the market reflects all known knowledge about the company and consequently excess risk-adjusted returns are possible only by chance. This implies that insiders cannot profit from dealing on inside information, *ie* insider trading is not profitable. [½]

[Total 5]

Solution 1.2

(i) ***Informational efficiency***

The market for a particular security is said to exhibit informational efficiency if new information is incorporated quickly and accurately into the price of the security. [1]

The main difficulties in testing for informational efficiency are:

- It is hard to establish exactly when information arrives. This may be because many events are widely rumoured prior to being announced officially. [1]

- The theory of market efficiency does not predict exactly how a particular piece of information should be incorporated into the price of a security, *ie* what effect it will have on the price. We therefore need to make an assumption about this. However, this means that if the price does not change in the way we expected, this may be due to an invalid assumption, rather than an inefficient market. [1]

[Total 3]

(ii) ***Discuss the empirical evidence***

The empirical evidence suggests that markets do over-react to certain events and under-react to other events and that the over/under-reaction is corrected over a long time period. [1]

If this is true then traders might be able to take advantage of the slow correction of the market to make excess risk-adjusted returns, and so efficiency would not hold. [1]

Some of the effects found by studies can be classified as over-reaction to events.

For example:

1. Past performance: past winners tend to be future losers and vice versa. The market appears to over-react to past performance. [1]

2. Certain accounting ratios appear to have predictive powers, *eg* companies with high earnings to price, cashflow to price and book value to market value – generally poor past performers – tend to have high future returns. Again, this is an example of the market apparently over-reacting to past growth. [1]

3. Firms coming to the market: in the US evidence appears to support the idea that stocks coming to the market by Initial Public Offerings and Seasoned Equity Offerings have poor subsequent performance. [1]

There are also well-documented examples of under-reaction to events:

1. Stock prices continue to respond to earnings announcements up to a year after their announcement. This is an example of under-reaction to information that is slowly corrected. [1]

2. Abnormal excess returns for both the parent and subsidiary firms following a de-merger. This is another example of the market being slow to recognise the benefits of an event. [1]

3. Abnormal negative returns following mergers: agreed takeovers lead to the poorest subsequent returns. The market appears to over-estimate the benefits from mergers, the stock price then slowly reacts as its optimistic view is proved to be wrong. [1]

Even in an efficient market, pure chance would be expected to throw up some apparent examples of mispricings. We would expect to see as many examples of over-reaction as under-reaction. This is broadly consistent with the empirical evidence to date. [1]

Even more important is that the reported effects do not appear to persist over prolonged time periods and so may not represent exploitable opportunities to make excess profits. This again is consistent with an efficient market. [1]

Other examples of pricing anomalies, for example the ability of accounting ratios to indicate out-performance, are arguably proxies for risk. Once these risks have been taken into account many empirical studies that claim to show evidence of inefficiency turn out to be compatible with the Efficient Markets Hypothesis. [1]

[Total 11]

Solution 1.3

This is Question 3 from the Certificate in Practical Financial Economics exam of April 2005.

Albert

Albert makes his recommendation based on the past price history of the investment. If weak form EMH holds, then the current share price already reflects the information contained in the past price history, so there would be no advantage in using this approach. [1]

Similarly, if the semi-strong or strong form of EMH holds, there is no advantage in using this approach. [½]

If the market was inefficient, Albert's strategy may be beneficial. [½]

Brian

Brian makes his recommendation based on company information that is in the public domain. If semi-strong form EMH holds, then the current share price already reflects relevant public information, so there would be no advantage in using this approach. [1]

Similarly, if the strong form of EMH holds, there is no advantage in using this approach. [½]

If the market is inefficient or only weak form efficient, Brian's strategy may be beneficial. [½]

Colin

The approach of choosing stocks at random provides no advantage, whatever the level of market efficiency. [½]

If strong form EMH holds, this strategy is no worse than any other. [½]

Dennis

Dennis makes his recommendation based on insider information. If strong form EMH holds, then the current share price already reflects all relevant information, so there would be no advantage in using this approach. [1]

If the market is inefficient or weak or semi-strong form efficient, Dennis's strategy may be beneficial (though it could be questionable on ethical grounds). [1]

[Total 7]

Solution 1.4

This question is taken from Subject 109, April 2002, Question 2.

(i) ***Excessively volatile markets***

An excessively volatile market is one in which the changes in the market values of stocks (the observed volatility) are greater than can be justified by the news arriving. This is claimed to be evidence of market over-reaction, which is not compatible with efficiency. [2]

(ii) ***Testing if a market is excessively volatile***

To test if a market is excessively volatile you need a long history of prices and cash flows for one of the securities in question – *eg* for the market in a particular equity, you would need many months or years of share prices and dividend payments. [1½]

A discounted cash flow model based on the actual dividends that were paid and some terminal value for the share could then be used to calculate a perfect foresight price for the equity. This would represent the "correct" equity price if market participants had been able to predict future dividends correctly. [1½]

The difference between the perfect foresight price and the actual price arises from the forecast errors of future dividends. If market participants are rational, there should be no systematic forecast errors. [1½]

Also if markets are efficient, then broad movements in the perfect foresight price should be correlated with moves in the actual price as both are reacting to the same news and hence the same changes in the anticipated future cash flows. [1½]

If instead the actual price changes are greater, then this would suggest that the market in the particular equity is excessively volatile. [1]

[Total 7]

This was the approach adopted by Shiller.

(iii) ***Practical and conceptual difficulties***

These include:

- the difficulty of choosing an appropriate terminal value for the share price [1]
- the difficulty of choosing an appropriate discount rate with which to discount future cash flows – in particular, should it be constant? [1]
- possible biases in the estimates of the variances because of autocorrelation in the time series data used [1]
- possible non-stationarity of the time series data used, *ie* it may have stochastic trends which invalidate the measurements obtained for the variance of the stock price [1]
- the distributional assumptions underlying the statistical tests used might not be satisfied [½]
- the distributional characteristics of the share prices and dividends are unlikely to remain constant over a long period of time. [½]

[Maximum 4]

Solution 1.5

(i) ***Range of wealth applicable***

Assuming *non-satiation*, which requires that $U'(w) > 0$, Colin's preferences can be modelled by this utility function provided that $0 < w < {}^3/_2$. [1]

(ii) ***How Colin's holdings of risky assets vary with his wealth***

Differentiating the expression given in the question yields $U''(w) = -2$.

Thus, over the relevant range of w:

$$A(w) = \frac{2}{3-2w} > 0\,,\ A'(w) = \frac{4}{(3-2w)^2} > 0$$ [½]

and $$R(w) = \frac{2w}{3-2w} > 0\,,\ R'(w) = \frac{6}{(3-2w)^2} > 0$$ [½]

Hence, as Colin's wealth *decreases* the:

- *absolute amount* of his investment in risky assets will *increase* (as his absolute risk aversion decreases as his wealth decreases) [1]

- *proportion* of his wealth that is invested in risky assets will *increase* (as his relative risk aversion decreases as his wealth decreases). [1]

[Total 3]

(iii) ***Colin's choice of investments***

Integrating the expression in the question gives Colin's utility function:

$$U(w) = a + 3w - w^2$$

As the properties of utility functions are invariant to linear transformations, we can set the arbitrary constant a equal to zero. [1]

His expected utility from each of the investments is therefore as follows.

$$EU_A = 0.3 \times U(.1) + 0.4 \times U(.3) + 0.3 \times U(.5)$$

$$= 0.3 \times 0.29 + 0.4 \times 0.81 + 0.3 \times 1.25$$

$$= 0.786$$ [½]

$$EU_B = 0.3 \times U(0) + 0.2 \times U(.2) + 0.5 \times U(.9)$$

$$= 1.057$$ [½]

$$EU_C = 0.45 \times U(.2) + 0.1 \times U(.3) + 0.45 \times U(.4)$$

$$= 0.801$$ [½]

Thus, Colin will choose Investment B to maximise his expected utility. [½]
[Total 3]

Solution 1.6

The relationship between absolute risk aversion $A(w)$ and relative risk aversion $R(w)$ is such that:

$$R(w) = w.A(w)$$

Differentiating with respect to wealth w gives:

$$\frac{\partial R}{\partial w} = A + w.\frac{\partial A}{\partial w} \qquad (1)$$ [1]

Considering the first statement. Equation 1 tells us that if $\frac{\partial R}{\partial w} < 0$ and so relative risk aversion is decreasing, then it must also be the case that $\frac{\partial A}{\partial w} < 0$ (given that w and $A(w)$ are both positive for a risk-averse individual), *ie* $\frac{\partial R}{\partial w} < 0 \Rightarrow \frac{\partial A}{\partial w} < 0$. [1]

An investor who displays decreasing relative risk aversion invests a larger proportion of his wealth in risky assets as his wealth increases. This also implies a larger monetary amount is invested in risky assets, ie decreasing absolute risk aversion.

Conversely, if we consider the second statement, then if $\frac{\partial A}{\partial w} < 0$, it does not follow that $\frac{\partial R}{\partial w}$ is necessarily negative. This will depend upon the relative magnitudes of $A(w)$, w and $\frac{\partial A}{\partial w}$. Thus, $\frac{\partial A}{\partial w} < 0$ does not imply that $\frac{\partial R}{\partial w} < 0$. [1]

An investor who displays decreasing absolute risk aversion invests a larger monetary amount in risky assets as his wealth increases. This does not necessarily equate to a larger percentage of his wealth.

Hence, the first statement is true, whereas the second statement is false. [1]
[Total 4]

Solution 1.7

- (b) will be preferred to (d) … [1]
- … because (b) contains more of both goods and we assume that consumers prefer more to less [1]
- (b) will be preferred to both of (a) and (c) … [1]
- … because (b) = ½(a) + ½(c) and we assume that a consumer's tastes exhibit a diminishing marginal rate of substitution [1]
- it is not possible to say whether (d) is preferred to (a) and (c) [1]
- (d) will either give higher utility than *both* (a) and (c) or lower utility than *both* (a) and (c) or the same utility as *both* (a) and (c) [1]

[Total 6]

Solution 1.8

This is Question 5 from the Subject 109 exam of September 2000.

The expected utility theorem can be derived formally from the following four axioms:

1. Comparability [½]

An investor can state a preference between all available certain outcomes. [1]

2. Transitivity [½]

If A is preferred to B and B is preferred to C, then A is preferred to C. [1]

3. Independence [½]

If an investor is indifferent between two certain outcomes, A and B, then he is also indifferent between the following two gambles:

(i) A with probability p and C with probability $(1-p)$; and

(ii) B with probability p and C with probability $(1-p)$. [1]

4. Certainty equivalence [½]

Suppose that A is preferred to B and B is preferred to C. Then there is a unique probability, p, such that the investor is indifferent between B and a gamble giving A with probability p and C with probability $(1-p)$.

B is known as the *certainty equivalent* of the above gamble. [1]
[Total 6]

Solution 1.9

(i) ***Salary range of utility function***

If: $U(w) = w - 10^{-5} w^2$

then: $U'(w) = 1 - 2w \times 10^{-5}$ [½]

and: $U''(w) = -2 \times 10^{-5}$ [½]

Now in order for Jenny to:

- prefer more to less, we require that $U'(w) > 0$, which in this case will be true for all $w < ½ \times 10^5$, *ie* $w < 50{,}000$
- be risk-averse, we require that $U''(w) < 0$, which in this case will be true for all $w > 0$.

Thus, the appropriate salary range is $w < \$50{,}000$. [1]
[Total 2]

(ii) ***Expected salary and expected utility***

Her expected salary is given by:

$$\tfrac{3}{4} \times 40{,}000 + \tfrac{1}{4} \times 30{,}000 = \$37{,}500$$ [1]

Her expected utility is given by:

$$\tfrac{3}{4} \times [(40{,}000 - 10^{-5}\times(40{,}000)^2] + \tfrac{1}{4} \times [(30{,}000 - 10^{-5}\times(30{,}000)^2] = 23{,}250$$ [1]

[Total 2]

(iii) ***Minimum fixed salary***

The minimum level of salary, x say, is equal to the certainty equivalent of the job offer from Company X. [½]

This is given by:

$$U(x) = 23,250$$

$$x - 10^{-5}x^2 = 23,250$$

$$-x + 10^{-5}x^2 + 23,250 = 0$$ [½]

Using the formula for solving for quadratic equations we find:

$$x = 63{,}229 \text{ or } 36{,}771$$ [1½]

As the first of these values is greater than the maximum salary available when Company X is successful it can be disregarded. Hence the minimum level of fixed salary that she would accept to work for Company Z is $36,771. [½]

[Total 3]

(iv) ***Should Company X offer a fixed salary?***

Yes – if they are risk-neutral, then they should offer Jenny a fixed salary in preference to a variable one. Jenny is risk-averse and therefore derives additional utility from the certainty offered by a fixed salary. [½]

Therefore, Company X will be able to entice Jenny to work for them in return for a salary of just (or strictly speaking slightly above) $36,771, instead of the expected salary of $37,500 in (i). [½]

[Total 1]

Solution 1.10

(i) ***Is the gamble fair?***

For any given initial level of wealth w, the expected value of the gamble is given by:

$$\tfrac{1}{4} \times 1.3w + \tfrac{3}{4} \times 0.9w - w = 0$$

Thus, the gamble is fair. [1]

(ii) ***Lance's certainty equivalent of the gamble alone***

Lance's expected utility should he undertake the gamble is given by:

$$EU = \tfrac{1}{4}\log(130) + \tfrac{3}{4}\log(90) = 4.59174$$ [½]

Thus, his certainty equivalent for the initial wealth and the gamble is given by:

$$U(c_w) = \log(c_w) = 4.59174$$

$$\Rightarrow \quad c_w = e^{4.59174} = 98.666$$

and the certainty equivalent for the gamble alone is given by:

$$c_x = c_w - w = -1.334$$ [1]

This is negative because he is risk-averse. [½]
[Total 2]

The negative value of c_x means that Lance would have to be paid to accept the gamble.

(iii) ***Allan's certainty equivalent of the gamble alone***

Allan's expected utility should he undertake the gamble is given by:

$$EU = \tfrac{1}{4}\log(260) + \tfrac{3}{4}\log(180) = 5.28489$$ [½]

His certainty equivalent for the gamble alone is:

$$c_x = e^{5.28489} - 200 = -2.668$$ [1]

Comparing the two answers, we can see that the two certainty equivalents are equal to the same proportion of each individual's initial wealth. This is because the log utility function is consistent with preferences that exhibit constant *relative* risk aversion. [½]
[Total 2]

(iv) ***Relative risk aversion***

The constancy of relative risk aversion with a log utility function can be confirmed by differentiating it, *ie*:

If: $U(w) = \log(w)$

then: $U'(w) = \dfrac{1}{w}$ and $U''(w) = -\dfrac{1}{w^2}$

Thus: $R(w) = -w\dfrac{U''(w)}{U'(w)} = 1$ and $R'(w) = 0$

So, the log utility function exhibits constant relative risk aversion irrespective of w – though the log utility function is of course defined only for $w > 0$. [2]

Solution 1.11

(i) ***First-order stochastic dominance***

A is preferred to B on the basis of first-order stochastic dominance if:

$F_A(w) \le F_B(w)$, for all $0 \le w \le 1$, and

$F_A(w) < F_B(w)$ for some value of w in this range. [½]

This is the case if:

$w \le w^{½}$ [½]

$\Leftrightarrow \quad w - w^{½} \le 0$

$\Leftrightarrow \quad w^{½}(w^{½} - 1) \le 0$

This clearly holds for all $0 \le w \le 1$, the equality being strict for $0 < w < 1$. Hence A first-order dominates B. [2]
[Total 3]

Alternatively, we could draw the graphs of $F_A(w)$ *and* $F_B(w)$ *over the range* $0 \le w \le 1$ *and note that the graph of* $F_A(w)$ *is below that of* $F_B(w)$ *for* $0 < w < 1$ *and equal for* $w = 0,1$.

(ii) ***Second-order stochastic dominance***

A second-order dominates B if $G_A(w) \le G_B(w)$ for all $0 \le w \le 1$, with the strict inequality holding for some value of *w*, where $G(w) = \int_0^w F(y)dy$. [½]

This is the case if:

$$\int_0^w y\, dy \le \int_0^w y^{½} dy$$ [½]

$$\Leftrightarrow \quad \left[½y^2\right]_0^w \le \left[\tfrac{2}{3}y^{\frac{3}{2}}\right]_0^w$$

$$\Leftrightarrow \quad ½w^2 - \tfrac{2}{3}w^{\frac{3}{2}} \le 0$$

$$\Leftrightarrow \quad \tfrac{2}{3}w^{\frac{3}{2}}(\,¾w^{½} - 1) \le 0$$

This is true for all $0 \le w \le 1$ and strictly true for $0 < w \le 1$. Hence A does second-order dominate B. [2]

[Total 3]

Alternatively, we note that $G(w) = \int_0^w F(y)dy$ *is the area under the graph of* $F(y)$. *We could plots graphs of* $G_A(w)$ *and* $G_B(w)$ *and note that the graph of* $G_A(w)$ *is below that of* $G_B(w)$ *for* $0 < w \le 1$ *and equal for* $w = 0$.

Solution 1.12

(i) ***Explain fully overconfidence***

People tend to overestimate their own abilities, knowledge and skills. [1]

Moreover, studies show that the discrepancy between accuracy and overconfidence increases (in all but the simplest tasks) as the respondent is more knowledgeable. (Accuracy increases to a modest degree but confidence increases to a much larger degree.) [1]

This may be a result of:

- *Hindsight bias* – events that happen will be thought of as having been predictable prior to the event; events that did not happen will be thought of as having been unlikely ever to happen. [1]

- *Confirmation bias* – people will tend to look for evidence that confirms their point of view (and will tend to dismiss evidence that does not justify it). [1]

[Total 4]

(ii) ***Possible actions to limit the extent of overconfidence bias***

The board could require that all investment decisions made by the investment manager are reviewed by a second investment manager before being implemented. [1]

Alternatively, the management of the investment trust could be split equally between two investment managers. [1]

Either of these should reduce the impact of overconfidence bias, although the views of the second investment manager could be subject to similar overconfidence biases as those of the first manager. [1]

The investment manager could be sent on a training course about behavioural finance, to make him aware of his possible overconfidence bias. [1]

However, highlighting his possible overconfidence may make him believe that he is no longer subject to overconfidence bias, making him even more confident in his abilities and more subject to overconfidence bias. [1]

The board could place tighter constraints on the investment decisions taken by the investment manager, *eg* limits could be placed on the size of any transactions and/or on the size of holdings in individual companies or sectors. [1]

Limiting his actions should limit the scope for biases in his investment decisions but will also reduce his scope for active investment management and possibly the returns he achieves. [1]

[Maximum 6]

Solution 1.13

(i) ***Define framing***

Framing refers to the fact that the way a choice is presented or "framed" and, particularly, the wording of a question in terms of gains and losses, can have an enormous impact on the answer given or the decision made. [1]

Changes in the way a question is framed of only a word or two can have a profound effect. [½]

In the same way, "structured response" questions are found to convey an implicit range of acceptable answers. [½]

Framing is related to:

- *anchoring and adjustment* – as the answer to a question will tend to be anchored by a suggested answer where one is presented [1]
- *prospect theory* – as decisions will be influenced by how information is presented in terms of gains or losses relative to some benchmark [1]
- the *effect of options* – as the ordering of a range of options can influence the choice of option or options selected. [1]

[Total 5]

(ii) ***How to present the past performance of the fund favourably***

The marketing team could choose to focus on past performance over a different time period, during which the unit trust has produced an annualised rate of return that is higher than 5%, *eg* 2 years or 5 years. [1]

They could highlight the performance of the unit trust against the FTSE 100, which the unit trust outperformed by 1% over the last calendar year. [1]

They could highlight the performance of the unit trust against a selection of other similar actively managed unit trusts that returned less than 5% over the last calendar year. [1]

They could highlight the performance of the unit trust against a selection of other similar but *passively* managed unit trusts that returned less than 5% over the last calendar year. [1]

They could highlight the performance of the unit trust against other asset classes that returned less than 5% over the last calendar year, *eg* cash. [1]

[Maximum 4]

(iii) ***Decisions with regard to the presentation of the possible investment strategies***

She will need to decide:

- how many possible investment strategies to present – remembering that the more options she presents, the more difficult it may be for the investor to reach a decision [1]

- which strategies to present – bearing in mind that the investor is more likely to choose an intermediate option than an outlying one. [1]

  In particular, she must decide whether to include the current strategy, and to highlight it as such, as the investor may tend to stick to that due to *status quo bias* / *regret aversion*. [1]

- how the strategies should be presented – remembering that the choice made may be influenced by how the options are *framed* [1]

- the order in which to present the strategies – bearing in mind that the investor may be more likely to choose the first strategy (*primary effect*) or the last strategy (the *recency effect*). [1]

[Maximum 4]

Solution 1.14

The expected shortfall below a benchmark level L is defined as:

$$\int_{-\infty}^{L}(L-x)f(x)\,dx \qquad [1]$$

For the risky asset, we have (working in percentage units):

$$f(x)=\begin{cases}\dfrac{1}{10-(-5)}=\dfrac{1}{15} & \text{if } -5<x<10\\[2ex] 0 & \text{otherwise}\end{cases} \qquad [1]$$

Adam

The expected shortfall of the *risky asset* is given by:

$$\int_{-5}^{-2}\frac{(-2-x)}{15}\,dx$$

$$=\frac{1}{15}\left[-2x-½x^2\right]_{-5}^{-2}$$

$$=\frac{1}{15}\left[(4-2)-(10-12½)\right]$$

$$=0.3\% \qquad [1]$$

Alternatively, you can note that there is a chance of $1/5$ *that he will earn less than the benchmark, and in this case, the average shortfall will be 1½%. So the expected shortfall will be* $1/5\times 1½\%=0.3\%$.

The expected shortfall of the *risk-free* asset is 0%. [½]

So Adam chooses the risk-free asset. [½]

Brett

The expected shortfall of the *risky asset* is given by:

$$\int_{-5}^{0} \frac{(-x)}{15}\, dx$$

$$= \frac{1}{15}\left[-½x^2\right]_{-5}^{-0}$$

$$= \frac{1}{15}\left[(0)-(-12½)\right]$$

$$= 0.833\%$$ [1]

The expected shortfall of the *risk-free* asset is again 0%. [½]

So Brett chooses the risk-free asset. [½]

Charlie

The expected shortfall of the *risky asset* is given by:

$$\int_{-5}^{2} \frac{(2-x)}{15}\, dx$$

$$= \frac{1}{15}\left[2x-½x^2\right]_{-5}^{2}$$

$$= \frac{1}{15}\left[(4-2)-(10-12½)\right]$$

$$= 1.633\%$$ [1]

The expected shortfall of the *risk-free* asset is 2% × 1 = 2%. [½]

So Charlie chooses the risky asset. [½]

Thus, the expected shortfall increases with the benchmark return. [1]
[Total 9]

Solution 1.15

(i) ***Definition***

The shortfall probability for a continuous random variable, X, is:

$$P(X < L) = \int_{-\infty}^{L} f(x)\,dx$$

where L is the chosen benchmark level of wealth. [1]

(ii)(a) ***Calculation based on lognormal distribution***

If $R \sim \log N(5,25)$, then the shortfall probability is:

$$P(R<1) = P\left(Z < \frac{\ln 1 - 5}{\sqrt{25}}\right) = P(Z < -1) \qquad [½]$$

where $Z \sim N(0,1)$. Therefore:

$$P(R<1) = \Phi(-1) = 1 - \Phi(1) = 1 - 0.84134 = 0.15866 \qquad [1]$$

(ii)(b) ***Calculation based on exponential distribution***

Since we know that $E(R) = 5$, this means that $R \sim Exp(0.2)$. [½]

So the shortfall probability is:

$$P(R<1) = \int_0^1 0.2e^{-0.2x}\,dx$$

$$= \left[-e^{-0.2x}\right]_0^1$$

$$= 1 - e^{-0.2} = 0.18127 \qquad [1]$$

[Total 3]

(iii) ***Two main limitations***

The two main limitations of the shortfall probability are:

1. It completely ignores *upside risk – ie* all levels of wealth greater than L. [1]

   Thus, if $L = 0$, then the investor will prefer a gamble that offers 0 with certainty to one that offers \$1,000,000 with probability 0.999, but –\$0.0001 with probability 0.001. This seems unlikely to be the case in practice. [1]

2. It ignores the *extent* of the shortfall below the benchmark L. [1]

   Thus, if $L = 0$, then the investor will prefer a gamble that offers +\$1 with probability 0.51 and –\$1,000,000 with probability 0.49 to one that that offers either \$1,000,000 or –\$2, each with probability of ½. This is also somewhat unrealistic. [1]

[Total 4]

Solution 1.16

(i)(a) ***Downside semi-variance***

The expected return on the bond is given by:

$$0.90 \times 10\% \ + \ 0.05 \times 5\% \ + 0.05 \times 0\% \ = \ 9.25\% \qquad [1]$$

So the downside semi-variance is equal to:

$$(9.25 - 5)^2 \times 0.05 \ + \ (9.25 - 0)^2 \times 0.05 \ = \ 5.18\%\% \qquad [1]$$

(i)(b) ***Shortfall probability***

The probability of receiving less than 6% is equal to the sum of the probabilities of receiving 5% and 0%, *ie* 0.10. [1]

(i)(c) ***Expected conditional shortfall***

The expected shortfall below the risk-free rate of 6% is given by:

$$(6 - 5) \times 0.05 \ + \ (6 - 0) \times 0.05 \ = \ 0.35\% \qquad [1]$$

The expected shortfall below the risk-free return *conditional on a shortfall occurring* is equal to:

$$\frac{\text{expected shortfall}}{\text{shortfall probability}} = \frac{0.35\%}{0.10} = 3.5\% \qquad [1]$$

[Total 5]

We can see this directly by noting that, given that there is a shortfall, it is equally likely to be 1% or 6%. So the expected conditional shortfall is 3½%.

(ii) ***Usefulness of downside semi-variance***

- It gives more weight to downside risk, *ie* variability of investment returns below the mean, than to upside risk. [½]
- In fact, it completely ignores risk above the mean. [½]
- This is consistent with the investor being risk-neutral above the mean, which is unlikely to be the case in practice. [½]
- The mean is an arbitrary benchmark, which might not be appropriate for the particular investor. [½]
- If investment returns are symmetrically distributed about the mean (as they would be, for example, with a normal distribution) then it will give equivalent results to the variance. [½]
- However, it is less mathematically tractable than the variance. [½]

[Total 3]

Solution 1.17

This is an extended version of Question 1 from the Subject CT8 exam of September 2006.

(a) ***Variance of return***

N has a Normal $[1,1]$ distribution, so R has a Normal distribution with mean $150{,}000$ and variance $100{,}000^2$, *ie* $R \sim N[150000, 10^{10}]$.

So, the variance of return is 10^{10}. [2]

(b) ***Downside semi-variance of return***

Any normal distribution is symmetrical about its mean, so that the downside semi-variance of return is equal to half of the variance, *ie* 5×10^9. [2]

(c) ***Shortfall probability, where the shortfall level is £50,000***

The shortfall probability below £50,000 is:

$$\begin{aligned}P(R<50,000) &= P\left(\frac{R-150,000}{100,000}<\frac{50,000-150,000}{100,000}\right)\\ &=\Phi\left(\frac{50,000-150,000}{100,000}\right)\\ &=\Phi(-1)\\ &=1-\Phi(1)=1-0.84134=0.15866\end{aligned}$$ [2]

(d) ***Value at risk***

From the *Tables*:

$$\Phi(-1.6449)=0.05$$

So, there is a 5% chance of the investment return *R* having a value less than:

$$\begin{aligned}R_{5\%} &= \mu_R-1.6449\sigma_R\\ &=150,000-1.6449\times100,000\\ &=-14,490\end{aligned}$$

So, the value at risk at the 5% level is £14,490. [2]

(e) ***Tail Value at Risk***

The VaR is £14,490. So, the formula for the conditional TVaR is:

$$\frac{1}{0.05}\int_{-\infty}^{-14,490}(-14,490-x)f(x)\,dx$$

where $f(x)$ is the *pdf* of a $N(150,000,\ 100,000^2)$ distribution. [1]

Splitting this into two integrals:

$$\frac{-14,490}{0.05}\int_{-\infty}^{-14,490} f(x)\,dx - \frac{1}{0.05}\int_{-\infty}^{-14,490} x\,f(x)\,dx$$ [½]

Evaluating the first of these integrals:

$$\begin{aligned}\int_{-\infty}^{-14,490} f(x)\,dx &= P\left(N(150,000,100,000^2) < -14,490\right)\\ &= P\left(Z < \frac{-14,490-150,000}{100,000}\right)\\ &= P(Z < -1.6449)\\ &= 0.05\end{aligned}$$ [½]

This is as expected since $-14,490$ is the VaR at the 5% level.

Evaluating the second integral, using the formula for the truncated first moment of a normal random variable on page 18 of the *Tables*:

$$\int_{-\infty}^{-14,490} x\,f(x)\,dx = 150,000\left[\Phi(U') - \Phi(L')\right] - 100,000\left[\phi(U') - \phi(L')\right]$$ [½]

where:

$$U' = \frac{-14,490-150,000}{100,000} = -1.6449 \quad \Rightarrow \quad \Phi(-1.6449) = 0.05$$ [¼]

and:

$$L' = \frac{-\infty-150,000}{100,000} = -\infty \quad \Rightarrow \quad \Phi(-\infty) = 0$$ [¼]

Now $\Phi(t)$ is the cumulative distribution function of the standard normal distribution, so:

$$\Phi(U') = \Phi(-1.6449) = 0.05$$ [¼]

and:

$$\Phi(L') = \Phi(-\infty) = 0$$ [¼]

Also, $\phi(t)$ is the probability density function of the standard normal distribution, which is stated on page 160 of the *Tables*. So:

$$\phi(U') = \phi(-1.6449) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(-1.6449)^2} = 0.10313$$ [¼]

and

$$\phi(L') = \phi(-\infty) = 0$$ [¼]

Putting these values into the formula:

$$\int_{-\infty}^{-14,490} x\, f(x)\, dx = 150,000\left[0.05 - 0\right] - 100,000\left[0.10313 - 0\right]$$
$$= -2,813$$ [½]

Putting all this together, the conditional TVaR is:

$$\frac{-14,490}{0.05} \times 0.05 - \frac{1}{0.05} \times (-2,813) = £41,770$$ [½]

So, the overall expected loss, given that the 5% VaR is exceeded is:

$$14,490 + 41,770 = £56,260$$

[Total 13]

Solution 1.18

(i) ***Efficient frontier***

The efficient frontier is the line that joins the points in expected return-standard deviation space that represent efficient portfolios. These portfolios will be combinations of one or more investments that give the highest expected rate of return for a given level of risk (in terms of the standard deviation of returns), or the lowest standard deviation of returns for a given expected return. [1]

(ii) ***Indifference curves***

Indifference curves in expected return-standard deviation space join together points representing all the portfolios that give the investor equal levels of expected utility, given the risk-return preferences of that particular investor. They slope upwards for a risk-averse investor. [1]

(iii) ***Optimal portfolio***

The investor's optimal portfolio is the portfolio on the efficient frontier that gives the highest possible level of expected utility, given the investor's particular indifference curves. [1]

It is represented by the point in expected return-standard deviation space where the efficient frontier is tangential to the highest attainable indifference curve. [1]
[Total 2]

Solution 1.19

(i) ***Expected return***

If the expected return on the portfolio be $E[R_P]$, then:

$$E[R_P] = 0.2 \times 0.1 + 0.8 \times 0.15 = 0.14 \text{ or } 14\%$$ [1]

(ii) ***Standard deviation of return***

If the standard deviation of return on the portfolio be σ_P, then:

$$\sigma_P^2 = x_A^2\sigma_A^2 + x_B^2\sigma_B^2 + 2x_A x_B \rho \sigma_A \sigma_B$$ [½]

where the x's denote the proportion of the portfolio invested in each asset, and ρ is the coefficient between the returns of Assets A and B.

Thus:

$$\sigma_P^2 = 0.2^2 0.2^2 + 0.8^2 0.3^2 + 2 \times 0.2 \times 0.8 \times 0.6 \times 0.2 \times 0.3$$
$$= 0.07072$$ [½]

So: $\sigma_P = 0.26593$ or 26.6%. [½]

(iii) ***Explain why the investor is risk-averse***

In modern portfolio theory, risk is measured by standard deviation of returns.

As the returns on the investments are not perfectly correlated, the investor can achieve an expected return that is proportionate to the proportions invested in each security, but with a standard deviation of returns that is lower than the arithmetical average of the individual standard deviations. [½]

An investor who does this rather than investing 100% in Asset B is demonstrating risk aversion, because he is choosing to reduce the level of risk at the cost of attaining a lower expected return. [1]

Solution 1.20

(i) ***Data items***

To work out $E(R_P)$, we would need to know the expected return of each of the N assets. So this requires N data items. [½]

To work out $\text{var}(R_P)$, we would need to know:

- the variance of each of the N assets (N data items) [½]
- the covariance of each different pair of assets (an additional $\dfrac{N(N-1)}{2}$ data items). [½]

The number of covariances required is the number of different ways of choosing 2 assets from N.

So the total number of data items needed is:

$$N+N+\frac{N(N-1)}{2}=\frac{N(N+3)}{2} \qquad [½]$$

[Total 2]

(ii) ***Variance***

The proportion of the portfolio invested in each asset is $\frac{1}{N}$. So the variance of the portfolio is given by:

$$\text{var}(R_P) = \text{var}\left(\sum_{i=1}^{N}\left(\frac{1}{N}\right)R_i\right) = \sum_{i=1}^{N}\left(\frac{1}{N}\right)^2 V_i + \sum_{j=1}^{N}\sum_{\substack{i=1 \\ i\neq j}}^{N}\left(\frac{1}{N}\right)^2 C_{ij}$$

where V_i is the variance of asset i and C_{ij} is the covariance of the return on asset i with that on asset j. [2]

This can alternatively be written as:

$$\text{var}(R_P) = \sum_{i=1}^{N}\left(\frac{1}{N}\right)^2 V_i + 2\sum_{j=1}^{N}\sum_{\substack{i=1 \\ i>j}}^{N}\left(\frac{1}{N}\right)^2 C_{ij}$$

(iii) ***Effect of diversification on specific risk***

The expression for the variance can be re-written as:

$$\text{var}(R_P) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{V_i}{N}\right) + \frac{N-1}{N}\sum_{j=1}^{N}\sum_{\substack{i=1 \\ i\neq j}}^{N}\left(\frac{C_{ij}}{N(N-1)}\right) \qquad [1]$$

Let V^* represent the average variance, and C^* represent the average covariance. Then:

$$V^* = \sum_{i=1}^{N}\left(\frac{V_i}{N}\right) \quad \text{and} \quad C^* = \sum_{j=1}^{N}\sum_{\substack{i=1 \\ i\neq j}}^{N}\left(\frac{C_{ij}}{N(N-1)}\right)$$

since there are N variances to average and $N(N-1)$ covariances in total. [½]

Note, from (i), that there are only $\frac{N(N-1)}{2}$ ***different*** *covariances since* $C_{ij} = C_{ji}$.

So:

$$\text{var}(R_P) = \frac{1}{N}V^* + \frac{N-1}{N}C^* \qquad [½]$$

As $N \to \infty$, the contribution to the overall portfolio variance of the individual variances (through V^*) tends to zero. So, the specific risk associated with the individual securities can be diversified away, but the contribution to the total risk (or variance) from the covariance terms cannot be diversified away. [1]

[Total 3]

Solution 1.21

(i) ***Assumptions underlying mean-variance portfolio theory***

- All expected returns, variances and covariances of pairs of assets are known. [½]
- Investors make their decisions purely on the basis of expected return and variance. [½]
- Investors are non-satiated. [½]
- Investors are risk-averse. [½]
- There is a fixed single-step time period. [½]
- There are no taxes or transaction costs. [½]
- Assets may be held in any amounts, *ie* short-selling, infinitely divisible holdings, no maximum investment limits. [½]

[Maximum 3]

(ii) ***Equation of the efficient frontier***

In order to find the efficient frontier, we can set up the Lagrangian function. When the correlation coefficient between the two security returns is equal to one, this is given by (using the usual definitions):

$$W = x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\sigma_1\sigma_2 - \lambda(x_1E_1 + x_2E_2 - E_p) - \mu(x_1 + x_2 - 1) \qquad [1]$$

Thus, the first-order conditions are:

$$\frac{\partial W}{\partial x_1} = 2x_1\sigma_1^2 + 2x_2\sigma_1\sigma_2 - \lambda E_1 - \mu = 0$$

$$\frac{\partial W}{\partial x_2} = 2x_2\sigma_2^2 + 2x_1\sigma_1\sigma_2 - \lambda E_2 - \mu = 0$$

$$\frac{\partial W}{\partial \lambda} = x_1E_1 + x_2E_2 - E_p = 0$$

$$\frac{\partial W}{\partial \mu} = x_1 + x_2 - 1 = 0$$ [2]

Combining the last two equations gives the optimal proportions of the two assets as:

$$x_1 = \frac{E_p - E_2}{E_1 - E_2} \quad \text{and} \quad x_2 = 1 - x_1 = \frac{E_1 - E_p}{E_1 - E_2}$$ [1]

Substituting these back into the expression for the variance of portfolio returns (*ie* the first three terms in the Lagrangian function) gives:

$$V_p = \left(\frac{1}{E_1 - E_2}\right)^2 \left[\left(E_p - E_2\right)^2 \sigma_1^2 + (E_1 - E_p)^2\sigma_2^2 + 2(E_p - E_2)(E_1 - E_p)\sigma_1\sigma_2\right]$$

$$= \left(\frac{1}{E_1 - E_2}\right)^2 \left[\left\{\left(E_p - E_2\right)\sigma_1 + (E_1 - E_p)\sigma_2\right\}^2\right]$$

Hence, the standard deviation of portfolio returns equals:

$$\sigma_p = \left(\frac{1}{E_1 - E_2}\right) \left\{\left(E_p - E_2\right)\sigma_1 + (E_1 - E_p)\sigma_2\right\}$$

Thus:

$$\sigma_p = aE_p + b$$

where:

$$a = \frac{\sigma_1 - \sigma_2}{E_1 - E_2} \quad \text{and} \quad b = \frac{\sigma_2 E_1 - \sigma_1 E_2}{E_1 - E_2}$$

This is therefore the equation of the efficient frontier in expected return-standard deviation space. [2]

[Total 6]

In fact, because we only have two risky assets, and they are perfectly positively correlated, we can use the following quicker method, which avoids Lagrangians:

- *first write the equation* $x_1E_1 + x_2E_2 = E_P$ *in the form* $x_1E_1 + (1-x_1)E_2 = E_P$
- *solve for* x_1 *in terms of the* E *'s and then find* $x_2 = 1 - x_1$
- *write down the variance, which is:*

$$V_P = x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\sigma_1\sigma_2 = (x_1\sigma_1 + x_2\sigma_2)^2$$

- *square root to get:*

$$\sigma_P = x_1\sigma_1 + x_2\sigma_2$$

- *substitute the expressions for* x_1 *and* x_2.

We have just shown that if two assets are perfectly positively correlated, then the efficient frontier is a straight line. This will also be the case if two assets are perfectly negatively correlated.

(iii)(a) ***Gradient***

The efficient frontier is normally plotted with expected return on the vertical axis, thus its gradient will be equal to $1/a$, *ie*:

gradient of the efficient frontier $= \dfrac{1}{a} = \dfrac{E_1 - E_2}{\sigma_1 - \sigma_2}$. [2]

(iii)(b) ***Location of the efficient frontier***

When $E_P = E_1$, we have:

$$\begin{aligned}\sigma_p &= aE_1 + b \\ &= \left[\frac{\sigma_1 - \sigma_2}{E_1 - E_2}\right]E_1 + \left[\frac{\sigma_2 E_1 - \sigma_1 E_2}{E_1 - E_2}\right] \\ &= \sigma_1 \frac{(E_1 - E_2)}{E_1 - E_2} + \sigma_2 \frac{(E_1 - E_1)}{E_1 - E_2} \\ &= \sigma_1\end{aligned}$$

Hence (σ_1, E_1) lies on the efficient frontier, which must therefore pass through Point 1. A similar argument shows that it also passes through Point 2. [2]

We know that it is a straight line, since the slope is $1/a$, which is constant. [1]

[Total 5]

Solution 1.22

(i) ***Opportunity set***

The variance of a portfolio consisting of Securities 1 and 2 is given by:

$$V = 10^2 x_1^2 + 20^2 x_2^2 + 2(10)(20)\rho x_1 x_2 \quad [1]$$

The portfolio expected return is given by:

$$E = 5x_1 + 10x_2 \quad [½]$$

Since the portfolio is fully invested, we require that:

$$1 = x_1 + x_2 \quad \Rightarrow x_2 = 1 - x_1 \quad [½]$$

Substituting the last equation into the previous one and rearranging gives:

$$x_1 = \frac{10-E}{5} \quad \Rightarrow x_2 = \frac{E-5}{5} \quad [1½]$$

Substituting these back into the expression for the variance and simplifying gives the equation of the opportunity set as:

$$V = (20-16\rho)E^2 - 240(1-\rho)E + 800(1-\rho) \quad [1½]$$

[Total 5]

(ii) ***Expressions for E and x_1***

A first-order condition for the point of global minimum variance (as a function of ρ) can be found by differentiating the expression for the variance found in (i) and setting it equal to zero. Thus:

$$\frac{\partial V}{\partial E} = 2(20-16\rho)E - 240(1-\rho) = 0 \quad [1]$$

Hence, we find that at this point, after simplifying:

$$E = \frac{30(1-\rho)}{5-4\rho} \qquad [1]$$

Substituting this into the expression for x_1 found in (i), and simplifying, then gives:

$$x_1 = \frac{4-2\rho}{5-4\rho} \qquad [1]$$

To see how E and x_1 vary with ρ, we can look at the derivatives, which can be found using the quotient rule. After simplifying, we get:

$$\frac{\partial E}{\partial \rho} = \frac{-30}{(5-4\rho)^2} \qquad [½]$$

$$\frac{\partial x_1}{\partial \rho} = \frac{6}{(5-4\rho)^2} \qquad [½]$$

Since $\frac{\partial E}{\partial \rho} < 0$, the portfolio expected return at the point of minimum global variance *decreases* as the correlation coefficient increases. In fact, E ranges from $6\frac{2}{3}\%$ when $\rho = -1$ through 6% when $\rho = 0$ to 0% when $\rho = 1$. [½]

Since $\frac{\partial x_1}{\partial \rho} > 0$, the portfolio proportion invested in Security 1 at the point of minimum global variance *increases* with the correlation coefficient. In fact, x_1 ranges from $66\frac{2}{3}\%$ when $\rho = -1$ through 80% when $\rho = 0$ to 200% when $\rho = 1$. [½]

[Total 5]

The portfolio proportion invested in Security 2 at the point of minimum global variance must correspondingly decrease with the correlation coefficient.

Solution 1.23

According to the single-index model, the return on security *i* is given by:

$$R_i = \alpha_i + \beta_i R_M + e_i \qquad [½]$$

where α_i and β_i are constants

R_M is the actual return on the market

e_i is a random variable representing the component of R_i not related to the market. [½]

By the linear additivity of expected values, we have:

$$E(R_i) = E(\alpha_i) + E(\beta_i R_M) + E(e_i) \qquad [½]$$

Since α_i and β_i are constants and α_i is chosen so that $E(e_i) = 0$, we have:

$$E(R_i) = \alpha_i + \beta_i E_M$$

as required. [½]

The variance of returns for Security i is:

$$V_i = \text{var}\left[\alpha_i + \beta_i R_M + e_i\right] \qquad [½]$$

As α_i and β_i are constant, this is equal to:

$$V_i = \text{var}\left[\beta_i R_M + e_i\right] \qquad [½]$$

Now, recall that the single-index model assumes that:

$$\text{cov}(e_i, R_M) = 0 \qquad [½]$$

Hence:

$$\Rightarrow \qquad V_i = \text{var}\left[\beta_i R_M\right] + \text{var}\left[e_i\right] \qquad [½]$$

ie $\quad V_i = \beta_i^2 . V_M + V_{ei}$

as required. [1]

The covariance between Securities i and j is given by:

$$C_{i,j} = \text{cov}\left[R_i, R_j\right]$$

$$= \text{cov}\left[\alpha_i + \beta_i R_M + e_i, \alpha_j + \beta_j R_M + e_j\right] \quad [½]$$

Again, since α_i and β_i are constant, this is equal to:

$$C_{i,j} = \text{cov}\left[\beta_i R_M + e_i, \beta_j R_M + e_j\right] \quad [½]$$

As before, recall that the single-index model assumes that:

$$\text{cov}(e_i, R_M) = 0 \quad [½]$$

Hence:

$$C_{i,j} = \text{cov}\left[\beta_i R_M, \beta_j R_M\right] + \text{cov}\left[e_i, e_j\right]$$

$$= \beta_i \beta_j \,\text{cov}\left[R_M, R_M\right] + \text{cov}\left[e_i, e_j\right] \quad [½]$$

We also have the assumption that $\text{cov}(e_i, e_j) = 0$ when $i \neq j$.

