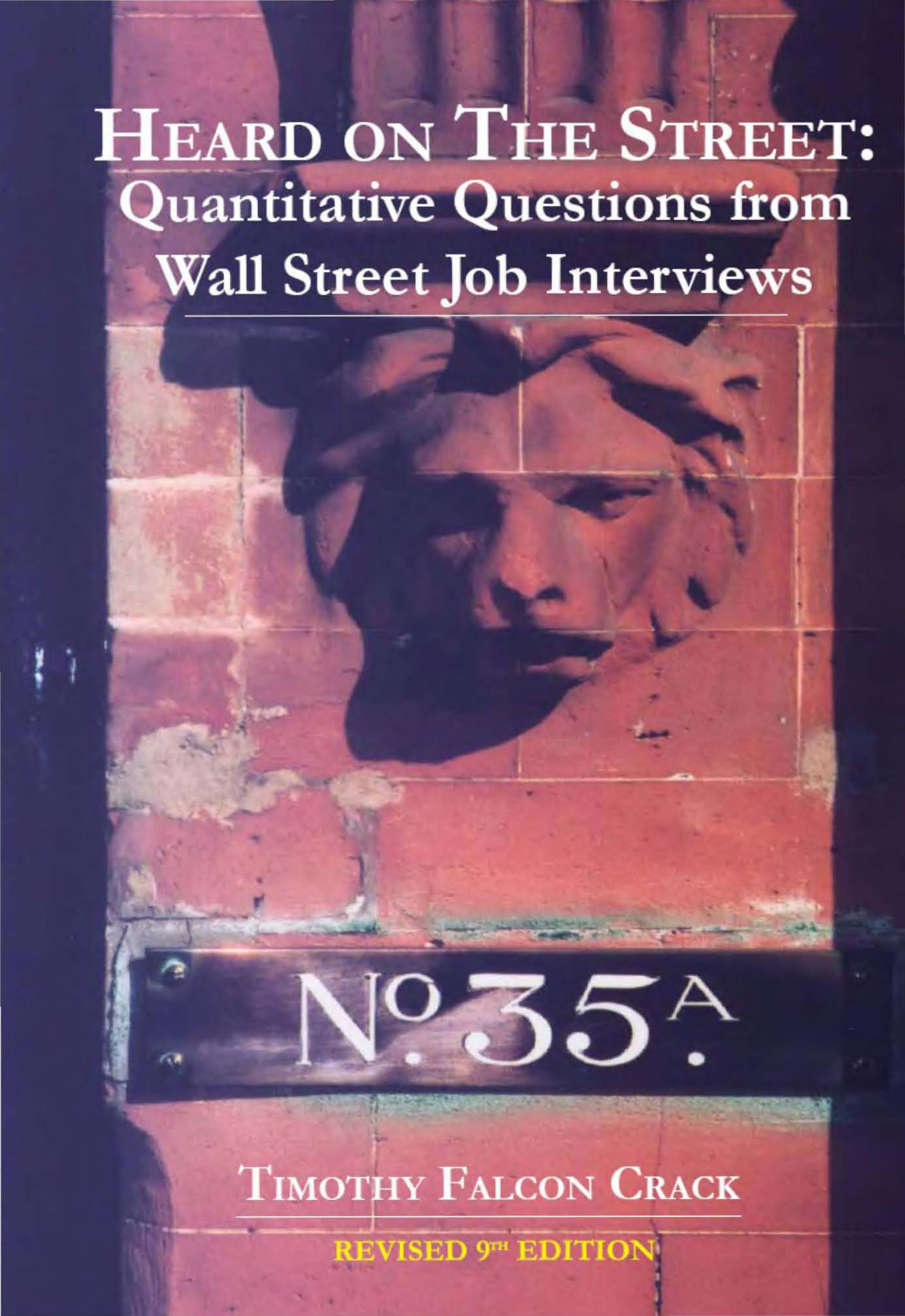
HEARD ON THE STREET: Quantitative Questions from Wall Street Job Interviews



TIMOTHY FALCON CRACK

REVISED 9TH EDITION

Heard on The Street: Quantitative Questions from Wall Street Job Interviews

Heard on The Street: Quantitative Questions from Wall Street Job Interviews

Timothy Falcon Crack
BSc (HONS 1st Class), PGDipCom, MCom, PhD (MIT), IMC

First published 1995. Revised 1997, 1998, 1999, 2000, 2002, 2004 (9th Edition).
First eBook version published 2004.

Typeset by the author. Printed in the United States of America.
www.InvestmentBankingJobInterviews.com
timcrack@alum.mit.edu

Preface

THIS BOOK BRIDGES THE CONSIDERABLE GAP between the typical finance MBA education and the knowledge required to successfully answer quantitative finance job interview questions.

The considerable gap arises because Wall Street interviewers must separate the "wolves" from the "sheep." The sheep are confined by the boundaries of their MBA education; the wolves are not. The interview questions reach beyond these boundaries in order to separate the two classes of interviewees. Hence the gap.

Of course, most interviewers are wolves. Unfortunately, many interviewees are sheep. The butchering that takes place has been described to me as "horrific." That is why you need this book.