So: $$C_{i,j} = \beta_i \beta_j \,\text{cov}\left[R_M, R_M\right] = \beta_i \beta_j V_M \quad [1]$$

[Total 8]

Solution 1.24

(i)(a) ***Calculate the expected return and standard deviation***

The expected return for each security is calculated using the equation:

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

Hence:

$$E_1 = 0 + 1.1 \times 10 = 11.0$$

$$E_2 = 2 + 0.6 \times 10 = 8.0$$

$$E_3 = -2.2 + 2.0 \times 10 = 17.8 \quad [1]$$

The variance for each security is calculated using the equation:

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

Hence, the standard deviation for each security is given by:

$$\sigma_1 = \left(\beta_1^2 V_M + V_{\varepsilon 1}\right)^{½} = \left(1.1^2 \times 25 + 2.2\right)^{½} = (32.45)^{½} = 5.70 \qquad [½]$$

$$\sigma_2 = \left(\beta_2^2 V_M + V_{\varepsilon 2}\right)^{½} = \left(0.6^2 \times 25 + 1.3\right)^{½} = (10.3)^{½} = 3.21 \qquad [½]$$

$$\sigma_3 = \left(\beta_3^2 V_M + V_{\varepsilon 3}\right)^{½} = \left(2.0^2 \times 25 + 1.2\right)^{½} = (101.2)^{½} = 10.06 \qquad [½]$$

(i)(b) ***Calculate the covariances***

The covariance of returns between securities i and j is calculated using the equation:

$$C_{i,j} = \beta_i \beta_j V_M$$

Hence:

$$C_{1,2} = \beta_1 \beta_2 V_M = 1.1 \times 0.6 \times 25 = 16.5 \qquad [½]$$

$$C_{1,3} = \beta_1 \beta_3 V_M = 1.1 \times 2.0 \times 25 = 55.0 \qquad [½]$$

$$C_{2,3} = \beta_2 \beta_3 V_M = 0.6 \times 2.0 \times 25 = 30.0 \qquad [½]$$

[Total 4]

(ii)(a) ***Variance of portfolio***

If R_P is the return on the portfolio and R_i is the return on security i, then:

$$R_P = \tfrac{1}{3}\left(R_1 + R_2 + R_3\right)$$

So, the variance is:

$$\begin{aligned}\text{var}(R_P) &= \tfrac{1}{9}\,\text{var}\left(R_1 + R_2 + R_3\right)\\ &= \tfrac{1}{9}\left\{\text{var}(R_1) + \text{var}(R_2) + \text{var}(R_3) + 2C_{1,2} + 2C_{1,3} + 2C_{2,3}\right\}\\ &= \tfrac{1}{9}\left\{32.45 + 10.3 + 101.2 + 2\times 16.5 + 2\times 55 + 2\times 30\right\}\\ &= 38.55 \qquad [1]\end{aligned}$$

(ii)(b) ***Systematic risk***

The beta of the portfolio, β_P, is the weighted average of the betas of the individual securities:

$$\beta_P = \tfrac{1}{3}(\beta_1 + \beta_2 + \beta_3) = \tfrac{1}{3}(1.1 + 0.6 + 2) = \tfrac{3.7}{3}$$ [½]

The systematic risk is:

$$\beta_P^2 V_M = \left(\tfrac{3.7}{3}\right)^2 \times 25 = 38.028$$ [½]

(ii)(c) ***Specific risk***

The specific risk, $V_{\varepsilon P}$, can be calculated as the total portfolio variance (or risk) minus the systematic risk:

$$38.55 - 38.028 = 0.522$$ [1]

Alternatively, the specific risk of the portfolio can be calculated directly using the specific risks of the individual securities:

$$V_{\varepsilon P} = \text{var}\left(\tfrac{1}{3}\varepsilon_1 + \tfrac{1}{3}\varepsilon_2 + \tfrac{1}{3}\varepsilon_2\right) = \left(\tfrac{1}{3}\right)^2 V_{\varepsilon 1} + \left(\tfrac{1}{3}\right)^2 V_{\varepsilon 2} + \left(\tfrac{1}{3}\right)^2 V_{\varepsilon 3}$$

$$= \left(\tfrac{1}{3}\right)^2 \times 2.2 + \left(\tfrac{1}{3}\right)^2 \times 1.3 + \left(\tfrac{1}{3}\right)^2 \times 1.2 = 0.522$$

[Total 3]

Solution 1.25

Macroeconomic factor models use observable economic time series as the factors. [½]

They therefore include factors such as:

- annual rates of inflation
- economic growth
- short-term interest rates
- the yields on long-term government bonds
- the yield margin on corporate bonds over government bonds. [1]

These are the macroeconomic variables that are assumed to influence security prices and returns in practice. [½]

A related class of model uses a market index plus a set of industry indices as the factors. [½]

Fundamental factor models are closely related to macroeconomic models but instead of (or in addition to) macroeconomic variables, they use company-specific variables. [½]

These may include such fundamental factors as:

- the level of gearing
- the price earnings ratio
- the level of R&D spending
- the industry group to which the company belongs. [1]

The commercially available fundamental factor models typically use many tens of factors. [½]

Statistical factor models do not rely on specifying the factors independently of the historical returns data. [½]

Instead a technique called principal components analysis can be used to determine a set of orthogonal indices that explain as much as possible of the observed variance. [½]

However, the resulting indices are unlikely to have any meaningful economic interpretation and may vary considerably between different data sets. [½]
[Total 6]

Part 2 – Questions

Question 2.1 (Bookwork)

(i) Explain what is meant by specific risk and systematic risk in the CAPM. [3]

(ii) Explain the meaning of the "beta" of a share, and describe how you would calculate it for:

(a) a company

(b) a portfolio. [7]

(iii) Explain how the beta for a portfolio can be used to determine the expected return for the portfolio. [4]

(iv) Why might the beta calculated in (ii)(a) be inappropriate for practical use? [3]
[Total 17]

Question 2.2 (Exam-style)

(i) State the assumptions of the capital asset pricing model (CAPM). [5]

(ii) An investment market consisting of a risk-free asset and a very large number of stocks is such that, for modelling purposes, the market capitalisation of the k th stock can be expressed as:

$$\frac{1}{2^k} \text{ where } k = 1, 2, 3, \ldots$$

The expected return on the k th stock (expressed as a percentage) is:

$$25e^{-(k-1)} + 5\left(1 - e^{-(k-1)}\right)$$

Assuming that the CAPM assumptions hold, find the expected return on the portfolio of risky assets held by each investor. [5]
[Total 10]

Question 2.3 (Exam-style)

(i) Within the context of the capital asset pricing model, explain what is meant by the "market price of risk". [3]

(ii) Show how the security market line relationship can be rearranged to give an expression for the expected return in terms of the market price of risk γ_M, and briefly interpret your answer. [3]

(iii) Show that the capital asset pricing model result can be written as a single-index model and hence that it is consistent with the arbitrage pricing theory. [4]

[Total 10]

Question 2.4 (Exam-style)

It is believed that three entirely uncorrelated factors give a satisfactory explanation of investment returns. The risk premiums on the three indices are 3% *pa*, 5% *pa* and 9% *pa* respectively. The sensitivities of Security A to each of these factors are 0, 0, and 2 respectively. The sensitivity of Security B to each factor is ¾. The risk-free rate of interest is 6% *pa*.

(i) Calculate the expected returns on each security assuming that the arbitrage pricing theory holds. [2]

(ii) Calculate the characteristics of a portfolio consisting of 75% Security A and 25% Security B. [4]

(iii) Explain what will happen if the risk-free asset does not exist. [2]

[Total 8]

Question 2.5 (Exam-style)

(i) Set down the equations for the expected returns based on the capital asset pricing model (CAPM) and the arbitrage pricing theory (APT). Define all symbols used. [3]

(ii) Briefly explain the major differences between these models. [5]

[Total 8]

Question 2.6 (Developmental)

"Brownian motion is the only process with <u>*stationary*</u> <u>*independent increments*</u> *and* <u>*continuous sample paths*</u>*."*

(i) Give mathematical definitions of each of the three underlined terms. [3]

(ii) What is the distribution of the increments for a standard Brownian motion? [1]
[Total 4]

Question 2.7 (Developmental)

(i) What is meant by saying that the process $\{Y_t\}$ is a martingale with respect to another process $\{X_t\}$? [2]

Let B_t $(t \geq 0)$ be a standard Brownian motion.

(ii) Show that B_t and $B_t^2 + kt$ are both martingales with respect to B_t, for a suitably chosen value of the constant k, which you should specify. [6]

(iii) Show that there is a value of the constant c, which you should specify, such that $(a + bB_t)^2 + ct$ is a martingale with respect to B_t, where a and b are constants.
[4]
[Total 12]

Question 2.8 (Developmental)

Let B_t $(t \geq 0)$ be a standard Brownian motion process starting with $B_0 = 0$.

(i) What is the probability that B_2 takes a positive value? [1]

(ii) What is the probability that B_2 takes a value in the interval $(-1,1)$? [2]

(iii) Show that the probability that B_1 and B_2 both take positive values is $\frac{3}{8}$. [4]

(iv) What is the probability that B_t takes a negative value at some time between $t = 0$ and $t = 2$? [1]
[Total 8]

Question 2.9 (Developmental)

A student has said: "If you want to find the variance of $X = B(s) + B(t)$, where $s < t$, for a standard Brownian motion process, you can use the fact that $B(s)$ and $B(t)$ are independent to get $\text{var}(X) = s + t$."

(i) Explain why the student's argument is not correct, and find a correct expression for $\text{var}(X)$. [4]

(ii) Hence show that the general formula for $\text{var}[B(t_1) + B(t_2)]$ when $t_1, t_2 > 0$ can be expressed as $t_1 + t_2 + 2\min(t_1, t_2)$. [2]

[Total 6]

Question 2.10 (Exam-style)

Let B_t $(t \geq 0)$ be a standard Brownian motion process starting with $B_0 = 0$.

(i) Show that, when $s < t$, $E(B_s B_t) = s$. [3]

(ii) Hence find a general formula for the correlation coefficient $\rho(B_{t_1}, B_{t_2})$. [2]

[Total 5]

Question 2.11 (Exam-style)

(i) Write down a formula for $E\left(e^{aX}\right)$ where $X \sim N\left(\mu, \sigma^2\right)$ and, by differentiating, or otherwise, derive an expression for $E\left(Xe^{aX}\right)$. [2]

(ii) Show that:

$$X_t = \left(B_t - at\right)e^{aB_t - \frac{1}{2}a^2 t}$$

is a martingale, where B_t is a standard Brownian motion, and a is an arbitrary constant. You may assume that $E\left[\left|X_t\right|\right] < \infty$. [5]

[Total 7]

Question 2.12 (Developmental)

(i) Write down Ito's Lemma as it applies to a function $f(X_t)$ of a stochastic process X_t that satisfies the stochastic differential equation $dX_t = \sigma_t dB_t + \mu_t dt$, where B_t is a standard Brownian motion. [2]

(ii) Hence find the stochastic differential equations for each of the following processes:

(a) $G_t = \exp(X_t)$

(b) $Q_t = X_t^2$

(c) $V_t = (1+X_t)^{-1}$

(d) $L_t = 100 + 10X_t$

(e) $J_t = \ln B_t$

(f) $K_t = 5B_t^3 + 2B_t$ [12]

[Total 14]

Question 2.13 (Developmental)

Let B_t ($t \geq 0$) be a standard Brownian motion with $B_0 = 0$.

(i) By first writing down an expression for $d(B_s^2)$, show that

$$\int_0^t B_s dB_s = \tfrac{1}{2}(B_t^2 - t)$$ [4]

(ii) What is the expected value at time 0 of $\int_0^t B_s dB_s$? [2]

(iii) What is the expected value at time u ($0 < u < t$) of $\int_0^t B_s dB_s$? [3]

(iv) What can you say about the process $I_t = \int_0^t B_s dB_s$, based on your results from (i) and (iii)? [1]

[Total 10]

Question 2.14 (Developmental)

Find the mean and variance of the stochastic integral $I = \int_0^1 t\,dB_t$. [3]

Question 2.15 (Developmental)

Let $\{X_t\}$ be a continuous time stochastic process defined by the equation $X_t = \alpha W_t^2 + \beta$, where $\{W_t\}$ is a standard Brownian motion and α and β are constants.

By applying Ito's Lemma, or otherwise, write down the stochastic differential equation satisfied by X_t. [3]

Question 2.16 (Developmental)

(i) Let X_t be a diffusion that satisfies $dX_t = \mu(X_t,t)\,dt + \sigma(X_t,t)\,dB_t$ where B_t is a standard Brownian motion.

Let $f(X_t,t)$ be a function of t and X_t. By considering Taylor's theorem, suggest a partial differential equation that must be satisfied by $f(X_t,t)$ in order that it is a martingale. [3]

(ii) Verify that your equation holds when $f(X_t,t) = B_t^2 - t$. [1]

(iii) Find $g(t)$ such that $B_t^3 + g(t)B_t$ is a martingale. [2]

[Total 6]

Question 2.17 (Exam-style)

In the following, B_t denotes a standard Brownian motion.

(i) Write down the general solution of the stochastic differential equation:

$$dX_t = -\gamma X_t dt + \sigma dB_t$$ [1]

(ii) Hence determine the solution of the stochastic differential equation:

$$dR_t = 0.8(4 - R_t)dt + dB_t$$

where $R_0 = 5$. [1]

(iii) Find the distribution of the process R_t at time t and in the long-term. [3]

[Total 5]

Question 2.18 (Exam-style)

The Ito process, X_t, is defined by the stochastic differential equation:

$$dX_t = 0.5X_t(1 - X_t)(1 - 2X_t)dt - X_t(1 - X_t)dB_t$$

where B_t is a standard Brownian motion, and $X_0 = 0.5$.

By considering the stochastic differential equation for the process $Y_t = \ln\left(\frac{1}{X_t} - 1\right)$, find X_t in terms of B_t. [7]

Question 2.19 (Exam-style)

The market price of a certain share is being modelled as a geometric Brownian motion. The price S_t at time $t \geq 0$ satisfies the equation:

$$\log_e \frac{S_t}{S_0} = \mu t + \sigma B_t$$

where $\{B_t, t \geq 0\}$ is a standard Brownian motion and μ and σ are constants.

(i) Show that the stochastic differential dS_t can be written in the form:

$$\frac{dS_t}{S_t} = c_1 dB_t + c_2 dt \,,$$

where c_1 and c_2 are constants you should specify. [4]

(ii) Derive expressions for $E[S_t]$ and $\text{var}[S_t]$. [4]

(iii) Derive expressions for $\text{cov}[S_{t_1}, S_{t_2}]$ and $E[S_{t_2} \mid S_{t_1}]$ where $0 < t_1 < t_2$. [6]

(iv) By using your expression for $E[S_{t_2} \mid S_{t_1}]$, write down a function of S_t that is a martingale. [1]

[Total 15]

Question 2.20 (Exam-style)

Let S_t be a geometric Brownian motion process defined by the equation $S_t = \exp(\mu t + \sigma W_t)$, where W_t is a standard Brownian motion and μ and σ are constants.

(i) Write down the stochastic differential equation satisfied by $X_t = \log_e S_t$. [1]

(ii) By applying Ito's Lemma, or otherwise, derive the stochastic differential equation satisfied by S_t. [3]

(iii) The price of a share follows a geometric Brownian motion with $\mu = 0.06$ and $\sigma = 0.25$ (both expressed in annual units). Find the probability that, over a given one-year period, the share price will fall. [3]
[Total 7]

Question 2.21 (Bookwork)

When valuing derivatives it is often assumed that the price of the underlying security follows an Ito process of the form:

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

where Z_t represents a standard Brownian motion.

What are the advantages and disadvantages of this assumption? [4]

Question 2.22 (Bookwork)

(i) Explain what is meant by the Wilkie model of investment returns. [4]

(ii) Explain what is meant by the continuous-time lognormal model of security prices. [4]

(iii) State two key differences between the Wilkie model and the continuous-time lognormal model. [2]
[Total 10]

Question 2.23 (Exam-style)

The regulators in Country K have recently issued new guidelines as to the solvency levels required from life office companies. Under the new guidelines the ratio of assets to liabilities of these companies must now always be at a level no less than 150%. The previous ratio was 125%. The new guidelines will not take effect for another five years.

You work for a life office company in Country K, whose liabilities over the next ten years are reasonably predictable. As asset manager you have been asked to carry out a modelling exercise to assess the company's ability to meet the new guidelines.

Some of the results of the modelling exercise are shown in the table below:

Assets/Liabilities									
	Simulation								
Year	1	2	…..	65	66	67	….	99	100
1	141%	143%	…..	135%	137%	135%	…..	147%	140%
2	139%	142%		131%	134%	135%		143%	137%
3	150%	149%		138%	142%	138%		148%	146%
4	160%	162%	…..	140%	144%	140%	…..	154%	154%
5	162%	173%		142%	144%	142%		162%	163%
6	168%	167%		144%	146%	146%		168%	165%
7	173%	174%	…..	146%	148%	148%	…..	169%	168%

(i) State what is meant by:

(a) longitudinal properties

(b) cross-sectional properties [2]

(ii) Describe which longitudinal properties and cross-sectional properties in the table are of particular interest. [3]

[Total 5]

Part 2 – Solutions

Solution 2.1

(i) ***Specific and systematic risk***

The fluctuation (both up and down) of returns from a security can be broken into two components according to the extent to which:

- company/industry specific events cause the returns to vary independently of movements in the investment market as a whole (*ie specific risk*) [1]
- the returns from the security move with the market as a whole (*ie systematic risk*). [1]

Specific risk is the risk unique to a particular security that can be eliminated from a portfolio if the portfolio is suitably diversified. In terms of portfolio theory, it is the unrewarded risk. [½]

Systematic risk cannot be diversified away. [½]
[Total 3]

(ii) ***Beta for a share***

A share's beta is a measure of its systematic risk. It is a coefficient that measures the extent to which the return from the security covaries with the return from the market of risky assets as a whole. It also indicates how the risk premium on the share compares to that for the market of risky assets as a whole. [1]

If the returns on a particular share move more aggressively than the market, the share has a high beta (greater than 1). A defensive share that does not fluctuate as much as the market would have a coefficient below 1. [1]

(a) ***To calculate it for a company***

Calculate the return on the share over a suitably large set of periods (say each month for 5 years) and also for the whole market. [1]

Plot these values with market returns on the horizontal axis, and the corresponding share returns on the vertical axis, and find the gradient of the line of best fit (by least-squares regression). The beta is estimated as the gradient of the line. [1]

You do, however, need to be wary of company- or industry-specific events that may have caused the historical beta of the company to change over the period of estimation. The estimate of the prospective beta may need to be adjusted accordingly. [1]

(b) ***To calculate it for a portfolio***

The beta is the weighted average of the betas for the individual shares, weighted by the value of the holding for each of the shares. So to calculate it, you would calculate the weighted average using the betas as worked out above. Alternatively, repeat (a) using the returns on the portfolio. [2]
[Total 7]

(iii) ***Expected return for the portfolio***

Assuming the capital asset pricing model holds, the beta for the portfolio gives a guide as to how the portfolio's return is expected to differ from the market as a whole. The following data is needed:

- β_P the beta for the portfolio
- r the risk-free rate of return (for short-term periods, take this to be treasury bill returns, for longer periods perhaps look at government bond yields)
- E_M the expected return on the market as a whole. [2]

The expected return for the portfolio is given by the security market line equation as:

$$E_P = r + (E_M - r)\beta_P$$ [1]

Thus, for a portfolio with a beta of 1, the expected return on the portfolio is equal to the expected market return. [½]

If the market is not efficient, the expected return may be higher or lower than this. [½]
[Total 4]

(iv) ***Why might the beta be inappropriate?***

1. An analysis over a limited time period may produce an estimate for the beta with some random bias. Empirical evidence suggests that betas for individual companies are not stable. [1]

2. A company's beta may change over time as the company may have a shift in emphasis and management.

   - So the beta based on an historical analysis may not be appropriate for the company as it currently is.
   - Similarly, the current beta may not be appropriate for the future. [1]

3. The assumptions underlying portfolio theory and CAPM may not hold exactly. So that even if we have an accurate estimate of beta we cannot use the equation in (iii) to estimate expected return. [1]

[Total 3]

Solution 2.2

(i) ***CAPM assumptions***

- investors make their decisions purely on the basis of expected return and variance. So all expected returns, variances and covariances of assets must be known. [1]
- investors are non-satiated [½]
- investors are risk-averse [½]
- there are no taxes or transaction costs [½]
- assets may be held in any amounts [½]
- all investors have the *same* fixed one-step time horizon [½]
- all investors make the *same* assumptions about the expected returns, variances and covariances of assets [½]
- all investors measure returns consistently (*eg* in the same currency or in the same real/nominal terms) [½]
- the market is perfect and in equilibrium [½]
- all investors may lend or borrow any amounts of a risk-free asset at the same risk-free rate r. [½]

[Maximum 5]

(ii) ***Market portfolio***

If CAPM holds then the portfolio of risky assets held by each investor will be the market portfolio. [½]

With a very large number of stocks, the expected return on the market portfolio is:

$$E_M = \sum_{k=1}^{\infty} x_k E_k$$

where x_k is the holding in stock k and E_k is the expected return on that stock. [½]

In this case we have:

$$E_M = \sum_{k=1}^{\infty} x_k E_k = \sum_{k=1}^{\infty} \frac{1}{2^k}\left[25e^{-(k-1)} + 5\left(1 - e^{-(k-1)}\right)\right] \quad [1]$$

Simplifying, we get:

$$E_M = \sum_{k=1}^{\infty} \frac{1}{2^k}\left[5 + 20e^{-(k-1)}\right] = 5\sum_{k=1}^{\infty} \frac{1}{2^k} + 20\sum_{k=1}^{\infty} \frac{1}{2^k} e^{-(k-1)}$$

$$= \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots\right) \times 5 + 20e\sum_{k=1}^{\infty}\left(\frac{1}{2e}\right)^k$$

$$= \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots\right) \times 5 + 20e\left[\left(\frac{1}{2e}\right) + \left(\frac{1}{2e}\right)^2 + \left(\frac{1}{2e}\right)^3 + \ldots\right] \quad [1½]$$

Using the formula for the infinite sum of a geometric progression, we get:

$$E_M = 5 \times \frac{\frac{1}{2}}{1 - \frac{1}{2}} + \frac{20e\left(\frac{1}{2e}\right)}{1 - \left(\frac{1}{2e}\right)} = 5 + \frac{20e}{2e - 1} = 17.25$$

So the expected return on the market is 17.25%. [1½]
[Total 5]

Solution 2.3

(i) ***Market price of risk***

Within the context of the capital asset pricing model, the *market price of risk* is defined as:

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

where:

- E_M = the expected return on market portfolio
- r = the risk-free rate of return
- σ_M = the standard deviation of market portfolio returns. [1]

It is the additional expected return that the market requires in order to accept an additional unit of risk, as measured by the portfolio standard deviation of return. [1]

It is equal to the gradient of the capital market line in $E - \sigma$ space. [1]

[Total 3]

(ii) ***Security market line***

Recall the security market line relationship $E_i - r = \beta_i(E_M - r)$ and also that the beta of a security is defined as:

$$\beta_i = \frac{\text{cov}(R_i, R_M)}{\sigma_M^2}$$

If we define $\sigma_{iM} = Cov(R_i, R_M)$, then the security market line can written as:

$$E_i = r + \frac{\sigma_{iM}}{\sigma_M^2}\left(E_M - r\right) = r + \frac{\sigma_{iM}}{\sigma_M}\left(\frac{E_M - r}{\sigma_M}\right) = r + \frac{\sigma_{iM}}{\sigma_M}\gamma_M \quad [2]$$

As $\dfrac{\sigma_{iM}}{\sigma_M} = \dfrac{\rho_{iM}\sigma_i\sigma_M}{\sigma_M} = \rho_{iM}\sigma_i$ is a measure of the risk of portfolio i, the security market line states that the expected return on any portfolio can be expressed as the sum of the risk-free rate plus the amount of risk multiplied by the market price of risk, *ie*:

expected return = risk-free rate + (market price of risk) × (amount of risk) [1]

[Total 3]

(iii) ***CAPM as a single-index model***

Consider a single-index model in which the single index is the return on the market portfolio. Then:

$$R_i = a_i + b_{i,1}R_M + c_i$$

Thus:

$$\text{cov}(R_i, R_M) = \text{cov}(a_i + b_{i,1}R_M + c_i, R_M) = b_{i,1}\,\sigma_M^2 \qquad [½]$$

So: $$\beta_i = \frac{\text{cov}(R_i, R_M)}{\sigma_M^2} = b_{i,1} \qquad [½]$$

Applying the same argument that is used to derive the arbitrage pricing theory result we can show that:

$$E_i = \lambda_0 + \lambda_1 b_{i,1} \qquad [½]$$

Now, for the risk-free asset, $\beta_i = 0$ and so:

$$E_i = \lambda_0 + \lambda_1 b_{i,1} = \lambda_0 + \lambda_1 \times 0 = \lambda_0 = r \qquad [½]$$

Thus, $\lambda_0 = r$. Additionally, for the market portfolio:

$$E_M = \lambda_0 + \lambda_1 b_{M,1} = \lambda_0 + \lambda_1 \times 1$$

So: $$\lambda_1 = E_M - \lambda_0 = E_M - r \qquad [½]$$

Hence:

$$E_i = \lambda_0 + \lambda_1 b_{i,1} = r + \beta_i\left(E_M - r\right) \qquad [½]$$

which is the security market line relationship in the capital asset pricing model. Thus, the capital asset pricing model and the arbitrage pricing theory are consistent if:

- returns are generated by a single-index model in which the market return is the index [½]
- there exists a risk-free rate of return. [½]

[Total 4]

Solution 2.4

(i) ***Expected returns under APT***

The expected return will be:

$$\text{risk-free return} + \sum (\text{factor sensitivities} \times \text{risk premiums})$$

So we have:

Security A: $0.06 + 2 \times 0.09 = 24\,\%$ [1]

Security B: $0.06 + \tfrac{3}{4} \times (0.03 + 0.05 + 0.09) = 18\tfrac{3}{4}\,\%$ [1]

(ii) ***Characteristics of the portfolio with 75% A and 25% B***

The expected return and sensitivities of the portfolio will be a linear combination of the expected return and sensitivities of the underlying securities:

Expected return: $0.75 \times 24\% + 0.25 \times 18.75\% = 22.69\%$ [1]

Sensitivity to Index 1: $0.25 \times 0.75 = 0.1875$ [1]

Sensitivity to Index 2: $0.25 \times 0.75 = 0.1875$ [1]

Sensitivity to Index 3: $0.75 \times 2 + 0.25 \times 0.75 = 1.6875$ [1]

[Total 4]

You can check the expected return result as follows, using the method from part (i):

$$0.06 + 0.1875 \times 0.03 + 0.1875 \times 0.05 + 1.6875 \times 0.09 = 22.69\%$$

(iii) ***No risk-free asset***

It will make little difference. APT assumes that there are a large number of securities and that short sales are possible. Consequently it should always be possible to construct a diversified portfolio that has a zero correlation with all three of the factors. This "zero-beta" portfolio would perform the role of a risk-free asset. (The 6% figure used above would be replaced by the expected return on this zero-beta portfolio.) [2]

Solution 2.5

This is Question 1 from the Subject 109 exam of April 2000.

(i) ***CAPM and APT equations***

The security market line from CAPM is given by:

$$E\left[R_i\right]=r_f+\beta_i\left(E\left[R_m\right]-r_f\right)$$

where r_f is the risk-free rate of return, R_i is the return on asset *i*, R_m is the return on the market, and $\beta_i=\dfrac{\operatorname{cov}\left(R_i,R_m\right)}{\operatorname{var}\left(R_m\right)}$. [1½]

The APT model tells us that:

$$E\left[R_i\right]=\lambda_0+\lambda_1 b_{i1}+\lambda_2 b_{i2}+\cdots+\lambda_n b_{in}$$

where $\lambda_0,\lambda_1,\ldots,\lambda_n$ are fixed constants (*ie* independent of the asset *i*). The parameters $b_{i1},b_{i2},\ldots,b_{in}$ measure the exposure of the return on asset *i* to the factor j. The constant λ_0 is the risk-free rate of return. For $j=1,2,\ldots,n$, the constant λ_j is the additional risk premium on asset *i* per unit of exposure to factor *j*. [1½]

[Total 3]

(ii) ***Differences between models***

APT assumes an underlying multifactor model, with more than one source of systematic risk, whereas CAPM assumes that the market is the only systematic factor. [1]

However, APT does not indicate how many factors are needed, nor what these underlying factors should be. [1]

CAPM makes several strong assumptions about investor preferences, (*eg* same estimates of means, variances and correlations over the same one-period time horizon). APT doesn't have to, and assumes only that all investors believe that the same multifactor model generates investment returns. [2]

CAPM requires identification of the market portfolio, whereas APT requires identification of the factors and estimates of both the b_{ij}'s and the λ's. [1]

[Total 5]

Solution 2.6

(i) ***Mathematical definitions***

"Stationary increments" means that the distribution of $B_t - B_s$ ($t > s$) depends only on $t - s$. [1]

"Independent increments" means that $B_t - B_s$ is independent of B_r whenever $r \le s < t$. [1]

"Continuous sample paths" means that the function $t \to B_t(\omega)$ for each particular realisation ω is a continuous function of t. [1]
[Total 3]

(ii) ***Distribution of the increments***

For a standard Brownian motion $B_t - B_s$ ($t > s$) has a $N(0, t-s)$ distribution. [1]

Solution 2.7

(i) ***What is a martingale?***

Strictly, we should say that the process $\{Y_t\}$ is a martingale with respect to the *filtration* $\{F_t\}$ of the process $\{X_t\}$, which means that:

$$E[Y_t \mid F_s] = Y_s \text{ for all } s < t$$

and $E[|Y_t|] < \infty$ [2]

(ii) ***Show that these processes are martingales***

If we use an s subscript to denote the expected value with respect to the filtration at time s, then we can write:

$$E_s[B_t] = E_s[(B_t - B_s) + B_s] = E_s[B_t - B_s] + E_s[B_s]$$

Since $B_t - B_s \sim N(0, t-s)$ and the value of B_s is known at time s, this gives:

$$E_s[B_t] = 0 + B_s = B_s \qquad [1]$$

We have shown that the future value of B_t is equal to its current value (at time s). We also need to show that $E\left[|B_t|\right] < \infty$.

One way to do this is to note that $|x| < 1 + x^2$ for all values of x.

So: $$E\left[|B_t|\right] < E\left[1 + B_t^2\right] = 1 + \text{var}(B_t) + \left[E(B_t)\right]^2 = 1 + t + 0^2 < \infty \qquad [1]$$

Therefore, B_t is a martingale (with respect to B_t).

Similarly:

$$\begin{aligned}
E_s[B_t^2] &= E_s[\{(B_t - B_s) + B_s\}^2] \\
&= E_s[(B_t - B_s)^2] + 2E_s[(B_t - B_s)B_s] + E_s[B_s^2] \\
&= \left\{\text{var}_s[B_t - B_s] + [E_s(B_t - B_s)]^2\right\} + 2B_s \times E_s[B_t - B_s] + B_s^2 \\
&= (t - s) + 0^2 + 0 + B_s^2 \\
&= t - s + B_s^2 \qquad [2]
\end{aligned}$$

So $$E_s[B_t^2 - t] = B_s^2 - s$$

We can show that $E\left[\left|B_t^2 - t\right|\right] < \infty$ for any value of *t*, by first noting that:

$$\left|x^2 - k\right| < x^2 + k \ \ (k > 0) \text{ for all values of } x.$$

So: $$E\left[\left|B_t^2 - t\right|\right] < E\left[B_t^2 + t\right] = \text{var}(B_t) + \left[E(B_t)\right]^2 + t = t + 0^2 + t = 2t < \infty \qquad [1]$$

Since the future value of $B_t^2 - t$ is equal to its current value (at time s) and the expected value of its modulus is finite, $B_t^2 - t$ is a martingale with respect to B_t, and the required constant is $k = -1$. [1]

[Total 6]

(iii) ***Value of c to make the process a martingale***

We know that B_t and $B_t^2 - t$ are both martingales with respect to B_t.

So, if we use an s subscript to denote the expected value with respect to the filtration at time s, then:

$$E_s[B_t] = B_s$$

and $$E_s[B_t{}^2 - t] = B_s{}^2 - s \;\Rightarrow\; E_s[B_t{}^2] = B_s{}^2 + t - s$$ [1]

Using these results, we then find that:

$$\begin{aligned} E_s[(a+bB_t)^2] &= E_s[a^2 + 2abB_t + b^2 B_t^2] \\ &= a^2 + 2abE_s[B_t] + b^2 E_s[B_t^2] \\ &= a^2 + 2abB_s + b^2(B_s^2 + t - s) \\ &= (a+bB_s)^2 + b^2(t-s) \end{aligned}$$

So: $$E_s[(a+bB_t)^2 - b^2 t] = (a+bB_s)^2 - b^2 s$$ [1½]

ie $(a+bB_t)^2 - b^2 t$ is a martingale with respect to B_t. So the required value of the constant is $c = -b^2$. [½]

The technical condition $E\left[\left|(a+bB_t)^2 + ct\right|\right] < \infty$ will hold as in (ii). [1]

[Total 4]

Solution 2.8

(i) ***Probability that B_2 takes a positive value***

$B_2 = B_2 - B_0 \sim N(0,2)$. Therefore, B_2 is equally likely to be positive or negative (and has zero probability of being exactly zero).

So: $P(B_2 > 0) = \frac{1}{2}$ [1]

(ii) ***Probability that B_2 takes a value in the interval $(-1,1)$***

$B_2 = B_2 - B_0 \sim N(0,2)$.

Standardising:

$$P(-1 < B_2 < 1) = \Phi\left(\frac{1}{\sqrt{2}}\right) - \Phi\left(-\frac{1}{\sqrt{2}}\right) = 0.760 - (1 - 0.760) = 0.520$$ [2]

(iii) ***Probability that B_1 and B_2 both take positive values***

We can write the required probability as:

$$p = P(B_1 > 0, B_2 > 0) = P(B_1 - B_0 > 0, B_2 - B_1 > -B_1)$$

If we now write $X = B_1 - B_0$ and $Y = B_2 - B_1$, then we know from the properties of Brownian motion that X and Y are independent, each with a $N(0,1)$ distribution.

So the required probability is:

$$p = P(X > 0, Y > -X)$$

Since the range of values of Y depends on the value of X, we must use a double integral to evaluate this:

$$p = \int_{x=0}^{\infty} \int_{y=-x}^{\infty} \phi(x)\phi(y)\, dy\, dx$$

where the joint density function is expressed as the product of the individual density functions by independence.

So:

$$p = \int_{x=0}^{\infty} \phi(x) \left\{ \int_{y=-x}^{\infty} \phi(y)\, dy \right\} dx$$

$$= \int_{x=0}^{\infty} \phi(x) \left\{ \left[\Phi(y) \right]_{-x}^{\infty} \right\} dx$$

$$= \int_{x=0}^{\infty} \phi(x) \left\{ 1 - \Phi(-x) \right\} dx$$

$$= \int_{x=0}^{\infty} \phi(x) \Phi(x) dx$$

Finally:

$$p = \int_0^{\infty} \phi(x)\, \Phi(x) dx = \left[\tfrac{1}{2} \{\Phi(x)\}^2 \right]_0^{\infty} = \tfrac{1}{2} \left[1^2 - \left(\tfrac{1}{2}\right)^2 \right] = \tfrac{3}{8}$$ [4]

(iv) ***Probability that B_t takes a negative value at some time between 0 and 2***

The probability is 1 because B_t will almost surely take a negative value at some point close to $t = 0$. [1]

Solution 2.9

(i) ***Explain why the student's argument is not correct***

The student's argument is not correct because $B(s)$ and $B(t)$, which represent the value of the process at two different times, are not independent. (In fact, they are positively correlated.) [1]

It is actually the increments $B(s)-B(0)$ and $B(t)-B(s)$ that are independent. The correct calculation can be done by expressing $X = B(s)+B(t)$ in terms of these increments:

$$\begin{aligned}
\text{var}(X) &= \text{var}[B(s)+B(t)] \\
&= \text{var}[2B(0)+2\{B(s)-B(0)\}+\{B(t)-B(s)\}] \\
&= 4\,\text{var}[B(0)]+4\,\text{var}[B(s)-B(0)]+\text{var}[B(t)-B(s)] \\
&= 0+4s+(t-s) = 3s+t
\end{aligned}$$

[3]

[Total 4]

(ii) ***Show the general formula***

This $3s+t$ formula works if $s<t$. If $t<s$, we can swap the letters to get $3t+s$. If $s=t$, we have:

$$\text{var}(X) = \text{var}[B(s)+B(t)] = \text{var}[2B(s)] = 4\,\text{var}[B(s)] = 4s \text{ (or } 4t)$$

which agrees with either formula.

So a general formula would be $s+t+2\min(s,t)$, or if we're using t_1 and t_2 to denote the times, $t_1+t_2+2\min(t_1,t_2)$. [2]

Solution 2.10

(i) ***Show that*** $\boldsymbol{E(B_s B_t) = s}$

If we express B_t in terms of the increment $B_t - B_s$, which is independent of the value of B_s, we get:

$$\begin{aligned} E[B_s B_t] &= E[B_s\{(B_t - B_s) + B_s\}] \\ &= E[B_s(B_t - B_s)] + E[B_s^2] \\ &= E(B_s)E(B_t - B_s) + \text{var}[B_s] + [E(B_s)]^2 = 0 + s + 0 = s \end{aligned}$$ [3]

(ii) ***Find a general formula for the correlation coefficient***

We can then calculate the covariance and correlation between these two values:

$$\text{cov}(B_s, B_t) = E(B_s B_t) - E(B_s)E(B_t) = s$$

and $$\rho(B_s, B_t) = \frac{\text{cov}(B_s, B_t)}{\sqrt{\text{var}(B_s)\,\text{var}(B_t)}} = \frac{s}{\sqrt{st}} = \sqrt{\frac{s}{t}}$$ [1]

This formula only applies when $s < t$. We can generalise this to cover any positive times t_1 and t_2, if we write it in the form:

$$\rho(B_{t_1}, B_{t_2}) = \sqrt{\frac{\min(t_1, t_2)}{\max(t_1, t_2)}}$$ [1]

[Total 2]

There are various alternative ways of writing this, eg $\rho(B_{t_1}, B_{t_2}) = \min\left(\sqrt{\frac{t_1}{t_2}}, \sqrt{\frac{t_2}{t_1}}\right)$.

Solution 2.11

(i) ***Formulae for expectations***

$$E\left(e^{aX}\right) = e^{\mu a + \frac{1}{2}\sigma^2 a^2}$$ [½]

This is the MGF of a normal random variable and can be found in the Tables.

We therefore have:

$$E\left(Xe^{aX}\right)=\frac{d}{da}E\left(e^{aX}\right)=\frac{d}{da}e^{\mu a+\frac{1}{2}\sigma^2a^2}=\left(\mu+\sigma^2a\right)e^{\mu a+\frac{1}{2}\sigma^2a^2}. \qquad [1½]$$

[Total 2]

(ii) ***Show that X_t is a martingale***

$$\begin{aligned}
E&\left[X_t|F_s\right]\\
&=E\left[\left(B_t-at\right)e^{aB_t-0.5a^2t}|F_s\right]\\
&=E\left[\left(B_s+B_t-B_s-at\right)e^{a\left(B_s+B_t-B_s\right)-0.5a^2t}|F_s\right]\\
&=E\left[\left(B_s-at\right)e^{aB_s-0.5a^2t}e^{a\left(B_t-B_s\right)}|F_s\right]+E\left[\left(B_t-B_s\right)e^{aB_s-0.5a^2t}e^{a\left(B_t-B_s\right)}|F_s\right] \qquad [2]
\end{aligned}$$

Now we can use the fact that we are conditioning on all the information known at time s. Any terms involving B_s can be taken outside the expectation, as can terms in s and t (which are fixed, not random points in time). This gives:

$$=\left(B_s-at\right)e^{aB_s-0.5a^2t}E\left[e^{a\left(B_t-B_s\right)}|F_s\right]+e^{aB_s-0.5a^2t}E\left[\left(B_t-B_s\right)e^{a\left(B_t-B_s\right)}|F_s\right] \qquad [1]$$

We can then drop the conditions since the increments are independent of the past:

$$=\left(B_s-at\right)e^{aB_s-0.5a^2t}E\left[e^{a\left(B_t-B_s\right)}\right]+e^{aB_s-0.5a^2t}E\left[\left(B_t-B_s\right)e^{a\left(B_t-B_s\right)}\right] \qquad [1]$$

Using part (i), and noting that $B_t-B_s\sim N(0,t-s)$ so that $\mu=0$ and $\sigma^2=t-s$, we therefore get:

$$\begin{aligned}
&=\left(B_s-at\right)e^{aB_s-0.5a^2t}e^{0.5a^2(t-s)}+e^{aB_s-0.5a^2t}a\left(t-s\right)e^{0.5a^2(t-s)}\\
&=\left(B_s-as\right)e^{aB_s-0.5a^2s}\\
&=X_s \qquad [1]
\end{aligned}$$

as required. [Total 5]

Note that the bounded condition is given in the question.

Solution 2.12

(i) ***Ito's Lemma***

Ito's Lemma states that $f(X_t)$ satisfies the SDE:

$$df(X_t) = \sigma_t f'(X_t)dB_t + [\mu_t f'(X_t) + \tfrac{1}{2}\sigma_t^2 f''(X_t)]dt \qquad [2]$$

As a "technicality", the function f has to be twice differentiable, in order for the RHS to make sense.

(ii)(a) ***Stochastic differential equation for G_t***

Here the function we are applying Ito's Lemma to is $f(x) = e^x$, with $f'(x) = e^x$ and $f''(x) = e^x$.

So we get:

$$\begin{aligned} dG_t &= \sigma_t e^{X_t} dB_t + [\mu_t e^{X_t} + \tfrac{1}{2}\sigma_t^2 e^{X_t}]dt \\ &= \sigma_t G_t dB_t + [\mu_t + \tfrac{1}{2}\sigma_t^2]G_t dt \end{aligned} \qquad [2]$$

(ii)(b) ***Stochastic differential equation for Q_t***

Here the function we are applying Ito's Lemma to is $f(x) = x^2$, with $f'(x) = 2x$ and $f''(x) = 2$.

So we get:

$$dQ_t = 2\sigma_t X_t dB_t + [2\mu_t X_t + \sigma_t^2]dt$$

We can write this entirely in terms of the new process as:

$$dQ_t = 2\sigma_t Q_t^{1/2} dB_t + [2\mu_t Q_t^{1/2} + \sigma_t^2]dt \qquad [2]$$

(ii)(c) ***Stochastic differential equation for V_t***

Here the function we are applying Ito's Lemma to is $f(x)=(1+x)^{-1}$, with $f'(x)=-(1+x)^{-2}$ and $f''(x)=2(1+x)^{-3}$. So we get:

$$\begin{aligned} dV_t &= -\sigma_t(1+X_t)^{-2}dB_t + [-\mu_t(1+X_t)^{-2} + \sigma_t^2(1+X_t)^{-3}]dt \\ &= -\sigma_t V_t^2 dB_t + [-\mu_t V_t^2 + \sigma_t^2 V_t^3]dt \end{aligned} \qquad [2]$$

(ii)(d) ***Stochastic differential equation for L_t***

Here the function we are applying Ito's Lemma to is $f(x)=100+10x$, with $f'(x)=10$ and $f''(x)=0$. So we get:

$$dL_t = 10\sigma_t dB_t + (10\mu_t + 0)dt = 10(\sigma_t dB_t + \mu_t dt) \qquad [2]$$

(ii)(e) ***Stochastic differential equation for J_t***

Since J_t is a function of standard Brownian motion, B_t, rather than X_t, when applying Ito's Lemma, we note:

$$dB_t = 1\times dB_t + 0\times dt$$

Equivalently, $\mu_t = 0$ and $\sigma_t = 1$.

Here the function we are applying Ito's Lemma to is $f(x)=\ln(x)$, with $f'(x)=x^{-1}$ and $f''(x)=-x^{-2}$. So we get:

$$\begin{aligned} dJ_t &= 1\times B_t^{-1}dB_t + \tfrac{1}{2}\times 1\times(-B_t)^{-2}dt \\ &= B_t^{-1}dB_t - \tfrac{1}{2}B_t^{-2}dt \end{aligned} \qquad [2]$$

(ii)(f) ***Stochastic differential equation for*** $\boldsymbol{K_t}$

K_t is a function of standard Brownian motion, B_t, so $\mu_t = 0$ and $\sigma_t = 1$.

Here the function we are applying Ito's Lemma to is $f(x) = 5x^3 + 2x$, with $f'(x) = 15x^2 + 2$ and $f''(x) = 30x$. So we get:

$$dJ_t = 1 \times (15B_t^2 + 2)dB_t + \tfrac{1}{2} \times 1 \times (30B_t)dt$$

$$= (15B_t^2 + 2)dB_t + 15B_t dt \qquad [2]$$

[Total 12]

Solution 2.13

(i) ***Expression for*** $\boldsymbol{d(B_s^2)}$

Either by applying Ito's Lemma or by simply expanding as a Taylor series, we find that:

$$d(B_s^2) = 2B_s dB_s + \tfrac{1}{2} \times 2(dB_s)^2 = 2B_s dB_s + ds$$

where, as usual, we have replaced the second order Brownian differential $(dB_s)^2$ with the time differential ds, using the 2×2 multiplication grid given in Core Reading. [2]

To show this, let $G(B_s) = B_s^2$, *then* $\dfrac{\partial G}{\partial s} = 0$, $\dfrac{\partial G}{\partial B_s} = 2B_s$ *and* $\dfrac{\partial^2 G}{\partial B_s^2} = 2$, *and since we can write* $dB_s = 0\,ds + 1\,dB_s$, *the drift and volatility functions in the stochastic differential equation for* dB_s *are 0 and 1 respectively. Hence, using Ito's Lemma from page 46 in the Tables we have:*

$$dG = \left[0 \times 2B_s + \tfrac{1}{2} \times 1^2 \times 2 + 0\right] ds + 1 \times 2B_s dB_s$$

ie $\quad d(B_s^2) = 2B_s dB_s + ds$

Alternatively, using a Taylor Series expansion we have:

$$dG = \frac{\partial G}{\partial s} ds + \frac{\partial G}{\partial B_s} dB_s + \tfrac{1}{2} \frac{\partial^2 G}{\partial B_s^2} (dB_s)^2$$

$$= 0 \times ds + 2B_s dB_s + \tfrac{1}{2} \times 2 \times (dB_s)^2$$

$$= 2B_s dB_s + ds$$

If we now integrate this equation between time 0 and time t, we get:

$$\int_0^t d(B_s^2) = 2\int_0^t B_s dB_s + \int_0^t ds$$

ie $$\left[B_s^2\right]_0^t = 2\int_0^t B_s dB_s + [s]_0^t$$

ie $$B_t^2 - B_0^2 = 2\int_0^t B_s dB_s + t \quad \Rightarrow \int_0^t B_s dB_s = \tfrac{1}{2}(B_t^2 - t)$$ [2]

[Total 4]

Note that the value of this integral is a random variable.

(ii) ***What is the expected value at time 0?***

At time 0 the future values of B_t are unknown, and we have:

$$E_0\left[\int_0^t B_s dB_s\right] = E_0\left[\tfrac{1}{2}(B_t^2 - t)\right] = \tfrac{1}{2}[E_0(B_t^2) - t]$$

But we know that $B_t - B_0 \sim N(0, t)$.

So: $$E_0(B_t^2) = \text{var}_0(B_t) + \left[E_0(B_t)\right]^2 = t + 0^2 = t$$

and $$E_0\left[\int_0^t B_s dB_s\right] = \tfrac{1}{2}[t - t] = 0$$ [2]

(iii) ***What is the expected value at time u?***

At time u (where $0 < u < t$) the values of B_t up to time u are known, but the values after time u are unknown. If we split the integral into two parts, we have:

$$\int_0^t B_s dB_s = \int_0^u B_s dB_s + \int_u^t B_s dB_s$$

The first integral on the RHS is known at time u and, from part (i), it is equal to:

$$\int_0^u B_s dB_s = \tfrac{1}{2}(B_u^2 - u)$$

The second integral on the RHS is a random quantity. Using the same method as in part (i), it is equal to:

$$\int_u^t B_s dB_s = \left[\tfrac{1}{2}(B_s^2 - s)\right]_u^t = \tfrac{1}{2}(B_t^2 - t) - \tfrac{1}{2}(B_u^2 - u)$$

The only random quantity in this expression is B_t^2 and we need to find the expected value of this based on the information we have at time u. We know that $B_t - B_u \sim N(0, t-u)$ or, since we know the value of B_u, $B_t \sim N(B_u, t-u)$. This tells us that:

$$E_u(B_t^2) = \text{var}_u(B_t) + [E_u(B_t)]^2 = t - u + B_u^2$$

So the expected value of our second integral is:

$$E_u\left[\int_u^t B_s dB_s\right] = 0$$

Combining the two parts of the integral, we get:

$$E_u\left[\int_0^t B_s dB_s\right] = \tfrac{1}{2}(B_u^2 - u) + 0 = \tfrac{1}{2}(B_u^2 - u) \qquad [3]$$

Alternatively, we could use the fact that $B_t^2 - t$ *is a martingale with respect to* B_t *to show this.*

(iv) ***What can you say about this process?***

From part (i), we know that $\int_0^t B_s dB_s$ is always equal to $\frac{1}{2}(B_t^2 - t)$. So, in particular, $\frac{1}{2}(B_u^2 - u) = I_u$.

So we can write the result in (iii) as:

$$E_u\left[I_t\right] = I_u \quad \text{whenever } 0 < u < t$$

This tells us that the process I_t is a martingale with respect to B_t. [1]

It is a general property that Ito integrals are martingales.

Solution 2.14

This integral can be thought of as the (limiting) sum of the small elements $t dB_t$.

The dB_t's are random quantities with mean 0 and variance dt. So the expected value of each element is 0 and the variance is $t^2 dt$.

So: $$E(I) = E\left(\int_0^1 t\, dB_t\right) = \int_0^1 t\, E(dB_t) = 0$$ [1]

Also:

$$\text{var}(I) = \text{var}\left(\int_0^1 t\, dB_t\right) = \int_0^1 \text{var}(t\, dB_t)$$

since the increments are independent. So:

$$\text{var}(I) = \int_0^1 t^2\, dt = \left[\tfrac{1}{3}t^3\right]_0^1 = \tfrac{1}{3}$$ [2]

[Total 3]

Solution 2.15

Since X_t is a function of standard Brownian motion, W_t, when applying Ito's Lemma, we note that the stochastic differential equation for the underlying stochastic process (standard Brownian motion) is:

$$dW_t = 1 \times dW_t + 0 \times dt \qquad [½]$$

Let $G(W_t) = X_t = \alpha W_t^2 + \beta$, then:

- $\dfrac{\partial G}{\partial t} = 0$
- $\dfrac{\partial G}{\partial W_t} = 2\alpha W_t$ [½]
- $\dfrac{\partial^2 G}{\partial W_t^2} = 2\alpha$ [½]

Hence, using Ito's Lemma from page 46 in the *Tables* we have:

$$dG = \left[0 \times 2\alpha W_t + ½ \times 1^2 \times 2\alpha + 0\right]dt + 1 \times 2\alpha W_t dW_t \qquad [½]$$

ie $\quad dX_t = 2\alpha W_t dW_t + \alpha dt$ [1]

[Total 3]

Alternatively, using a Taylor Series expansion we have:

$$\begin{aligned} dG &= \frac{\partial G}{\partial t}dt + \frac{\partial G}{\partial W_t}dW_t + ½\frac{\partial^2 G}{\partial W_t^2}(dW_t)^2 \\ &= 0 \times dt + 2\alpha W_t dW_t + ½ \times 2\alpha \times (dW_t)^2 \\ &= 2\alpha W_t dW_t + \alpha dt \end{aligned}$$

Remembering that $(dW_t)^2 = dt$ *from the multiplication table for increments in Chapter 9.*

Solution 2.16

(i) ***PDE for a martingale***

Taylor's theorem (ignoring higher order terms) is given by

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

Substituting in for dX_t gives:

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

*Note that in the function f we have **explicit** time dependence. This gives the extra term* $\frac{\partial f}{\partial t}dt$ *in the Taylor expansion.*

For a martingale we require zero drift and hence

$$\frac{\partial f}{\partial t}+\frac{1}{2}\sigma^2\left(X_t,t\right)\frac{\partial^2 f}{\partial X_t{}^2}+\mu\left(X_t,t\right)\frac{\partial f}{\partial X_t}=0 \qquad [1]$$

[Total 3]

(ii) ***Verify that the equation holds***

Here $X_t = B_t$ and we are looking at a function of standard Brownian motion.

Note that $dB_t = 0\times dt + 1\times dB_t$, so $\mu = 0$ and $\sigma = 1$, and the derivatives are:

$$\frac{\partial f}{\partial t}=-1,\ \frac{\partial f}{\partial X_t}=2B_t,\ \frac{\partial^2 f}{\partial X_t^2}=2 \qquad [½]$$

Since the terms on the left-hand side of the equation in part (i) sum to zero, the equation holds. [½]

[Total 1]

(iii) ***Find $g(t)$***

Again $\mu = 0$ and $\sigma = 1$. Using the equation from part (i), we require:

$$g'(t)B_t + ½ \times 6B_t = 0 \qquad [1]$$

So it follows that $g'(t) = -3$ and hence $g(t) = -3t$ will do, *ie* $B_t^3 - 3tB_t$ is a martingale. [1]

We could also have used $g(t) = -3t + c$, where c is any constant.

Solution 2.17

(i) The process X_t is an Ornstein-Uhlenbeck process so that:

$$X_t = X_0 e^{-\gamma t} + \sigma \int_0^t e^{-\gamma(t-s)} dB_s \qquad [1]$$

(ii) The process $R_t - 4$ is an Ornstein-Uhlenbeck process so that:

$$R_t - 4 = (R_0 - 4)e^{-\gamma t} + \sigma \int_0^t e^{-\gamma(t-s)} dB_s \qquad [½]$$

which in this case becomes:

$$R_t = 4 + e^{-0.8t} + \int_0^t e^{-0.8(t-s)} dB_s \qquad [½]$$

[Total 1]

(iii) Since $dB_s \sim N(0, ds)$, and these increments are independent, it follows that:

$$R_t \sim N\left(4 + e^{-0.8t}, \int_0^t e^{-1.6(t-s)} ds\right) = N\left(4 + e^{-0.8t}, \frac{1 - e^{-1.6t}}{1.6}\right) \qquad [2]$$

In the long-term $\lim_{t \to \infty} R_t \sim N(4, 0.625)$. [1]

[Total 3]

Solution 2.18

Starting with a Taylor Series expansion for dY_t:

$$dY_t = d\ln\left(\frac{1}{X_t}-1\right) = f'\left(X_t\right)dX_t + 0.5f''\left(X_t\right)\left(dX_t\right)^2 \qquad (1) \qquad [1]$$

where $f(x)=\ln\left(\frac{1}{x}-1\right)$. We have:

$$f'(x)=\frac{-1/x^2}{\left(\frac{1}{x}-1\right)}=\frac{1}{x^2-x} \qquad \text{and} \qquad f''(x)=-\frac{2x-1}{\left(x^2-x\right)^2}=\frac{1-2x}{x^2\left(1-x\right)^2} \qquad [2]$$

We also have:

$$\left(dX_t\right)^2 = X_t^2\left(1-X_t\right)^2 dt \qquad [1]$$

by using the 2×2 multiplication grid for increments given in the Core Reading, which tells us that $(dt)^2=0$ and $dB_t \times dt = 0$. Substituting this into (1), along with the derivatives of f gives:

$$\begin{aligned} dY_t &= \left[\frac{-1}{X_t\left(1-X_t\right)}\right]\left[0.5X_t\left(1-X_t\right)\left(1-2X_t\right)dt - X_t\left(1-X_t\right)dB_t\right] \\ &\qquad + 0.5\left[\frac{1-2X_t}{X_t^{\,2}\left(1-X_t\right)^2}\right]\left[X_t^2\left(1-X_t\right)^2 dt\right] \\ &= dB_t \end{aligned} \qquad [1]$$

Alternatively, substituting the same partial derivatives as above into Ito's Lemma on page 46 in the Tables, together with the drift and volatility functions:

$$a\left(X_t\right) = 0.5X_t\left(1-X_t\right)\left(1-2X_t\right)$$

$$b\left(X_t\right) = -X_t\left(1-X_t\right)$$

gives:

$$dY_t = \left[0.5X_t\left(1-X_t\right)\left(1-2X_t\right)\times\frac{1}{X_t^2-X_t} + \tfrac{1}{2}X_t^2\left(1-X_t\right)^2\frac{1-2X_t}{X_t{}^2\left(1-X_t\right)^2} \right]dt$$

$$-X_t\left(1-X_t\right)\times\frac{1}{X_t^2-X_t}dB_t$$

$$= dB_t \qquad [2]$$

It follows by integrating and taking the initial condition $X_0 = 0.5$ (which implies $Y_0 = 0$) into account that:

$$Y_t = B_t \qquad ie: \qquad \ln\left(\frac{1}{X_t}-1\right) = B_t \qquad [1]$$

Rearranging gives: $X_t = \dfrac{1}{1+e^{B_t}}$ [1]

[Total 7]

Solution 2.19

(i) ***Stochastic differential dS_t***

Rearranging the relationship given, we get:

$$S_t = S_0 e^{\mu t+\sigma B_t} \qquad [½]$$

Since S_t is a function of standard Brownian motion, B_t, when applying Ito's Lemma, we note that the stochastic differential equation for the underlying stochastic process (standard Brownian motion) is:

$$dB_t = 0\times dt + 1\times dB_t \qquad [½]$$

Let $G(t, B_t) = S_t = S_0 e^{\mu t+\sigma B_t}$, then:

- $\dfrac{\partial G}{\partial t} = \mu S_0 e^{\mu t+\sigma B_t} = \mu S_t$ [½]
- $\dfrac{\partial G}{\partial B_t} = \sigma S_0 e^{\mu t+\sigma B_t} = \sigma S_t$ [½]

- $$\frac{\partial^2 G}{\partial B_t^2} = \sigma^2 S_0 e^{\mu t + \sigma B_t} = \sigma^2 S_t$$ [½]

Hence, using Ito's Lemma from page 46 in the *Tables* we have:

$$dG = \left[0 \times \sigma S_t + ½ \times 1^2 \times \sigma^2 S_t + \mu S_t\right] dt + 1 \times \sigma S_t dB_t$$ [½]

ie $$dS_t = \left(\mu + ½\sigma^2\right) S_t dt + \sigma S_t dB_t$$ [½]

Alternatively, using a Taylor Series expansion we have:

$$\begin{aligned} dG &= \frac{\partial G}{\partial t} dt + \frac{\partial G}{\partial B_t} dB_t + ½ \frac{\partial^2 G}{\partial B_t^2} (dB_t)^2 \\ &= \mu S_t dt + \sigma S_t dB_t + ½\sigma^2 S_t \times (dB_t)^2 \\ &= \left(\mu + ½\sigma^2\right) S_t dt + \sigma S_t dB_t \end{aligned}$$

Remember that $(dB_t)^2 = dt$ *from the multiplication table for increments in Chapter 9.*

This can be written as:

$$\frac{dS_t}{S_t} = \sigma dB_t + \left(\mu + \tfrac{1}{2}\sigma^2\right) dt$$

So: $c_1 = \sigma$ and $c_2 = \mu + \tfrac{1}{2}\sigma^2$ [1]

[Maximum 4]

(ii) ***Expressions for the mean and variance***

The expected value of S_t is:

$$E[S_t] = E[S_0 e^{\mu t + \sigma B_t}] = S_0 e^{\mu t} E[e^{\sigma B_t}]$$

Since $B_t \sim N(0,t)$, its MGF is $E[e^{\theta B_t}] = e^{½\theta^2 t}$.

So: $$E[S_t] = S_0 e^{\mu t} \times e^{½\sigma^2 t} = S_0 e^{\mu t + ½\sigma^2 t}$$ [2]

The variance of S_t is:

$$\begin{aligned}\text{var}[S_t] &= E[S_t^2] - (E[S_t])^2 \\ &= E[S_0^2 e^{2\mu t + 2\sigma B_t}] - (S_0 e^{\mu t + \frac{1}{2}\sigma^2 t})^2 \\ &= S_0^2 e^{2\mu t} E[e^{2\sigma B_t}] - S_0^2 e^{2\mu t + \sigma^2 t} \\ &= S_0^2 e^{2\mu t + 2\sigma^2 t} - S_0^2 e^{2\mu t + \sigma^2 t} = S_0^2 e^{2\mu t}(e^{2\sigma^2 t} - e^{\sigma^2 t})\end{aligned}$$ [2]

[Total 4]

An alternative approach is to use the formulae for the mean and variance of the lognormal distribution.

(iii) ***Expressions for the covariance and conditional expectation***

The covariance of S_{t_1} and S_{t_2} is:

$$\text{cov}[S_{t_1}, S_{t_2}] = E[S_{t_1} S_{t_2}] - E[S_{t_1}]E[S_{t_2}]$$

From above:

$$E[S_{t_1}] = S_0 e^{\mu t_1 + \frac{1}{2}\sigma^2 t_1} \quad \text{and} \quad E[S_{t_2}] = S_0 e^{\mu t_2 + \frac{1}{2}\sigma^2 t_2}$$

The expected value of the product is:

$$\begin{aligned}E[S_{t_1} S_{t_2}] &= E[S_0 \exp(\mu t_1 + \sigma B_{t_1}) S_0 \exp(\mu t_2 + \sigma B_{t_2})] \\ &= S_0^2 e^{\mu(t_1+t_2)} E[\exp(\sigma B_{t_1} + \sigma B_{t_2})]\end{aligned}$$

To evaluate this we need to split B_{t_2} into two independent components:

$$B_{t_2} = B_{t_1} + (B_{t_2} - B_{t_1}) \quad \text{where} \quad B_{t_2} - B_{t_1} \sim N(0, t_2 - t_1)$$

We then get:

$$E[S_{t_1}S_{t_2}] = S_0^2 e^{\mu(t_1+t_2)} E[\exp(\sigma B_{t_1} + \sigma\{B_{t_1} + (B_{t_2} - B_{t_1})\})]$$

$$= S_0^2 e^{\mu(t_1+t_2)} E[\exp(2\sigma B_{t_1} + \sigma\{B_{t_2} - B_{t_1}\})]$$

$$= S_0^2 e^{\mu(t_1+t_2)} E[\exp(2\sigma B_{t_1})]\, E[\exp(\sigma\{B_{t_2} - B_{t_1}\})]$$

$$= S_0^2 e^{\mu(t_1+t_2)} \exp(2\sigma^2 t_1)\, \exp[½\sigma^2(t_2 - t_1)]$$

$$= S_0^2 e^{\mu(t_1+t_2)} e^{\frac{3}{2}\sigma^2 t_1 + \frac{1}{2}\sigma^2 t_2}$$

Putting these together gives:

$$\text{cov}[S_{t_1}, S_{t_2}] = S_0^2 e^{\mu(t_1+t_2)} e^{\frac{3}{2}\sigma^2 t_1 + \frac{1}{2}\sigma^2 t_2} - S_0 e^{\mu t_1 + ½\sigma^2 t_1} . S_0 e^{\mu t_2 + ½\sigma^2 t_2}$$

$$= S_0^2 e^{\mu(t_1+t_2)} \left(e^{\frac{3}{2}\sigma^2 t_1} - e^{\frac{1}{2}\sigma^2 t_1} \right) e^{\frac{1}{2}\sigma^2 t_2}$$ [4]

The conditional expectation is:

$$E[S_{t_2} \mid S_{t_1}] = E[S_0 \exp(\mu t_2 + \sigma B_{t_2}) \mid B_{t_1}]$$

$$= S_0 e^{\mu t_2} E[\exp(\sigma\{B_{t_1} + (B_{t_2} - B_{t_1})\}) \mid B_{t_1}]$$

$$= S_0 e^{\mu t_1 + \mu(t_2 - t_1)} \exp(\sigma B_{t_1}) E[\exp(\sigma\{B_{t_2} - B_{t_1}\}) \mid B_{t_1}]$$

$$= S_{t_1} e^{\mu(t_2 - t_1)} e^{½\sigma^2(t_2 - t_1)} = S_{t_1} e^{(\mu + ½\sigma^2)(t_2 - t_1)}$$ [2]

[Total 6]

(iv) ***Martingale***

We can rearrange the last result in the form:

$$E[e^{-(\mu+½\sigma^2)t_2} S_{t_2} \mid S_{t_1}] = e^{-(\mu+½\sigma^2)t_1} S_{t_1}$$

which shows that, subject to the convergence criterion, the process $e^{-(\mu+½\sigma^2)t} S_t$ is a martingale with respect to $\{S_t\}$. [1]

This is an example of a "discounted" security price process and it is a martingale. Such processes are very important in the theory of derivative pricing.

Solution 2.20

(i) ***Stochastic differential equation***

We know that:

$$X_t = \log S_t = \mu t + \sigma W_t$$

So X_t satisfies the SDE:

$$dX_t = \mu dt + \sigma dW_t \qquad [1]$$

(ii) ***Apply Ito's Lemma***

We can now apply Ito's Lemma to find the SDE for S_t, which is a function of X_t, namely $S_t = \exp(X_t)$.

Now:

- $\dfrac{\partial S_t}{\partial X_t} = \exp(X_t)$
- $\dfrac{\partial^2 S_t}{\partial X_t^2} = \exp(X_t)$

and from (i) above , $dX_t = \mu dt + \sigma dW_t$, so the drift and volatility functions in the stochastic differential equation for X_t are μ and σ respectively. Hence, Ito's Lemma gives:

$$\begin{aligned} dS_t &= [\mu \exp(X_t) + \tfrac{1}{2}\sigma^2 \exp(X_t)]\, dt + \sigma \exp(X_t)\, dW_t \\ &= S_t[(\mu + \tfrac{1}{2}\sigma^2)dt + \sigma dW_t] \end{aligned}$$

"Otherwise"

The "otherwise" approach uses a Taylor Series expansion:

$$dS_t = \frac{\partial S_t}{\partial X_t} dX_t + \tfrac{1}{2} \frac{\partial^2 S_t}{\partial X_t^2} (dX_t)^2$$

$$= \exp(X_t)\ dX_t + \tfrac{1}{2} \exp(X_t)\ (dX_t)^2$$

$$= \exp(X_t)[\mu dt + \sigma dW_t] + \tfrac{1}{2} \exp(X_t)[\mu dt + \sigma dW_t]^2$$

$$= \exp(X_t)[\mu dt + \sigma dW_t] + \tfrac{1}{2} \exp(X_t)[\mu^2 (dt)^2 + 2\mu\sigma\, dt dW_t + \sigma^2 (dW_t)^2]$$

We can then disregard the second order terms, except for $(dW_t)^2$, which we replace with dt. This gives:

$$dS_t = \exp(X_t)[\mu dt + \sigma dW_t] + \tfrac{1}{2} \exp(X_t)\sigma^2 dt$$

$$= S_t[(\mu + \tfrac{1}{2}\sigma^2)dt + \sigma dW_t]$$

[3]

(iii) ***Find the probability***

The probability that the share price will fall during a given year – say from time $t-1$ to time t – is:

$$P(S_t < S_{t-1}) = P\left[\frac{S_t}{S_{t-1}} < 1\right] \quad [½]$$

From the equation given in the question, we know that:

$$\frac{S_t}{S_{t-1}} = \frac{\exp[\mu t + \sigma W_t]}{\exp[\mu(t-1) + \sigma W_{t-1}]} = \exp\left[\mu + \sigma(W_t - W_{t-1})\right] \quad [½]$$

Using the fact that $W_t - W_{t-1} \sim N(0,1)$, we then find:

$$\begin{aligned} P(S_t < S_{t-1}) &= P\left[\exp\left[\mu + \sigma(W_t - W_{t-1})\right] < 1\right] \\ &= P\left[\mu + \sigma(W_t - W_{t-1}) < 0\right] \\ &= P\left[W_t - W_{t-1} < -\frac{\mu}{\sigma}\right] \\ &= \Phi\left(-\frac{0.06}{0.25}\right) = \Phi(-0.24) = 0.405 \end{aligned}$$ [2]

[Total 3]

Solution 2.21

Advantages

- This assumption makes the mathematics more tractable than other more complex models.
- The variance of the returns in a particular period is proportional to the length of that period, which seems intuitively reasonable.
- An Ito process is consistent with the assumption that markets are efficient (at least, in weak form).
- The model implies that the return and risk characteristics of the underlying share (when expressed relative to the share price) are independent of the absolute level of the share price and the currency unit adopted.
- The geometric Brownian motion model underlies other, widely-used models, such as the Black-Scholes option pricing model.

Disadvantages

- In reality share price movements may not be consistent with such a process. For example, prices may "jump", *ie* change suddenly by a significant amount.
- Also, in reality, μ and σ are not constant. For instance, the drift parameter might depend on interest rates, and studies of historical option prices suggest that investors' expectations of volatility fluctuate markedly over time.

- Historical evidence suggests that large movements in prices are more common than the lognormal distribution implied by an Ito process would suggest.

- The assumption of an efficient market may be invalid.

- There is evidence that share prices exhibit momentum effects in the short-term and mean-reversion in the long-term. This contradicts the property of independent increments and hence weak form efficiency.

[½ each, maximum 4]

Solution 2.22

(i) ***Explain what is meant by the Wilkie model***

The Wilkie model attempts to model the processes generating investment returns for several different types of asset. [½]

It can therefore be used to simulate the possible future development of investment returns, *eg* as part of an asset liability modelling exercise. [½]

It is primarily statistically-based – having been estimated using historical data for the relevant time series involved. [½]

It has a cascade structure. [½]

The key variable is the force of inflation. [½]

This is modelled as a first-order autoregressive process with normally distributed innovations. [½]

Inflation is assumed to be the driving force behind the other variables, such as: [½]

- share dividends
- share dividend yields
- wage inflation. [½]

It is therefore a particular case of a vector autoregressive moving average (VARMA) model. [½]

[Maximum 4]

(ii) ***Explain what is meant by the continuous-time lognormal model***

The continuous-time lognormal model of security prices assumes that log prices form a random walk. [½]

If S_t denotes the market price of an investment at time t, then the model states that, for $u > t$, log returns are given by:

$$\log(S_u) - \log(S_t) \sim N\left[\mu(u-t), \sigma^2(u-t)\right]$$

where μ is the parameter associated with the drift and σ is the parameter associated with the volatility. [1]

The values of μ and σ are specific to the investment considered. [½]

Under the continuous-time lognormal model, the proportional change in the price is lognormally distributed, so that returns over any interval do not depend on the initial value of the investment, S_t. [1]

The mean and variance of the log returns are proportional to the length of the interval considered $(u-t)$. [½]

It is assumed that returns over non-overlapping intervals are independent of each other. [½]
[Total 4]

(iii) ***Two key differences between the models***

The lognormal model assumes that investment returns in non-overlapping intervals are independent and that the expected return does not change over time. In contrast, the Wilkie model models key investment variables – *eg* yields – as autoregressive processes that tend to revert to a long-term mean value. [1]

The Wilkie model models the yields and/or prices and/or returns produced by several different classes of asset (including equities, conventional and index-linked bonds, and property), whereas the lognormal model is usually applied only to equity prices. [1]
[Total 2]

Solution 2.23

(i) ***Definitions***

(a) A *longitudinal property* picks one simulation and looks at a statistic sampled repeatedly from that simulation over a long period of time. [1]

(b) A *cross-sectional property* fixes a time horizon and looks at the distribution over all the simulations. [1]
[Total 2]

(ii) ***Properties of interest in the table***

Cross-sectional properties of interest:

- The most important property is that of the distribution of the ratio of assets to liabilities after five years. This is when the new guidelines take effect. [1]

- All simulations show a fall in the ratio from year 1 to year 2 and subsequent rise from year 2 to year 3. This needs to be investigated. It may be that the company employs a risky strategy early on investing heavily in order to increase the ratio in future years. [1]

Longitudinal property of interest:

- Simulations 65, 66 and 67 are producing ratios of less than 150% which is not desirable. It would be a good idea to investigate the input parameters causing this and hopefully plan to avoid these simulations occurring in practice. [1]
[Total 3]

Part 3 – Questions

Question 3.1 (Developmental)

(i) Write down expressions for the payoff functions for the option holder for each of the following derivative contracts in terms of the current time U, the expiry time T, the price S_t of the underlying at time t and the exercise price K, indicating clearly the time(s) at which exercise may occur:

(a) a European call option

(b) a European put option

(c) an American call option

(d) an American put option [4]

(ii) Identify the profit or loss (ignoring dealing costs) in each of the following scenarios:

(a) A European call option with an exercise price of 480p is bought for a premium of 37p. The price of the underlying share is 495p at the expiry date.

(b) A European put option with an exercise price of 180p is bought for a premium of 12p. The price of the underlying share is 150p at the expiry date.

(c) A European put option with an exercise price of 250p is written for a premium of 22p. The price of the underlying share is 272p at the expiry date. [3]

[Total 7]

Question 3.2 (Developmental)

A speculator has a portfolio consisting of one call option and one put option on the same underlying security. The two options have the same expiry date and the same strike price k. The prices paid for the options were c and P respectively.

(i) Sketch the position diagram (a diagram of the profit/loss at expiry against the security price at expiry) for this portfolio, marking the coordinates of the key points on your graph. [4]

(ii) What opinion concerning the share price would have led the speculator to set up this portfolio? [2]

[Total 6]

Question 3.3 (Developmental)

A speculator has a portfolio consisting of one short holding of a call option on a share. Explain what this means and sketch the position diagram (a diagram of the profit/loss at expiry against the security price at expiry) for this portfolio assuming that no dividends are payable. [4]

Question 3.4 (Bookwork)

(i) What is the difference between a "European" and an "American" option? [1]

(ii) Explain why it is never optimal for the holder of an American call option on a non-dividend-paying share to exercise the option before the expiry date if there is an active trading market for the option. [2]

(iii) Explain why, even if there is no active trading market, it is never optimal for the holder of an American call option on a non-dividend-paying share, who does not intend to sell the underlying stock during the life of the option, to exercise the option before the expiry date. [3]

[Total 6]

Question 3.5 (Developmental)

"*The longer the time to expiry, the greater the chance that the underlying share price can move significantly.*"

(i) Explain how this observation implies that the value of a call option is higher for a 6-month option than for a 3-month option. [2]

(ii) What is the corresponding observation for a put option and what is the explanation in this case? [3]

[Total 5]

Question 3.6 (Exam-style)

(i) By considering a suitably chosen notional portfolio or portfolios (which you should specify carefully), show that the price p_t of a European put option exercisable at time T with a strike price K on an underlying non-dividend-paying share with price S_t at time t satisfies the inequality:

$$p_t \geq Ke^{-r(T-t)} - S_t \qquad [4]$$

(ii) Explain how you would modify your inequality if you knew that holders of the share on the day before the option expires are entitled to receive a cash dividend of $0.02S_T$ payable at time T. [2]

[Total 6]

Question 3.7 (Exam-style)

Consider an asset with price S_t at time t, paying a dividend at a constant dividend yield, D. Dividends are paid at the end of each year and are immediately reinvested in the asset. The continuously-compounded risk-free rate of interest is r *pa.*

Derive the forward price, K, of a contract issued at time t, with maturity at time T, to purchase one unit of the asset, where $T-t$ is an integer number of years. State any assumptions you make. [6]

Question 3.8 (Developmental)

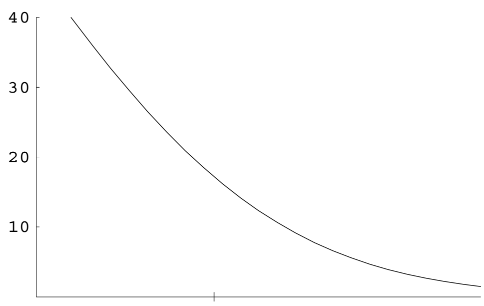
A call option on a stock that does not pay dividends has the following parameter values (in the usual notation):

$$S = 240,\ K = 250,\ T = 0.5,\ r = 0.06,\ \sigma = 0.2$$

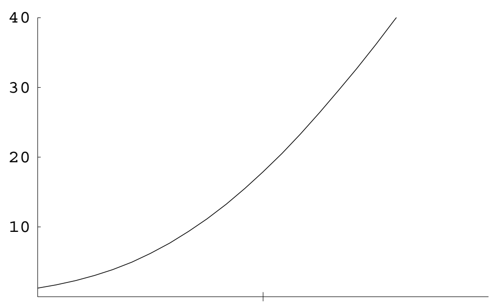
The graphs below show the theoretical price of this option at time $t = 0$ when each of the parameters S, K, T and r is varied without changing the values of the other parameters.

Identify which parameter has been plotted along the x-axis of each graph. [4]

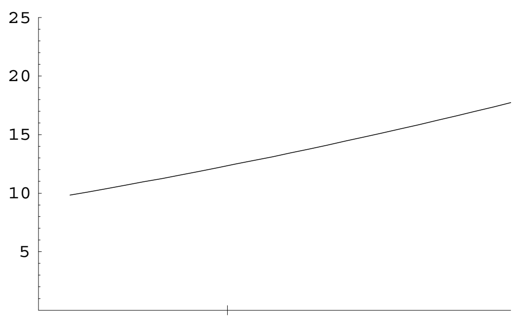
Graph 1



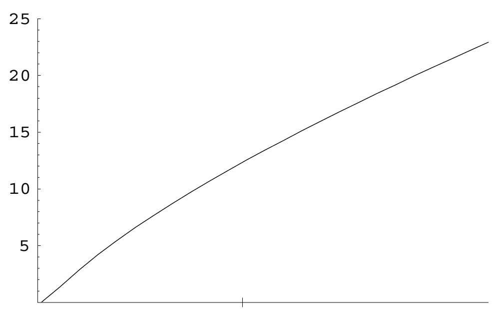
Graph 2



Graph 3



Graph 4



Question 3.9 (Exam-style)

(i) Define the delta, gamma and theta of an option. [3]

(ii) Describe, using a numerical example, the concept of delta hedging. [6]

[Total 9]

Question 3.10 (Exam-style)

Give definitions of the "Greeks" that could be used as an aid to management in each of the following situations. State also the desired ranges for their numerical values and define any notation you use.

(a) A hedge fund manager wishes to establish a delta-neutral position that would not need frequent rebalancing for part of his portfolio.

(b) A derivatives trader is concerned that a change in the distribution of the daily price movements of particular shares might affect the values of the options he holds on those shares.

(c) The trustee of a pension fund that purchased a large number of options last year as a means of hedging is concerned about changes in the value of the fund as the options approach their expiry date. [6]

Question 3.11 (Developmental)

A call option, on an underlying share, has a price of 20.15p and a delta of 0.558 at time t. Determine the hedging portfolio of shares and cash for this option at time t, given that the price of the underlying share, $S_t = 240\text{p}$. [4]

Question 3.12 (Developmental)

In a one-step binomial tree model it is assumed that the initial share price of 260 will either increase to 285 or decrease to 250 at the end of one year. Assume that the annual force of interest is 0.05 and that no dividends are payable.

(i) Calculate the price of a one year European call option with a strike price of 275, using each of the following:

(a) a replicating portfolio method

(b) risk-neutral valuation

(c) a risk-free portfolio method. [6]

(ii) Repeat your calculations in (i) for a one year European put option with a strike price of 275. [6]

(iii) Verify numerically that the put-call parity relationship holds in this case. [2]

[Total 14]

Question 3.13 (Developmental)

The market price of a non-dividend-paying security with current market price S is being modelled using a one-step binomial tree in which the proportionate changes in the security price following an up- and a down-movement are denoted by u and d. The risk-free force of interest over the period is r.

Show that if an option on this security has a payoff of z_u following an up-movement and a payoff of z_d following a down-movement, then the option can be replicated exactly using a portfolio consisting of Δ securities, where $\Delta S = \dfrac{z_u - z_d}{u-d}$, and an amount of cash, ψ, which you should specify. [5]

Question 3.14 (Developmental)

The increase in the price of a share over the next year is believed to have a mean of 10% and a standard deviation of 10%.

(i) Determine the values of u and d for a one-step binomial tree model that are consistent with the mean and standard deviation of the return on the underlying share, assuming that the share price is twice as likely to go up than to go down. [4]

(ii) Hence calculate the value of each of the following options, given that the current share price is 250, the risk-free force of interest is 7½% per annum and dividends can be ignored:

(a) a one-year European call option with a strike price of 275

(b) a one-year European put option with a strike price of 300. [7]

[Total 11]

Question 3.15 (Exam-style)

(i) Show that in a one-step binomial tree model of the price of a non-dividend-paying share, the risk-neutral probability q of an up movement is given by:

$$q = \frac{e^{r\delta t} - d}{u - d}$$

where d, u, r and δt are quantities you should define. [5]

(ii) Explain briefly why it must be assumed that $d < e^{r\delta t} < u$. [2]

(iii) (a) Write down a formula for θ, the expected one-step rate of return on the share based on the real-world probability p of an up-movement.

(b) Show that the real-world variance, σ^2, of the one-step rate of return on the share is $p(1-p)(u-d)^2$. [4]

(iv) Show that $p > q$ if and only if $1+\theta > e^{r\delta t}$ and interpret this result. [4]
[Total 15]

Question 3.16 (Exam-style)

A European put on a share traded in an overseas market has one month to expiry and an exercise price of \$2.00. It is assumed that the share price at expiry will be either \$1.50 or \$2.50. The risk-free interest rate is 1% per month (continuously-compounded).

(i) Use a risk-free portfolio approach to value the put option assuming that the underlying share is currently priced at \$2.19. [4]

(ii) Use risk-neutral valuation to value the put option assuming that the underlying share is currently priced at \$2.21. [2]

(iii) Hence estimate the option's delta when the share is priced at \$2.20. [1]
[Total 7]

Question 3.17 (Exam-style)

The movement of a share price over the next two months is to be modelled using a two-period recombining binomial model. Over each month, it is assumed that the share price will either increase or decrease by 10%.

(i) Over each month, the risk-neutral probability of an up-step is $q = 0.55$. Calculate the monthly risk-free force of interest r that has been used to arrive at this figure. [1]

(ii) The current share price is 1. The annualised expected force of return on the share is $\mu = 30\%$. Calculate the state-price deflators in each of the three possible final states of the share price. [4]

(iii) Calculate the value of each of the following two-month derivatives:

(a) a derivative with payoff profile $(1, 0, 0)$

(b) a derivative with payoff profile $(0, 1, 0)$

(c) a derivative with payoff profile $(0, 0, 1)$

(d) a European call option with a strike price of $K = 0.95$

(e) a European put option with a strike price of $K = 1.05$

(f) a derivative whose payoff is $2 \times |S - 0.98|$, where S is the share price at the end of the two months. [5]

A payoff profile of (x, y, z) means that the derivative returns x if the share price goes up twice, y if the share price goes up once and down once, and z if the share price goes down twice. [Total 10]

Question 3.18 (Exam-style)

An investor believes that each month the price of gold can either move up by 10% or down by 20%.

(i) By using a risk-neutral valuation, calculate the price of a two-month at-the-money European style put option on gold (expressed as a percentage of the price of gold) if the risk-free rate of return is 2% per month and storage costs can be ignored. [4]

(ii) An actuary calculates the same figure for the value of the option using a similar binomial gold price model. However, the actuary arrives at her figure by discounting the expected payoff from the option at a "risk discount rate". Assuming that the actuary's estimate of the probability of a 10% up-movement is 0.8, calculate the risk discount rate used by the actuary. [2]

(iii) With reference to the riskiness (as defined in the capital asset pricing model) of investing in gold explain why the actuary's probability is plausible. [3]

(iv) By considering the sign of the beta of this put option, explain how the risk discount rate used by the actuary can be justified. [2]

[Total 11]

Question 3.19 (Bookwork)

(i) One of the assumptions underlying the Black-Scholes model is that the price of the underlying follows a geometric Brownian motion. Explain briefly what this means and why this assumption may not be valid in practice. [4]

(ii) State the other assumptions underlying the Black-Scholes model. [5]

[Total 9]

Question 3.20 (Bookwork)

One form of the Black-Scholes partial differential equation is:

$$\theta + rs\Delta + \tfrac{1}{2}\sigma^2 s^2 \Gamma = rf$$

(i) State the context in which this formula applies, indicating what s and f represent. (You are not required to state the Black-Scholes assumptions.) [2]

(ii) What do r and σ represent? What assumptions are made about these quantities in this equation? [3]

(iii) State the names and give the mathematical definitions of the "Greeks" that appear in this equation. [3]

(iv) What boundary condition would you need to use in order to solve this equation when applied to a European call option with a strike price K? [1]

[Total 9]

Question 3.21 (Exam-style)

An investor buys, for a premium of 187.06, a call option on a non-dividend-paying stock whose current price is 5,000. The strike price of the call is 5,250 and the time to expiry is 6 months. The risk-free rate of return is 5% *pa* continuously-compounded.

The Black-Scholes formula for the price of a call option on a non-dividend-paying share is assumed to hold.

(i) Calculate the price of a put option with the same time to maturity and strike price as the call. [2]

(ii) The investor buys a put option with strike price 4,750 with the same time to maturity. Calculate the price of the put option if the implied volatility were the same as that in (i).

[You need to estimate the implied volatility to within 1% pa of the correct value.] [7]

[Total 9]

Question 3.22 (Exam-style)

An option is described as "at-the-money-forward" if the price of the underlying asset equals the strike price discounted at the risk-free rate over the remaining life of the option.

(i) Show that, according to the Black-Scholes model, the price of an at-the-money-forward T-year call option on a non-dividend-paying share is approximately a proportion $0.4\sigma\sqrt{T}$ of the price of the underlying asset, where σ is the annual volatility of the underlying asset. [5]

Hint: For a differentiable function: $f(x) \approx f(0) + xf'(0)$

(ii) Hence estimate the price of a 3-month at-the-money-forward call option on shares worth £1m with a volatility of 20% per annum. [2]

[Total 7]

Question 3.23 (Exam-style)

The solution to the Black-Scholes equation for the price V (assuming a risk-free force of interest r) of a European put option maturing u years from now with strike price K on a stock that pays dividends at force q whose current spot price is S is:

$$V = Ke^{-ru}\Phi(-d_2) - Se^{-qu}\Phi(-d_1)$$

where $d_1, d_2 = \dfrac{\log(S/K) + (r - q \pm ½\sigma^2)u}{\sigma\sqrt{u}}$.

(i) Show that the hedge ratio $\Delta = \dfrac{\partial V}{\partial S}$ is given by $\Delta = -e^{-qu}\Phi(-d_1)$. [8]

(ii) Hence find a formula for $\Gamma = \dfrac{\partial^2 V}{\partial S^2}$. [2]

[Total 10]

Question 3.24 (Exam-style)

An investment bank has issued a special derivative security which provides a payoff in one year of:

$S_1-(S_0-15)$	if	$S_0-15\le S_1\le S_0-5$
10	if	$S_0-5\le S_1\le S_0+5$
$(S_0+15)-S_1$	if	$S_0+5\le S_1\le S_0+15$
0	otherwise	

where S_t is the price of the underlying share at time t.

An investor purchases one of these special derivatives on a share with initial price £50.

(i) Write down the investor's payoff from this special derivative in one year's time. [1]

(ii) Explain how this payoff can be written in terms of two long and two short call options with different strike prices. [4]

(iii) Calculate the fair price for this special derivative by the investor, using the following basis:

- volatility of the share price, $\sigma = 15\%$ *pa*
- risk-free interest rate, $r = 3\%$ *pa* (continuously-compounded)
- no dividends are paid on the underlying share [5]

[Total 10]

Part 3 – Solutions

Solution 3.1

(i) ***Payoff functions***

(a) Payoff $= \max(S_T - K, 0)$ defined at time T only [1]

(b) Payoff $= \max(K - S_T, 0)$ defined at time T only [1]

(c) Payoff $= \max(S_U - K, 0)$ defined for $0 < U \le T$ [1]

(d) Payoff $= \max(K - S_U, 0)$ defined for $0 < U \le T$ [1]

[Total 4]

(ii) ***Profit or loss***

(a) Profit $= -37 + (495 - 480) = -22$ *ie* a loss of 22p per contract purchased [1]

(b) Profit $= -12 + (180 - 150) = +18$ *ie* a profit of 18p per contract purchased [1]

(c) Profit $= +22 + 0 = +22$ *ie* a profit of 22p per contract written [1]

[Total 3]

In (c) you have ***collected*** *the premium and the option has expired worthless.*

Solution 3.2

(i) ***Position diagram***

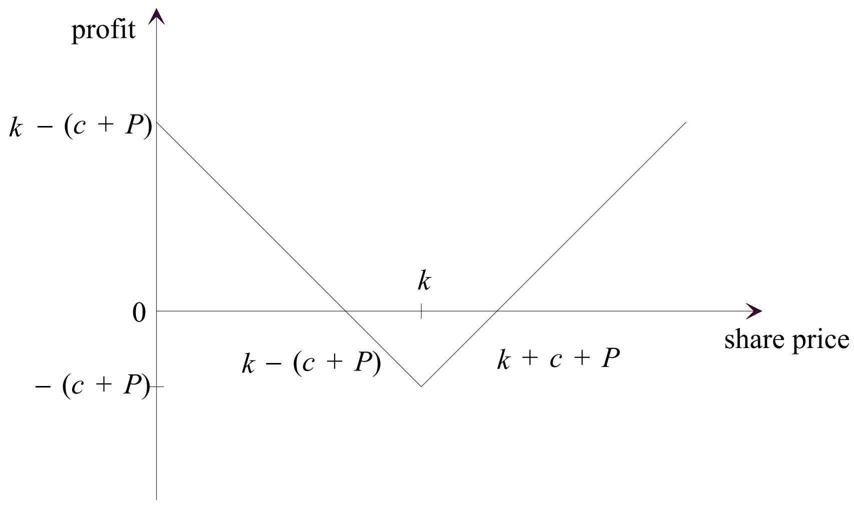
The profit for the call option is $\max(S - k, 0) - c$. [¼]

The profit for the put option is $\max(k - S, 0) - P$. [¼]

The profit for a portfolio containing both of these options is equal to the sum of these, which can be written as:

$$\max(S - k, 0) + \max(k - S, 0) - c - P = |S - k| - (c + P)$$ [1½]

So the position diagram for the holder of this portfolio looks like this:



[2 for a correct diagram]
[Total 4]

A portfolio that has a payoff with this shape is actually called a "*straddle*".

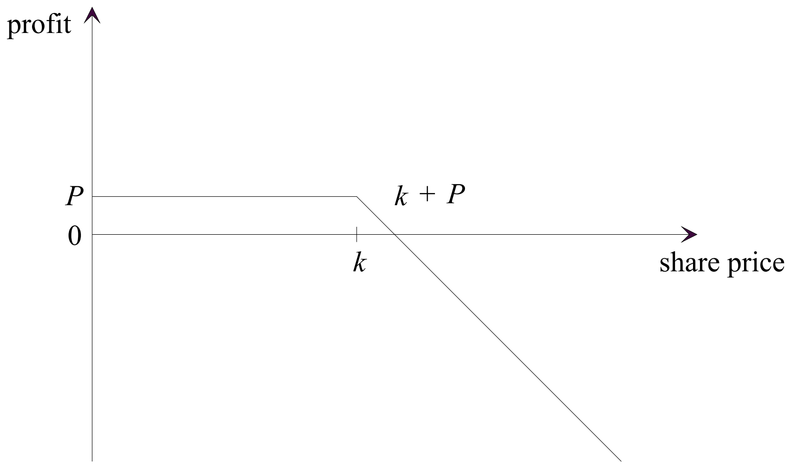
(ii) ***What opinion?***

The speculator would have set up this portfolio if he or she believed the share price at the expiry date would have moved an unexpectedly long way from the strike price k, but this movement could be in either direction. This situation can arise when the market is waiting for news (*eg* an announcement of the company's latest results or a court ruling) that will have a definite effect on the share price, but no one knows which way it will go. [2]

Solution 3.3

Having a short holding of one call option means that the speculator has sold a call option to someone else (when he didn't actually own any). The speculator will therefore have received an initial premium ($=P$, say) and will have to deliver one share to the other party in exchange for a payment equal to the strike price ($=k$, say) on the expiry date if the other party elects to exercise the option. [2]

The position diagram looks like this:



[2]
[Total 4]

Solution 3.4

(i) ***Difference between a European and an American option***

A "European" option allows the holder to exercise the option at the end of the life of the contract only, whereas an "American" option allows the holder to exercise the option at any time during the life of the contract. [1]

(ii) ***Explain why it is never optimal (when there is an active trading market)***

The price at which the option trades on the market will be the sum of the intrinsic value and the time value. [½]

If the option is exercised early, the holder will receive the intrinsic value, and will receive no other benefit before the expiry date (since there are no dividends). [½]

So it would be better instead for the holder to sell the option in the market. [1]
[Total 2]

(iii) ***Explain why it is never optimal (when the option cannot be traded)***

If the holder of an American option exercises the option early he or she will receive some underlying stock (in exchange for cash) at that time. Since we are assuming that the investor does not sell the underlying stock before the original expiry date, he or she will still be just holding the underlying stock at the expiry date. [1]

However, this investor could have done better by waiting till the expiry date and then using the call option, allowing the cash required (which is a constant amount) to earn interest in the meantime. [1]

This second strategy also allows the investor to benefit if the share price happens to have fallen to below the exercise price by the expiry date. In this case the investor can allow the call option to lapse and will have avoided having his or her "fingers burnt" by a stock that has fallen in value. [1]

[Total 3]

Solution 3.5

(i) ***Call option***

The payoff for a call option is $\max(S-K,0)$. Since S cannot fall below zero but has no theoretical upper limit, the downside risk is capped but the upside potential is unlimited, *ie* the intrinsic value can increase without limit but it cannot fall below zero. This asymmetry introduces an upward bias on the value of the call option when the share price is allowed to move. [1½]

The longer the life of the option, the more opportunity there is for the share price to move and hence the value of the option increases. [½]

[Total 2]

(ii) ***Put option***

The value of a put option is also higher for a 6-month option than for a 3-month option, for a similar reason. [½]

As before, the longer the time to expiry, the greater is the chance that the share price will fall (further) below the strike price, thus increasing the intrinsic value of the option and the payoff at expiry. Whilst there is also more chance that the share price will increase further above the strike price, the upside risk is again capped and the intrinsic value of the option cannot fall below zero – no matter how high the share price goes. [1]

Thus, there is again an asymmetry that introduces an upward bias on the value of the put option when the share price is allowed to move. [1]

The longer the life of the option, the more opportunity there is for the share price to move and so the value of the option again increases. [½]
[Total 3]

In fact, with a put option, the potential payoff is also capped above – it cannot exceed K*, the payoff if the share price falls to zero. However, in normal circumstances, this effect is of lesser importance.*

Solution 3.6

(i) ***Show that the price satisfies the inequality***

Consider the following portfolios set up at the current time t:

Portfolio A	*Portfolio B*
Cash of $Ke^{-r(T-t)}$	1 European put option 1 share

Portfolio A will be worth K at the expiry time T, in any event.

Now consider the value of Portfolio B at the expiry time T. If the share is below the strike price, we can exercise the put (selling the share in the process) and obtain an amount K. If the share is above the strike price, we have a share worth more than K.

So, either way, Portfolio B is worth at least K.

We can show this mathematically by writing down the portfolio payoff, which is:

$$\max(K - S_T, 0) + S_T = \max(K, S_T) \geq K$$

Hence, at the expiry date, Portfolio B is worth at least as much as Portfolio A.

It follows (from the principle of no-arbitrage) that this must also be true at all earlier times t. In symbols:

$$p_t + S_t \geq Ke^{-r(T-t)}$$

ie $$p_t \geq Ke^{-r(T-t)} - S_t$$ [4]

(ii) ***Modifying the inequality***

Consider an alternative Portfolio B* consisting of 1 put option and $\frac{1}{1.02}$ shares. The (cash) dividend payable on these shares at time T will be worth $0.02\times\frac{1}{1.02}$ times the share price at that time, which would allow us to buy an extra $\frac{0.02}{1.02}$ shares. So we would end up with $\frac{1}{1.02}+\frac{0.02}{1.02}=1$ share and 1 put option in our portfolio.

By the same argument as before, this is worth at least as much as Portfolio A. So we now have the inequality:

$$p_t+\frac{1}{1.02}S_t \geq Ke^{-r(T-t)}$$

ie $$p_t \geq Ke^{-r(T-t)}-\frac{1}{1.02}S_t$$ [2]

Solution 3.7

Recall that the forward price, K, is the price that the investor agrees to pay for one unit of the underlying asset at time T.

If we hold $(1+D)^{-(T-t)}$ units of the asset at time t, and reinvest the dividend income as it is received, then at time T we would hold 1 unit of the asset.

Consider the following two portfolios, set up at time t:

Portfolio A: A forward contract to buy one unit of the asset at time T for forward price K; simultaneously invest an amount $Ke^{-r(T-t)}$ in the risk-free investment.

Portfolio B: Buy $(1+D)^{-(T-t)}$ units of the asset, reinvesting the dividend income in the asset immediately as it is received.

[2]

At time T, the risk-free investment in Portfolio A has grown to amount K, which is used to buy one unit of the asset using the forward contract.

At time T, the amount of the asset in Portfolio B has grown with the reinvested dividend income to one unit.

So, the outcome of both of these portfolios is that one unit of the underlying asset is held at time T. [1]

Assuming no arbitrage, the value of these portfolios must therefore also be equal at time t. [1]

The cost of setting up Portfolio A is $Ke^{-r(T-t)}$. [½]

The cost of setting up Portfolio B is $S_t(1+D)^{-(T-t)}$ [½]

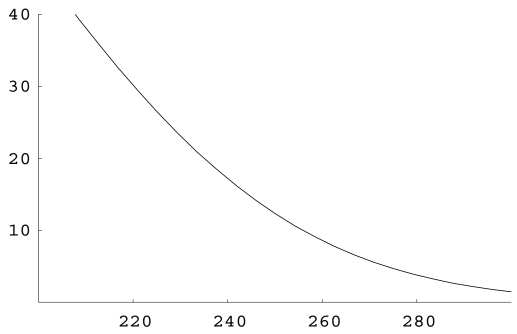
Equating these:

$$Ke^{-r(T-t)} = S_t(1+D)^{-(T-t)} \quad \Rightarrow \quad K = S_t e^{r(T-t)}(1+D)^{-(T-t)}$$ [1]

[Total 6]

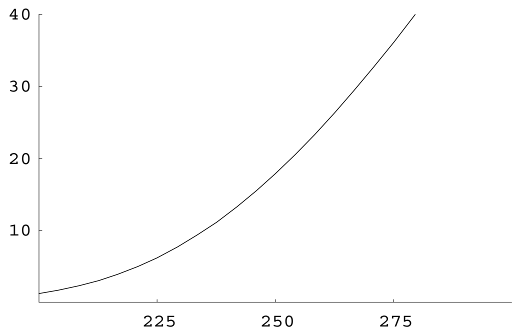
Solution 3.8

Graph 1



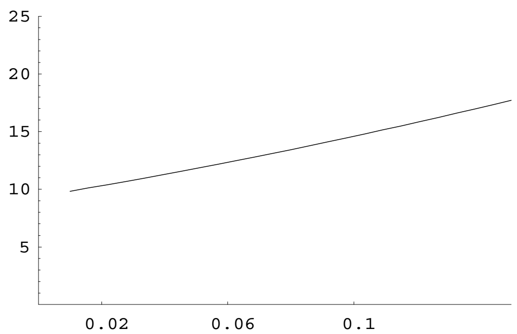
This is a graph of the option price against the strike price K.

Graph 2



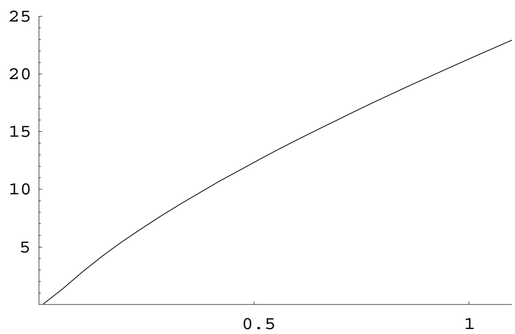
This is a graph of the option price against the current share price S.

Graph 3



This is a graph of the option price against the risk-free interest rate r.

Graph 4



This is a graph of the option price against the time to expiry T.

The clues to look for were:

- The value of a call option decreases as the strike price increases, and Graph 1 is the only graph that goes down.
- Since the payoff for a call option depends on the difference $S-K$, the graph against S should be similar to the graph against K, but reflected in the vertical axis (because increasing K by 10 would have a similar effect to reducing S by 10). So this is Graph 2.
- If $T=0$, this call option would be at its maturity date and it would be out-of-the-money with no value (since $S<K$). So the graph against T must go through the origin, as in Graph 4.
- The value of a call option increases with the interest rate, but would not have zero value when $r=0$. This is consistent with Graph 3.

[1 for each correct answer, total 4]

Solution 3.9

This is Question 8 from the Subject 109 exam of April 2000.

(i) ***Define delta, gamma and theta***

Δ measures the sensitivity of the price f of a derivative to changes in the price S of the underlying asset:

$$\Delta = \frac{\partial f}{\partial S} \qquad [1]$$

Γ measures the sensitivity of the Δ of a derivative to changes in the price S of the underlying asset:

$$\Gamma = \frac{\partial \Delta}{\partial S} = \frac{\partial^2 f}{\partial S^2} \qquad [1]$$

Θ measures the sensitivity of the price f of a derivative to changes in time:

$$\Theta = \frac{\partial f}{\partial t} \qquad [1]$$

[Total 3]

(ii) ***Describe delta hedging and give a numerical example***

Delta hedging involves creating a portfolio consisting of a holding of a derivative and the underlying asset, so that the delta for the portfolio is zero. [1]

This means that the value of the portfolio will not change if the price of the underlying changes by a small amount (all other factors remaining unchanged). [1]

The delta of a portfolio can be calculated as a weighted sum of the deltas of the constituents of the portfolio. [1]

Suppose, for example, that an institution has sold 1,000,000 call options on a share, each with a delta of 0.5. The delta for this portfolio would be $-0.5 \times 1,000,000 = -500,000$, and so this portfolio is not delta-hedged. [1]

If, however, the institution also purchased 500,000 shares, the portfolio would now be delta-hedged, because the shares themselves have a delta of 1, so that the delta for the portfolio would be $1,000,000 \times (-0.5) + 500,000 \times 1 = 0$. A small change in the value of the shares would now make no difference to the value of the portfolio as a whole. [2]

[Total 6]

Solution 3.10

Let f denote the value of the part of the portfolio containing the relevant shares and derivatives on those shares. [½]

Let S denote the share price. [½]

Let σ denote the market volatility of the share price, *ie* the annualised standard deviation of the increments in the log-share price. [½]

Let t denote calendar time. [½]

(a) ***Hedge fund manager***

For a delta-neutral position, the hedge fund manager will want to have an overall delta of zero:

$$\Delta = \frac{\partial f}{\partial S} = 0 \qquad [1]$$

To minimise the need for rebalancing to maintain a delta-neutral position, he will also want to have a low gamma:

$$\Gamma = \frac{\partial^2 f}{\partial S^2} \approx 0 \qquad [1]$$

(b) ***Derivatives trader***

The derivatives trader will be primarily concerned about the volatility. If he doesn't want changes in the volatility to affect the value of the options, he will want to have a vega close to zero:

$$V = \frac{\partial f}{\partial \sigma} \approx 0 \qquad [1]$$

Note that the question refers to a change in the distribution of the price movements, not the values themselves.

(c) ***Pension fund trustee***

The pension fund trustee will be primarily concerned about the effect of calendar time on the value of the options. If he wants to avoid the fund value falling, he will prefer to have a non-negative theta:

$$\Theta = \frac{\partial f}{\partial t} \geq 0$$ [1]

[Total 6]

Solution 3.11

Let there be x shares and y units of cash in the hedging portfolio at time t.

The value of the hedging portfolio must be equal to the value of the call option at time t. This gives:

$$240x + y = 20.15 \qquad (1)$$ [1]

The delta of the hedging portfolio must be equal to the delta of the call option at time t. This gives:

$$x \times \Delta_{share} + y \times \Delta_{cash} = 0.558$$ [1]

Now:

$$\Delta_{share} = \frac{\partial S_t}{\partial S_t} = 1 \quad \text{and} \quad \Delta_{cash} = 0$$

giving $x = 0.558$. [1]

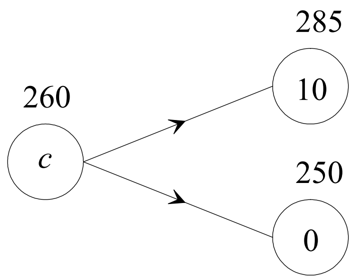
Using equation (1), this gives $y = -113.77$. [1]

So the hedging portfolio consists of 0.558 shares and cash of -113.77p. [Total 4]

Solution 3.12

(i)(a) ***Price of call option using a replicating portfolio method***

The diagram for this call option (with the option values / payoffs inside the circles and the share prices on the top) is:



A replicating portfolio, consisting of ϕ_1 shares and ψ_1 invested in cash, is set up at $t=0$ to give the same payoff as the derivative at $t=1$, irrespective of whether the share price increases or decreases over the year.

If the share price increases:

$$285\phi_1 + \psi_1 e^{0.05} = 10 \qquad (1)$$

and if the share price decreases:

$$250\phi_1 + \psi_1 e^{0.05} = 0 \qquad (2)$$

Solving equations (1) and (2) simultaneously gives:

$$\phi_1 = \tfrac{2}{7} \quad \text{and} \quad \psi_1 = -250 \times \tfrac{2}{7} \times e^{-0.05} = -67.945$$ [1]

This portfolio gives the same payoff as the call option at $t=1$, so assuming no arbitrage, the value of the call option at $t=0$, c, must equal the value of the portfolio at $t=0$. So:

$$c = 260\phi_1 + \psi_1 = 6.34$$ [1]

(i)(b) ***Price of call option using a risk-neutral valuation***

The risk-neutral probability of an up-movement is:

$$q = \frac{e^{0.05} - 250/260}{285/260 - 250/260} = 0.66659 \qquad [1]$$

So the value of the call option is:

$$c = \left[10 \times 0.66659 + 0 \times (1 - 0.66659)\right] e^{-0.05}$$
$$= 6.34 \qquad [1]$$

(i)(c) ***Price of call option using a risk-free portfolio method***

A risk-free portfolio, consisting of 1 call option and x shares, is set up at $t = 0$ so that the value of the portfolio at $t = 1$ will be the same irrespective of whether the share price increases or decreases over the year.

Considering the possibilities at $t = 1$, this means that:

$$10 + 285x = 0 + 250x \quad \Rightarrow \quad x = -\tfrac{2}{7} \qquad [½]$$

The value of this portfolio at $t = 1$ is:

$$10 + 285\left(-\tfrac{2}{7}\right) = 0 + 250\left(-\tfrac{2}{7}\right) = -\tfrac{500}{7} \qquad [½]$$

This portfolio has been set up to be risk-free (it has the same value at $t = 1$ no matter what happens to the share price), so we can use the risk-free force of interest to calculate its value at $t = 0$:

$$-\tfrac{500}{7} e^{-0.05}$$

Since this is the cost of setting up the risk-free portfolio at $t = 0$:

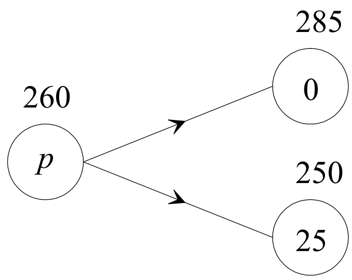
$$c + 260\left(-\tfrac{2}{7}\right) = -\tfrac{500}{7} e^{-0.05} \quad \Rightarrow \quad c = 6.34 \qquad [1]$$

[Total 6]

Note that we obtain the same value of the call option, whichever approach we take.

(ii)(a) ***Price of put option using a replicating portfolio method***

The diagram for this put option (with the option values / payoffs inside the circles and the share prices on the top) is:



A replicating portfolio, consisting of ϕ_2 shares and ψ_2 invested in cash, is set up at $t=0$ to give the same payoff as the derivative at $t=1$, irrespective of whether the share price increases or decreases over the year.

If the share price increases:

$$285\phi_2 + \psi_2 e^{0.05} = 0 \qquad (3)$$

and if the share price decreases:

$$250\phi_2 + \psi_2 e^{0.05} = 25 \qquad (4)$$

Solving equations (3) and (4) simultaneously gives:

$$\phi_2 = -\tfrac{5}{7} \quad \text{and} \quad \psi_2 = 285 \times \tfrac{5}{7} \times e^{-0.05} = 193.643 \qquad [1]$$

This portfolio gives the same payoff as the put option at $t=1$, so assuming no arbitrage, the value of the put option at $t=0$, p, must equal the value of the portfolio at $t=0$. So:

$$p = 260\phi_2 + \psi_2 = 7.93 \qquad [1]$$

(ii)(b) ***Price of put option using a risk-neutral valuation***

The risk-neutral probability of an up-movement is:

$$q = \frac{e^{0.05} - 250/260}{285/260 - 250/260} = 0.66659 \quad [1]$$

This is the same as for the call option in (i).

So the value of the put option is:

$$\begin{aligned} p &= \left[0 \times 0.66659 + 25 \times (1 - 0.66659)\right] e^{-0.05} \\ &= 7.93 \end{aligned} \quad [1]$$

(ii)(c) ***Price of put option using a risk-free portfolio method***

A risk-free portfolio, consisting of 1 put option and y shares, is set up at $t = 0$ so that the value of the portfolio at $t = 1$ will be the same irrespective of whether the share price increases or decreases over the year.

Considering the possibilities at $t = 1$, this means that:

$$0 + 285y = 25 + 250y \quad \Rightarrow \quad y = \tfrac{5}{7} \quad [½]$$

The value of this portfolio at $t = 1$ is:

$$0 + 285\left(\tfrac{5}{7}\right) = 25 + 250\left(\tfrac{5}{7}\right) = \tfrac{1425}{7} \quad [½]$$

This portfolio has been set up to be risk-free (it has the same value at $t = 1$ no matter what happens to the share price), so we can use the risk-free force of interest to calculate its value at $t = 0$:

$$\tfrac{1425}{7} e^{-0.05}$$

Since this is the cost of setting up the risk-free portfolio at $t = 0$:

$$p + 260\left(\tfrac{5}{7}\right) = \tfrac{1425}{7} e^{-0.05} \quad \Rightarrow \quad p = 7.93 \quad [1]$$

[Total 6]

Note that we obtain the same value of the put option, whichever approach we take.

(iii) ***Verify put-call parity***

The put-call parity relationship states that:

value of put + share price = value of call + discounted strike price [½]

The LHS is: $7.93 + 260 = 267.93$ [½]

The RHS is: $6.34 + 275e^{-0.05} = 6.34 + 261.59 = 267.93$ [½]

Since these are equal, the put-call parity relationship holds in this case. [½]
[Total 2]

Solution 3.13

Suppose the portfolio consists of Δ securities and an initial amount of cash ψ.

If the security price moves up, the value of the portfolio will be:

$$\psi e^r + \Delta Su = \psi e^r + \frac{z_u - z_d}{u - d}u$$ [1]

We want this to equal the payoff following an upward movement in the security price:

ie $$\psi e^r + \frac{z_u - z_d}{u - d}u = z_u$$

So: $$\psi = e^{-r}\left(z_u - \frac{z_u - z_d}{u - d}u\right) = e^{-r}\left(\frac{uz_d - dz_u}{u - d}\right)$$ [2]

If the security price moves down, the value of the portfolio will be:

$$\psi e^r + \Delta Sd = \psi e^r + \frac{z_u - z_d}{u - d}d = \left(\frac{uz_d - dz_u}{u - d}\right) + \frac{z_u - z_d}{u - d}d = z_d$$ [1]

So, with an initial cash holding of $\psi = e^{-r}\left(\frac{uz_d - dz_u}{u - d}\right)$, the portfolio replicates the derivative payoff, irrespective of the actual price movement. [1]
[Total 5]

Note that Δ here is calculated as the proportionate change in the derivative price relative to the price of the underlying security. This corresponds to the definition of the "Greek" delta, namely $\Delta = \dfrac{\partial V}{\partial S}$.

Solution 3.14

(i) ***Values of u and d***

Equating the mean and variance of the returns gives:

$$\tfrac{2}{3}u + \tfrac{1}{3}d = 1.1 \quad \Rightarrow d = 3.3 - 2u$$ [1]

and $$\tfrac{2}{3}u^2 + \tfrac{1}{3}d^2 - 1.1^2 = 0.1^2 \quad \Rightarrow 2u^2 + d^2 = 3.66$$ [1]

Eliminating d from these simultaneous equations gives:

$$2u^2 + (3.3 - 2u)^2 = 3.66$$

ie $$6u^2 - 13.2u + 7.23 = 0$$

Solving this using the quadratic formula gives:

$$u = \frac{13.2 \pm \sqrt{13.2^2 - 4(6)(7.23)}}{2(6)} = \frac{13.2 \pm \sqrt{0.72}}{12}$$ [1]

So:

$$u = 1.17071 \text{ and } d = 0.95858$$ [¼]

or:

$$u = 1.02929 \text{ and } d = 1.24142$$ [¼]

Since we need $u > d$, we can eliminate the second pair of values and conclude that the appropriate parameter values are $u = 1.17071$ and $d = 0.95858$. [½]

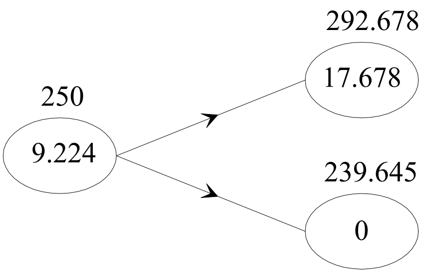
[Total 4]

(ii)(a) ***Call option***

The risk-neutral probability of an up-movement is:

$$q = \frac{e^{0.075} - 0.95858}{1.17071 - 0.95858} = 0.56241 \qquad [1]$$

The tree diagram for the call option looks like this:

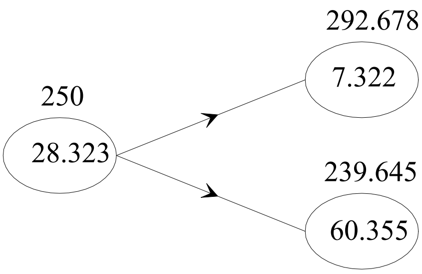


So the value of the call option is:

$$(0.56241 \times 17.678 + 0)\, e^{-0.075} = 9.224 \qquad [3]$$

(ii)(b) ***Put option***

The tree diagram for the put option looks like this:



So the value of the put option is:

$$(0.56241 \times 7.322 + 0.43759 \times 60.355)\, e^{-0.075} = 28.323 \qquad [3]$$

[Total 7]

Solution 3.15

(i) ***Risk-neutral probability***

Let u and d be the assumed proportionate changes in the price of the underlying if it goes up and down respectively and let r be the risk-free interest rate (continuously-compounded). δt is the length of the one-step time period. [2]

Let S be the current price of the share.

If q and $1-q$ are the risk-neutral probabilities for the tree, the expected final value of the share should be the same as if it had been invested in risk-free cash. [1]

So we need:

$$qSu+(1-q)Sd = Se^{r\delta t} \qquad [1]$$

Cancelling the S 's gives:

$$qu+(1-q)d = e^{r\delta t}$$

Rearranging gives:

$$q(u-d)+d = e^{r\delta t} \quad \Rightarrow q = \frac{e^{r\delta t}-d}{u-d} \qquad [1]$$

[Total 5]

(ii) ***Explain the inequality***

Our derivation in part (i) required the assumption that the market was arbitrage-free. This requires $d < e^{r\delta t} < u$. Otherwise we could make a guaranteed profit. [1]

For example, if $d < u < e^{r\delta t}$, the cash investment would outperform the share in all circumstances. So we could make a guaranteed profit by selling the share at the start and investing the proceeds in cash. When we buy back the share at the end, we would have a positive profit of either $Se^{r\delta t} - Su$ or $Se^{r\delta t} - Sd$. [1]

Alternatively, if this inequality did not hold, the value of q calculated using the formula derived in part (i) would lie outside the range $[0,1]$ and hence could not represent a probability. [2]

[Maximum 2]

(iii)(a) ***Formula for θ***

Let R be the one-step return on the share. Then

$$1+R=\begin{cases} u & \text{with probability } p \\ d & \text{with probability } 1-p \end{cases}$$

So:

$$E(1+R)=pu+(1-p)d \quad \Rightarrow \quad \theta=E(R)=pu+(1-p)d-1 \qquad [1]$$

(iii)(b) ***Real-world variance, σ^2***

Similarly, the variance of the rate of return is:

$$\begin{aligned}\sigma^2 &= \text{var}(R)\\ &= \text{var}(1+R)\\ &= E\left[(1+R)^2\right]-\left(E[1+R]\right)^2\\ &= pu^2+(1-p)d^2-[pu+(1-p)d]^2\end{aligned}$$

If we collect the p's together, expand and simplify, we get:

$$\begin{aligned}\sigma^2 &= p(u^2-d^2)+d^2-[p(u-d)+d]^2\\ &= p(u^2-d^2)-p^2(u-d)^2-2pd(u-d)\\ &= p(u-d)[(u+d)-p(u-d)-2d]\\ &= p(u-d)[(u-d)-p(u-d)]\\ &= p(u-d)^2[1-p]\\ &= p(1-p)(u-d)^2 \qquad [3]\end{aligned}$$

[Total 4]

(iv) ***Show the inequality***

Since $u > d$ by definition, the inequality $p > q$ is equivalent to:

$$p(u-d)+d > q(u-d)+d$$

ie $$1+\theta > \frac{e^{r\delta t}-d}{u-d}(u-d)+d = e^{r\delta t}-d+d = e^{r\delta t}$$ [2]

In words, this says that the real-world probability of an up movement is greater than the risk-neutral probability whenever the expected increase in the value of the underlying security exceeds the risk-free interest rate *ie* whenever the underlying security is risky.

[2]
[Total 4]

Solution 3.16

(i) ***Value the put option using a risk-free portfolio approach***

A risk-free portfolio, consisting of 1 put option and n shares is set up at time 0, so that the value at time 1 month will be the same irrespective of whether the share price increases or decreases over the month.

The option will be worth \$0.50 if the share price falls to \$1.50. It will be worthless if the share price rises to \$2.50. So, considering the possibilities at time 1 month, we have:

$$50+150\times n = 0+250\times n \Rightarrow n=\tfrac{1}{2}$$

With this value of n, both sides of this equation equal 125 *ie* the portfolio would be worth \$1.25 whichever way the share price moves. As this portfolio has been set up to be risk-free, its value at time 0 satisfies:

$$p+219\times\tfrac{1}{2} = 125e^{-0.01} \Rightarrow p=14.26$$

So the value of the put option is \$0.14. [4]

(ii) ***Value the put option using risk-neutral valuation***

The risk-neutral probability q of an upward movement in the share price can be determined from the equation:

$$q \times 250 + (1-q) \times 150 = 221e^{0.01}$$

$$\Rightarrow \qquad q = \frac{221e^{0.01} - 150}{250 - 150} = 0.73221$$

So the value of the option is:

$$p' = e^{-0.01}[q \times 0 + (1-q) \times 50] = 13.26 \qquad [2]$$

(iii) ***Estimate the delta***

Using the results from parts (i) and (ii), the option's delta is therefore approximately:

$$\Delta = \frac{\partial f}{\partial S} \approx \frac{13.26 - 14.26}{221 - 219} = -\tfrac{1}{2} \qquad [1]$$

This is the value we would expect for delta since creating a delta-hedged position, as we did in part (i) where we found that $n = ½$, involves purchasing or selling Δ shares.

Solution 3.17

(i) ***The risk-free force of interest r***

Under the risk-neutral probability measure we have:

$$E_Q[S_1] = S_0 e^r$$

$$\Leftrightarrow \qquad q \times 1.1S_0 + (1-q) \times 0.9S_0 = S_0 e^r \qquad [½]$$

Dividing through by S_0, substituting in the given value of q and solving for r gives:

$$r = \log(0.55 \times 1.1 + 0.45 \times 0.9)$$
$$= \log(1.01) = 0.995\% \qquad [½]$$

[Total 1]

(ii) ***State price deflators***

The state price deflators for each of the three final possible states can be derived from the formulae:

$$A_2(1) = e^{-2r}\frac{q^2}{p^2} \qquad A_2(2) = e^{-2r}\frac{2q(1-q)}{2p(1-p)} \qquad A_2(3) = e^{-2r}\frac{(1-q)^2}{(1-p)^2} \qquad [1]$$

We can use the annualised expected force of return on the share, μ, to calculate the real-world probability of an up-step p:

$$E_P[S_1] = S_0 e^{\mu/12}$$

$$\Leftrightarrow \qquad p\times 1.1 + (1-p)\times 0.9 = e^{\mu/12} = e^{0.3/12}$$

$$\Leftrightarrow \qquad p = \frac{e^{0.3/12}-0.9}{1.1-0.9} = 0.62658 \qquad [1]$$

Using this value of p, the state price deflators are:

$$A_2(1) = \frac{1}{1.01^2}\frac{0.55^2}{0.62658^2} = 0.75533$$

$$A_2(2) = \frac{1}{1.01^2}\frac{2\times 0.55\times 0.45}{2\times 0.62658\times 0.37342} = 1.03695$$

$$A_2(3) = \frac{1}{1.01^2}\frac{0.45^2}{0.37342^2} = 1.42356 \qquad [2]$$

[Total 4]

(iii)(a) ***Payoff*** $\boldsymbol{(1,0,0)}$

$$e^{-2r}q^2 = \frac{0.55^2}{1.01^2} = 0.2965 \qquad [½]$$

(iii)(b) ***Payoff*** $\boldsymbol{(0,1,0)}$

$$e^{-2r}2q(1-q) = \frac{2\times 0.55\times 0.45}{1.01^2} = 0.4852 \qquad [½]$$

(iii)(c) ***Payoff*** $\boldsymbol{(0,0,1)}$

$$e^{-2r}(1-q)^2 = \frac{0.45^2}{1.01^2} = 0.1985 \qquad [½]$$

Note that the values calculated in parts (a), (b) and (c) are known as "state prices". We can use them to calculate the value of more complicated derivatives as we now do for the rest of part (iii).

We could alternatively have calculated the values in (a), (b) and (c) using:

$$A_2(1) \times p^2 \ , \ A_2(2) \times 2p(1-p) \ , \ A_2(3) \times (1-p)^2$$

(iii)(d) ***Call option, strike price K = 0.95***

Depending on the final price of the share, the payoff for this option will be:

$\max(1 \times 1.1^2 - 0.95, 0) = 0.26$,
$\max(1 \times 1.1 \times 0.9 - 0.95, 0) = 0.04$, or
$\max(1 \times 0.9^2 - 0.95, 0) = 0$.

So, the value is:

$$0.26 \times 0.2965 + 0.04 \times 0.4853 = 0.0965 \qquad [1]$$

(iii)(e) ***Put option, strike price K = 1.05***

Depending on the final price of the share, the payoff for this option will be:

$\max(1.05 - 1 \times 1.1^2, 0) = 0$,
$\max(1.05 - 1 \times 1.1 \times 0.9, 0) = 0.06$, or
$\max(1.05 - 1 \times 0.9^2, 0) = 0.24$.

So, the value is:

$$0.06 \times 0.4853 + 0.24 \times 0.1985 = 0.07676 \qquad [1]$$

(iii)(f) ***Payoff*** $\boldsymbol{2\times|S-0.98|}$

Depending on the final price of the share, the payoff for this option will be:

$$2\times\left|1\times1.1^2-0.98\right|=0.46\,,$$
$$2\times\left|1\times1.1\times0.9-0.98\right|=0.02\text{ , or}$$
$$2\times\left|1\times0.9^2-0.98\right|=0.34\,.$$

So, the value is:

$$0.46\times0.2965+0.02\times0.4853+0.34\times0.1985=0.2136$$ [1½]
[Total 5]

Solution 3.18

(i) ***Option value***

Without loss of generality, assume that the current price of gold (and the exercise price of the option, since the option is at-the-money) is 100. We can then find the risk-neutral probability (q) of an upward movement in the price of gold from:

$$110q+80(1-q)=102 \quad\Rightarrow\quad q=0.7333$$ [1]

Now consider what might happen over the two months to the exercise date:

- If the gold price goes up in each of the two months, then the option's payoff is:

  $$\max\left\{100-(100\times1.1\times1.1),0\right\}=0$$ [½]

- If the gold price goes up in one month and down in the other month, then the option's payoff is:

  $$\max\left\{100-(100\times1.1\times0.8),0\right\}=12$$ [½]

- If the gold price goes down in each of the two months, then the option's payoff is:

  $$\max\left\{100-(100\times0.8\times0.8),0\right\}=36$$ [½]

Thus the value of the option at time 0 is given by:

$$V_0 = (1.02)^{-2}\left(0 \times q^2 + 12 \times 2q(1-q) + 36 \times (1-q)^2\right) = 6.972 \qquad [1]$$

ie the price of an at-the-money European style put option on gold, expressed as a percentage of the current price of gold, is 6.97%. [½]

[Total 4]

(ii) ***Actuary's risk discount rate***

Using the actuary's estimate of the probability an up-movement of 0.8, the probability that the gold price:

- goes up both months, and hence the payoff is equal to 0, is 0.8^2
- goes up once and down once, and hence the payoff is equal to 12, is $2 \times 0.8 \times (1-0.8)$
- goes down both months, and hence the payoff is equal to 36, is $(1-0.8)^2$.

Thus, the expected payoff at the exercise date is:

$$0 + 12 \times 2 \times 0.8 \times (1-0.8) + 36 \times (1-0.8)^2 = 5.28 \qquad [1]$$

We can therefore find the risk discount rate by solving the equation of value:

$$6.97 = \frac{5.28}{(1+i)^2} \quad \Rightarrow \quad i = -13\% \text{ per month} \qquad [1]$$

[Total 2]

(iii) ***Why the actuary's probability is plausible***

Gold is risky and in particular, the total risk will include an element of systematic risk, that cannot be diversified away. [½]

Therefore, investors in gold should expect a rate of return higher than the risk-free rate. [½]

For gold to give a higher expected return than a risk-free asset we must have:

$$110p + 80(1-p) > 102$$

where p is the "true" probability of an up-movement. [½]

Hence, comparing this inequality with the equality we used to derive the risk-neutral probability in (i) above, it must be the case that:

$$110p + 80(1-p) > 110q + 80(1-q) \quad \Rightarrow \quad p > q$$ [1]

The actuary's choice of $p = 0.8$ satisfies this condition. It is therefore plausible. [½]
[Total 3]

(iv) ***Justification of risk discount rate***

The CAPM gives the following relationship:

$$E_i = r + \beta_i (E_M - r)$$

where β_i is a measure of the correlation (or covariance) of the return on asset i with the return on the market for all risky assets. [½]

If we assume that gold is positively correlated with the market for all risky assets, a put option on gold will have negative beta. This is because an increase in the price of gold would lead to a reduction in the value of the put option. [1]

An asset with a negative beta will have an expected return of less than the risk-free rate.
[½]
[Total 2]

Solution 3.19

(i) ***Explain why the assumption might not be valid***

This means that S_t, the share price at time t, can be considered to be a random variable that obeys the stochastic differential equation:

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

where Z_t represents a standard Brownian motion.

Another way of expressing this is to say that the distribution of $\log \dfrac{S_t}{S_0}$ is $N\left((\mu - \tfrac{1}{2}\sigma^2)t, \sigma^2 t\right)$ and that movements in S_t in non-overlapping time intervals are statistically independent. [3 for either formulation]

This will not be realistic in practice for any underlying asset that experiences sudden changes (discontinuous jumps) in price or whose drift or volatility is not constant. [1]

[Total 4]

(ii) ***Other assumptions***

The other Black-Scholes assumptions are:

- There are no risk-free arbitrage opportunities.
- The risk-free interest rate is constant, *ie* the same for all maturity dates and the same for borrowing or lending.
- Unlimited short selling (*ie* having a negative holding) is allowed.
- There are no taxes or transaction costs.
- The underlying asset can be traded continuously and in arbitrarily small amounts.

[1 each, total 5]

Solution 3.20

(i) ***State the context, and indicate what s and f represent***

This PDE describes the relationship between the theoretical price f of a derivative and the current price s of the underlying stock on whose value the derivative payoff depends. [2]

(ii) ***What do r and σ represent?***

r represents the force of interest earned on a risk-free asset (usually expressed as an annualised rate). [1]

σ represents the volatility of the underlying stock *ie* the standard deviation of the log of the price ratio movements (again usually expressed as an annualised rate). [1]

This PDE assumes that r and σ are constant. [1]

[Total 3]

(iii) ***Names of the Greeks***

Theta is defined as $\theta = \dfrac{\partial f}{\partial t}$. [1]

Delta is defined as $\Delta = \dfrac{\partial f}{\partial s}$. [1]

Gamma is defined as $\Gamma = \dfrac{\partial^2 f}{\partial s^2}$. [1]

[Total 3]

(iv) ***Boundary condition***

The boundary condition would specify the payoff at the end of the contract (time T):

$$f(T,s) = \max(s-K,0)$$ [1]

Solution 3.21

This is part of Question 5 from the Subject 109 exam of April 2001.

(i) ***Calculate the price of a put option with strike price 5,250***

The put-call parity relationship states that:

$$c_t + Ke^{-r(T-t)} = p_t + S_t$$ [1]

Substituting in the values given:

$$187.06 + 5{,}250e^{-0.05\times½} = p_t + 5{,}000$$

$\Rightarrow \quad p_t = 307.44$ [1]

[Total 2]

(ii) ***Calculate the price of a put option with strike price 4,750***

We first need to estimate the implied volatility of the stock using the information given in the question. This can be done by trial and improvement. We start by substituting the parameter values into the Black-Scholes formula:

$$c_t = 5{,}000\Phi(d_1) - 5{,}250e^{-0.05\times½}\Phi(d_2)$$

where $d_1 = \dfrac{\log\left(\dfrac{5{,}000}{5{,}250}\right) + \left(0.05 + ½\sigma^2\right)\times ½}{\sigma\sqrt{½}}$ and $d_2 = d_1 - \sigma\sqrt{½}$ [1]

In most instances, $\sigma = 0.2$ is a reasonable starting point for the interpolation. If $\sigma = 0.2$, then substituting both this value and the other parameter values into the Black-Scholes formula gives:

$$d_1 = -0.0975 \text{ and } d_2 = -0.2389 \quad [½]$$

$$\Rightarrow \quad c_t = 5{,}000\Phi(-0.0975) - 5{,}250e^{-0.05\times½}\Phi(-0.2389) = 228.18 \quad [1]$$

This is above the actual price of 187.06, so we need to try a lower value of σ. If we try $\sigma = 0.1$, then we obtain:

$$d_1 = -0.3011 \text{ and } d_2 = -0.3718 \quad [½]$$

$$\Rightarrow \quad c_t = 5{,}000\Phi(-0.3011) - 5{,}250e^{-0.05\times½}\Phi(-0.3718) = 91.27 \quad [1]$$

As the two call option prices straddle the actual price of 187.06, we can interpolate between the two values of σ to obtain an estimate for the implied volatility:

$$\frac{187.06 - 91.27}{228.18 - 91.27} \approx \frac{\sigma - 0.1}{0.2 - 0.1} \quad \Rightarrow \quad \sigma \simeq 17\% \quad [1]$$

We can now use this estimate of σ to determine the price of a put option with a strike price of 4,750:

$$d_1 = \frac{\log\left(\dfrac{5{,}000}{4{,}750}\right) + \left(0.05 + ½\times 0.17^2\right)\times ½}{0.17\times\sqrt{½}} = 0.6948 \quad [½]$$

and:

$$d_2 = d_1 - 0.17 \times \sqrt{½} = 0.5746 \qquad [½]$$

So the price of the 4750 put option is:

$$p_t = 4,750e^{-0.05 \times ½}\Phi(-0.5746) - 5,000\Phi(-0.6948) = 92.11 \qquad [1]$$

[Total 7]

Solution 3.22

(i) ***Approximate formula***

The Black-Scholes formula (which is the Garman-Kohlhagen formula on page 47 of the *Tables* with $q = 0$) for the price of a T-year European call option at time 0 is:

$$c_0 = S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2)$$

The option is "at-the-money-forward", which means that $S_0 = Ke^{-rT}$. [½]

So: $c_0 = S_0\left[\Phi(d_1) - \Phi(d_2)\right]$ [½]

Since $\Phi(x)$ is a differentiable function with $\Phi'(x) = \phi(x) = \dfrac{1}{\sqrt{2\pi}}e^{-½x^2}$ (from page 10 of the *Tables*), we can apply the Taylor approximation given in the hint in the question. [½]

This gives:

$$\Phi(d_1) \approx \Phi(0) + d_1\phi(0) = \frac{1}{2} + \frac{d_1}{\sqrt{2\pi}} \quad \text{and} \quad \Phi(d_2) \approx \frac{1}{2} + \frac{d_2}{\sqrt{2\pi}} \qquad [1]$$

So we find that:

$$c_0 \approx S_0\left[\left(\frac{1}{2} + \frac{d_1}{\sqrt{2\pi}}\right) - \left(\frac{1}{2} + \frac{d_2}{\sqrt{2\pi}}\right)\right] = \frac{S_0(d_1 - d_2)}{\sqrt{2\pi}} \qquad [1]$$

But since $d_2 = d_1 - \sigma\sqrt{T}$, this simplifies to:

$$c_0 \approx \frac{S_0\sigma\sqrt{T}}{\sqrt{2\pi}} \quad \text{or} \quad \frac{c_0}{S_0} \approx \frac{\sigma\sqrt{T}}{\sqrt{2\pi}} \qquad [1]$$

Since $\frac{1}{\sqrt{2\pi}} = 0.3989 \approx 0.4$, we can approximate further to get the required result:

$$\frac{c_0}{S_0} \approx 0.4\sigma\sqrt{T} \qquad [½]$$

Alternatively, you can work out the values of d_1 and d_2 directly, which are $d_1 = ½\sigma\sqrt{T}$ and $d_2 = -½\sigma\sqrt{T}$. You can then substitute into the formula for the call option price and simplify to get the required result.

[Total 5]

(ii) ***Estimate the price***

Here we have:

$$T = 0.25, \quad \sigma = 20\% \qquad [½]$$

So the approximation gives:

$$\frac{c_0}{S_0} \approx 0.4\sigma\sqrt{T} = 0.4 \times 20\% \times \sqrt{0.25} = 4\% \qquad [1]$$

So the approximate option value on shares worth £1m, will be £40,000. [½]

[Total 2]

For comparison, the accurate answer is £39,878.

Solution 3.23

(i) ***Formula for delta***

Differentiating the formula given for V with respect to S, using the product rule for the second term, we get:

$$\Delta = Ke^{-ru}\frac{\partial}{\partial S}\Phi(-d_2) - e^{-qu}\Phi(-d_1) - Se^{-qu}\frac{\partial}{\partial S}\Phi(-d_1) \qquad [1]$$

Using the function-of-a-function rule, this can be written as:

$$\Delta = -Ke^{-ru}\phi(-d_2)\frac{\partial d_2}{\partial S} - e^{-qu}\Phi(-d_1) + Se^{-qu}\phi(-d_1)\frac{\partial d_1}{\partial S} \qquad [1]$$

Here $\phi(\ )$ denotes the derivative of $\Phi(\ )$ *ie* it is the probability density function of the standard normal distribution, which is:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad [½]$$

From the definitions of d_1 and d_2, we find that:

$$\frac{\partial d_1}{\partial S} = \frac{\partial d_2}{\partial S} = \frac{1}{S\sigma\sqrt{u}} \qquad [½]$$

Putting these together and regrouping gives:

$$\Delta = \frac{-Ke^{-ru}e^{-\frac{1}{2}d_2^2} + Se^{-qu}e^{-\frac{1}{2}d_1^2}}{\sqrt{2\pi}S\sigma\sqrt{u}} - e^{-qu}\Phi(-d_1) \qquad [1]$$

In fact, the first term is zero. To show this, we add and subtract the definitions of d_1 and d_2 to get:

$$d_1 + d_2 = \frac{2\log(S/K) + 2(r-q)u}{\sigma\sqrt{u}} \qquad [½]$$

and $$d_1 - d_2 = \frac{\sigma^2 u}{\sigma\sqrt{u}} = \sigma\sqrt{u} \qquad [½]$$

Multiplying these two equations gives us a difference of two squares:

$$d_1^2 - d_2^2 = 2\log(S/K) + 2(r-q)u \qquad [1]$$

Halving gives:

$$\tfrac{1}{2}d_1^2 - \tfrac{1}{2}d_2^2 = \log(S/K) + (r-q)u$$

Exponentiating gives:

$$e^{\frac{1}{2}d_1^2}e^{-\frac{1}{2}d_2^2}=\frac{S}{K}\times e^{ru}e^{-qu}$$

Rearranging then gives:

$$Ke^{-ru}e^{-\frac{1}{2}d_2^2}=Se^{-qu}e^{-\frac{1}{2}d_1^2} \qquad [1]$$

ie the numerator in the expression for Δ is zero.

So: $\Delta=-e^{-qu}\Phi(-d_1)$ [1]

[Total 8]

It is would not be acceptable here to differentiate the expression for V ignoring the fact that d_1 and d_2 are functions of S. Although this happens to give the correct answer, it is not a valid method.

(ii) ***Formula for gamma***

Finding the second derivative, starting from the formula in (i), is more straightforward:

$$\Gamma=\frac{\partial^2 V}{\partial S^2}=\frac{\partial\Delta}{\partial S}=\frac{\partial}{\partial S}\left[-e^{-qu}\Phi(-d_1)\right] \qquad [½]$$

So:

$$\Gamma=-e^{-qu}\frac{\partial}{\partial S}\Phi(-d_1)=e^{-qu}\phi(-d_1)\frac{\partial d_1}{\partial S}$$

$$=e^{-qu}\frac{e^{-\frac{1}{2}d_1^2}}{\sqrt{2\pi}}\frac{1}{S\sigma\sqrt{u}}=\frac{e^{-\frac{1}{2}d_1^2-qu}}{S\sigma\sqrt{2\pi u}} \qquad [1½]$$

[Total 2]

Solution 3.24

(i) ***Investor's payoff***

If $S_0 = 50$, then the payoff is:

$S_1 - 35$	if $35 \le S_1 \le 45$	
10	if $45 \le S_1 \le 55$	
$65 - S_1$	if $55 \le S_1 \le 65$	
0	otherwise	[1]

Alternatively, this may be expressed:

0	if $S_1 \le 35$
$S_1 - 35$	if $35 \le S_1 \le 45$
10	if $45 \le S_1 \le 55$
$65 - S_1$	if $55 \le S_1 \le 65$
0	if $S_1 \ge 65$

(ii) ***Express the payoff in terms of call options***

The payoff from the special derivative can be replicated with the following combination of call options:

- a long call option with strike price £35 [1]
- a short call option with strike price £45 [1]
- a short call option with strike price £55 [1]
- a long call option with strike price £65 [1]

[Total 4]

With this combination of options the overall payoff is:

$$\max(S_1 - 35, 0) - \max(S_1 - 45, 0) - \max(S_1 - 55, 0) + \max(S_1 - 65, 0)$$

We can check that this combination of options does indeed give the same payoff as the special derivative, by considering the payoff for different ranges of the share price.

If $S_1 \le 35$, no option is exercised, so the payoff is 0.

If $35 \le S_1 \le 45$*, only the call option with strike price £35 is exercised, so the payoff is:*

$$S_1 - 35$$

If $45 \le S_1 \le 55$*, the call options with strike prices £35 and £45 will be exercised, giving a payoff of:*

$$S_1 - 35 - (S_1 - 45) = 10$$

If $55 \le S_1 \le 65$*, the call options with strike prices £35, £45 and £55 will be exercised, giving a payoff of:*

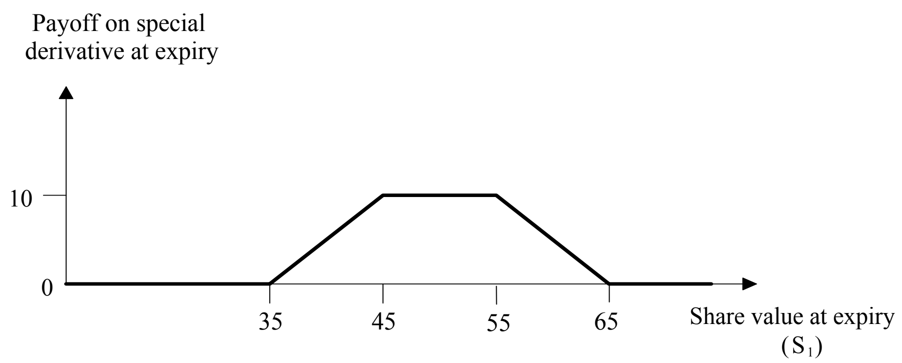
$$S_1 - 35 - (S_1 - 45) - (S_1 - 55) = 65 - S_1$$

If $S_1 \ge 65$*, all call options will be exercised, giving a payoff of:*

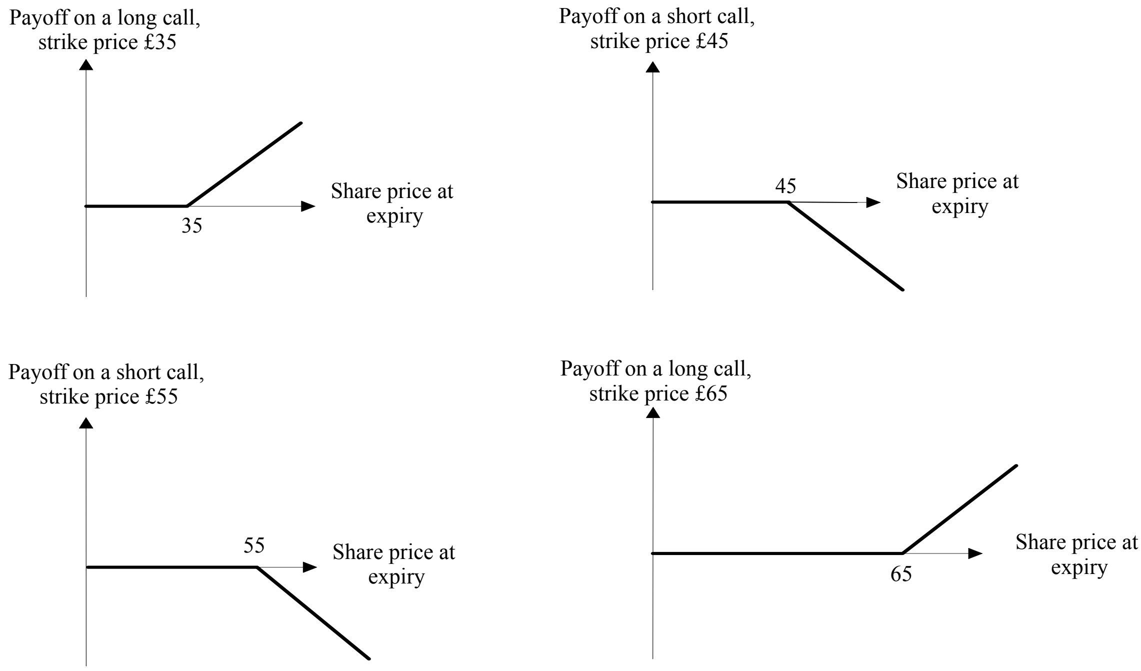
$$S_1 - 35 - (S_1 - 45) - (S_1 - 55) + S_1 - 65 = 0$$

So, in all cases, this combination of call options gives the same payoff as the special derivative.

Alternatively, it may help to consider the graph of the payoff from the special derivative:



This can be replicated by superimposing the following four payoff graphs relating to the four call options:



(iii) ***Cost of special derivative***

To work out the cost of the special derivative, we can calculate the total cost of the combination of the four call options that replicate its payoff.

In this case we have $S_0 = 50, \sigma = 15\%, r = 3\%, T - t = 1$ and $q = 0$, so the Black-Scholes formula for the price of a call option is:

$$50\Phi(d_1) - Ke^{-0.03}\Phi(d_2)$$

where:

$$d_1 = \frac{\ln\left(50/K\right) + 0.03 + \tfrac{1}{2}(0.15)^2}{0.15} \quad \text{and} \quad d_2 = d_1 - 0.15$$

For the call option with strike price £35:

$$d_1 = 2.6528 \quad \Rightarrow \quad \Phi(d_1) = 0.99601$$
$$d_2 = 2.5028 \quad \Rightarrow \quad \Phi(d_2) = 0.99384$$

So, the first call option has value:

$$c_1 = 50 \times 0.99601 - 35e^{-0.03} \times 0.99384 = £16.04$$ [1]

For the call option with strike price £45:

$$d_1 = 0.9774 \quad \Rightarrow \quad \Phi(d_1) = 0.83582$$
$$d_2 = 0.8274 \quad \Rightarrow \quad \Phi(d_2) = 0.79600$$

So, the second call option has value:

$$c_2 = 50 \times 0.83582 - 45e^{-0.03} \times 0.79600 = £7.03$$ [1]

For the call option with strike price £55:

$$d_1 = -0.3604 \quad \Rightarrow \quad \Phi(d_1) = 0.35927$$
$$d_2 = -0.5104 \quad \Rightarrow \quad \Phi(d_2) = 0.30489$$

So, the third call option has value:

$$c_3 = 50 \times 0.35927 - 55e^{-0.03} \times 0.30489 = £1.69$$ [1]

For the call option with strike price £65:

$$d_1 = -1.4741 \quad \Rightarrow \quad \Phi(d_1) = 0.07023$$
$$d_2 = -1.6241 \quad \Rightarrow \quad \Phi(d_2) = 0.05218$$

So, the fourth call option has value:

$$c_4 = 50 \times 0.07023 - 65e^{-0.03} \times 0.05218 = £0.22$$ [1]

The overall cost of the special derivative is therefore:

$$16.04 - 7.03 - 1.69 + 0.22 = £7.54$$ [1]

[Total 5]

Part 4 – Questions

Question 4.1 (Bookwork)

(i) Explain what is meant by a "complete market". [2]

(ii) Give two reasons why in practice financial markets may not be complete. [2]

(iii) State the relevance of the concept of complete markets in derivative pricing. [2]
[Total 6]

Question 4.2 (Developmental)

S_t denotes the price of a security at time t. The discounted security process $e^{-rt}S_t$, where r denotes the continuously-compounded risk-free interest rate, is a martingale under the risk-neutral measure Q.

(i) Express mathematically the fact that the discounted security process is a Q-martingale. [1]

B_t denotes the accumulated value at time t of an initial investment of 1 unit of cash.

(ii) (a) Write down an expression for B_t.

(b) Show that the discounted cash process is also a Q-martingale.

(c) Deduce that the discounted value of any self-financing portfolio (where transactions are made only by switching funds between the security and cash, with no injections or withdrawals of funds from the portfolio) will also be a Q-martingale. [4]

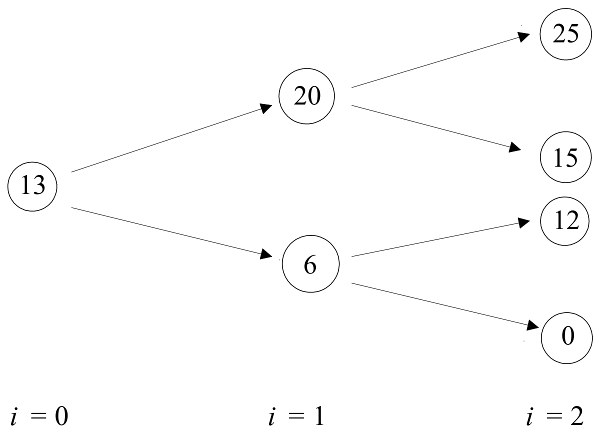
V_t is a process defined by $V_t = e^{-r(T-t)}E_Q[X \mid F_t]$, where X is a function of S_T, T is a fixed time, and F_t denotes the filtration representing the history of the security price up to time t.

(iii) Show that the discounted process $e^{-rt}V_t$ is also a Q-martingale. [3]

(iv) Explain the significance of these results in derivative pricing. [3]
[Total 11]

Question 4.3 (Developmental)

The diagram shows a two-step non-recombining binomial tree. The numerical values shown are X_i, the possible values of a particular derivative at times $i = 0,1,2$, based on a probability measure P that attributes equal probability to the two branches at each step. F_i denotes the filtration of the derivative value process at time i.



(i) (a) If $X_1 = 20$, what are the realised values of F_0 and F_1?

(b) What is the value of $E_P(X_2 \mid F_1)$ in this case?

(c) What is the value of $E_P(X_2 \mid F_1)$ in the case where $X_1 = 6$?

(d) Hence calculate $E_P[E_P(X_2 \mid F_1) \mid F_0]$.

(e) Calculate $E_P(X_2 \mid F_0)$ and comment on your answer. [6]

(ii) The risk-neutral probability measure Q attributes probabilities of 0.4 and 0.6 to the up-paths and down-paths at each branch of this tree.

(a) Does your conclusion in (i)(e) still apply when the probability measure Q is used in place of P?

(b) State briefly why this type of result is useful.

(c) Calculate the value of the derivative at time 0, presenting your calculations in the form of a tree. Ignore interest. [4]

[Total 10]

Question 4.4 (Exam-style)

The process S_t is defined by $S_t = S_0 e^{(\alpha - \frac{1}{2}\sigma^2)t + \sigma W_t}$, where W_t is standard Brownian motion under a probability measure P, and α and σ^2 are constants.

(i) (a) State the name given to the process S_t.

(b) Give two real-world quantities that are commonly modelled using such a process. [3]

(ii) (a) State what is meant by "equivalent probability measures".

(b) State how the Cameron-Martin-Girsanov Theorem could be applied here if we wished to work with a process of the form $S_t = S_0 e^{(r - \frac{1}{2}\sigma^2)t + \sigma \tilde{W}_t}$, where $\tilde{W}_t$ is standard Brownian motion and r is the risk-free rate of interest. [3]

(iii) (a) Determine the stochastic differential equation for dS_t in terms of dt and $d\tilde{W}_t$.

(b) State the drift of the process in (iii)(a) and comment on your answer. [5]

[Total 11]

Question 4.5 (Exam-style)

The random variable S_T is defined by $S_T = S_0 e^{(r-½\sigma^2)T+\sigma\tilde{Z}_T}$, where $\tilde{Z}_T \sim N(0,T)$ under a probability measure Q. The random variable X is defined by $X = \max(S_T - K, 0)$, where K is a positive constant.

(i) Show that:

$$E_Q[X|F_0] = S_0 e^{rT}\Phi\left(d_2 + \sigma\sqrt{T}\right) - K\Phi\left(d_2\right)$$

where $d_2 = \dfrac{\log(S_0/K) + (r - ½\sigma^2)T}{\sigma\sqrt{T}}$. [7]

Hint: You may wish to use the lognormal integral formulae given on page 18 of the Tables.

(ii) Explain carefully the relevance of this result in financial mathematics. [3]

[Total 10]

Question 4.6 (Exam-style)

(i) State the general risk-neutral pricing formula for the price of a derivative at time t in terms of the derivative payoff X_T at the maturity date T and the constant risk-free force of interest r. [1]

Assume that the price of a share, which pays a constant force of dividend yield q, follows geometric Brownian motion.

(ii) (a) Derive the formula for the price at time t of Derivative 1, which pays one at time T provided the share price at that time is less than K.

(b) Derive the formula for the price at time t of Derivative 2, which pays the share price at time T provided the share price at that time is less than K.

(c) Hence derive the formula for the price of a European put option with strike price K. [12]

Hint for (ii)(a) and (ii)(b): You may wish to use the lognormal integral formulae given on page 18 of the Tables.

(iii) An exotic forward contract provides a payoff equal to the value of the square of the share price at the maturity date T in return for a payment equal to the square of the forward price. Derive the formula for the value of this contract at time t.

[3]

[Total 16]

Question 4.7 (Bookwork)

Explain the similarities and differences in the following three interest rate models:

- the Hull & White model
- the Cox-Ingersoll-Ross model
- the Vasicek model [8]

Question 4.8 (Developmental)

Explain the following formulae as they are used in interest rate modelling:

(a) $B(t,T) = E_Q\left[\exp\left(-\int_t^T r(u)du\right)\middle| r(t)\right]$

(b) $B(t,T) = \dfrac{1}{A(t)} E_P[A(T) \mid F_t]$

(c) $\gamma(t,T) = \dfrac{m(t,T) - r(t)}{S(t,T)}$

[Total 9]

Question 4.9 (Developmental)

The Redington model of interest rates assumes that the force of interest takes a value that is independent of the term of the investment.

(i) Explain briefly the advantages and disadvantages of this model. [4]

An investor in zero-coupon bonds (ZCBs) uses this model for his portfolios. At time 0 he purchases holdings of a 1-year and a 3-year ZCB. The maturity values for these holdings (in units of $million) are 1 and $e^{2\delta_0}$ respectively, where δ_0 denote the force of interest applicable at time 0.

(ii) At the same time the investor wishes to short sell a 2-year ZCB. Determine the maturity value of this holding if the total net cash flow from the three transactions is to equal zero. [2]

(iii) After 1 year the force of interest is assumed to be δ_1 for investments of all terms. Assuming that the investor has set up the portfolio determined in (ii), show that the value of the portfolio will now be:

$$\left(1 - e^{\delta_0 - \delta_1}\right)^2$$ [3]

(iv) Comment on the formula derived in part (iii). [2]

[Total 11]

Question 4.10 (Bookwork)

The stochastic differential equations defining the short-rate process assumed in three commonly used models for the term structure of interest rates are shown below:

Model 1: $dr(t) = \alpha[\mu - r(t)]dt + \sigma d\tilde{W}(t)$

Model 2: $dr(t) = \alpha[\mu - r(t)]dt + \sigma\sqrt{r(t)}d\tilde{W}(t)$

Model 3: $dr(t) = \alpha[\mu(t) - r(t)]dt + \sigma d\tilde{W}(t)$

In each case, $\tilde{W}(t)$ denotes a standard Brownian motion under the risk-neutral probability measure.

(i) Identify these three models. [1]

(ii) Outline the key statistical properties of the short-rate processes for each of these models. [6]

(iii) The dynamics of a fourth model are defined by:

Model 4: $dr(t) = \theta\, dt + \sigma d\tilde{W}(t)$

where θ and σ are constants.

State the limitations of this model. [4]

[Total 11]

Question 4.11 (Exam-style)

(i) List the desirable characteristics of interest rate models. [4]

(ii) Discuss the limitations of one-factor models of interest rates. [5]

(iii) The short rate of interest is governed by the stochastic differential equation (SDE):

$$dr_t = 0.6\left(0.04 - r_t\right)dt + 0.006dB_t$$

where B_t is a standard Brownian motion.

By considering the function $f\left(r_t, t\right) = r_t e^{0.6t}$, or otherwise, solve this SDE. [5]

[Total 14]

Question 4.12 (Exam-style)

A bond trader assumes that $f(t,T)$, the instantaneous forward rate of interest at time T implied by the market prices of bonds at the current time t, can be modelled by:

$$f(t,T) = 0.04e^{-0.2\tau} + 0.06(1 - e^{-0.2\tau}) + 0.1(1 - e^{-0.2\tau})e^{-0.2\tau}$$

where $\tau = T - t$.

(i) Sketch a graph of $f(t,T)$ as a function of τ. [3]

(ii) This model is a time-homogeneous, Markov model. Explain what this means. [2]

(iii) Calculate the following quantities using this model:

(a) the instantaneous forward rate of interest in two years' time

(b) the current price of a 10 year zero-coupon bond

(c) the current 10-year spot rate

You should express your answers to (a) and (c) as annualised continuously-compounded rates. [6]

[Total 11]

Question 4.13 (Exam-style)

If $A(t)$ is a strictly positive supermartingale, then zero-coupon bond prices can be modelled using the formula $B(t,T) = \dfrac{E_P[A(T) \mid F_t]}{A(t)}$, where P is a suitably-chosen probability measure.

(i) (a) Express mathematically the fact that $A(t)$ is a strictly positive supermartingale.

(b) Verify that the function $A(t) = e^{-0.05t+0.02W(t)}$, where $W(t)$ denotes standard Brownian motion, satisfies the properties in (i)(a).

(c) State why the supermartingale property is required.

(d) Write down the name given to this type of process. [7]

(ii) By writing $A(t)$ in the form $A(t) = e^{X(t)}$, or otherwise, show that $A(t)$ satisfies a stochastic differential equation of the form:

$$dA(t) = A(t)[\mu_A(t)dt + \sigma_A(t)dW(t)]$$

State the forms of the functions $\mu_A(t)$ and $\sigma_A(t)$. [4]

(iii) (a) Write down or derive a formula for $B(t,T)$ based on the process $A(t)$ specified in (i)(b).

(b) Write down expressions for the instantaneous forward rate $f(t,T)$ and the spot rate $R(t,T)$ based on this model.

(c) State one problem that this model of interest rates has. [4]

(iv) Calculate the prices at time 5 according to the model in (ii) of the following risk-free bonds:

(a) a 10-year zero-coupon bond

(b) a 10-year bond that pays a coupon of 5% at the end of each year. [4]

[Total 19]

Question 4.14 (Bookwork)

(i) Explain what is meant by a default-free bond. [1]

(ii) State the possible outcomes of a default. [2]

(iii) List four types of credit event. [2]

(iv) Explain what is meant by the recovery rate for a bond. [1]
[Total 6]

Question 4.15 (Bookwork)

Describe, in general terms, three different approaches to modelling credit risk. [6]

Question 4.16 (Developmental)

Company X has the following financial structure at time 0:

Debt	£3m	(current book value)
Equity	£6m	(issued share capital)

The debt is a zero-coupon bond with face value £5m that is repayable at par at time 10.

There are 400,000 shares in circulation.

(i) Explain how the Merton model could be used to value shares in Company X. [3]

(ii) Assuming that the debt is repaid directly from the company's funds at that time, state the share price at time 10 if the total value of Company X at that time is:

(a) £15m

(b) £4m [2]
[Total 5]

Question 4.17 (Developmental)

A two-state model is to be used to model the probability that a bond defaults:

$$\boxed{\text{No default, N}} \xrightarrow{\lambda(t)} \boxed{\text{Default, D}}$$

where $\lambda(t) = \dfrac{5 + 20t - t^2}{500}$, $0 \le t \le 20$.

(i) Calculate the probability that the bond does not default between times 5 and 10. [2]

(ii) Explain how the model may be modified to allow the default intensity $\lambda(t)$ to depend on future unforeseen events such as a sudden downturn in the economy. [2]

[Total 4]

Question 4.18 (Exam-style)

A company has just issued 4-year zero-coupon bonds with a nominal value of £4 million. The total value of the company now stands at £7.5 million. A constant risk-free rate of return of 2% *pa* continuously-compounded is available in the market.

(i) Use the Merton model to calculate the theoretical price of £100 nominal of the company's bonds, assuming that the annual volatility of the value of the company's assets is 30%. [4]

(ii) Estimate the risk-neutral probability of default on the company's bonds. [3]

[Total 7]

Question 4.19 (Exam-style)

An analyst is using a two-state continuous-time model to study the credit risk of zero-coupon bonds issued by different companies.

The risk-neutral transition intensity function is:

- $\tilde{\lambda}_A(s) = 0.0148$ for Company A, and
- $\tilde{\lambda}_B(s) = 0.01s^2$ for Company B

where s measures time in years from now.

The analyst observes that the credit spread on a 3-year zero-coupon bond just issued by Company B is twice that on a 3-year zero-coupon bond just issued by Company A.

(i) Given that the risk-free force of interest is 5% *pa,* and that the average recovery rate in the event of default, δ, where $0 < \delta < 1$, is the same for both companies, calculate δ. [7]

(ii) Explain how the two-state model for credit risk can be generalised to give the Jarrow-Lando-Turnbull model. [3]

[Total 10]

Question 4.20 (Exam-style)

The credit-worthiness of debt issued by companies is assessed at the end of each year by a credit rating agency. The ratings are A (the most credit-worthy), B and D (debt defaulted). Historical evidence supports the view that the credit rating of a debt can be modelled as a Markov chain with the following matrix of one-year transition probabilities:

$$\mathbf{X} = \begin{pmatrix} 0.92 & 0.05 & 0.03 \\ 0.05 & 0.85 & 0.1 \\ 0 & 0 & 1 \end{pmatrix}$$

(i) Determine the probability that a company rated A will never be rated B in the future. [2]

(ii) (a) Calculate the second order transition probabilities of the Markov chain.

(b) Hence calculate the expected number of defaults within the next two years from a group of 100 companies, all initially rated A. [2]

The manager of a portfolio investing in company debt follows a "downgrade trigger" strategy. Under this strategy, any debt in a company whose rating has fallen to B at the end of a year is sold and replaced with debt in an A-rated company.

(iii) Calculate the expected number of defaults for this investment manager over the next two years, given that the portfolio initially consists of 100 A-rated bonds. [2]

(iv) Comment on the suggestion that the downgrade trigger strategy will improve the return on the portfolio. [2]

[Total 8]

Part 4 – Solutions

Solution 4.1

(i) ***Explain what is meant by a "complete market"***

A financial market is said to be complete if any contingent claim (*ie* derivative payoff) can be replicated using the underlying asset and cash. [1]

For example, we can replicate the payoff from a "plain vanilla" call option on a share using a long holding in the underlying share and a short holding of cash. [1]
[Total 2]

(ii) ***Give two reasons why markets might not be complete in practice***

A replicating strategy will typically involve holding fractions of shares and fractions of units of cash. Since these are indivisible it will not be possible to replicate exactly. [1]

A replicating strategy for a put option will require a short holding of the underlying security. If the financial markets do not allow short selling (or if there are restrictions on it), then replication may not be possible. [1]
[Total 2]

(iii) ***Relevance for derivative pricing***

Although it is usually just a technicality, the concept of a complete market is important because, without it, we could not be sure that we could replicate the payoff from a derivative. This means that the steps in the derivation of the risk-neutral pricing formula cannot be applied. So we could not use this formula to price the derivative. [2]

Solution 4.2

(i) ***Express the martingale property mathematically***

The martingale property tells us that, whenever $t < T$:

$$E_Q\left[e^{-rT}S_T \,\middle|\, F_t\right] = e^{-rt}S_t \qquad [1]$$

(ii)(a) ***Expression for B_t***

B_t is the accumulated value at time t of an initial investment of 1 unit of cash.

So: $B_t = e^{rt}$ [1]

(ii)(b) ***Show that the discounted cash process is a martingale***

The discounted cash process is therefore:

$$e^{-rt}B_t = e^{-rt}e^{rt} = 1$$

which trivially satisfies the martingale equation $E_Q\left[e^{-rT}B_T \,\middle|\, F_t\right] = e^{-rt}B_t$. [1]

(ii)(c) ***Deduce that any self-financing portfolio is also a martingale***

We have established that the discounted values of both of the components (the shares and the cash) of such a portfolio are martingales. So any multiple of these will also be a martingale. [1]

Also, if we "rebalance" the portfolio by making switches from cash to shares or vice versa, this will not affect the martingale property, provided that we don't put any money in or take any out. [1]
[Total 4]

A good intuitive way to think of martingales here is to think that, on average, they don't drift up or down. So, if our (discounted) cash and shares are not drifting up or down, neither will any combination of them.

(iii) ***Show that $e^{-rt}V_t$ is also a martingale***

To show that the process $e^{-rt}V_t$ is a martingale, we need to prove that, if $t_1 < t_2$ then:

$$E_Q\left[e^{-rt_2}V_{t_2} \,\middle|\, F_{t_1}\right] = e^{-rt_1}V_{t_1} \qquad (1)$$ [½]

We've used t_1 and t_2 here, rather than t and T, to avoid confusion, because there's already a T in the definition of V_t.

Substituting the definition of V_{t_2} into the left-hand side (*LHS*) of (1), we have:

$$LHS = E_Q\left[e^{-rt_2}e^{-r(T-t_2)}E_Q(X \mid F_{t_2})\middle|F_{t_1}\right]$$

$$= e^{-rT}E_Q\left[E_Q(X \mid F_{t_2})\middle|F_{t_1}\right] \qquad [½]$$

We can simplify this nested expectation using the Tower Law to get:

$$LHS = e^{-rT}E_Q(X \mid F_{t_1}) \qquad [1]$$

From the definition of V_{t_1}, the right-hand side (*RHS*) of (1) equals:

$$RHS = e^{-rt_1}e^{-r(T-t_1)}E_Q(X \mid F_{t_1}) = e^{-rT}E_Q(X \mid F_{t_1}) \qquad [½]$$

So: $LHS = RHS$

This shows that $e^{-rt}V_t$ is also a Q-martingale. [½]

[Total 3]

(iv) ***Explain the significance***

We have shown that, if we have a self-financing portfolio consisting of shares and cash, its discounted value will be a Q-martingale, *ie* it will have no drift. [½]

If it is possible to rebalance such a portfolio so that it will always replicate over the next instant the value of a derivative based on the share, then the discounted value of the derivative must equal the discounted value of this portfolio. [½]

It turns out that such a replicating strategy *is* possible. The Martingale Representation Theorem guarantees this. [½]

But the discounted process $e^{-rt}V_t$ behaves in precisely this way, and gives the correct payoff X when $t = T$. [½]

So V_t must equal the value of the derivative. [½]

This gives us the derivative pricing formula $V_t = e^{-r(T-t)}E_Q[X \mid F_t]$. [½]

[Total 3]

Solution 4.3

(i)(a) The filtration F_i is just the set of values representing the path followed up to time i. So, in this case (where we are told what has happened up to time 1), we have:

$$F_0 = \{13\} \quad \text{and} \quad F_1 = \{13, 20\} \qquad [1]$$

(i)(b) If $F_1 = \{13, 20\}$, there are two possible values for X_2, namely 25 and 15, which we are assuming are equally likely.

So: $E_P(X_2 \mid F_1) = \frac{1}{2} \times 25 + \frac{1}{2} \times 15 = 20$ [1]

This figure has already been written in on the tree in the question at time 1.

(i)(c) If $X_1 = 6$, then $F_1 = \{13, 6\}$.

Here: $E_P(X_2 \mid F_1) = \frac{1}{2} \times 12 + \frac{1}{2} \times 0 = 6$ [1]

Again, this matches the figure shown in the tree at time 1.

(i)(d) We have established that, if we start at the "13" node (corresponding to the only possible value of F_0), the conditional expectation $E_P(X_2 \mid F_1)$ can take two possible values, and these are equally likely under the probability measure P.

So: $E_P[E_P(X_2 \mid F_1) \mid F_0] = \frac{1}{2} \times 20 + \frac{1}{2} \times 6 = 13$ [1]

This matches the figure shown in the tree at time 0.

(i)(e) If we start at the "13" node, X_2 can take four possible values (25, 15, 12 or 0), each with probability $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

So: $E_P(X_2 \mid F_0) = \frac{1}{4} \times 20 + \frac{1}{4} \times 15 + \frac{1}{4} \times 12 + \frac{1}{4} \times 0 = 13$ [1]

This expectation is the same as the expectation in (i)(d), *ie* $E_P[E_P(X_2 \mid F_1) \mid F_0] = E_P(X_2 \mid F_0)$. This is an example of the "tower property" of conditional expectations. [1]

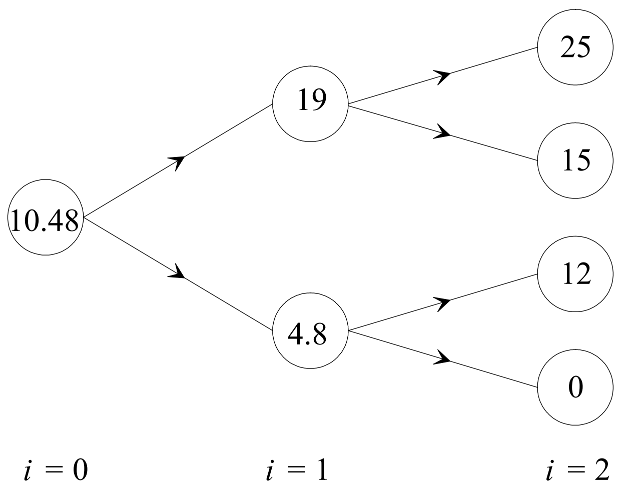
[Total 6]

(ii)(a) Yes. The tower property applies equally well in this case *ie* it is also true that:

$$E_Q[E_Q(X_2 \mid F_1) \mid F_0] = E_Q(X_2 \mid F_0) \qquad [1]$$

(ii)(b) Equalities of the form $E_Q(X_n \mid F_{i-1}) = E_Q[E_Q(X_n \mid F_i) \mid F_{i-1}]$ are useful because they enable us to work backwards through a binomial tree, calculating the value of the derivative at time $i-1$ from the values at time i. [1]

(ii)(c) We need to calculate $E_Q(X_2 \mid F_0)$. Calculating the values of the conditional expectations by working backwards using the risk-neutral probability measure *Q*, leads to the following tree:



So the value of the derivative at time 0 is 10.48. [2]

[Total 4]

Solution 4.4

(i)(a) ***Name of the process***

This process is *geometric Brownian motion* (also known as the lognormal model). [1]

(i)(b) ***Two real-world quantities***

This process is commonly used to model

- share prices (with or without dividends)
- currency exchange rates [2]

[Total 3]

(ii)(a) ***State what is meant by "equivalent probability measures"***

In words, two probability measures are equivalent if they are defined on the same sample space and have the same null sets (*ie* sets that have probability zero).

Mathematically, P and Q are equivalent if $P(A)>0 \iff Q(A)>0$, where $P(A)$ denotes the probability under measure P and $Q(A)$ denotes the probability under measure Q. [1]

This is sometimes expressed as "Equivalent measures agree on what is possible".

(ii)(b) ***State how the CMG Theorem could be applied***

The CMG Theorem tells us that, for any constant γ, there is a probability measure Q (equivalent to P) such that $\tilde{W}_t = W_t + \gamma t$ is standard Brownian motion under Q.

So we could change probability measures and work with Q.

Since $W_t = \tilde{W}_t - \gamma t$, under the measure Q, the formula for S_t in terms of $\tilde{W}_t$ becomes:

$$S_t = S_0 e^{(\alpha - \frac{1}{2}\sigma^2)t + \sigma(\tilde{W}_t - \gamma t)} = S_0 e^{(\alpha - \gamma\sigma - \frac{1}{2}\sigma^2)t + \sigma\tilde{W}_t}$$

If we want this to equal $S_0 e^{(r - \frac{1}{2}\sigma^2)t + \sigma\tilde{W}_t}$, we need to set $\alpha - \gamma\sigma = r$, *ie* $\gamma = \dfrac{\alpha - r}{\sigma}$. [2]

[Total 3]

(iii)(a) ***SDE for dS_t***

We can write the equation $S_t = S_0 e^{(r-½\sigma^2)t+\sigma\tilde{W}_t}$ in the form:

$$S_t = f(X_t)$$

where $X_t = (r - ½\sigma^2)t + \sigma\tilde{W}_t$ and $f(x) = S_0 e^x$

So: $dX_t = (r - ½\sigma^2)dt + \sigma d\tilde{W}_t$ and $f'(x) = f''(x) = S_0 e^x$

Using a Taylor Series expansion, we can write:

$$\begin{aligned} dS_t = df(X_t) &= f'(X_t)dX_t + \tfrac{1}{2}f''(X_t)(dX_t)^2 \\ &= S_0 e^{X_t}\left\{dX_t + \tfrac{1}{2}(dX_t)^2\right\} \\ &= S_t\left\{dX_t + \tfrac{1}{2}(dX_t)^2\right\} \end{aligned}$$

Substituting the SDE for X_t gives:

$$dS_t = S_t\left\{[(r-½\sigma^2)dt + \sigma d\tilde{W}_t] + \tfrac{1}{2}[(r-½\sigma^2)dt + \sigma d\tilde{W}_t]^2\right\}$$

Simplifying using the 2×2 multiplication grid given in the Core Reading, we get:

$$\begin{aligned} dS_t &= S_t\left\{[(r-½\sigma^2)dt + \sigma d\tilde{W}_t] + \tfrac{1}{2}\sigma^2 dt\right\} \\ &= S_t(rdt + \sigma d\tilde{W}_t) \end{aligned}$$ [3]

(iii)(b) ***SDE for dS_t***

The drift in this equation is rS_t, or just r, if we're thinking in units of S_t. [1]

Since $\tilde{W}_t$ is standard Brownian motion, the increment $d\tilde{W}_t$ has mean zero. This means that the expected value of dS_t under the measure Q is $rS_t dt$, *ie* the share price is drifting upwards at the risk-free rate. This means that Q is the risk-neutral probability measure for the process S_t. [1]

[Total 5]

Solution 4.5

(i) ***Formula for $E_Q[X]$***

If we take logs of the equation given, we get:

$$\log S_T = \log S_0 + (r - ½\sigma^2)T + \sigma \tilde{Z}_T$$

Since $\tilde{Z}_T \sim N(0,T)$, we see that the distribution of S_T given F_0 under the probability measure Q is:

$$S_T \left| F_0 \sim logN[\log S_0 + (r - ½\sigma^2)T, \sigma^2 T] \right. \qquad [1]$$

The expectation we require is:

$$E_Q[X|F_0] = E_Q[\max(S_T - K, 0)|F_0]$$

$$= \int_K^\infty (x-K) f(x)dx$$

$$= \int_K^\infty x\, f(x)dx - K\int_K^\infty f(x)dx \qquad [1]$$

where $f(x)$ is the density function of the above lognormal distribution.

We can evaluate these integrals using the formula on page 18 of the Tables, with $L = K$ and $U = \infty$, and with μ and σ^2 replaced with $\log S_0 + (r - ½\sigma^2)T$ and $\sigma^2 T$.

If we put $k = 1$, we get:

$$\int_K^\infty xf(x)dx = e^{\log S_0 + (r-½\sigma^2)T + ½\sigma^2 T}\left[\Phi(\infty) - \Phi\left(\frac{\log K - [\log S_0 + (r - ½\sigma^2)T]}{\sigma\sqrt{T}} - \sigma\sqrt{T}\right)\right]$$

$$= S_0 e^{rT}\left[1 - \Phi\left(\frac{\log K - [\log S_0 + (r - ½\sigma^2)T]}{\sigma\sqrt{T}} - \sigma\sqrt{T}\right)\right]$$

We can simplify this using the identity $1-\Phi(x)=\Phi(-x)$ to get:

$$\int_K^\infty xf(x)dx = S_0e^{rT}\Phi\left(\frac{-\log K+\log S_0+(r-\tfrac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}+\sigma\sqrt{T}\right)$$

$$=S_0e^{rT}\Phi\left(\frac{\log(S_0/K)+(r-\tfrac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}+\sigma\sqrt{T}\right)$$

$$=S_0e^{rT}\Phi\left(d_2+\sigma\sqrt{T}\right) \qquad [3]$$

where d_2 is as defined in the question.

Similarly, if we put $k=0$, we get:

$$\int_K^\infty f(x)dx = \Phi\left(d_2\right) \qquad [1]$$

So the expectation is:

$$E_Q[X|F_0]=S_0e^{rT}\Phi\left(d_2+\sigma\sqrt{T}\right)-K\Phi\left(d_2\right) \qquad [1]$$

[Total 7]

(ii) ***Relevance of result***

If S_T denotes the value of a share at time T, and Q denotes the risk-neutral measure for this share, this formula allows us to work out the fair price of a European call option on the share with strike price K and time to expiry T. [1]

This price would be calculated as the discounted value of the expectation in part (i), namely:

$$e^{-rT}E_Q[X|F_0]=e^{-rT}\left[S_0e^{rT}\Phi\left(d_2+\sigma\sqrt{T}\right)-K\Phi\left(d_2\right)\right]$$

$$=S_0\Phi\left(d_1\right)-Ke^{-rT}\Phi\left(d_2\right) \qquad [1]$$

where $d_1=d_2+\sigma\sqrt{T}$.

This matches the Black-Scholes formula for valuing a call option on a non-dividend-paying share. [1]

[Total 3]

Solution 4.6

(i) ***State the general risk-neutral pricing formula***

$$V_t = e^{-r(T-t)} E_Q\left[X_T \left| F_t \right.\right] \qquad [1]$$

where:

- r = constant risk-free force of interest
- T = maturity date of the derivative
- t = today's date
- X_T = derivative payoff at maturity date T
- Q = risk-neutral probability measure
- F_t = filtration at time t

(ii)(a) ***Derive the formula for the price of Derivative 1***

The payoff function of this non-standard derivative is:

$$X_{1T} = \begin{cases} 1 & S_T < K \\ 0 & S_T \geq K \end{cases} \qquad [½]$$

To derive the pricing formula of the derivative, we substitute the payoff function into the general risk-neutral pricing formula in part (i), *ie*:

$$V_{1t} = e^{-r(T-t)} E_Q\left[X_{1T} \left| F_t \right.\right]$$

which, as the derivative pays 1 provided the share price is between 0 and the strike price K, is equal to:

$$V_{1t} = e^{-r(T-t)} \int_0^K 1\, f(S_T \left| S_t \right.)\, dS_T \qquad (1) \qquad [1]$$

where $f(S_T \left| S_t \right.)$ is the probability density function of the share price at the maturity date, S_T, given the current share price, S_t.

Note that as the share price is assumed to follow geometric Brownian motion, the independence of increments means we do not have to condition on the full past history of the share price – only the current share price, S_t.

The distribution of S_T given S_t is:

$$S_T \big| S_t \sim \log N\left[\ln S_t + (r - q - ½\sigma^2)(T-t), \sigma^2(T-t)\right] \qquad [½]$$

We can simply state this result or obtain it using the following reasoning:

Under the risk-neutral probability measure, Q, the tradable asset is expected to grow at the risk-free force of interest, r. Here, the tradable asset is the share plus the dividends earned. Assuming the dividends are immediately reinvested in the asset, they will give a rate of growth of the tradable asset of q. So this means the share price alone must grow at a rate $r-q$.

Therefore, if S_t denotes the share price at time t, then it has SDE:

$$dS_t = S_t\left((r-q)dt + \sigma dZ_t\right)$$

where Z_t denotes standard Brownian motion. This has solution:

$$S_T = S_t \exp\left(\left(r - q - ½\sigma^2\right)(T-t) + \sigma(Z_T - Z_t)\right)$$

giving the distribution:

$$\ln S_T \big| S_t \sim N\left[\ln S_t + (r - q - ½\sigma^2)(T-t), \sigma^2(T-t)\right]$$
$$\Leftrightarrow \quad S_T \big| S_t \sim \log N\left[\ln S_t + (r - q - ½\sigma^2)(T-t), \sigma^2(T-t)\right]$$

To work out the integral (1) above, we need to use the formula for evaluating the truncated moments of a lognormal distribution, which appears on page 18 of the *Tables.*

We have, $1 = (S_T)^0$ and so $k = 0$ in the formula on page 18 of the *Tables*.

Thus, (1) above is equal to:

$$V_{1t} = e^{-r(T-t)}\left[e^0\right]\left[\Phi(U_0) - \Phi(L_0)\right] \qquad (2) \qquad [½]$$

where:

$$\begin{aligned} U_0 &= \frac{\ln K - \left\{\ln S_t + (r - q - ½\sigma^2)(T-t)\right\}}{\sigma\sqrt{T-t}} - 0 \\ &= -\left[\frac{\ln(S_t/K) + (r - q - ½\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right] \\ &= -d_2 \end{aligned} \qquad [1]$$

$$L_0 = \frac{\ln 0 - \left\{\ln S_t + (r - q - ½\sigma^2)(T-t)\right\}}{\sigma\sqrt{T-t}} - 0 = -\infty \qquad [½]$$

So, (2) above becomes:

$$\begin{aligned} V_{1t} &= e^{-r(T-t)}\left[\Phi(-d_2) - \Phi(-\infty)\right] \\ &= e^{-r(T-t)}\left[\Phi(-d_2) - 0\right] \\ &= e^{-r(T-t)}\Phi(-d_2) \end{aligned} \qquad [1]$$

(ii)(b) ***Derive the formula for the price of Derivative 2***

The payoff function of this non-standard derivative is:

$$X_{2T} = \begin{cases} S_T & S_T < K \\ 0 & S_T \geq K \end{cases} \qquad [½]$$

To derive the pricing formula of this derivative, we substitute the payoff function into the risk-neutral pricing formula in part (i), *ie*:

$$V_{2t} = e^{-r(T-t)}E_Q\left[X_{2T} \middle| F_t\right]$$

Substituting in for X_{2T} gives:

$$V_{2t} = e^{-r(T-t)} \int_0^K S_T \; f(S_T | S_t) \, dS_T \qquad (3) \qquad [1]$$

To evaluate this, we again use the formula for evaluating the truncated moments of a lognormal distribution on page 18 of the *Tables*, where:

$$S_T | S_t \sim \log N\left[\ln S_t + (r - q - ½\sigma^2)(T-t), \sigma^2(T-t) \right] \qquad [½]$$

In this instance, $S_T = (S_T)^1$, so $k = 1$ in the formula on page 18 of the *Tables*. Thus, (3) above is equal to:

$$V_{2t} = e^{-r(T-t)} \left[e^{\ln S_t + (r-q-½\sigma^2)(T-t) + ½\sigma^2(T-t)} \right] \left[\Phi(U_1) - \Phi(L_1) \right] \qquad (4) \qquad [½]$$

where:

$$\begin{aligned} U_1 &= \frac{\ln K - \left\{ \ln S_t + (r-q-½\sigma^2)(T-t) \right\}}{\sigma\sqrt{T-t}} - \sigma\sqrt{T-t} \\ &= \frac{\ln K - \ln S_t - (r-q-½\sigma^2)(T-t) - \sigma^2(T-t)}{\sigma\sqrt{T-t}} \\ &= \frac{\ln K - \ln S_t - (r-q+½\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ &= -\left[\frac{\ln(S_t/K) + (r-q+½\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \right] \\ &= -d_1 \end{aligned} \qquad [1]$$

$$L_1 = \frac{\ln 0 - \left\{ \ln S_t + (r-q-½\sigma^2)(T-t) \right\}}{\sigma\sqrt{T-t}} - \sigma\sqrt{T-t} = -\infty \qquad [½]$$

So, after some cancelling of terms involving r and σ^2, (4) above becomes:

$$\begin{aligned} V_{2t} &= S_t e^{-q(T-t)} \left[\Phi(-d_1) - \Phi(-\infty) \right] \\ &= S_t e^{-q(T-t)} \left[\Phi(-d_1) - 0 \right] \\ &= S_t e^{-q(T-t)} \Phi(-d_1) \end{aligned} \qquad [1]$$

(ii)(c) ***Derive the pricing formula for a European put option***

The payoff function for a European put option can be written as:

$$X_{PT} = \begin{cases} K - S_T & S_T < K \\ 0 & S_T \geq K \end{cases}$$ [½]

This payoff function can be replicated using a combination of $+K$ of the derivatives in part (ii)(a) and –1 of the derivatives in part (ii)(b). [½]

Consequently, and assuming that markets are arbitrage-free, the price of a European put option must be given by:

$$\begin{aligned} p_t &= K \times V_{1t} - V_{2t} \\ &= Ke^{-r(T-t)}\Phi(-d_2) - S_t e^{-q(T-t)}\Phi(-d_1) \end{aligned}$$ [1]

[Total 12]

(iii) ***Price of exotic forward contract***

Here the payoff function is equal to:

$$X_T = S_T^2 - K^2$$ [1]

So, once again substituting this into the general risk-neutral pricing formula in part (i) gives:

$$\begin{aligned} V_t &= e^{-r(T-t)} E_Q\left[S_T^2 - K^2 \middle| F_t \right] \\ &= e^{-r(T-t)} \left(E_Q\left[S_T^2 \middle| F_t \right] - K^2 \right) \end{aligned}$$

where:

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

As this derivative contract provides a non-zero payoff regardless of the share price at maturity, we can evaluate $E_Q\left[S_T^2 \middle| F_t\right]$ using the formula for the moments of a non-truncated lognormal distribution on page 14 of the *Tables*, *ie*:

$$V_t = e^{-r(T-t)}\left[e^{2(\ln S_t+(r-q-½\sigma^2)(T-t))+½\times 4\sigma^2(T-t)} - K^2\right]$$

$$= S_t^2 e^{(r-2q+\sigma^2)(T-t)} - K^2 e^{-r(T-t)} \qquad [1]$$

[Total 3]

Solution 4.7

These three models are all one-factor models used for modelling the short rate of interest $r(t)$. [½]

All three models assume that $r(t)$ has the dynamics of an Ito process under the risk-neutral probability measure Q. [½]

The equations defining the three models are:

Vasicek: $dr(t) = \alpha[\mu - r(t)]dt + \sigma d\tilde{W}(t)$ [1]

Cox-Ingersoll-Ross: $dr(t) = \alpha[\mu - r(t)]dt + \sigma\sqrt{r(t)}d\tilde{W}(t)$ [1]

Hull & White: $dr(t) = \alpha[\mu(t) - r(t)]dt + \sigma d\tilde{W}(t)$ [1]

All three models are mean reverting to μ, where… [½]

… μ is time-dependent for the Hull & White model, but constant for the other two. [½]

The Hull & White model is easier to fit to past and current data than the Vasicek or Cox-Ingersoll-Ross models due to the fact that there is more choice of parameters (since μ is time-dependent). [½]

All three models generate arbitrage-free zero-coupon bond prices. [½]

All three models can be used to price simple options on zero-coupon bonds. [½]

The Cox-Ingersoll-Ross model includes the factor $\sqrt{r(t)}$ in the volatility coefficient. This prevents $r(t)$ taking negative values. [½]

The Vasicek model is much more tractable mathematically than the other two. [½]

Over long periods the distribution of $r(t)$ under the Cox-Ingersoll-Ross model involves the non-central chi-square distribution, whereas the distribution under the other two models is normal. [½]

Since these are all one-factor models:

- they cannot be used to price derivatives whose payoffs depend on more than one interest rate. [½]

- there will be positive correlation between bond prices of all durations, which is unrealistic. [½]

[Maximum 8]

Solution 4.8

(a) $$\boldsymbol{B(t,T) = E_Q\left[\exp\left(-\int_t^T r(u)du\right)\middle| r(t)\right]}$$

This formula is used to find the price at time t of a zero-coupon bond maturing at time T when a short rate model is used. [1]

The short rate of interest $r(t)$ is assumed to be a one-factor, Markov, Ito (diffusion) process. [1]

"Q" denotes the risk-neutral probability measure for this process. [1]

(b) $$\boldsymbol{B(t,T) = \frac{1}{A(t)} E_P[A(T) \mid F_t]}$$

This formula is used to find the price at time t of a zero-coupon bond maturing at time T when a state-price deflator model is used. [1]

$A(t)$ is the state-price deflator, which is a stochastic process. [½]

$A(t)$ must be a strictly positive supermartingale. [½]

"P" denotes a suitably-chosen probability measure. [½]

The stochastic differential equation for $A(t)$ has the following form under P:

$$dA(t) = A(t)[\mu_A(t)dt + \sigma_A(t)dW(t)]$$

where $\mu_A(t)$ and $\sigma_A(t)$ are appropriately-chosen stochastic processes. [½]

(c) $\boldsymbol{\gamma(t,T) = \dfrac{m(t,T) - r(t)}{S(t,T)}}$

$\gamma(t,T)$ is the "market price of risk", which represents the excess expected return over the risk-free rate per unit of volatility. [1]

This formula assumes that $B(t,T)$, the price at time t of a zero-coupon bond maturing at time T, is an Ito process. [½]

Under the real-world probability measure P, the stochastic differential equation for $B(t,T)$ is:

$$dB(t,T) = B(t,T)[m(t,T)dt + S(t,T)dW(t)] \quad [1]$$

With this one-factor model the value of $\gamma(t,T)$ is independent of T, the maturity date of the bond. [½]
[Total 9]

Solution 4.9

(i) ***Advantages / disadvantages of the Redington model***

The advantages of this model are:

- It is mathematically tractable (easy to do calculations). [1]

- It provides a sufficiently accurate approximation in situations when the term of the investments is not crucial, *eg* for long-term investors such as pension funds. [1]

The disadvantages of this model are:

- It is not arbitrage-free. [1]

- It does not accurately represent real-life yield curves, which are not flat. [1]
[Total 4]

(ii) ***Determine the maturity value***

The purchase price of the 1-year bond will be $e^{-\delta_0}$. [½]

The purchase price of the 3-year bond will be $e^{2\delta_0} \times e^{-3\delta_0} = e^{-\delta_0}$. [½]

So, if the total net cash flow is to be zero, the investor will need to generate $2e^{-\delta_0}$ from the sale of the 2-year bonds. So their maturity value must be $2e^{-\delta_0} \times e^{2\delta_0} = 2e^{\delta_0}$. [1]
[Total 2]

(iii) ***Value after 1 year***

After one year the interest rate has changed to δ_1 for all terms. So the values of the three bonds will be as follows:

Bond	*Remaining life*	*Value*	
1-year bond	(matured)	1	[½]
3-year bond	2 years	$e^{2\delta_0} \times e^{-2\delta_1} = e^{2(\delta_0-\delta_1)}$	[½]
2-year bond	1 year	$-2e^{\delta_0} \times e^{-\delta_1} = -2e^{\delta_0-\delta_1}$	[1]

The total value of the portfolio will be:

$$1 + e^{2(\delta_0-\delta_1)} - 2e^{\delta_0-\delta_1} = \left(1 - e^{\delta_0-\delta_1}\right)^2$$ [1]

[Total 3]

(iv) ***Comment***

Since the expression derived for the value at time 1 of the portfolio is a square, its value is non-negative. [½]

This means that the investor can set up a portfolio at no cost at time 0 whose value at time 1 cannot be negative, but will be strictly positive if the interest rate changes. In other words, the investor can make an arbitrage profit. [1]

This shows that the Redington "flat" interest rate model is not arbitrage-free. [½]
[Total 2]

Solution 4.10

(i) ***Identify the models***

These are the Vasicek, Cox-Ingersoll-Ross and Hull & White models, respectively. [1]

(ii) ***Key statistical properties***

In each case the short rate of interest $r(t)$ is modelled as an Ito process. [1]

The process therefore operates in continuous time and has normally-distributed increments over short time intervals. [1]

The volatility parameter σ controls the size of the random movements. [½]

The coefficient of dt represents the expected annual drift under the risk-neutral probability measure (since the increments $d\tilde{W}(t)$ have zero mean under this measure). But under the real-world probability measure the drift will have a different value. [½]

All three models exhibit mean reversion ... [½]

... provided the parameter α (which is positive) takes a value less than 1. [½]

In Model 1 and Model 2 the long-term "target" rate μ is constant, while in Model 3 it is a function of the time t. [½]

With Model 1 and Model 3 it is possible to get a negative value for $r(t)$, which is unrealistic. [½]

The inclusion of the $\sqrt{r(t)}$ factor in Model 2 prevents $r(t)$ taking negative values. [½]

Over longer periods the distribution of $r(t)$ under Model 2 involves the non-central chi-square distribution, whereas the distribution under the other two models is normal. [½]

Model 3 is easier to fit to past and current data than the other models due to the fact that there is more choice of parameters (since μ is time-dependent). [½]

[Maximum 6]

(iii) ***Limitations of Model 4***

Model 4 does not exhibit mean reversion. [1]

If θ is non-zero, the trend value of $r(t)$ will increase or decrease steadily, which is unrealistic. [1]

The model allows negative interest rates, which is not necessarily realistic. [1]

The model involves only two parameters, θ and σ, and so may be difficult to calibrate to past and current data. [1]

Since this is a one-factor model:

- it cannot be used to price derivatives whose payoffs depend on more than one interest rate. [½]
- there will be positive correlation between bond prices of all durations, which is unrealistic. [½]

[Maximum 4]

Solution 4.11

(i) ***Desirable characteristics of interest rate models***

- The model should be arbitrage-free.
- Interest rates should be positive.
- Interest rates should be mean-reverting over the long term.
- Bonds and derivative contracts should be easy to price.
- The model should produce realistic interest rate dynamics.
- It should fit historical interest rate data adequately.
- It should be easy to calibrate to current market data.
- It should be flexible enough to cope with a range of derivatives.

[½ each, Total 4]

(ii) ***Limitations of one-factor models***

A result of one-factor models is that yields on bonds of different durations are perfectly correlated. This is not realistic. [1]

In fact, they need not even be positively correlated. Sometimes we see, for example, that short-dated bonds fall in price while long-dated bonds increase in price. [½]

If we look at the long run of historical data we find that there have been sustained periods of both high and low interest rates with periods of both high and low volatility. These are features that are difficult to capture in one-factor models. [1]

This issue is especially important for two types of problem in insurance:

1. the pricing and hedging of long-dated insurance contracts with interest rate guarantees [½]

2. asset-liability modelling and long-term risk management. [½]

One-factor models struggle to cope with valuing derivative contracts that are more complex than, say, standard European call options. We need more complex models to deal effectively with these. For example, any contract that makes reference to more than one interest rate should allow these rates to be less than perfectly correlated. [1½]
[Total 5]

(iii) ***Short rate of interest***

Using a Taylor Series expansion for the given function we get:

$$d\left(r_t e^{0.6t}\right) = df\left(r_t, t\right)$$

$$= \frac{\partial f}{\partial r_t} dr_t + \frac{1}{2}\frac{\partial^2 f}{\partial r_t^2}\left(dr_t\right)^2 + \frac{\partial f}{\partial t} dt$$

$$= e^{0.6t} \times \left[0.6\left(0.04 - r_t\right)dt + 0.006 dB_t\right] + \frac{1}{2} \times 0 \times \left(dr_t\right)^2 + 0.6 r_t e^{0.6t} dt$$

$$= 0.024 e^{0.6t} dt + 0.006 e^{0.6t} dB_t - 0.6 r_t e^{0.6t} dt + 0.6 r_t e^{0.6t} dt$$

$$= 0.024 e^{0.6t} dt + 0.006 e^{0.6t} dB_t \qquad [2]$$

Alternatively, you can use the general form of Ito's formula, the product rule or an integrating factor to derive this equation.

Substituting s for t and integrating between 0 and t we get:

$$\int_0^t d\left(r_s e^{0.6s}\right) = \int_0^t 0.024 e^{0.6s} ds + \int_0^t 0.006 e^{0.6s} dB_s$$

$$\Leftrightarrow \quad \left[r_s e^{0.6s}\right]_0^t = 0.04\left[e^{0.6s}\right]_0^t + 0.006\int_0^t e^{0.6s} dB_s \qquad [2]$$

$$\Leftrightarrow \quad r_t e^{0.6t} - r_0 = 0.04\left(e^{0.6t} - 1\right) + 0.006\int_0^t e^{0.6s} dB_s$$

We can then rearrange to get the required solution:

$$\Leftrightarrow \quad r_t = r_0 e^{-0.6t} + 0.04\left(1 - e^{-0.6t}\right) + 0.006\int_0^t e^{-0.6(t-s)} dB_s \qquad [1]$$

[Total 5]

Note that this is an example of the Vasicek model, which is also an Ornstein-Uhlenbeck process.

Solution 4.12

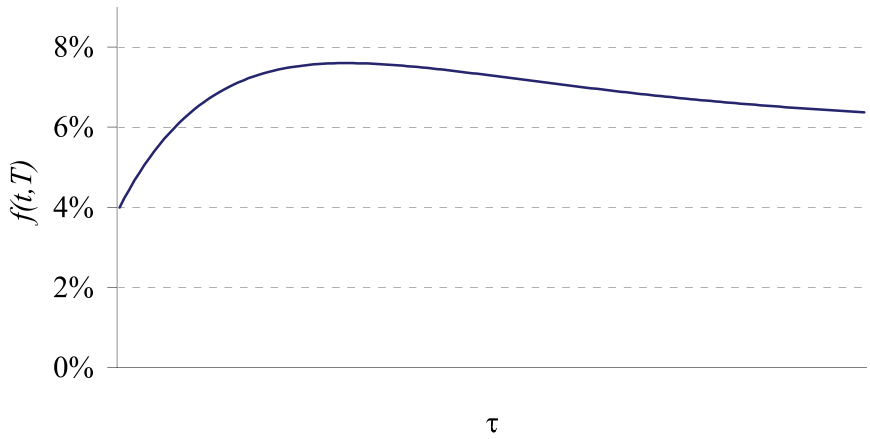
(i) ***Graph***

The factor $e^{-0.2\tau}$ reduces from 1 to zero as the term τ increases from zero to infinity. So the first two terms represent a weighted average of 0.04 (the short rate) and 0.06 (the long rate). [½]

The final term is zero when $\tau = 0$ or $\tau = \infty$, but positive in between. So this adds a "hump" to the graph. [½]

In fact, if you differentiate the function given, you will find that the maximum occurs when $e^{-0.2\tau} = 0.4$*, which corresponds to the point where* $\tau = 4.58$ *and* $f = 7.6\%$.

The graph looks like this:



[2]
[Total 3]

(ii) ***Explain what is meant by "time homogeneous" and "Markov"***

"Time homogeneous" means that the rates do not depend on the calendar time t, just on the term τ. [1]

"Markov" means that the distribution of the future values of the process depends only on the current value. Knowledge of previous values provides no additional useful information. [1]
[Total 2]

(iii)(a) ***Calculate the instantaneous forward rate in 2 years' time***

This is:

$$f(t,t+2) = 0.04e^{-0.2\times 2} + 0.06(1-e^{-0.2\times 2}) + 0.1(1-e^{-0.2\times 2})e^{-0.2\times 2} = 0.0687$$

ie 6.87% [1]

(iii)(b) ***Calculate the price of a 10-year zero-coupon bond***

This is:

$$B(t,t+10)=\exp\left[-\int_t^{t+10} f(t,u)du\right]$$

The integral, using the substitution $\tau = u-t$, is:

$$\begin{aligned}\int_t^{t+10} f(t,u)du &= \int_0^{10}[0.04e^{-0.2\tau}+0.06(1-e^{-0.2\tau})+0.1(e^{-0.2\tau}-e^{-0.4\tau})]d\tau \\ &= \int_0^{10}[0.06+0.08e^{-0.2\tau}-0.1e^{-0.4\tau}]d\tau \\ &= \Big[0.06\tau-0.4e^{-0.2\tau}+0.25e^{-0.4\tau}\Big]_0^{10} \\ &= 0.55044-(-0.15)=0.70044\end{aligned}$$

So: $B(t,t+10)=e^{-0.70044}=0.4964$

ie £49.64 per £100 nominal [3]

(iii)(c) ***Calculate the 10-year spot rate***

This can be calculated as the average of the forward rates:

$$R(t,t+10)=\frac{1}{10}\int_t^{t+10} f(t,u)du=\frac{1}{10}\times 0.70044=0.070044$$

ie 7.00% [2]

Alternatively, we could solve:

$$B(t,t+10)=e^{-0.70044}=1\times e^{-10R(t,t+10)}$$

[Total 6]

Solution 4.13

(i)(a) ***Express these properties mathematically***

"Strictly positive" simply means that:

$$A(t) > 0 \text{ for all times } t \qquad [1]$$

The "supermartingale" property means that, whenever $t < T$:

$$E_P[A(T) \mid F_t] < A(t) \qquad [1]$$

This is also correct with "$\leq$" in place of "<".

(i)(b) ***Verify that this function has these properties***

The presence of the exponential function ensures that this function is strictly positive. [½]

Since $A(t)$ is strictly positive, the supermartingale property is equivalent to:

$$\frac{E_P[A(T) \mid F_t]}{A(t)} < 1 \qquad [½]$$

With the definition given for $A(t)$, the left-hand side is:

$$\begin{aligned} LHS &= \frac{E_P[A(T) \mid F_t]}{A(t)} \\ &= e^{0.05t - 0.02W(t)} E_P\left[e^{-0.05T + 0.02W(T)} \,\middle|\, F_t \right] \\ &= e^{-0.05(T-t)} E_P\left[e^{0.02[W(T) - W(t)]} \,\middle|\, F_t \right] \end{aligned} \qquad [1]$$

Because of the independent increments property of Brownian motion, we can drop the F_t. We can then use the fact that $W(T) - W(t) \sim N(0, T-t)$ under P to evaluate the expectation on the right-hand side, which corresponds to an MGF based on a normal distribution. Using the formula given on page 11 of the Tables, we get:

$$LHS = e^{-0.05(T-t)} e^{½(0.02)^2 (T-t)} = e^{-0.0498(T-t)} \qquad [½]$$

When $t < T$ (which we have assumed throughout), this is indeed less than 1. So the supermartingale property is satisfied. [½] (i)(c) ***State why the supermartingale property is required***

The supermartingale property is equivalent to:

$$\frac{E_P[A(T)\,|\,F_t]}{A(t)} < 1$$

The left-hand side matches the formula for the bond price $B(t,T)$. So this property ensures that the price of a zero-coupon bond is always less than 1. This is equivalent to prohibiting negative interest rates. [1]

(i)(d) ***Name of the process***

The process $A(t)$ is a state price deflator. [1]

[Total 7]

(ii) ***Stochastic differential equation for A(t)***

We can write:

$$A(t) = e^{X(t)}$$

where $X(t) = -0.05t + 0.02W(t)$, so that $dX(t) = -0.05dt + 0.02dW(t)$. [1]

So, using a Taylor Series expansion, we can write:

$$\begin{aligned} dA(t) = d\left[e^{X(t)}\right] &= e^{X(t)}dX(t) + \tfrac{1}{2}e^{X(t)}[dX(t)]^2 \\ &= A(t)\left\{dX(t) + \tfrac{1}{2}[dX(t)]^2\right\} \end{aligned}$$ [1]

Substituting the SDE for $X(t)$ gives:

$$dA(t) = A(t)\left\{-0.05dt + 0.02dW(t) + \tfrac{1}{2}[-0.05dt + 0.02dW(t)]^2\right\}$$

Simplifying using the 2×2 multiplication grid given in the Core Reading, we get:

$$\begin{aligned} dA(t) &= A(t)\left\{-0.05dt + 0.02dW(t) + \tfrac{1}{2}(0.02)^2 dt\right\} \\ &= A(t)\left\{-0.0498dt + 0.02dW(t)\right\} \end{aligned}$$ [1]

So, in this case, the drift and volatility coefficients are:

$$\mu_A(t) = -0.0498 \quad \text{and} \quad \sigma_A(t) = 0.02$$ [1]

[Total 4]

(iii)(a) ***Formula for $B(t,T)$***

We have already evaluated the formula for $B(t,T)$ in part (i)(b), which gave:

$$B(t,T) = \frac{E_P[A(T)|F_t]}{A(t)} = e^{-0.0498(T-t)}$$ [1]

(iii)(b) ***Expressions for $f(t,T)$ and $R(t,T)$***

The instantaneous forward rate is:

$$f(t,T) = -\frac{\partial}{\partial T}\log B(t,T) = -\frac{\partial}{\partial T}[-0.0498(T-t)] = 0.0498$$

ie a constant rate of 4.98%. [1]

The spot rate therefore also takes a constant value of 4.98%. [1]

(iii)(c) ***One problem with this model***

A model with constant interest rates over all terms is not arbitrage-free. This would be a serious problem if the model was used in practical applications. [1]

[Total 4]

(iv)(a) ***Price of a zero-coupon bond***

According to this model, the price at time 5 (or indeed, at *any* time) of a 10-year zero-coupon bond is:

$$B(5,15) = e^{-0.0498(15-5)} = e^{-0.498} = 0.6077$$

ie 60.77 per 100 nominal. [2]

(iv)(b) ***Price of a 5% annual coupon bond***

The price of a 5% annual coupon bond per 100 nominal is:

$$P = 5[v + v^2 + \cdots + v^{10}] + 100v^{10}$$

where $v = e^{-0.0498} = 0.95142$.

Evaluating the sum as a geometric progression, we get:

$$P = 5v\left(\frac{1-v^{10}}{1-v}\right) + 60.77 = 5(7.682) + 60.77 = 99.18$$ [2]

[Total 4]

Solution 4.14

(i) ***Meaning of default-free***

A bond is default-free if the stream of payments due from the bond will definitely be paid in full and on time. [1]

(ii) ***Outcomes of a default***

The outcome of a default may be that the contracted payment stream is:

- rescheduled [½]
- cancelled by the payment of an amount which is less than the default-free value of the original contract [½]
- continues but at a reduced rate [½]
- totally wiped out. [½]

[Total 2]

(iii) ***Types of credit events***

A credit event is an event that will trigger the default of a bond and includes the following:

- failure to pay either capital or a coupon [½]
- loss event [½]
- bankruptcy [½]
- rating downgrade of the bond by a rating agency such as Standard and Poor's or Moody's. [½]

[Total 2]

A "loss event" is when it becomes clear that the borrower is not going to make a full payment on time, eg if they announce that they have suspended future payments because of cashflow problems or if a court rules that they are not legally required to pay the full amount.

(iv) ***Meaning of recovery rate***

In the event of a default, the fraction of the defaulted amount that can be recovered through bankruptcy proceedings or some other form of settlement is known as the recovery rate. [1]

Solution 4.15

Approaches to modelling credit risk include structural models, reduced form models and intensity-based models. [1]

Structural models are explicit models for a company issuing both shares and bonds. They aim to link default events explicitly to the fortunes of the issuing company. Whilst the models are simple and less realistic than the others, studying them does give an insight into the nature of default and the interaction between bond holders and shareholders. An example is the Merton model. [2]

Reduced-form models are statistical models that use observed market statistics such as credit ratings, as opposed to specific data relating to the company. The credit ratings agencies will use detailed data specific to the issuing company when setting the rating, and will regularly review the data to ensure the rating remains appropriate. [1]

Reduced-form models use market statistics along with data on the default-free market to model the movement of the credit rating of the bonds issued by a company over time. The output of such a model is a distribution of the time to default of the bond. [1]

An intensity-based model is a particular type of continuous-time reduced form model. It typically models the "jumps" between different states (usually credit ratings) using transition intensities. [1]

[Total 6]

Solution 4.16

(i) ***Expressing the equity as a call option***

The Merton model values shares as call options on the company's assets with a strike price equal to the face value of Company X's debt. [½]

The equity value at time 10 will be:

$$\max\{F(10)-5,0\}$$

where $F(10)=$ total value of Company X at time 10. [1]

Company X will default on payment of the debt if the total value of its assets at time 10 is less than the promised debt repayment at that time. [½]

There are 400,000 shares, so the (theoretical) share price at time 10 in £m will be:

$$\frac{\max\{F(10)-5,0\}}{400,000}$$ [½]

An appropriate option pricing formula can then be used to value this "call option" at time 0. [½]

[Total 3]

(ii) ***Share price at time 10***

The share price at time 10 will be:

(a) $\dfrac{\max\{15-5,0\}}{400,000}=\dfrac{£10\text{m}}{400,000}=£25$ per share [1]

(b) Here the share price at time 10 will be 0 because the value of the outstanding debt exceeds the total value of the company. [1]

[Total 2]

Solution 4.17

(i) ***Probability the bond does not default between times 5 and 10***

The probability that the bond does not default between times 5 and 10 is:

$$\exp\left\{-\int_5^{10} \lambda(t)\,dt\right\} = \exp\left\{-\frac{1}{500}\int_5^{10}\left(5+20t-t^2\right)dt\right\}$$

$$= \exp\left\{-\frac{1}{500}\left[5t+10t^2-\frac{t^3}{3}\right]_5^{10}\right\}$$

$$= \exp\left\{-\frac{1}{500}\left[716\tfrac{2}{3}-233\tfrac{1}{3}\right]\right\}$$

$$= e^{-0.9667} = 0.3803$$ [2]

(ii) ***Incorporating unforeseen events***

Unforeseen events can be considered as random and so a stochastic approach would be needed. [½]

By using a stochastic approach, $\lambda(t)$ can be allowed to vary with company fortunes and other economic factors. [½]

This is done by introducing an additional stochastic process $X(t)$, which models a suitable economic indicator, such as interest rates. [½]

For example, a downturn in the economy may make default more likely and so $\lambda(t)$ could include appropriate allowance for this possibility. [½]

[Total 2]

Solution 4.18

(i) ***Theoretical price of the bonds***

The approach we will take is to value the shareholders' funds and then subtract this from the total value of the company to determine the bondholders' funds.

Under the Merton model, the shareholders in the company receive a payoff after 4 years equivalent to that from a call option with strike price equal to the amount to be repaid to the bondholders. [1]

The current value of the shareholding can be assessed using the Black-Scholes formula for the value of a call option, with parameters:

$$S_0 = 7.5,\ K = 4,\ \sigma = 30\%,\ r = 2\%,\ T - t = 4,\ q = 0 \qquad [½]$$

Letting $E(0)$ represent the value of the shareholding at time 0:

$$E(0) = 7.5\Phi(d_1) - 4e^{-0.02\times 4}\Phi(d_2)$$

where:

$$d_1 = \frac{\ln\left(7.5\big/4\right) + \left(0.02 + \frac{1}{2}(0.3)^2\right)\times 4}{0.3\times 2} = 1.4810 \quad \Rightarrow \quad \Phi(d_1) = 0.93069$$

and: $d_2 = d_1 - 0.3\times 2 = 0.8810 \quad \Rightarrow \quad \Phi(d_2) = 0.81084$

So:

$$E(0) = 7.5\times 0.93069 - 4e^{-0.02\times 4}\times 0.81084 = £3.986 \text{ million} \qquad [1½]$$

Therefore the value of £4 million nominal of bonds at time 0, $B(0)$, is:

$$B(0) = 7.5 - 3.986 = £3.514 \text{ million} \qquad [½]$$

The theoretical price of £100 nominal of these bonds is:

$$\frac{3.514}{4}\times 100 = £87.85 \qquad [½]$$

[Total 4]

Another way of doing this question is to note that the payoff on the bonds in 4 years' time is:

$$\min(4, F(4))$$

where $F(t)$ *is the value of the company's assets at time* t*.*

This can be re-expressed as:

$$4 - \max(4 - F(4), 0)$$

where $\max(4 - F(4), 0)$ *is the payoff on a 4-year put option on the company's assets with strike price of 4.*

The value of this payoff at time 0 is:

$$B(0) = 4e^{-0.02\times 4} - p_0$$

where:

$$p_0 = 4e^{-0.02\times 4}\Phi(-d_2) - 7.5\Phi(-d_1)$$

As before, $d_1 = 1.4810$ *and* $d_2 = 0.8810$*, so:*

$$\Phi(-d_1) = \Phi(-1.4810) = 1 - \Phi(1.4810) = 0.06931$$

and:

$$\Phi(-d_2) = \Phi(-0.8810) = 1 - \Phi(0.8810) = 0.18916$$

This gives:

$$p_0 = 4e^{-0.02\times 4} \times 0.18916 - 7.5 \times 0.06931 = 0.17866$$

so:

$$B(0) = 4e^{-0.02\times 4} - 0.17866 = 3.514$$

as before.

(ii) ***Risk-neutral probability of default***

In the standard Black-Scholes formula for the price of a call option, $\Phi(d_2)$ represents the risk-neutral probability that the option will be exercised, or, equivalently, the risk-neutral probability that the share price at expiry exceeds the strike price. Under the Merton model approach, the call option replicates the shareholders' position and is exercised if the shareholders repay the bondholders in full. [1]

So, $\Phi(d_2)$ is equal to the probability that the bondholders are repaid in full, or, equivalently, the probability that the company does not default. This means that the probability of default is:

$$1-\Phi(d_2) \qquad [1]$$

In this case, the probability of default is:

$$1-\Phi(0.8810)=1-0.81084=0.18916 \qquad [1]$$

[Total 3]

Alternatively, this can be derived from first principles.

Under the assumptions of the Black-Scholes formula, the value of the company's assets at time t, $F(t)$, given $F(0)$, follows a lognormal distribution:

$$F(t)\big|F(0)\sim \log N\left(\ln F(0)+(r-\tfrac{1}{2}\sigma^2)t,\sigma^2 t\right)$$

Here:

$$F(4)\big|F(0)\sim \log N\left(\ln 7.5+(0.02-\tfrac{1}{2}0.3^2)\times 4,0.3^2\times 4\right)$$
$$\Rightarrow\quad F(4)\big|F(0)\sim \log N\left(\ln 7.5-0.1,0.36\right)$$

The company will default on the bonds if the value of the company at time 4, $F(4)$, is less than the amount to be repaid of £4 million.

So, the probability of default is:

$$\begin{aligned} P(F(4)<4) &= P(\ln F(4)<\ln 4) \\ &= P\left(Z<\frac{\ln 4-(\ln 7.5-0.1)}{\sqrt{0.36}}\right) \\ &= P(Z<-0.8810) \\ &= 1-\Phi(0.8810) \\ &= 0.18916 \end{aligned}$$

Solution 4.19

(i) ***Value of δ***

The general formula for the price of a zero-coupon bond under the two-state model for credit risk using a risk-neutral probability measure is:

$$B(t,T)=e^{-r(T-t)}\left[1-(1-\delta)\left(1-\exp\left(-\int_t^T \tilde{\lambda}(s)\,ds\right)\right)\right] \qquad [1]$$

For the bond issued by Company A:

$$\begin{aligned} B_A &= e^{-0.05\times 3}\left[1-(1-\delta)\left(1-\exp\left(-\int_0^3 0.0148\,ds\right)\right)\right] \\ &= e^{-0.15}\left[1-(1-\delta)\left(1-e^{-0.0444}\right)\right] \end{aligned} \qquad [1]$$

For the bond issued by Company B:

$$\begin{aligned} B_B &= e^{-0.05\times 3}\left[1-(1-\delta)\left(1-\exp\left(-\int_0^3 0.01s^2\,ds\right)\right)\right] \\ &= e^{-0.15}\left[1-(1-\delta)\left(1-\exp\left(-\left[\frac{0.01}{3}s^3\right]_0^3\right)\right)\right] \\ &= e^{-0.15}\left[1-(1-\delta)\left(1-e^{-0.09}\right)\right] \end{aligned} \qquad [1]$$

The credit spread on a zero-coupon bond is the difference between the yield on the bond and the yield on a similar bond issued by the government, which we take here to be the risk-free force of interest of 5% *pa*.

If C_i is the credit spread on the zero-coupon bond issued by Company i and r_i is the continuously-compounded yield on the zero-coupon bond issued by Company i, then:

$$C_A = r_A - 0.05 \quad \Rightarrow \quad r_A = C_A + 0.05$$
$$C_B = r_B - 0.05 = 2C_A \quad \Rightarrow \quad r_B = 2C_A + 0.05 \qquad [½]$$

We can express the price of each zero-coupon bond in terms of the continuously-compounded yield on the bond, so:

$$B_A = e^{-3r_A} = e^{-3(C_A+0.05)}$$
$$B_B = e^{-3r_B} = e^{-3(2C_A+0.05)} \qquad [½]$$

This gives the simultaneous equations:

$$e^{-3(C_A+0.05)} = e^{-0.15}\left[1-(1-\delta)\left(1-e^{-0.0444}\right)\right]$$
$$e^{-3(2C_A+0.05)} = e^{-0.15}\left[1-(1-\delta)\left(1-e^{-0.09}\right)\right]$$

Cancelling the $e^{-0.15}$ terms gives:

$$e^{-3C_A} = 1-(1-\delta)\left(1-e^{-0.0444}\right) \qquad (1)$$
$$e^{-6C_A} = 1-(1-\delta)\left(1-e^{-0.09}\right) \qquad (2) \qquad [1]$$

Squaring Equation (1) and substituting it into Equation (2):

$$\left(1-(1-\delta)\left(1-e^{-0.0444}\right)\right)^2 = 1-(1-\delta)\left(1-e^{-0.09}\right)$$

Expanding the left-hand side:

$$1-2(1-\delta)\left(1-e^{-0.0444}\right)+(1-\delta)^2\left(1-e^{-0.0444}\right)^2 = 1-(1-\delta)\left(1-e^{-0.09}\right)$$
$$\Rightarrow -2(1-\delta)\left(1-e^{-0.0444}\right)+(1-\delta)^2\left(1-e^{-0.0444}\right)^2 = -(1-\delta)\left(1-e^{-0.09}\right)$$

Cancelling $(1-\delta)$ on both sides, since $0<\delta<1$, and solving for δ:

$$-2\left(1-e^{-0.0444}\right)+(1-\delta)\left(1-e^{-0.0444}\right)^2=-\left(1-e^{-0.09}\right)$$

$$\Rightarrow \quad 1-\delta=\frac{2\left(1-e^{-0.0444}\right)-\left(1-e^{-0.09}\right)}{\left(1-e^{-0.0444}\right)^2}=\frac{0.0007887}{0.0018861}=0.418164$$

$$\Rightarrow \quad \delta=0.5818$$ [2]

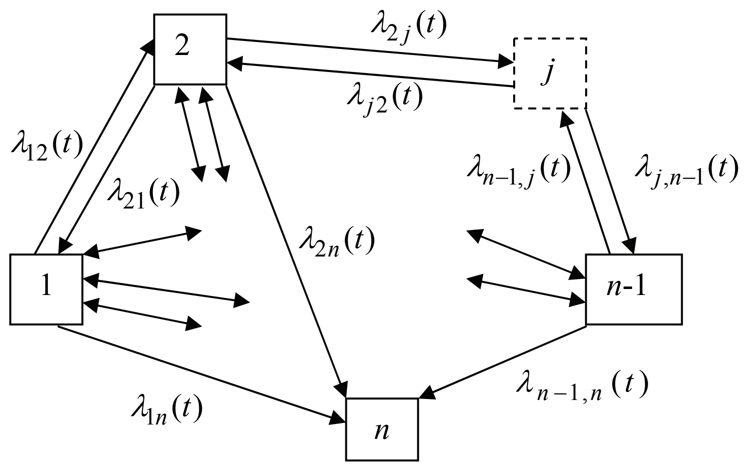
[Total 7]

(ii) ***Jarrow-Lando-Turnbull model***

Instead of the simple default / no default two-state model, a more general model has been developed by Jarrow, Lando and Turnbull, in which there are n states. The n states relate to $n-1$ possible credit ratings for a non-defaulted company, and one default state. [1]

Transitions are possible between all states, except for the default state, which is absorbing (*ie* once a company has entered the default state, it cannot leave it). [½]

If the transition rate from state i to state j at time t is denoted $\lambda_{ij}(t)$, where $\lambda_{ij}(t)$ is assumed to be deterministic, then the model can be represented by the following diagram:



[1½]

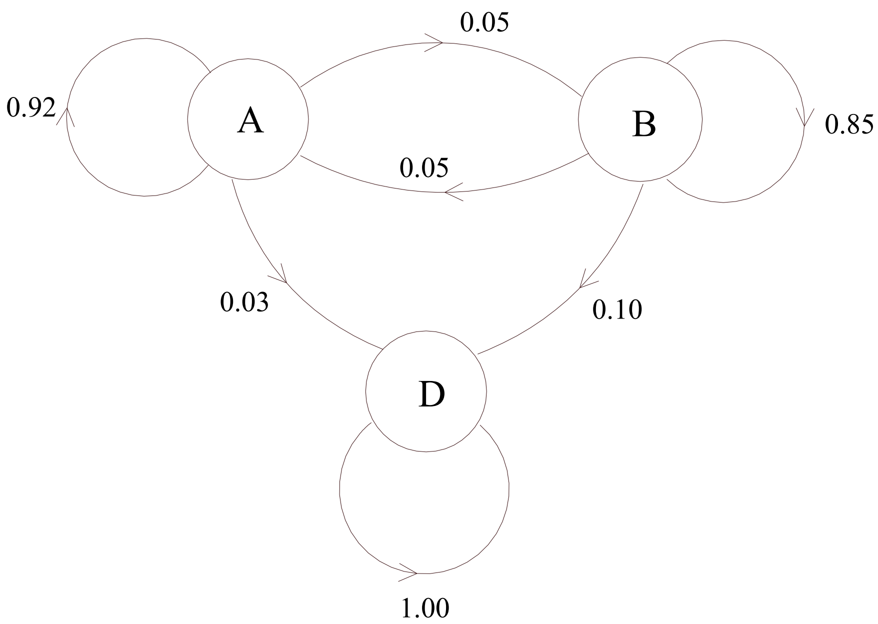
[Total 3]

Solution 4.20

This is Question A4 from the Subject CT4 exam of September 2006.

(i) ***Probability A never rated B in the future***

We have the following diagram of the one-year transition probabilities:



A company that is never rated B in the future will

(a) remain in State A for some period of time, and

(b) will then move to State D and remain there.

So we can sum over all the times at which the single transition from State A to State D can take place. This gives us the following expression:

$$0.03 + 0.92 \times 0.03 + (0.92)^2 \times 0.03 + (0.92)^3 \times 0.03 + \cdots \qquad [1]$$

This is an infinite geometric progression, whose sum is:

$$\frac{0.03}{1-0.92} = 0.375$$

So the probability that a company is never rated B in the future is 0.375. [1]

[Total 2]

(ii)(a) ***Second-order transition probabilities***

The second-order transition probabilities are given by:

$$\mathbf{X}^2 = \begin{pmatrix} 0.92 & 0.05 & 0.03 \\ 0.05 & 0.85 & 0.1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.92 & 0.05 & 0.03 \\ 0.05 & 0.85 & 0.1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0.8489 & 0.0885 & 0.0626 \\ 0.0885 & 0.7250 & 0.1865 \\ 0 & 0 & 1 \end{pmatrix} \quad [1]$$

(ii)(b) ***Expected number of defaults***

The probability that a company rated A at time zero is in State D at time 2 is 0.0626. So the expected number of companies in this state out of 100 is 6.26. [1]

[Total 2]

(iii) ***Expected number of defaults***

For this manager we use the original matrix $\mathbf{X}$. After one year, the expected number of companies in each state will be:

$$\begin{pmatrix} 100 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0.92 & 0.05 & 0.03 \\ 0.05 & 0.85 & 0.1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 92 & 5 & 3 \end{pmatrix} \quad [1]$$

If the five state B's are replaced with State A's and the process repeated, we have:

$$\begin{pmatrix} 97 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0.92 & 0.05 & 0.03 \\ 0.05 & 0.85 & 0.1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 89.24 & 4.85 & 5.91 \end{pmatrix}$$

So the expected number of defaults by the end of the second year under this arrangement is 5.91. [1]

[Total 2]

Alternatively we could have used the revised transition matrix:

$$X = \begin{pmatrix} 0.97 & 0 & 0.03 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and calculated X^2 *as before.*

Note that the question does not say that the companies in default at the end of the first year are replaced by grade A companies in the second year. If it had done so, then there would have been three expected defaults each year, and the answer would be 6.

(iv) ***Comment***

The downgrade trigger strategy will reduce the expected number of defaults, as we have seen. However, the return on the portfolio will also be a function of the yields on the debt. Companies rated B are likely to have bonds with a higher yield (because of the higher risk), so excluding these may in fact reduce the yield on the portfolio. [1½]

Also, the actual number of defaults may not match the expected number. The return depends on the actual progress of the portfolio, rather than the expected outcome. [½]
[Total 2]

Part 5 – Questions

Question 5.1

A zero-coupon bond (ZCB) promises to pay a return of 8% when it matures in a year's time, assuming that it doesn't default. In the event of default, which has a probability of 10%, it will yield a return of either 6% or 0%, with equal likelihood.

A 1-year risk-free government ZCB provides a return of 5%.

(i) Calculate the following measures of investment risk for the risky ZCB:

(a) downside semi-variance

(b) shortfall probability relative to the risk-free rate

(c) the expected shortfall relative to the risk-free rate conditional on a shortfall occurring. [4]

A bond analyst is going to use a two-state intensity-based model to derive risk-neutral transition intensities.

(ii) (a) State the general risk-neutral pricing formula for a ZCB subject to default risk in a two-state intensity-based model with a deterministic transition intensity. Define all notation used.

(b) Starting from the formula in (ii)(a), derive the corresponding formula for the risk-neutral transition intensity function in terms of the ZCB price. [6]

[Total 10]

Question 5.2

(i) Explain what is meant by a short rate model of interest rates and how such models can be used to price interest rate derivatives. [3]

(ii) Explain what is meant by a one-factor short rate model of interest rates. [2]

[Total 5]

Question 5.3

An exotic forward provides a payoff equal to the square root of the share price at maturity time T less the square root of the delivery price, K.

(i) (a) Assuming that the Black-Scholes assumptions apply, use risk-neutral pricing to derive a formula for the price at time t of the forward on a non-dividend-paying share.

(b) Derive the corresponding formula for the vega of the forward. [6]

(ii) (a) Explain why an investor might want to vega hedge their portfolio.

(b) Use the result that $S_t \phi(d_1) - Ke^{-r(T-t)} \phi(d_2) = 0$, where d_1 and d_2 are defined as on page 47 in the *Tables*, to show that the formula for the vega of a European call option is $\nu_{call} = S_t \phi(d_1)\sqrt{T-t}$ [6]

The current price of the share is \$1, which is also the delivery price of the forward. The risk-free force of interest is 5%, the volatility of the underlying share, which pays no dividends, is 20% and the forward has one year to delivery.

(iii) An investor has a long position in 1,000 exotic forwards. Find the vega-hedged portfolio for this position involving standard European call options on the underlying share and also the underlying share itself. [8]

[Total 20]

Question 5.4

(i) Describe briefly the Efficient Markets Hypothesis. [3]

(ii) Explain the implications of the Efficient Markets Hypothesis for investment trading strategies. [3]

(iii) Explain why investors will still wish to have as much information as possible concerning a company and its securities before investing in it even if the Efficient Markets Hypothesis applies? [2]

[Total 8]

Question 5.5

(i) (a) State the equation of the security market line and, assuming that the market portfolio offers a return in excess of the risk-free rate, use it to derive the betas of the market portfolio and the risk-free asset.

(b) Draw a diagram of the security market line relationship.

(c) What does the security market line indicate about the relationship between risk and return? [9]

(ii) (a) Derive the capital market line relationship and comment briefly upon its applicability.

(b) Briefly interpret each of the terms in the relationship. [6]

[Total 15]

Question 5.6

The table below shows the closing prices (represented by letters) on a particular day for a series of European call options with different strike prices and expiry dates on a particular risky non-dividend-paying security.

Call option prices	*Strike price* 125	150
3 months	W	Y
6 months	X	Z

(i) Write down, with reasons, the strictest inequalities that can be deduced for the relative values of W, X, Y, Z, assuming that the market is arbitrage-free. (Your inequalities should not involve any other quantities.) [4]

(ii) Write down numerical values for a lower and an upper bound for X, given that the current share price is 120 and the continuously-compounded annual risk-free interest rate is 6%. [2]

[Total 6]

Question 5.7

(i) Explain the difference between a recombining and a non-recombining binomial tree. [2]

(ii) A researcher is using a two-step binomial tree to determine the value of a 6-month European put option on a non-dividend-paying share. The put option has a strike price of 450.

During the first 3 months it is assumed that the share price of 400 will either increase by 10% or decrease by 5% and that the continuously compounded risk-free rate (per 3 months) is 0.01. During the following 3 months it is assumed that the share price will either increase by 20% or decrease by 10% and that the risk-free rate is 0.015 (per 3 months).

Calculate the value of the put option. [6]

(iii) The researcher is considering subdividing the option term into months. Explain the advantages and disadvantages of this modification of the model and suggest an alternative model based on months that might be more efficient numerically. [5]

[Total 13]

Question 5.8

A binomial lattice is used to model the price of a non-dividend-paying share up to time T. The interval $(0,T)$ is subdivided into a large number of intervals of length $\delta t = T/n$.

It is assumed that, at each node in the lattice, the share price is equally likely to increase by a factor u or decrease by a factor d, where $u = e^{\mu\delta t+\sigma\sqrt{\delta t}}$ and $d = e^{\mu\delta t-\sigma\sqrt{\delta t}}$. The movements at each step are assumed to be independent.

(i) Show that, if the share price makes a total of X_n "up jumps", the share price at time T will be:

$$S_T = S_0 \exp\left\{\mu T + \sigma\sqrt{T}\left(\frac{2X_n - n}{\sqrt{n}}\right)\right\}$$

where S_0 denotes the initial share price. [4]

(ii) Write down the distribution of X_n and state how this distribution can be approximated when n is large. [2]

(iii) Hence determine the asymptotic distribution of $\dfrac{S_T}{S_0}$ for large n. [4]

[Total 10]

Question 5.9

Consider the single index model of investment returns in which for any security i:

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

where $E(\varepsilon_i) = 0$, $E(\varepsilon_i \varepsilon_j) = 0$ for $i \neq j$, $E(R_m \varepsilon_i) = 0$ and R_m is the return on the market.

(i) Assuming that this model applies, derive expressions for the mean investment return on security i, and the mean investment return on a portfolio P, containing n securities, with a proportion x_i invested in security i. [3]

(ii) Show that $C_{iP} = \sum_{j=1}^{n} x_j C_{ij}$, where C_{iP} and C_{ij} are the covariance of investment returns between security i and portfolio P and securities i and j respectively. [2]

(iii) State a general expression for the variance σ_P^2 of portfolio P in terms of the covariances C_{ij}. [1]

(iv) Use your results from (ii) and (iii) to show that:

$$\beta_{iP} = \frac{\partial \sigma_P}{\partial x_i} \frac{1}{\sigma_P}$$

where $\beta_{iP} = \dfrac{C_{iP}}{\sigma_P^2}$ and comment briefly on this result. [7]

[Total 13]

Part 5 – Solutions

Solution 5.1

(i)(a) ***Downside semi-variance***

The expected return on the risky ZCB is equal to:

$$0.9 \times 8\% + 0.05 \times 6\% + 0.05 \times 0\% = 7.5\% \qquad [1]$$

Consequently, the downside semi-variance is given by:

$$(7.5-6)^2 \times 0.05 + (7.5-0)^2 \times 0.05 = 2.925\%\% \qquad [1]$$

(i)(b) ***Shortfall probability relative to risk-free rate***

The only outcome that provides a lower return than the risk-free rate of 5% is when the ZCB defaults and returns 0%. The shortfall probability is therefore 5%. [1]

(i)(c) ***Expected shortfall conditional on a shortfall occurring***

This is equal to the unconditional expected shortfall divided by the shortfall probability.

Again, there is only one outcome that is worse than the risk-free rate of 5% and so this is equal to:

$$\frac{(5-0) \times 0.05}{0.05} = 5\% \qquad [1]$$

[Total 4]

(ii)(a) ***General risk-neutral pricing formula for a ZCB***

$$B(t,T) = e^{-r(T-t)} \left[1 - (1-\delta) \left\{ 1 - \exp\left(-\int_t^T \tilde{\lambda}(s)\,ds \right) \right\} \right] \qquad [1]$$

where:

- t and T are the current time and the maturity date of the ZCB
- r is the constant risk-free force of interest
- δ is the assumed (constant) recovery rate
- $\tilde{\lambda}(s)$ is the risk-neutral transition intensity. [1]

(ii)(b) ***Derive formula for risk-neutral transition intensity function***

Multiplying both sides of the ZCB pricing formula by $e^{r(T-t)}$ and then multiplying out the right-hand side, gives:

$$\begin{aligned} e^{r(T-t)}B(t,T) &= 1-\left\{1-\exp\left(-\int_t^T \tilde{\lambda}(s)ds\right)\right\}+\delta\left\{1-\exp\left(-\int_t^T \tilde{\lambda}(s)ds\right)\right\} \\ &= \exp\left(-\int_t^T \tilde{\lambda}(s)ds\right)+\delta-\delta\exp\left(-\int_t^T \tilde{\lambda}(s)ds\right) \\ &= (1-\delta)\exp\left(-\int_t^T \tilde{\lambda}(s)ds\right)+\delta \qquad [1] \end{aligned}$$

Moving the constant δ to the left-hand side, then taking logs, we get:

$$e^{r(T-t)}B(t,T)-\delta=(1-\delta)\exp\left(-\int_t^T \tilde{\lambda}(s)ds\right)$$

$$\Rightarrow \qquad \log\left[e^{r(T-t)}B(t,T)-\delta\right]=\log(1-\delta)-\int_t^T \tilde{\lambda}(s)ds \qquad [1]$$

Differentiating both sides with respect to T then gives:

$$\frac{\partial}{\partial T}\log\left[e^{r(T-t)}B(t,T)-\delta\right]=-\tilde{\lambda}(T) \qquad [1]$$

Remember that differentiating an integral (with respect to the upper limit) takes you back to the original function.

Finally, renaming the variable T as s, swapping the two sides and flipping the signs gives:

$$\tilde{\lambda}(s)=-\frac{\partial}{\partial s}\log\left[e^{r(s-t)}B(t,s)-\delta\right] \qquad [1]$$

[Total 6]

Solution 5.2

(i) ***Short rate model of interest rates***

Short rate models are used to model the term structure of interest rates as a function of the short rate $r(t)$. [½]

In theory, this is the interest rate that applies over the next instant of time, of length dt. [½]

In practice, the short rate is an overnight rate, *ie* the force of interest earned when money is lent today and received back with interest the following day. [½]

$r(t)$ itself is usually assumed to be a diffusion or Ito process, with a stochastic differential equation of the form:

$$dr(t) = a(t, r(t))\ dt + b(t, r(t))\ d\tilde{W}_t$$

where $\tilde{W}_t$ is a standard Brownian motion under the risk-neutral probability measure, Q. [1]

The price at time t of an interest rate derivative that pays X_T at maturity date $T > t$ can then be found from:

$$V_t = E_Q\left[\exp\left(-\int_t^T r(u)du\right)X_T \,\middle|\, F_t\right] \qquad [1]$$

[Maximum 3]

(ii) ***Explain what is meant by a one-factor short rate model***

A *one-factor short rate model* is one in which the short rate, and hence the term structure as a whole, is assumed to be influenced by a single source of randomness. [½]

The prices of all bonds (of all maturities) and interest-rate derivatives must therefore move together. [½]

The single source of randomness is typically assumed to be a standard Brownian motion process. [½]

The short rate is therefore usually modelled as an Ito process. [½]
[Total 2]

Solution 5.3

(i)(a) ***Derive formula for price of forward***

The general risk-neutral formula for pricing a derivative at time $t < T$ is:

$$V_t = e^{-r(T-t)} E_Q\left[X_T \middle| F_t\right] \qquad [½]$$

In this instance, the payoff function of the derivative is:

$$X_T = S_T^{½} - K^{½}$$

So, the price will be given by:

$$V_t = e^{-r(T-t)} E_Q\left[S_T^{½} - K^{½} \middle| F_t\right] \qquad [1]$$

Black-Scholes assumes that under the risk-neutral measure Q, the underlying share price follows geometric Brownian motion with drift r and volatility σ and so:

$$\log S_T \sim N\left[\log S_t + \left(r - ½\sigma^2\right)(T-t), \sigma^2(T-t)\right] \qquad [½]$$

Hence, using the formula for the moments of a lognormal distribution from page 14 in the *Tables*, with $r = ½$, the price will be given by:

$$V_t = e^{-r(T-t)}\left[e^{½\left(\log S_t + (r - ½\sigma^2)(T-t)\right) + ½ \times ½^2 \sigma^2 (T-t)} - K^{½}\right] \qquad [1]$$

Note that the formula in the Tables also works for non-integer values of r.

This simplifies to:

$$V_t = S_t^{½} e^{-½\left(r + ¼\sigma^2\right)(T-t)} - K^{½} e^{-r(T-t)} \qquad [1]$$

[Total 4]

(i)(b) ***Formula for the vega of the forward***

The vega of a derivative with price f based on an underlying share with volatility σ is defined as:

$$\nu = \frac{\partial f}{\partial \sigma}$$ [½]

So, here:

$$\nu = -\tfrac{1}{4}\sigma(T-t)\, S_t^{1/2} e^{-\frac{1}{2}\left(r+\frac{1}{4}\sigma^2\right)(T-t)}$$ [1½]

[Total 2]

(ii)(a) ***Why vega hedge a portfolio?***

A vega-hedged portfolio is one whose overall vega, which is equal to the sum of the vegas of the constituent securities, is close to zero. [½]

Consequently, the value of such a portfolio will be relatively insensitive to changes in the volatility of the underlying share. [½]

An investor might therefore wish to vega hedge in order to:

- protect the value of a portfolio against (small) changes in the volatility of the underlying share. [½]
- compensate for the fact that the volatility of the share is unknown, as it cannot be observed directly. It is less important to have an accurate estimate of the volatility if vega is low and hence has little effect on the portfolio's value. [½]

[Total 2]

(ii)(b) ***Show that vega of European call option is $\nu_{call} = S_t\phi(d_1)\sqrt{T-t}$***

Differentiating the Black-Scholes formula for a European call option with respect to σ using the product and chain rules for differentiation gives:

$$\nu_{call} = S_t\phi(d_1)\frac{\partial d_1}{\partial \sigma} - Ke^{-r(T-t)}\phi(d_2)\frac{\partial d_2}{\partial \sigma}$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right)$ is the probability density function of the standard normal distribution (from page 11 of the *Tables*). [1]

Now:

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

Thus:

$$\frac{\partial d_2}{\partial \sigma} = \frac{\partial d_1}{\partial \sigma} - \sqrt{T-t} \qquad [½]$$

Hence:

$$\begin{aligned} \nu_{call} &= S_t\phi(d_1)\frac{\partial d_1}{\partial \sigma} - Ke^{-r(T-t)}\phi(d_2)\times\left[\frac{\partial d_1}{\partial \sigma} - \sqrt{T-t}\right] \\ &= \frac{\partial d_1}{\partial \sigma}\left[S_t\phi(d_1) - Ke^{-r(T-t)}\phi(d_2)\right] + Ke^{-r(T-t)}\phi(d_2)\sqrt{T-t} \qquad [1] \end{aligned}$$

So, using the result that:

$$S_t\phi(d_1) - Ke^{-r(T-t)}\phi(d_2) = 0$$

we have:

$$\begin{aligned} \nu_{call} &= \frac{\partial d_1}{\partial \sigma}\times 0 + Ke^{-r(T-t)}\phi(d_2)\sqrt{T-t} \\ &= Ke^{-r(T-t)}\phi(d_2)\sqrt{T-t} \qquad [1] \end{aligned}$$

Or equally, given that:

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

this can be written as:

$$\nu_{call} = S_t\phi(d_1)\sqrt{T-t} \qquad [½]$$

[Total 4]

(iii) ***Find vega-hedged portfolio***

Using the information given in the question, together with the formulae found earlier in the question, the price and vega of the forward are:

$$V_t = S_t^{\frac{1}{2}} e^{-\frac{1}{2}\left(r + \frac{1}{4}\sigma^2\right)(T-t)} - K^{\frac{1}{2}} e^{-r(T-t)}$$

$$= 1 \times e^{-\frac{1}{2}\left(0.05 + \frac{1}{4} \times 0.2^2\right) \times 1} - 1 \times e^{-0.05 \times 1}$$

$$= 0.019216 \qquad [1]$$

$$\nu = -\tfrac{1}{4}\sigma(T-t)\, S_t^{\frac{1}{2}} e^{-\frac{1}{2}\left(r + \frac{1}{4}\sigma^2\right)(T-t)}$$

$$= -\tfrac{1}{4} \times 0.2 \times 1 \times 1 \times e^{-\frac{1}{2}\left(0.05 + \frac{1}{4} \times 0.2^2\right) \times 1}$$

$$= -0.048522 \qquad [1]$$

Using the Black-Scholes formula on page 47 in the *Tables*:

$$d_1 = \frac{\log(1) + \left(0.05 + \frac{1}{2} \times 0.2^2\right) \times 1}{0.2 \times 1} = 0.35 \qquad [½]$$

$$d_2 = 0.35 - 0.2 \times 1 = 0.15 \qquad [½]$$

So, the price of the European call option is:

$$c_t = 1 \times \Phi(0.35) - 1 \times e^{-0.05 \times 1} \Phi(0.15)$$

$$= 0.63683 - 0.951229 \times 0.55962$$

$$= 0.104503 \qquad [1]$$

and its vega is equal to:

$$\nu_{call} = S_t \phi(d_1) \sqrt{T-t}$$

$$= 1 \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \times 0.35^2} \times 1$$

$$= 0.375240 \qquad [1]$$

Finally, the vega of a share is equal to zero. [½]

So, to vega hedge the long position in 1,000 forwards, we need to find the number of European calls x and shares y such that:

$$1,000 \times 0.019216 = 0.104503x + 1 \times y$$ [½]

$$-1,000 \times 0.048522 = 0.375240x + 0 \times y$$ [½]

Solving these equations gives:

$$x = -129.31$$ [½]

$$y = 32.729$$ [½]

ie we need to short sell 129.31 calls and buy 32.729 shares. [½]
[Total 8]

This portfolio will respond in a similar way to the exotic forward to (small) changes in the volatility of the underlying share. Strictly speaking, in order to vega hedge the exotic forward we need to take the opposite positions to those found above, ie buy 129.31 calls and short sell 32.729 shares.

Solution 5.4

(i) ***Describe briefly the Efficient Markets Hypothesis***

The Efficient Markets Hypothesis states that security markets *are* efficient, *ie* the price of every security fully reflects all available information and hence is equal to its true investment value. [1]

Three different forms of market efficiency have been distinguished, corresponding to different levels of information: [½]

- Strong form: market prices reflect all information, whether or not it is publicly available. [½]
- Semi-strong form: market prices reflect all publicly available information. [½]
- Weak form: market prices reflect all the information contained in historical price data. [½]

[Total 3]

(ii) ***Implications of the Efficient Markets Hypothesis***

The Efficient Markets Hypothesis implies that it is impossible, except by chance, to make abnormal profits using trading strategies that are based on only past share prices (weak form), publicly available information (semi-strong form) or any information (strong form). [1]

In practice, however, the definition has sometimes been refined to preclude the possibility of systematically higher returns after allowing for taxes and transaction costs. [1]

Market efficiency also implies that active investment management (which aims to enhance returns by identifying under- or over-priced securities) cannot be justified and consequently provides a rationale for passive investment management strategies, such as index tracking. [1]
[Total 3]

(iii) ***Information***

Even if markets are efficient, investors will still wish to have as much information as possible concerning a company and its securities in order to identify the characteristics of the shares, *eg* the volatility of returns, systematic risk, income and capital growth *etc.* An appreciation of these will enable the investor to make an informed decision whether or not to hold the security as part of a portfolio designed to meet their investment objectives. [2]

Solution 5.5

(i)(a) ***Equation of the security market line***

The security market line for any portfolio *P* is:

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

where:

- E_P is the expected return on portfolio *P*
- r is the risk-free rate of return
- E_M is the expected return on the market portfolio
- β_P is the beta of the portfolio with respect to the market portfolio [1]

The security market line holds for all securities and portfolios. Thus, applying it to the market portfolio gives:

$$E_M = r + (E_M - r)\beta_M$$

ie $$E_M(1-\beta_M) = r(1-\beta_M)$$

Given that $E_M \neq r$, it must be the case that $\beta_M = 1$. [1]

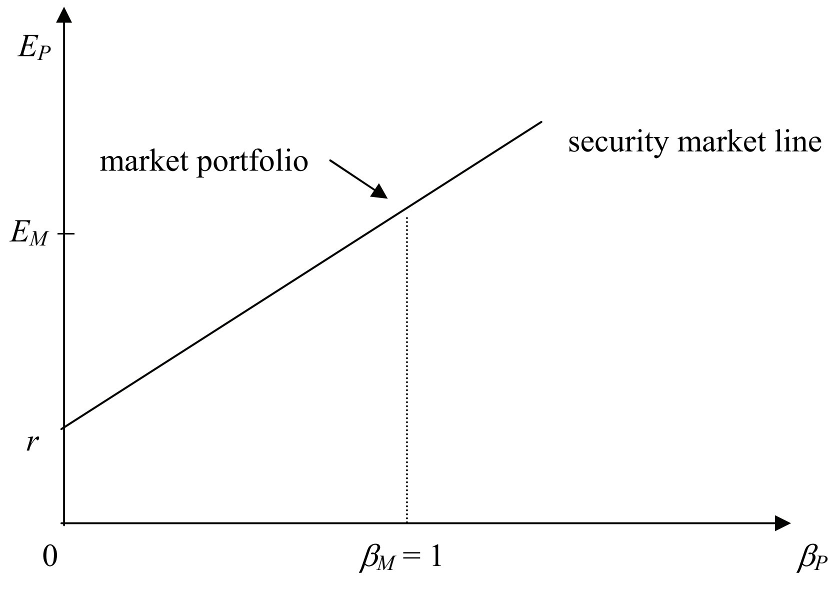
Similarly, applying the security market line relationship to the risk-free asset (with a beta of β_r) gives:

$$r = r + (E_M - r)\beta_r$$

ie $$0 = (E_M - r)\beta_r$$

Given that $E_M \neq r$, then it must be the case that $\beta_r = 0$, *ie* the risk-free asset has a beta of zero – which must be the case as it involves zero risk – systematic or otherwise. [1]

(i)(b) ***Diagram***



[2 for correct diagram]

(i)(c) ***What does the security market line indicate?***

The security market line relationship is of interest because:

- it enables us to determine the expected return on *any* asset or portfolio. This can be done if we can estimate the risk-free rate, the expected return on the market portfolio and the beta of the individual asset or portfolio. [1]

- it tells us that the expected return on any asset is equal to the risk-free rate plus a risk premium, which is a linear function of the systematic risk of the asset as measured by the beta factor. [1]

- it tells us that expected return does not depend on any other factors and in particular it is independent of the specific risk of an asset, which can be eliminated by diversification. [1]

The above results do of course depend upon the appropriateness or otherwise of the capital asset pricing model. [1]
[Total 9]

(ii)(a) ***Capital market line relationship***

The capital market line is the equation of the efficient frontier in (E,σ) space, which is a straight line. [½]

It passes through the risk-free asset with coordinates $(0,r)$ and the market portfolio, which has coordinates (σ_M, E_M). [½]

Thus, the gradient of the capital market line is equal to:

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

It has an intercept on the vertical axis at the risk-free rate r and consequently, for any efficient portfolio P, its equation must be:

$$E_P = r + \left(\frac{E_M - r}{\sigma_M}\right)\sigma_P$$ [½]

The capital market line relationship only holds for efficient portfolios – those for which there is no other portfolio that offers either a higher expected return for a given risk or a lower risk for a given expected return – assuming that the capital asset pricing model itself applies. Efficient portfolios are always combinations of the risk-free asset and the market portfolio. [1½]

[Total 3]

(ii)(b) ***Interpret the relationship***

r is the risk-free rate of return, *ie* the rate of return on a security that has a zero standard deviation of return. This is sometimes interpreted as the return on a Treasury bill. [1]

The quantity $\dfrac{E_M - R}{\sigma_M}$ is the *market price of risk*.

It can be interpreted as the extra expected return that can be gained by increasing the level of risk of an efficient portfolio by one unit. In this context, risk strictly means the standard deviation of investment returns. [1]

The second term in the relationship, $\left(\dfrac{E_M - R}{\sigma_M}\right)\sigma_P$, is known as the risk premium. It represents the additional return over and above the risk-free rate that can be obtained on a portfolio P by accepting risk, *ie* a non-zero portfolio standard deviation. [1]

[Total 3]

Solution 5.6

(i) ***Inequalities***

The value of an option is greater if the remaining life is longer.

So: $X > W$ and $Z > Y$ [1]

The value of a call option is smaller if the exercise price is greater.

So: $Y < W$ and $Z < X$ [1]

Also, the value of an option on a risky asset will be strictly positive. [1]

Combining these results, we have:

$$0 < Y < (W, Z) < X$$

ie W and Z must have values between Y and X. [1]

[Total 4]

Based on the information given in the question we can't determine the order of W and Z. They could go "either way".

(ii) ***Lower and upper bounds***

The lower bound for a European call option is given by the inequality:

$$c_t \geq S_t - Ke^{-r(T-t)}$$

Using the parameter values for X, this gives:

$$c_t \;\geq\; S_t - Ke^{-r(T-t)} \;=\; 120 - 125\,e^{-0.06\times 0.5} \;=\; -1.31$$

Since this is negative, we can take $c_t \geq 0$ instead. [1]

The upper bound for a call option is given by the inequality:

$$c_t \leq S_t \quad ie \quad c_t \leq 120$$ [1]

[Total 2]

Solution 5.7

(i) ***Explain the difference***

In a *recombining* binomial tree u and d, the proportionate increase and decrease in the underlying security price at each step, are assumed constant throughout the tree. As a result the security price after a specified number of up- and down- movements is the same, irrespective of the order in which the movements occurred. [1]

In a *non-recombining* binomial tree the values of u and d can change at each stage. As a result, each node in the tree will in general generate two new nodes, making the tree much larger than a combining tree. Consequently, an n-period tree will have 2^n, rather than just $n+1$, possible states at time n. [1]

[Total 2]

(ii) ***Calculate the value of the put option***

The risk-neutral probabilities of an up-movement at the first and second step are:

$$q_1 = \frac{e^{0.01} - 0.95}{1.10 - 0.95} = 0.40033 \qquad \text{and} \qquad q_2 = \frac{e^{0.015} - 0.90}{1.20 - 0.90} = 0.38371 \qquad [2]$$

And the put option payoffs at each of the four possible states at expiry are 0, 54, 0 and 108. [1]

Working backwards through the tree, we can then find the option value $V_1(1)$ following an up-step over the first 3 months from:

$$V_1(1)\,e^{0.015} = 0.38371 \times 0 + (1 - 0.38371) \times 54$$

ie $V_1(1) = 32.784$ [1]

Similarly, the option value $V_1(2)$ following a down-step over the first 3 months is found from:

$$V_1(2)\,e^{0.015} = 0.38371 \times 0 + (1 - 0.38371) \times 108$$

ie $V_1(1) = 65.568$ [1]

Finally, the current value V_0 of the put option is found from:

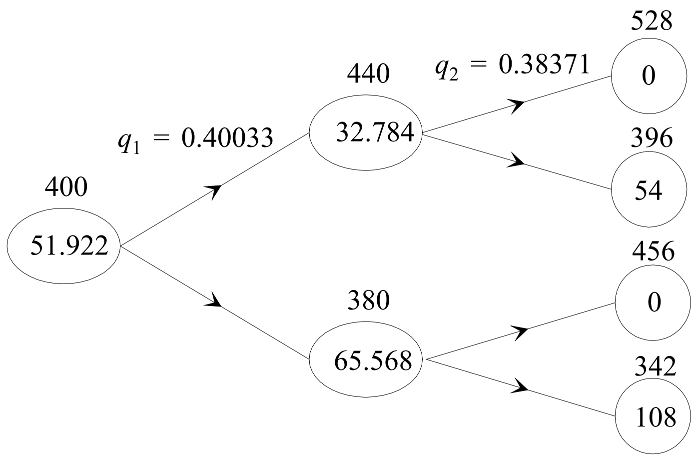
$$V_0 e^{0.01} = 0.40033 \times 32.784 + (1 - 0.40033) \times 65.568$$

ie $V_1(1) = 51.922$

So the value of the put option is 51.922. [1]

[Total 6]

In fact, the tree diagram looks like this:



(iii) ***Suggest an alternative model***

The researcher is proposing using a 6-step non-recombining tree.

This would result in a model that was much less crude than the two-step tree and should be capable of producing a more accurate valuation. [1]

However, there would be a lot more parameter values to specify (although some of these may be assumed to be equal). Appropriate values of u and d would be required for each branch of the tree and values of r for each month. [2]

The new tree would be big, having $2^6 = 64$ nodes in the expiry column. This would make the calculations prohibitive to do manually and would require more programming and calculation time on a computer. [1]

An alternative model that might be more efficient numerically would be a 6-step recombining tree (lattice), which would have only 7 nodes in the final column. [1]

[Total 5]

Solution 5.8

(i) ***Formula for S_T***

If the share price makes X_n up jumps (and hence $n-X_n$ down jumps) its value at time T will be:

$$S_T = S_0 \times u^{X_n} d^{n-X_n} \qquad [1]$$

Using the expressions given for u and d, this is:

$$S_T = S_0 \times e^{(\mu\delta t+\sigma\sqrt{\delta t})X_n} e^{(\mu\delta t-\sigma\sqrt{\delta t})(n-X_n)} \qquad [½]$$

Simplifying the exponent and separating the terms involving δt and $\sqrt{\delta t}$, we get:

$$S_T = S_0 \times \exp\left\{n\mu\delta t + (2X_n - n)\sigma\sqrt{\delta t}\right\} \qquad [1]$$

Using the fact that $\delta t = \dfrac{T}{n}$, we get the required result: [½]

$$S_T = S_0 \exp\left\{\mu T + \sigma\sqrt{T}\left(\frac{2X_n - n}{\sqrt{n}}\right)\right\} \qquad [1]$$

[Total 4]

(ii) ***Distribution of X_n***

Since there are n independent price movements, each equally likely to go up or down, X_n will have a binomial distribution with parameters n and ½. [1]

If n is large this can be approximated by a normal distribution with the same mean and variance, *ie* $N(\frac{1}{2}n, \frac{1}{4}n)$. [1]

[Total 2]

(iii) ***Asymptotic distribution***

Using the result in (ii), we know that asymptotically the distribution of the expression $\dfrac{2X_n - n}{\sqrt{n}}$ is normal with mean $\dfrac{2\times\frac{1}{2}n - n}{\sqrt{n}} = 0$ and variance $\dfrac{2^2\times\frac{1}{4}n}{n} = 1$, *ie* it is approximately standard normal. [2]

Taking logs of the result in part (i), we see that:

$$\log \frac{S_T}{S_0} = \mu T + \sigma \sqrt{T} \left(\frac{2X_n - n}{\sqrt{n}} \right)$$

Since the expression in brackets has a standard normal distribution, the RHS has a $N(\mu T, \sigma^2 T)$ distribution (approximately). Hence, $\frac{S_T}{S_0}$ has a lognormal distribution with parameters μT and $\sigma^2 T$. [2]

[Total 4]

Solution 5.9

(i) ***Mean investment returns***

The mean investment return on security *i*, E_i, is found from:

$$E_i = E[R_i] = E\left[\alpha_i + \beta_i R_m + \varepsilon_i\right]$$

Since α_i and β_i are both constants, $E\left[\alpha_i\right] = \alpha_i$ and $E\left[\beta_i\right] = \beta_i$. [½]

Also, by assumption, $E(\varepsilon_i) = 0$. [½]

So: $E_i = E\left[\alpha_i + \beta_i R_m + \varepsilon_i\right] = \alpha_i + \beta_i E_m$ [1]

where E_m is the mean return on the market.

For a portfolio *P* with portfolio weightings x_i, $i = 1, \ldots, n$, the mean investment return is given by:

$$E_P = E\left[\sum_{i=1}^{n} x_i R_i\right] = \sum_{i=1}^{n} x_i E_i = \sum_{i=1}^{n} x_i \left(\alpha_i + \beta_i E_m\right)$$ [½]

If we define:

$$\alpha_P = \sum_{i=1}^{n} x_i \, \alpha_i \quad \text{and} \quad \beta_P = \sum_{i=1}^{n} x_i \, \beta_i$$

then this can be written as:

$$E_P = \alpha_P + \beta_P \, E_m$$ [½]

[Total 3]

(ii) ***Covariance***

The covariance between security i and the portfolio is:

$$\begin{aligned} C_{iP} &= \text{cov}(R_i, R_P) \\ &= \text{cov}\left(R_i, \sum_{j=1}^{n} x_j R_j \right) \\ &= \sum_{j=1}^{n} x_j \, \text{cov}\left(R_i, R_j \right) \\ &= \sum_{j=1}^{n} x_j \, C_{ij} \end{aligned}$$ [2]

(iii) ***Variance of portfolio P***

We can write the portfolio variance in terms of the covariances, as follows:

$$\sigma_P^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j C_{ij}$$ [1]

(iv) ***Show the formula for β_{iP}***

Differentiating the left-hand side of the expression in (iii) with respect to x_i gives:

$$2\sigma_P \frac{\partial \sigma_P}{\partial x_i} \quad \dots (1)$$ [1]

The right-hand side of the expression in (iii) can be written as:

$$\sum_{i=1}^{n} x_i^2 C_{ii} + 2\sum_{i=1}^{n}\sum_{j\neq i}^{n} x_i x_j C_{ij}$$ [1]

Differentiating this with respect to x_i gives:

$$2x_i C_{ii} + 2\sum_{j\neq i}^{n} x_j C_{ij} = 2\sum_{j=1}^{n} x_j C_{ij}$$ [1½]

Using the result from (ii) above, this can be written as:

$$2\, C_{iP} \quad \ldots (2)$$ [½]

Hence, equating Equations (1) and (2) and cancelling the 2's gives:

$$\sigma_P \frac{\partial \sigma_P}{\partial x_i} = C_{iP}$$ [1]

Dividing both sides of this equation by σ_P^2:

$$\frac{1}{\sigma_P}\frac{\partial \sigma_P}{\partial x_i} = \frac{C_{iP}}{\sigma_P^2}$$ [½]

The left-hand side now matches the definition of β_{iP} given in the question. So we have:

$$\beta_{iP} = \frac{\partial \sigma_P}{\partial x_i}\frac{1}{\sigma_P}$$ [½]

The definition of β_{iP} given here is:

$$\beta_{iP} = \frac{C_{iP}}{\sigma_P^2} = \frac{\text{cov}(R_i, R_P)}{V_P}$$

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]

Subject CT8: Assignment X1

2015 Examinations

Time allowed: 2½ hours

Instructions to the candidate

1. *Please note that assignment marking is not included in the price of the Course Materials. Please purchase Series Marking or Marking Vouchers before submitting your script.*
2. *We only accept the current version of assignments for marking, ie you can only submit this assignment in the sessions leading to the 2015 exams.*
3. *Attempt all of the questions, leaving space in the margin and beginning your answer to each question on a new page.*
4. *Write in black ink using a medium-sized nib because we will be unable to mark illegible scripts.*
5. ***Leave at least 2cm margin on all borders.***
6. *Attempt the questions as far as possible under exam conditions.*
7. *You should aim to submit this script for marking by the recommended submission date. The recommended and deadline dates for submission of this assignment are listed on the summary page at the back of this pack and on our website at **www.ActEd.co.uk**.*

   Scripts received after the deadline date will not be marked, unless you are using a Marking Voucher. It is your responsibility to ensure that scripts reach ActEd in good time. ActEd will not be responsible for scripts lost or damaged in the post or for scripts received after the deadline date. If you are using Marking Vouchers, then please make sure that your script reaches us by the Marking Voucher deadline date to give us enough time to mark and return the script before the exam.

At the end of the assignment

If your script is being marked by ActEd, please follow the instructions on the reverse of this page.

In addition to this paper, you should have available actuarial tables and an electronic calculator.

Submission for marking

If you are submitting by **email**:

- complete the cover sheet, including the checklist
- scan your script and cover sheet (and Marking Voucher if applicable) to a pdf document, then email it to: **ActEdMarking@bpp.com**.

Please note the following:

- Please title the email to ensure that the subject and assignment are clear *eg* "CT8 Assignment X1 No. 12345", inserting your ActEd Student Number for 12345.
- The assignment should be scanned the **right way up** (so that it can be read normally without rotation) and as a single document. We cannot accept individual files for each page.
- Please set the resolution so that the script is legible and the resulting PDF is less than 3 MB in size. **The file size cannot exceed 4 MB**.
- Do not protect the PDF in any way (otherwise the marker cannot return the script to ActEd, which causes delays) and include the "feedback from marker" sheet when scanning.
- Before emailing to ActEd, please check that your scanned assignment includes all pages and conforms to the above.

If you are submitting by **fax**:

- only write on one side of the paper when completing the assignment
- complete the cover sheet, including the checklist
- fax your script and cover sheet (and Marking Voucher if applicable) to **0844 583 4501**.

In addition:

- We recommend that you stay by the fax machine until the fax has been sent so that you can deal with any problems immediately. (If an error occurs, please re-fax the whole script.)
- An email will be sent by the end of the next working day to confirm that we have processed your script. Please do not phone to check progress before then. If the fax was sent without error then it's very unlikely that there will be a problem.

We will **not** accept:

- scripts submitted to other ActEd fax numbers – please use **0844 583 4501**
- scripts that have been split over a number of faxes. (If an error occurs, please re-fax the whole script.)
- more than one script per fax
- jumbled scripts – please fax the pages in the correct order.

If you are submitting by **post**:

- complete the cover sheet, including the checklist.
- we recommend that you photocopy your script before posting, in case your script is lost in the post.
- please put the correct postage on the envelope and post your script to: **First Floor, McTimoney House, 1 Kimber Road, Abingdon, Oxfordshire, OX14 1BZ.**
- please staple the cover sheet (and Marking Voucher if applicable) to the front of your assignment
- please do not staple more than one assignment together.

Subject CT8: Assignment X1

2015 Examinations

<table>
<tr><td colspan="2">Please complete the following information:</td></tr>
<tr><td>Name:
Address:

ActEd Student Number (see Note below):
[ ][ ][ ][ ][ ]

Note: Your ActEd Student Number is printed on all personal correspondence from ActEd. Quoting this number will help us to process your scripts quickly. If you do not complete this box, your script may be delayed. If you do not know your ActEd Student Number, please email ActEd@bpp.com. Your ActEd Student Number is not the same as your Faculty/Institute Actuarial Reference Number or ARN.</td><td>Number of following pages: _______

Please put a tick in this box if you have solutions and a cross if you do not: [ ]

Please tick here if you are allowed extra time or other special conditions in the profession's exams: [ ]

Time to do assignment
(see Note below):
_____ hrs _____ mins

Under exam conditions
(delete as applicable): yes / nearly / no

Note: If you spend more than 2½ hours on the assignment, you should indicate on the assignment how much you completed within this time so that the marker can provide useful feedback on your chances of success in the exam.</td></tr>
</table>

<table>
<tr><td colspan="11">Score and grade for this assignment (to be completed by marker):</td></tr>
<tr><td>Q1</td><td>Q2</td><td>Q3</td><td>Q4</td><td>Q5</td><td>Q6</td><td>Q7</td><td>Q8</td><td>Q9</td><td colspan="2">Total</td></tr>
<tr><td>__
16</td><td>_
6</td><td>_
4</td><td>_
9</td><td>__
12</td><td>_
7</td><td>__
10</td><td>_
9</td><td>_
7</td><td>__
80</td><td>=_____%</td></tr>
<tr><td colspan="7">Grade: A B C D E</td><td colspan="4">Marker's initials: ________</td></tr>
</table>

Please tick the following checklist so that your script can be marked quickly. Have you:

[ ] Checked that you are using the latest version of the assignments, *eg* 2015 for the sessions leading to the 2015 exams?
[ ] Written your full name and postal address in the box above?
[ ] Completed your ActEd Student Number in the box above?
[ ] Recorded your attempt conditions?
[ ] Numbered all pages of your script (excluding this cover sheet)?
[ ] Written the total number of pages (excluding the cover sheet) in the space above?
[ ] Attached your Marking Voucher or ordered Series X Marking?

Please follow the instructions on the previous page when submitting your script for marking.

This page has been left blank in case you wish to submit your script by fax.

Feedback from marker

Notes on marker's section

The marker's main objective is to give you advice on how to improve your answers. The marker will also assess your script quantitatively and qualitatively. The percentage score gives you a quantitative assessment. The grade is a qualitative assessment of how your script might be classified in the exam. The grades are as follows:

A = Clear Pass   B = Probable Pass   C = Borderline   D = Probable Fail   E = Clear Fail

Please note that you can provide feedback on the marking of this assignment at:

www.ActEd.co.uk/marking

or when you submit your next script.

This page has been left blank in case you wish to submit your script by fax.

Question X1.1

(i) Explain what is meant by the multifactor model. You should define any notation you use. [4]

(ii) Briefly describe three different types of factor that can be used in a multifactor model. [3]

(iii) Define the three forms of the Efficient Markets Hypothesis. [3]

(iv) An insurance company has set up a two-factor model in order to explain the returns experienced by its shareholders on the company's own shares. The indices are described below:

- $I_1 = f(a,n)$, a function of a and n, where n is the number of times the share price has jumped by more than ten pence on a single day and a is the average share price over the last three years.

- I_2 is based on an experienced investment analyst's view of the company. The analyst has no access to information outside the public domain.

Discuss the implications of this model for the company's beliefs with regard to the three forms of the Efficient Markets Hypothesis. Your answer should address the plausibility of this particular model. [6]

[Total 16]

Question X1.2

(i) *Technical analysis* is the study of chart patterns of various asset prices. Explain whether this can be used to an investor's advantage, if the Efficient Markets Hypothesis (EMH) holds. [2]

(ii) *Insider trading* is illegal in the UK stock market. Explain what this suggests about the EMH. [2]

(iii) *Fundamental analysis* includes the analysis of balance sheets, consideration of company strategy, the environment in which the company operates *etc*. Explain how this relates to the EMH. [2]

[Total 6]

Question X1.3

(i) An investor has the utility function $U(w) = -\exp\left(-\frac{w}{100}\right)$.

Determine whether the investor exhibits increasing, constant or decreasing absolute and relative risk aversion. [2]

(ii) The investor has an initial wealth of $1,000$ and is offered a gamble with a payoff described by a random variable:

$$X = \begin{cases} +100 & \textit{with probability } ½ \\ -50 & \textit{with probability } ½ \end{cases}$$

Find the investor's certainty equivalent of this gamble. [2]
[Total 4]

Question X1.4

Two investments are available. A risk-free investment B that returns 1%, and an investment A whose return is given by:

$$R_A = \begin{cases} -1\% & \text{probability } 0.5 \\ 3\% & \text{probability } 0.5 \end{cases}$$

(i) Explain why Asset B must be second-order stochastically dominant over Asset A in terms of investors and utility functions. [2]

(ii) Verify numerically the second-order stochastic dominance expressed in part (i). [2½]

(iii) What can be said about dominance if Asset A offers instead a return of:

(a) $$R_A = \begin{cases} -1\% & \text{probability } 0.5 \\ 4\% & \text{probability } 0.5 \end{cases}$$ [2½]

(b) $$R_A = \begin{cases} -1\% & \text{probability } 0.5 \\ 1\% & \text{probability } 0.5 \end{cases}$$ [1]

(iv) Can an asset that allows the possibility of a return less than 1% ever dominate Asset B? [1]
[Total 9]

Question X1.5

An investor is trying to choose between the investments whose distributions of returns are described below:

Investment A:	0.4 probability that it will return 10% 0.2 probability that it will return 15% 0.4 probability that it will return 20%
Investment B:	0.25 probability that it will return 10% 0.70 probability that it will return 15% 0.05 probability that it will return 40%
Investment C:	A uniform distribution on the range 10% to 20%

Calculate the following for each investment:

(i) expected return [1½]

(ii) variance of return [3½]

(iii) semi-variance [3½]

(iv) expected shortfall below 12% [2½]

(v) shortfall probability below 15%. [1]

[Total 12]

Question X1.6

(i) Explain the following terms in the context of mean-variance portfolio theory:

(a) opportunity set

(b) efficient frontier for a portfolio of risky assets

(c) indifference curves

(d) optimal portfolio [5]

(ii) Describe, using a sketch, the effect on the efficient frontier of introducing a risk-free asset that can be bought or sold in unlimited quantities. [2]

[Total 7]

Question X1.7

Assets A and B have the following distributions of returns in various states:

State	*Asset A*	*Asset B*	*Probability*
1	10%	–12%	0.1
2	8%	0%	0.2
3	6%	3%	0.3
4	4%	16%	0.4
			1.0

(i) Calculate the correlation coefficient between the returns on asset A and asset B and comment on your answer. [4]

(ii) An investor is going to set up a portfolio consisting entirely of assets A and B. Calculate the proportion of assets that should be invested in asset A to obtain the portfolio with the smallest possible variance. [2]

(iii) Assume that the means and the variances of the returns on assets A and B remain unchanged, but that the correlation ρ_{AB} between assets A and B does change. The investor decides to hold 80% of her wealth in asset A and 20% in asset B. Calculate the range of values of ρ_{AB} such that her portfolio has a smaller variance than if she were to invest everything in asset A. [4]
[Total 10]

Question X1.8

(i) State the assumptions of mean variance portfolio theory (MVPT). [3]

(ii) You are given the choice of only two assets, A and B. The expected returns and variances of return of the two assets are:

$$E_A = 13\% \qquad E_B = 5\%$$

$$V_A = 36\%\% \qquad V_B = 4\%\%$$

Find the equation of the efficient frontier in (E,σ) space in the special case where the returns on Assets A and B are perfectly correlated. Comment on your result. [3]

(iii) State, giving the relevant equations, how your approach in (ii) would be modified if there were more than two assets? [*Numerical calculations are not required.*] [3]
[Total 9]

Question X1.9

(i) State the principal theme of behavioural finance. State the assumption of expected utility theory that is challenged by this theme. [2]

(ii) Outline what is meant by Prospect Theory. [2]

(iii) Briefly describe the behavioural finance theme categorising the behaviour of the people in each of the following cases.

(a) Short-term interest rates have remained at historically low levels for several years now and there is no compelling reason for them to change. However, Saver A continues to expect them to increase anytime soon.

(b) In 2011, Investor B invests his spare cash in Share X, which produces a return of 5%. In 2012, he splits his money equally between Shares Y and Z, which return +15% and –5% respectively. He is less happy with the performance of his share portfolio in 2012 than in 2011. [3]

[Total 7]

Subject CT8: Assignment X2

2015 Examinations

Time allowed: 2½ hours

Instructions to the candidate

1. *Please note that we only accept the current version of assignments for marking, ie you can only submit this assignment in the sessions leading to the 2015 exams.*

2. *Attempt all of the questions, leaving space in the margin and beginning your answer to each question on a new page.*

3. *Write in black ink using a medium-sized nib because we will be unable to mark illegible scripts.*

4. ***Leave at least 2cm margin on all borders.***

5. *Attempt the questions as far as possible under exam conditions.*

6. *You should aim to submit this script for marking by the recommended submission date. The recommended and deadline dates for submission of this assignment are listed in the Study Guide for the 2015 exams, on the summary page at the back of this pack and on our website at **www.ActEd.co.uk**.*

   Scripts received after the deadline date will not be marked, unless you are using a Marking Voucher. It is your responsibility to ensure that scripts reach ActEd in good time. ActEd will not be responsible for scripts lost or damaged in the post or for scripts received after the deadline date. If you are using Marking Vouchers, then please make sure that your script reaches us by the Marking Voucher deadline date to give us enough time to mark and return the script before the exam.

At the end of the assignment

If your script is being marked by ActEd, please follow the instructions on the reverse of this page.

In addition to this paper, you should have available actuarial tables and an electronic calculator.

Submission for marking

If you are submitting by **email**:

- complete the cover sheet, including the checklist
- scan your script and cover sheet (and Marking Voucher if applicable) to a pdf document, then email it to: **ActEdMarking@bpp.com**.

Please note the following:

- Please title the email to ensure that the subject and assignment are clear *eg* "CT8 Assignment X2 No. 12345", inserting your ActEd Student Number for 12345.
- The assignment should be scanned the **right way up** (so that it can be read normally without rotation) and as a single document. We cannot accept individual files for each page.
- Please set the resolution so that the script is legible and the resulting PDF is less than 3 MB in size. **The file size cannot exceed 4 MB**.
- Do not protect the PDF in any way (otherwise the marker cannot return the script to ActEd, which causes delays) and include the "feedback from marker" sheet when scanning.
- Before emailing to ActEd, please check that your scanned assignment includes all pages and conforms to the above.

If you are submitting by **fax**:

- only write on one side of the paper when completing the assignment
- complete the cover sheet, including the checklist
- fax your script and cover sheet (and Marking Voucher if applicable) to **0844 583 4501**.

In addition:

- We recommend that you stay by the fax machine until the fax has been sent so that you can deal with any problems immediately. (If an error occurs, please re-fax the whole script.)
- An email will be sent by the end of the next working day to confirm that we have processed your script. Please do not phone to check progress before then. If the fax was sent without error then it's very unlikely that there will be a problem.

We will **not** accept:

- scripts submitted to other ActEd fax numbers – please use **0844 583 4501**
- scripts that have been split over a number of faxes. (If an error occurs, please re-fax the whole script.)
- more than one script per fax
- jumbled scripts – please fax the pages in the correct order.

If you are submitting by **post**:

- complete the cover sheet, including the checklist.
- we recommend that you photocopy your script before posting, in case your script is lost in the post.
- please put the correct postage on the envelope and post your script to: **First Floor, McTimoney House, 1 Kimber Road, Abingdon, Oxfordshire, OX14 1BZ.**
- please staple the cover sheet (and Marking Voucher if applicable) to the front of your assignment
- please do not staple more than one assignment together.

Subject CT8: Assignment X2
2015 Examinations

<table>
<tr><td colspan="2">Please complete the following information:</td></tr>
<tr><td>Name:
Address:

ActEd Student Number (see Note below):
[ ][ ][ ][ ][ ]

Note: Your ActEd Student Number is printed on all personal correspondence from ActEd. Quoting this number will help us to process your scripts quickly. If you do not complete this box, your script may be delayed. If you do not know your ActEd Student Number, please email ActEd@bpp.com. Your ActEd Student Number is not the same as your Faculty/Institute Actuarial Reference Number or ARN.</td>
<td>Number of following pages: _______

Please put a tick in this box if you have solutions and a cross if you do not: [ ]

Please tick here if you are allowed extra time or other special conditions in the profession’s exams: [ ]

Time to do assignment
(see Note below):
_____ hrs _____ mins

Under exam conditions
(delete as applicable): yes / nearly / no

Note: If you spend more than 2½ hours on the assignment, you should indicate on the assignment how much you completed within this time so that the marker can provide useful feedback on your chances of success in the exam.</td></tr>
</table>

Score and grade for this assignment (to be completed by marker):

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Total	
__/10	_/8	_/8	_/9	__/12	__/14	_/8	_/6	_/5	__/80	=_____%

Grade: A B C D E **Marker’s initials:** ________

Please grade your Assignment X1 marker by ticking the appropriate box.

[ ] **Excellent** – the marker’s comments were thorough and very helpful
[ ] **Good** – the marker’s comments were generally helpful
[ ] **Acceptable** – please explain below how the marker could have been more helpful
[ ] **Poor** – the marker’s comments were generally unhelpful; please give details below

Please give any additional comments here (especially if you rate the marker less than good):

Note: Giving feedback on your marker helps us to improve the quality of marking.

Please follow the instructions on the previous page when submitting your script for marking.

This page has been left blank in case you wish to submit your script by fax.

Please tick the following checklist so that your script can be marked quickly. Have you:	
[ ]	Checked that you are using the latest version of the assignments, *eg* 2015 for the sessions leading to the 2015 exams?
[ ]	Written your full name and postal address in the appropriate box?
[ ]	Completed your ActEd Student Number in the appropriate box?
[ ]	Recorded your attempt conditions?
[ ]	Numbered all pages of your script (excluding this cover sheet)?
[ ]	Written the total number of pages (excluding the cover sheet) in the space above?
[ ]	Attached your Marking Voucher or ordered Series X Marking?
[ ]	Rated your Assignment X1 marker?

Feedback from marker

Notes on marker's section

The marker's main objective is to give you advice on how to improve your answers. The marker will also assess your script quantitatively and qualitatively. The percentage score gives you a quantitative assessment. The grade is a qualitative assessment of how your script might be classified in the exam. The grades are as follows:

A = Clear Pass B = Probable Pass C = Borderline D = Probable Fail E = Clear Fail

Please note that you can provide feedback on the marking of this assignment at:

www.ActEd.co.uk/marking

or when you submit your next script.

This page has been left blank in case you wish to submit your script by fax.

Question X2.1

(i) Explain what the separation theorem implies about optimal investment strategies. [2]

(ii) Explain why an individual investor wouldn't hold the market portfolio as part of her investment portfolio in practice. [2]

You are given the following historical information for a share in Company ABC and for a portfolio of 100 shares.

	Return (% *pa*)	Standard deviation of return (% *pa*)	beta
ABC	8.5	20	0.7
Portfolio	10.5	16	1.1

(iii) Use these results to derive the expected return on the market portfolio and the risk-free rate of return assuming the CAPM applies. [3]

A student has commented that ABC's lower return and higher standard deviation, relative to the 100-share portfolio, contradicts the predictions of the CAPM.

(iv) Discuss the student's comment. [3]
[Total 10]

Question X2.2

(i) Outline the main difficulties with testing the arbitrage pricing theory (APT) in practice. [3]

A modeller is developing a 2-factor APT model and has estimated the following parameter values in respect of three securities:

Security	Expected return %	Exposure to factor 1, $b_{i,1}$	Exposure to factor 2 $b_{i,2}$
1	8	5	7.5
2	12	12.5	10
3	15	12.5	15

(ii) Use the data in the table to estimate numerical values for λ_0, λ_1 and λ_2, and explain what your answers represent. [5]
[Total 8]

Question X2.3

(i) Use Taylor's formula to derive Ito's Lemma for a function $f(X_t)$. [2]

An oil trader uses the following model for the short-term behaviour of the oil price X_t, measured in terms of US $100 per barrel:

$$X_t = 0.05\,t + 0.10\,B_t$$

where B_t is a standard Brownian motion.

(ii) Use Ito's Lemma for a function $f(B_t, t)$ to derive the stochastic differential equation for X_t. [3]

A bank offers an exotic derivative whose value is given by:

$$G(X_t) = X_t^2$$

(iii) Use Ito's Lemma to derive the stochastic differential equation for the exotic derivative. [3]

[Total 8]

Question X2.4

Let B_t be a standard Brownian motion, and let $F_t = \sigma\left(B_s, 0 \le s \le t\right)$ be its natural filtration.

(i) Derive the conditional expectations $E\left[B_t^2 \middle| F_s\right]$ and $E\left[B_t^4 \middle| F_s\right]$, where $s \le t$.

You may assume that the third and fourth moments of a random variable with distribution $N\left(0, \sigma^2\right)$ are 0 and $3\sigma^4$ respectively. [4]

(ii) Hence construct a martingale out of B_t^4. [5]

[Total 9]

Question X2.5

An investment banker wishes to model exchange rate movements between US Dollars and Euros as a geometric Brownian motion. Suppose she decides to use the following stochastic differential equation for this purpose:

$$dX_t = (r_d - r_e)X_t dt + \sigma X_t dB_t$$

where:

- X_t represents the value of US Dollars in terms of Euros
- r_d and r_e are the (constant) short-term interest rates in the US and the Eurozone respectively
- B_t is a standard Brownian motion.

(i) Explain why, in economic terms, the value of the US Dollar is likely to increase when $r_d > r_e$. [2]

(ii) By considering the function $f(X_t) = \log X_t$, use Ito's Lemma to solve the above stochastic differential equation for X_t. [5]

(iii) Let $G_t = \dfrac{1}{X_t}$ denote the value of the Euro in terms of US Dollars. Derive the stochastic differential equation for G_t and comment on your answer. [5]

[Total 12]

Question X2.6

(i) (a) State the defining properties of standard Brownian motion B_t.

(b) Write down the probability density function, for an increment over a time lag $t-s$, of general Brownian motion $W_t = \sigma B_t + \mu t$. [5]

(ii) By first obtaining the stochastic differential equation for the function $f(S_t) = \log S_t$, solve the stochastic differential equation defining geometric Brownian motion:

$$dS_t = \mu S_t \, dt + \sigma S_t \, dB_t$$ [5]

(iii) S_t, the price of a share at time *t*, is modelled as geometric Brownian motion. If $\mu = 20\%$ *pa* and $\sigma = 10\%$ *pa*, calculate the probability that the share price will exceed 110 in six months' time given that its current price is 100. [4]

[Total 14]

Question X2.7

(i) Describe what is meant by the lognormal model of security prices. [2]

(ii) If X_t is defined to be the deviation of the log of the security price S_t from its trend value, show that changes in X_t over a time interval *h* are stationary. [2]

(iii) A modeller uses the lognormal model to run 10,000 projections of a share price over the next 120 months. Explain whether the longitudinal and cross-sectional properties will be the same or different for the following two data sets:

(a) the monthly share prices

(b) the log returns over the next month. [4]

[Total 8]

Question X2.8

The stochastic differential equation that implies that the price of an asset at time t, S_t, follows a geometric Brownian motion, is:

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

where Z_t is a standard Brownian motion process.

Describe and discuss the plausibility of the assumptions behind this equation when it is used as a model of share prices. [6]

Question X2.9

Consider an investment market in which:

- the risk-free rate of return on Treasury bills is 4% *pa*
- the expected return on the market as a whole is 8% *pa*
- the standard deviation of the return on the market as a whole is 30% *pa*
- the assumptions of the capital asset pricing model (CAPM) hold.

(i) Consider an efficient portfolio Z that consists entirely of Treasury bills and non-dividend-paying shares, there being no other types of investment. If Z yields an expected return of 7% *pa*, what is its beta? [1]

(ii) Calculate the standard deviation of returns for Portfolio Z. [2]

(iii) Split the total standard deviation for Portfolio Z into the amounts attributable to systematic risk and specific risk. [1]

(iv) Calculate the market value of Portfolio Z assuming that its constituent securities are expected to realise a total sum of $100 in one period from now. [1]

[Total 5]

Subject CT8: Assignment X3

2015 Examinations

Time allowed: 3 hours

Instructions to the candidate

1. *Please note that we only accept the current version of assignments for marking, ie you can only submit this assignment in the sessions leading to the 2015 exams.*

2. *Attempt all of the questions, leaving space in the margin and beginning your answer to each question on a new page.*

3. *Write in black ink using a medium-sized nib because we will be unable to mark illegible scripts.*

4. ***Leave at least 2cm margin on all borders.***

5. *Attempt the questions as far as possible under exam conditions.*

6. *You should aim to submit this script for marking by the recommended submission date. The recommended and deadline dates for submission of this assignment are listed in the Study Guide for the 2015 exams, on the summary page at the back of this pack and on our website at* ***www.ActEd.co.uk****.*

   Scripts received after the deadline date will not be marked, unless you are using a Marking Voucher. It is your responsibility to ensure that scripts reach ActEd in good time. ActEd will not be responsible for scripts lost or damaged in the post or for scripts received after the deadline date. If you are using Marking Vouchers, then please make sure that your script reaches us by the Marking Voucher deadline date to give us enough time to mark and return the script before the exam.

At the end of the assignment

If your script is being marked by ActEd, please follow the instructions on the reverse of this page.

In addition to this paper, you should have available actuarial tables and an electronic calculator.

Submission for marking

If you are submitting by **email**:

- complete the cover sheet, including the checklist
- scan your script and cover sheet (and Marking Voucher if applicable) to a pdf document, then email it to: **ActEdMarking@bpp.com**.

Please note the following:

- Please title the email to ensure that the subject and assignment are clear *eg* "CT8 Assignment X3 No. 12345", inserting your ActEd Student Number for 12345.
- The assignment should be scanned the **right way up** (so that it can be read normally without rotation) and as a single document. We cannot accept individual files for each page.
- Please set the resolution so that the script is legible and the resulting PDF is less than 3 MB in size. **The file size cannot exceed 4 MB**.
- Do not protect the PDF in any way (otherwise the marker cannot return the script to ActEd, which causes delays) and include the "feedback from marker" sheet when scanning.
- Before emailing to ActEd, please check that your scanned assignment includes all pages and conforms to the above.

If you are submitting by **fax**:

- only write on one side of the paper when completing the assignment
- complete the cover sheet, including the checklist
- fax your script and cover sheet (and Marking Voucher if applicable) to **0844 583 4501**.

In addition:

- We recommend that you stay by the fax machine until the fax has been sent so that you can deal with any problems immediately. (If an error occurs, please re-fax the whole script.)
- An email will be sent by the end of the next working day to confirm that we have processed your script. Please do not phone to check progress before then. If the fax was sent without error then it's very unlikely that there will be a problem.

We will **not** accept:

- scripts submitted to other ActEd fax numbers – please use **0844 583 4501**
- scripts that have been split over a number of faxes. (If an error occurs, please re-fax the whole script.)
- more than one script per fax
- jumbled scripts – please fax the pages in the correct order.

If you are submitting by **post**:

- complete the cover sheet, including the checklist.
- we recommend that you photocopy your script before posting, in case your script is lost in the post.
- please put the correct postage on the envelope and post your script to: **First Floor, McTimoney House, 1 Kimber Road, Abingdon, Oxfordshire, OX14 1BZ.**
- please staple the cover sheet (and Marking Voucher if applicable) to the front of your assignment
- please do not staple more than one assignment together.

Subject CT8: Assignment X3

2015 Examinations

Please complete the following information:

Name:

Address:

ActEd Student Number (see Note below):

Note: Your ActEd Student Number is printed on all personal correspondence from ActEd. Quoting this number will help us to process your scripts quickly. If you do not complete this box, your script may be delayed. If you do not know your ActEd Student Number, please email ActEd@bpp.com. **Your ActEd Student Number is not the same as your Faculty/Institute Actuarial Reference Number or ARN.**

Number of following pages: _______

Please put a tick in this box if you have solutions and a cross if you do not: ☐

Please tick here if you are allowed extra time or other special conditions in the profession's exams: ☐

Time to do assignment
(see Note below):

_____ hrs _____ mins

Under exam conditions
(delete as applicable): yes / nearly / no

Note: If you spend more than 3 hours on the assignment, you should indicate on the assignment how much you completed within this time so that the marker can provide useful feedback on your chances of success in the exam.

Score and grade for this assignment (to be completed by marker):

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	**Total**	
__/10	__/10	_/6	__/10	__/12	_/9	__/12	__/12	__/19	___/100	=_____%

Grade: A B C D E **Marker's initials:** ________

Please grade your Assignment X2 marker by ticking the appropriate box.

[ ] **Excellent** – the marker's comments were thorough and very helpful

[ ] **Good** – the marker's comments were generally helpful

[ ] **Acceptable** – please explain below how the marker could have been more helpful

[ ] **Poor** – the marker's comments were generally unhelpful; please give details below

Please give any additional comments here (especially if you rate the marker less than good):

Note: Giving feedback on your marker helps us to improve the quality of marking.

Please follow the instructions on the previous page when submitting your script for marking.

This page has been left blank in case you wish to submit your script by fax.

Please tick the following checklist so that your script can be marked quickly. Have you:	
[ ]	Checked that you are using the latest version of the assignments, *eg* 2015 for the sessions leading to the 2015 exams?
[ ]	Written your full name and postal address in the appropriate box?
[ ]	Completed your ActEd Student Number in the appropriate box?
[ ]	Recorded your attempt conditions?
[ ]	Numbered all pages of your script (excluding this cover sheet)?
[ ]	Written the total number of pages (excluding the cover sheet) in the space above?
[ ]	Attached your Marking Voucher or ordered Series X Marking?
[ ]	Rated your Assignment X2 marker?

Feedback from marker

Notes on marker's section

The marker's main objective is to give you advice on how to improve your answers. The marker will also assess your script quantitatively and qualitatively. The percentage score gives you a quantitative assessment. The grade is a qualitative assessment of how your script might be classified in the exam. The grades are as follows:

A = Clear Pass B = Probable Pass C = Borderline D = Probable Fail E = Clear Fail

Please note that you can provide feedback on the marking of this assignment at:

www.ActEd.co.uk/marking

or when you submit your next script.

This page has been left blank in case you wish to submit your script by fax.

Question X3.1

(i) What is the difference between a European option and an American option? [1]

(ii) What is meant by the "intrinsic value" of an option? [1]

(iii) A *position diagram* is a diagram showing the profit or loss at expiry on an option against the underlying share price. Sketch the position diagram for a long position in a European call option with premium c and exercise price K. You should mark clearly the regions where the option is in-the-money and out-of-the-money. [2]

(iv) Calculate the value of a put option on a share at time t using the Garman-Kohlhagen formula and the following parameter values:

$$S_t = 450\,,\ \ K = 400\,,\ \ r = 0.06\,,\ \ \sigma = 0.2\,,\ \ q = 0.03\,,\ \ T - t = 0.25 \qquad [4]$$

(v) Explain why the calculated value in (i) is very low relative to the current share price. [2]

[Total 10]

Question X3.2

(i) In the context of a non-dividend-paying security, define the Greeks (in both words and using formulae) and state whether each has a positive or negative value for a call option and a put option. [5]

(ii) Consider a 30-day, at-the-money call option on a non-dividend-paying share currently valued at £8. The volatility of the share is 30% and the continuously compounded risk-free rate of interest is 4%.

The outputs of a computer model used to value the option are:

Option value	28.7p	Theta	-0.499pday^{-1}
Delta	0.532	Vega	$0.91\text{p}\%^{-1}$
Gamma	0.00578p^{-1}	Rho	$0.33\text{p}\%^{-1}$

After one day the share price increases by 50p, volatility is reassessed to be 35% *pa* and the risk-free rate moves to 3.5% *pa*. Estimate the new price of the option. [5]

[Total 10]

Question X3.3

(i) Consider a call option and a put option on a dividend-paying security, each with the same term and exercise price. By considering the put-call parity relationship or otherwise, state the value of n such that:

$$\Delta_c = \Delta_p + n$$

(Δ_c is the delta for the call option and Δ_p is the delta for the put option.) [1]

(ii) Derive similar relationships for the other five Greeks. [3]

(iii) Hence, or otherwise, decide whether or not the following relationship holds:

$$r\rho_c + q\lambda_c + (T-t)\theta_c = r\rho_p + q\lambda_p + (T-t)\theta_p$$ [2]

[Total 6]

Question X3.4

A non-dividend-paying stock has a current price of £100. In any unit of time the price of the stock is expected to increase by 10% or decrease by 5%. The continuously compounded risk-free interest rate is 4% per unit of time.

A European call option is written with a strike price of £103 and is exercisable after two units of time, at $t = 2$.

Establish, using a binomial tree, the replicating portfolio for the option at the start and end of the first unit of time, *ie* at $t = 0, 1$. Hence, calculate the value of the option at $t = 0$. [10]

Question X3.5

(i) Explain what is meant by a "replicating portfolio". [2]

(ii) Explain what is meant by a "risk-neutral probability measure" and state mathematically what it implies about the pricing of derivatives relative to the price of the underlying asset. [2]

(iii) Consider a one-period model of a non-dividend paying-stock, currently priced at S_0 and which may move up or down to give $S_1 = S_0 u$ or $S_1 = S_0 d$. Consider a derivative that pays c_u or c_d following an up or down event. The risk-free rate of return (continuously compounded) is r.

(a) Use a replicating portfolio to derive an equation for the price of the derivative at time $t = 0$.

(b) Hence find the price of a derivative whose payoff is defined as $|S_1 - S_0|$, assuming $d < 1$ and $u > 1$.

(c) Explain how to synthesise the derivative in (iii)(b) from simpler options. [8]

[Total 12]

Question X3.6

Using the Black-Scholes formula for the value of a European call option on a non-dividend-paying stock, show that the call price, c, tends to the maximum of $S - Ke^{-r(T-t)}$ and zero (depending on the strike price) as σ tends to zero. [9]

Question X3.7

The price of a non-dividend-paying stock at time 1, S_1, is related to the price at time 0, S_0, as follows:

$$S_1 = S_0 u \text{ with probability } p \text{ and}$$
$$S_0 d \text{ with probability } 1-p.$$

The continuously compounded rate of return on a risk-free asset is *r*.

(i) (a) Determine the replicating portfolio for a European call option written on the stock that expires at time 1 and has a strike price of *k*, where $dS_0 < k < uS_0$. You should give expressions for the number of units for each constituent in the portfolio.

(b) Use your expressions in (i)(a) to find a formula for the price of the European call option.

(c) Use put-call parity to derive a formula for the price of the corresponding European put option, with the same strike price and strike date.

(d) Show that the price of the European call option in (i)(b) can be written as the discounted expected payoff under a probability measure *Q*. Hence find an expression for the probability, *q*, of an upward move in the stock price under *Q*. [7]

(ii) Explain the relationship between the probability measure *Q* in (ii) and the real-world probability measure *P*. Explain what relationship you would expect *q* and *p* to have if all investors are:

(a) risk-averse

(b) risk-seeking

(c) risk-neutral. [5]

[Total 12]

Question X3.8

(i) Explain why $c_t + Ke^{-r(T-t)} = p_t + S_t$ must hold for a non-dividend-paying share with value S_t, explaining any assumptions and all the notation used. [4]

(ii) Using the result in part (i) and the Black-Scholes formula for the value of a European call option on a non-dividend-paying share as given in the Tables, derive an expression for p_t in terms of K, r, $T-t$, d_1 and d_2. [2]

(iii) The underlying share pays no dividends and has a current value of £20 and a volatility of 0.3. An investor who has £100 to invest has a choice between investing in either a one-year zero-coupon bond (redeemable at par) with a current market value of £94.18 or in one-year put options with a strike price of £17.50. If the investor chooses to allocate all of his money to the options, how many can he buy?

[Ignore tax and investment expenses and assume that the bond market and the options market are both arbitrage-free. Assume that the option price is quoted to the nearest penny.] [6]

[Total 12]

Question X3.9

(i) State the assumptions of Black-Scholes pricing analysis. [3]

(ii) State the Black-Scholes formula for a call option, defining all the notation used. [3]

(iii) Explain what is meant by implied volatility and how an investor might use it to make investment decisions. [3]

Consider the following one-month option prices on a non-dividend-paying share whose current value is 330. Assume that the continuously compounded risk-free rate is 4% *pa*.

K	c	p
325	15.33	9.25
350	4.95	23.79
375	1.68	45.43

(iv) Verify that the principle of no-arbitrage holds. [2]

(v) Use Black-Scholes pricing to calculate, to the nearest 1% *pa*, the implied volatility of the call option whose strike price is 325. [4]

(vi) Using Black-Scholes pricing, the implied volatilities obtained from the other two strike prices are 31% *pa* for the call option whose strike price is 350 and 34% *pa* for the call option whose strike price is 375. Explain which of the three sets of options can be considered to be relatively cheap. [2]

(vii) An investment strategy is proposed to consistently buy options that are relatively cheap and sell options that are relatively expensive. Explain the main reasons why this strategy may not be beneficial in the long run. [2]

[Total 19]

Subject CT8: Assignment X4

2015 Examinations

Time allowed: 3 hours

Instructions to the candidate

1. *Please note that we only accept the current version of assignments for marking, ie you can only submit this assignment in the sessions leading to the 2015 exams.*
2. *Attempt all of the questions, leaving space in the margin and beginning your answer to each question on a new page.*
3. *Write in black ink using a medium-sized nib because we will be unable to mark illegible scripts.*
4. ***Leave at least 2cm margin on all borders.***
5. *Attempt the questions as far as possible under exam conditions.*
6. *You should aim to submit this script for marking by the recommended submission date. The recommended and deadline dates for submission of this assignment are listed in the Study Guide for the 2015 exams, on the summary page at the back of this pack and on our website at **www.ActEd.co.uk**.*

   Scripts received after the deadline date will not be marked, unless you are using a Marking Voucher. It is your responsibility to ensure that scripts reach ActEd in good time. ActEd will not be responsible for scripts lost or damaged in the post or for scripts received after the deadline date. If you are using Marking Vouchers, then please make sure that your script reaches us by the Marking Voucher deadline date to give us enough time to mark and return the script before the exam.

At the end of the assignment

If your script is being marked by ActEd, please follow the instructions on the reverse of this page.

In addition to this paper, you should have available actuarial tables and an electronic calculator.

Submission for marking

If you are submitting by **email**:

- complete the cover sheet, including the checklist
- scan your script and cover sheet (and Marking Voucher if applicable) to a pdf document, then email it to: **ActEdMarking@bpp.com**.

Please note the following:

- Please title the email to ensure that the subject and assignment are clear *eg* "CT8 Assignment X4 No. 12345", inserting your ActEd Student Number for 12345.
- The assignment should be scanned the **right way up** (so that it can be read normally without rotation) and as a single document. We cannot accept individual files for each page.
- Please set the resolution so that the script is legible and the resulting PDF is less than 3 MB in size. **The file size cannot exceed 4 MB**.
- Do not protect the PDF in any way (otherwise the marker cannot return the script to ActEd, which causes delays) and include the "feedback from marker" sheet when scanning.
- Before emailing to ActEd, please check that your scanned assignment includes all pages and conforms to the above.

If you are submitting by **fax**:

- only write on one side of the paper when completing the assignment
- complete the cover sheet, including the checklist
- fax your script and cover sheet (and Marking Voucher if applicable) to **0844 583 4501**.

In addition:

- We recommend that you stay by the fax machine until the fax has been sent so that you can deal with any problems immediately. (If an error occurs, please re-fax the whole script.)
- An email will be sent by the end of the next working day to confirm that we have processed your script. Please do not phone to check progress before then. If the fax was sent without error then it's very unlikely that there will be a problem.

We will **not** accept:

- scripts submitted to other ActEd fax numbers – please use **0844 583 4501**
- scripts that have been split over a number of faxes. (If an error occurs, please re-fax the whole script.)
- more than one script per fax
- jumbled scripts – please fax the pages in the correct order.

If you are submitting by **post**:

- complete the cover sheet, including the checklist.
- we recommend that you photocopy your script before posting, in case your script is lost in the post.
- please put the correct postage on the envelope and post your script to: **First Floor, McTimoney House, 1 Kimber Road, Abingdon, Oxfordshire, OX14 1BZ.**
- please staple the cover sheet (and Marking Voucher if applicable) to the front of your assignment
- please do not staple more than one assignment together.

Subject CT8: Assignment X4
2015 Examinations

Please complete the following information:	
Name: **Address:** **ActEd Student Number** (see Note below):	**Number of following pages:** _______ **Please put a tick in this box if you have solutions and a cross if you do not:** ☐ **Please tick here if you are allowed extra time or other special conditions in the profession's exams:** ☐ Time to do assignment (see Note below): _____ hrs _____ mins
Note: Your ActEd Student Number is printed on all personal correspondence from ActEd. Quoting this number will help us to process your scripts quickly. If you do not complete this box, your script may be delayed. If you do not know your ActEd Student Number, please email ActEd@bpp.com. **Your ActEd Student Number is not the same as your Faculty/Institute Actuarial Reference Number or ARN.**	Under exam conditions (delete as applicable): yes / nearly / no **Note:** If you spend more than 3 hours on the assignment, you should indicate on the assignment how much you completed within this time so that the marker can provide useful feedback on your chances of success in the exam.

Score and grade for this assignment (to be completed by marker):

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	**Total**	
__ / 13	__ / 11	__ / 11	__ / 12	__ / 17	__ / 9	__ / 14	__ / 13	__ / 100	=_____%

Grade: A B C D E **Marker's initials:** ________

Please grade your Assignment X3 marker by ticking the appropriate box.

[ ] **Excellent** – the marker's comments were thorough and very helpful
[ ] **Good** – the marker's comments were generally helpful
[ ] **Acceptable** – please explain below how the marker could have been more helpful
[ ] **Poor** – the marker's comments were generally unhelpful; please give details below

Please give any additional comments here (especially if you rate the marker less than good):

Note: Giving feedback on your marker helps us to improve the quality of marking.

Please follow the instructions on the previous page when submitting your script for marking.

This page has been left blank in case you wish to submit your script by fax.

Please tick the following checklist so that your script can be marked quickly. Have you:	
[ ]	Checked that you are using the latest version of the assignments, *eg* 2015 for the sessions leading to the 2015 exams?
[ ]	Written your full name and postal address in the appropriate box?
[ ]	Completed your ActEd Student Number in the appropriate box?
[ ]	Recorded your attempt conditions?
[ ]	Numbered all pages of your script (excluding this cover sheet)?
[ ]	Written the total number of pages (excluding the cover sheet) in the space above?
[ ]	Attached your Marking Voucher or ordered Series X Marking?
[ ]	Rated your Assignment X3 marker?

Feedback from marker

Notes on marker's section

The marker's main objective is to give you advice on how to improve your answers. The marker will also assess your script quantitatively and qualitatively. The percentage score gives you a quantitative assessment. The grade is a qualitative assessment of how your script might be classified in the exam. The grades are as follows:

A = Clear Pass B = Probable Pass C = Borderline D = Probable Fail E = Clear Fail

Please note that you can provide feedback on the marking of this assignment at:

www.ActEd.co.uk/marking

This page has been left blank in case you wish to submit your script by fax.

Question X4.1

The price S_t of a particular share follows a geometric random walk:

$$S_t = S_{t-1} Z_t$$

where $\{Z_t\}$ is a sequence of independent, identically distributed random variables:

$$Z_t = \begin{cases} 1.1 & \text{with probability } 0.6 \\ 0.95 & \text{with probability } 0.4 \end{cases}$$

and t denotes the time in months.

A 1-month European call option is available on the share with a strike price of £10.50. The current market price of the share is £10. No dividends are to be paid over the next 6 months. An annualised risk-free force of interest of 4% is available.

(i) Find the expected payoff of the call option. [1]

(ii) Construct a replicating portfolio for the derivative out of shares and cash, and hence find the fair price of the derivative. [3]

(iii) Hence state how many options need to be bought (or sold) per share, in order to construct a risk-free portfolio out of shares and these options. [1]

(iv) Describe quantitatively the arbitrage opportunity that would arise if the price of the option in the market was equal to the discounted value of the expected payoff. [4]

(v) Another derivative is available on this share. It gives the purchaser the right (but not the obligation) to buy two shares at a price of £12.00 each, or sell one share at a price of £9.00, both in 4 months' time. Determine the possible payoffs for this derivative, and hence find the fair price of this derivative using risk-neutral valuation. [4]

[Total 13]

Question X4.2

An analyst is using the Merton model, together with the following information, to value the five-year zero-coupon bonds (ZCBs) issued by a company:

- The nominal value of ZCBs issued is \$100 million.
- The company's shares have a market capitalisation of \$118.46 million.
- The volatility of the company's underlying assets has been estimated to be 25% *pa.*
- The five-year risk-free force of interest is 5% *pa*.

(i) Calculate the price per \$100 nominal of a five-year risk-free ZCB. [1]

(ii) Using your answer to (i) to obtain an initial estimate, and then applying linear interpolation, estimate the value of the company's assets (to the nearest \$10,000) and hence show that the value of its ZCBs is \$76.47 million. [5]

(iii) (a) State the formula for the delta of a European call option based on the Black-Scholes formula (assuming no dividends) and use it to derive a formula for the delta of the ZCBs with respect to the value of the company's assets.

(b) Estimate the numerical value of delta using your calculations in part (ii) and use it to estimate the new value of the ZCBs following a \$10 million fall in the value of the company's assets.

(c) The actual value of the ZCBs following a \$10 million fall in the value of the company's assets is \$76.16 million. Give a possible reason for the discrepancy between your estimated value of the ZCBs and the actual value. [5]

[Total 11]

Question X4.3

An investment bank has developed a new exotic derivative, which will pay an amount equal to the share price at maturity multiplied by the share price at maturity less one dollar. Let T be the maturity date of the derivative and r be the risk-free force of interest and assume that the Black-Scholes analysis applies.

(i) Use risk-neutral valuation to derive the pricing formula for this derivative at time $t < T$, based on a share that pays no dividends. [7]

(ii) (a) Derive the corresponding formula for the delta of the derivative.

(b) Derive a condition for the range of values for the current share price for which delta is positive and comment on what your answer suggests for derivatives of this type with differing terms.

(c) Derive the corresponding formula for the gamma of the derivative and comment on the sign of gamma. [4]

[Total 11]

Question X4.4

The price S_t of a share that pays no dividends follows a geometric Brownian motion:

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

where Z_t is a standard Brownian motion. A derivative is available on this share that can only be exercised at time *T*. The price of the derivative at time *t*, $f\left(t, S_t\right)$, depends on the time and the current share price. A cash bond is also available that offers a risk-free rate of return of *r* (continuously compounded). The price of the bond is B_t. You wish to set up a replicating portfolio for the derivative made out of shares and cash, so that:

$$\phi_t S_t + \psi_t B_t = f\left(t, S_t\right) \qquad (*)$$

(i) Write down the differential equation that is satisfied by B_t. [1]

(ii) What does it mean for the portfolio to be self-financing? Give a differential equation that must be satisfied by the portfolio considered above in order that this is the case. [2]

(iii) What does it mean for a process to be previsible? [1]

(iv) By applying Ito's Lemma to the right-hand side of equation (*) and using your answers to parts (i) and (ii), deduce that:

(a) $\phi_t = \dfrac{\partial}{\partial S_t} f\left(t, S_t\right)$

(b) $\theta + rS_t\Delta + \frac{1}{2}\sigma^2 S_t^2 \Gamma = rf$

where θ, Δ and Γ are symbols that you should define. [8]

[Total 12]

Question X4.5

(i) State the martingale representation theorem in continuous time. [2]

(ii) Define a replicating portfolio. [1]

(iii) It is proposed that $V_t = e^{-r(T-t)} E_Q[X \mid F_t]$ is the fair price at time t for a derivative with a random payoff X at time T. The derivative is based on an underlying non-dividend-paying share with value S_t, $t \geq 0$. Let $B_t = e^{rt}$ denote the value at time t of a simple cash process.

In this formula Q is the risk-neutral probability measure. Let $D_t = e^{-rt} S_t$. Show that D_t is a martingale under Q. [2]

(iv) Let $E_t = e^{-rT} E_Q[X \mid F_t]$. Show that E_t is also a martingale under Q. [2]

(v) Let ϕ_t be defined by $dE_t = \phi_t dD_t$ and $\psi_t = E_t - \phi_t D_t$.

Explain how it is known that ϕ_t is previsible and why this property is important in this context. [3]

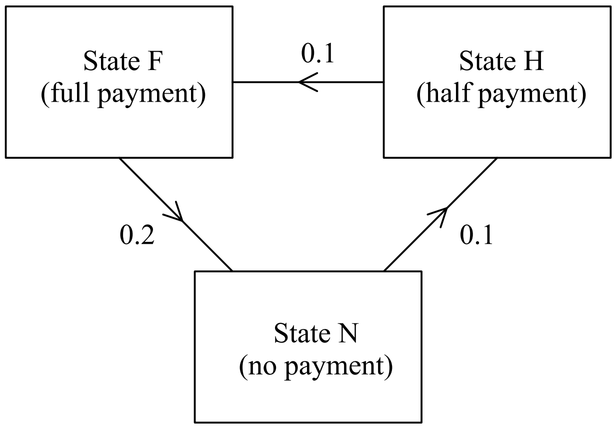
(vi) It is decided to hold the portfolio (ϕ_t, ψ_t) (*ie* ϕ_t shares and ψ_t cash) during the infinitesimal time interval $[t, t+dt)$. Show that the change in the value of this portfolio over this time interval equals dV_t and deduce that it is self-financing. [5]

(vii) Deduce that $V_t = e^{-r(T-t)} E_Q[X \mid F_t]$ is the fair price to pay for the derivative at time t. [2]

[Total 17]

Question X4.6

A bank is using a three-state discrete-time Markov chain model to value its bond portfolio.



On 1 January each year the bank assigns each of its client companies to one of the following categories:

- State F: The bank expects to receive any payments due that year in full.
- State H: The bank expects to receive only 50% of any payments due that year.
- State N: The bank expects to receive no payments from the company that year.

The diagram shows the risk-neutral probabilities that each company will move from its current rating level to another level at the time of each review. These probabilities are independent of the company's previous ratings and the behaviour of other companies.

Let $p_{ij}(0,t)$ denote the probability that a company initially in State i will be in State j t years later.

(i) Calculate $p_{Fj}(0,t)$ for $j=F,H$ and $t=1,2,3$. [3]

The bank is considering purchasing at par a 3-year bond issued by a company currently rated as F. Under the terms of the bond, interest of 10% of the face value of the bond will be paid at the end of each year, and the bond will be redeemed at par at the end of the 3 years.

The annual effective yields on 1-year, 2-year and 3-year government bonds are all 5%.

(ii) (a) Calculate the risk-neutral expected present value of the payments from the bond per £100 face value.

(b) Comment on your answer in (ii)(a). [3]

After negotiations, the bank agrees to purchase the bonds at a price of £95.20.

(iii) Calculate the credit spread for this bond. [3]
[Total 9]

Question X4.7

In the Vasicek model, the spot rate of interest is governed by the stochastic differential equation:

$$dr_t = a\left(b - r_t\right)dt + \sigma\, dB_t$$

where B_t is a standard Brownian motion and a, $b > 0$ are constants.

(i) A stochastic process $\{U_t : t \geq 0\}$ is defined by $U_t = e^{at} r_t$.

(a) Derive an equation for dU_t.

(b) Hence solve the equation to find U_t.

(c) Hence show that:

$$r_t = b + \left(r_0 - b\right)e^{-at} + \sigma \int_0^t e^{a(s-t)} dB_s \qquad [5]$$

(ii) Determine the probability distribution of r_t and the limiting distribution for large t. [4]

(iii) Derive, in the case where $s < t$, the conditional expectation $E\left[r_t \middle| F_s\right]$, where $\left\{F_s : s \geq 0\right\}$ is the filtration generated by the Brownian motion B_s. [5]
[Total 14]

Question X4.8

(i) Describe briefly the Vasicek one-factor model of interest rates and its key statistical properties. [4]

(ii) According to a particular parameterisation of this model, the instantaneous forward rate applicable at a fixed time $T > t$ implied by the market prices at time t is found to be:

$$f(t,T) = r_t e^{-\alpha\tau} + r_\infty(1-e^{-\alpha\tau}) + k(1-e^{-\alpha\tau})e^{-\alpha\tau}$$

where $\tau = T - t$ and $\alpha > 0$.

Show that, if a humped curve is required for $f(t,T)$, the parameter values must satisfy the condition $k > \left| r_\infty - r_t \right|$. [5]

Here a "humped curve" means one where the value of the function for some intermediate values of τ exceeds the values for both $\tau = 0$ and $\tau = \infty$. In other words, there will be a maximum value for some positive value of τ.

(iii) Describe briefly the main advantages and limitations of the Vasicek model. [4]
[Total 13]

For the session leading to the April 2015 exams – CT Subjects

Marking vouchers

Subjects	*Assignments*	*Mocks*
CT1, CT3, CT4, CT8	25 March 2015	6 April 2015
CT2, CT5, CT6, CT7	1 April 2015	13 April 2015

Series X Assignments

Subjects	*Assignment*	*Recommended submission date*	*Final deadline date*
CT1, CT3, CT4, CT8	**X1**	**12 November 2014**	21 January 2015
CT2, CT5, CT6, CT7		**19 November 2014**	28 January 2015
CT1, CT3, CT4, CT8	**X2**	**26 November 2014**	11 February 2015
CT2, CT5, CT6, CT7		**3 December 2014**	18 February 2015
CT1, CT3, CT4, CT8	**X3**	**28 January 2015**	4 March 2015
CT2, CT5, CT6, CT7		**4 February 2015**	11 March 2015
CT1, CT3, CT4, CT8	**X4**	**18 February 2015**	18 March 2015
CT2, CT5, CT6, CT7		**25 February 2015**	25 March 2015

Mock Exams

Subjects	*Recommended submission date*	*Final deadline date*
CT1, CT3, CT4, CT8	**25 March 2015**	6 April 2015
CT2, CT5, CT6, CT7	**1 April 2015**	13 April 2015

We encourage you to work to the recommended submission dates where possible.

We strongly recommend that you submit your mock exam electronically, by email or fax, in order for us to return your marked script to you in plenty of time before your exam. If you submit your mock by post to arrive with us on the final deadline date, you are likely to receive your script back less than a week before your exam.

In general, the turnaround of all scripts is likely to be quicker if you submit it electronically and well before the final deadline date.

For the session leading to the September/October 2015 exams – CT Subjects

Marking vouchers

Subjects	***Assignments***	***Mocks***
CT1, CT4	2 September 2015	14 September 2015
CT2, CT3, CT5, CT8	9 September 2015	21 September 2015
CT6, CT7	16 September 2015	28 September 2015

Series X Assignments

Subjects	***Assignment***	***Recommended submission date***	***Final deadline date***
CT1, CT4	**X1**	**17 June 2015**	8 July 2015
CT2, CT3, CT5, CT6, CT7, CT8		**24 June 2015**	15 July 2015
CT1, CT4	**X2**	**8 July 2015**	29 July 2015
CT2, CT3, CT5, CT6, CT7, CT8		**15 July 2015**	5 August 2015
CT1, CT4	**X3**	**29 July 2015**	12 August 2015
CT2, CT3, CT5, CT6, CT7, CT8		**5 August 2015**	19 August 2015
CT1, CT4	**X4**	**12 August 2015**	26 August 2015
CT2, CT3, CT5, CT8		**19 August 2015**	2 September 2015
CT6, CT7		**26 August 2015**	9 September 2015

Mock Exams

Subjects	***Recommended submission date***	***Final deadline date***
CT1, CT4	**26 August 2015**	14 September 2015
CT2, CT3, CT5, CT8	**2 September 2015**	21 September 2015
CT6, CT7	**9 September 2015**	28 September 2015

We encourage you to work to the recommended submission dates where possible.

We strongly recommend that you submit your mock exam electronically, by email or fax, in order for us to return your marked script to you in plenty of time before your exam. If you submit your mock by post to arrive with us on the final deadline date, you are likely to receive your script back less than a week before your exam.

In general, the turnaround of all scripts is likely to be quicker if you submit it electronically and well before the final deadline date.