I bridge the above-mentioned gap by presenting over 140 quantitative questions from actual finance job interviews. I could not find even one of these questions in any of the three-dozen other "interview books" at a large US bookstore. My solutions and advice are carefully designed to sharpen your quantitative skills. My advice is based on my experiences as a frontline teaching assistant for MBA students at MIT, as a finance professor at Indiana University, and as the head of a quant research team for the world's largest institutional asset manager.

My intended audience includes interviewees (wolves and sheep alike) seeking employment at Wall Street or other finance-related firms; their interviewers, who need to weed out the hapless sheep; university professors who want to "spice up" finance courses with Wall Street job interview questions (both for fun and to show the importance of the basic concepts on The Street); students of finance who want to fill in some gaps; and finally, doctoral students in need of entertainment during periods of downtime.

Many of the questions collected and presented here are "classics" that appear year-after-year without fail. However, this book is definitely not for people who just want "The Answers" to such questions. Such people are the archetypal sheep in wolves' clothing, and they are quickly identified as such in an interview. To benefit from this book, you must make a serious investment of your time.

I thank MIT students, MIT faculty, and people on "The Street" who have supplied me with information. I thank Olivier Ledoit, Cecily Lown, Bingjian Ni, Eva Porro, and Juan Tenorio for their constructive criticism. This book was written and edited in 1995 while commuting to and from MIT on the subways and buses of the Massachusetts Bay Transit Authority (MBTA).

TFC/MIT/1995

I updated this book while working as a professor at Indiana University (IU). I thank all the people thanked above (especially Olivier Ledoit). I also thank Sean Curry and The MathWorks Inc for a free copy of MATLAB (used to check many answers and draw most figures), MBA Style Magazine (`www.mbastyle.com`) for horror stories, Andres Almazan, Tom Arnold, Mary Chris Bates, Klara Buff, Alex Butler, Victor W. Goodman, Tim Hoel, Taras Klymchuk, Victor H. Lin, Marianne Lown, Alan J. Marcus, David Maslen, Marc Rakotomalala, Jason Roth, Yi Shen, Valeri Smelyansky, Dahn Tamir, Paul Turner, and students (MBA and undergraduate) at each of MIT, UCLA, and IU.

TFC/IU/1996–2000

I updated this book while working as Head of Quantitative Active Equity Research (UK/Europe) at Barclays Global Investors in London. I now also thank Jinpeng Chang, Mark Rubinstein, Alex Vigodner, and Nick Vivian.

TFC/BGI/2001–2003

I updated this book in December 2003, after accepting a position as a full professor of finance at Otago University in New Zealand. I now also thank Giulio Agostini, Scott Chaput, Allesio Farhadi, Vince Moshkevich, Katie Price, Wolfgang Prymas, Naoki Sato, Mikhail Voropaev, and Thomas C. Watson.

TFC/OU/2004–

Contents

Introduction **1**
Questions in This Book 2
The Interview 4
Will They Ask me These Questions? 4
ATQ! 4
Other Advice 8

1 Purely Quantitative & Logic Questions **11**

2 Derivatives Questions **25**

3 Other Financial Economics Questions **37**

4 Statistics Questions **41**

5 Non-Quantitative Questions **47**
5.1 Questions about You 48
5.2 Questions about Your Job Awareness 51
5.3 Questions about the Markets or the Economy 53
5.4 Financial Management Questions 54
5.5 Thinking Questions 56

A Purely Quantitative & Logic Answers **59**

B Derivatives Answers **115**

C Other Financial Economics Answers **185**

D Statistics Answers **207**

E Non-Quantitative Answers (Selected) 227

References for Further Research 231

Index 249

List of Tables

A.1 Weighings Needed to Find Bad Coin 71
A.2 Weighings Needed to Find Bad Coin 72
A.3 Trigonometrical Functions: Definitions 97
A.4 Trigonometrical Functions: Calculus 98
A.5 Sums of k, k^2, and k^3 102
A.6 The Red/Black Card Game 107
A.7 E(Payoff) in Red/Black Card Games ($2n$ cards, n red, n black) 108

B.1 Straddle Prices when the Stock Price Jumps 149
B.2 Pricing Methods Summary: Plain Vanilla Options 169
B.3 Pricing Methods Summary: Exotic Options 170

C.1 Duration/Convexity Summary 199

D.1 Distribution of Payoff to Third Roll of a Die 208
D.2 Distribution of Maximum Payoff in Three Rolls of a Die . . . 210
D.3 The Monty Hall Problem 216

E.1 Alphabets and Numerical Equivalences 248

List of Figures

1.1 Number of Cubes on Each Square of a 20×20 Chessboard (Q) 20

A.1 A Road Race Analogy for the LCM Problem 81
A.2 Two Possible Triangle Configurations 83
A.3 The Lighthouse Problem 92
A.4 Number of Cubes on Each Square of a 20×20 Chessboard (A) 94
A.5 S-E-N Problem: The Earth 111

B.1 Standard Call: Price, Delta, and Gamma. 119
B.2 Sensitivity of Option Prices to Volatility 129
B.3 Time Value of a European Call Option 134
B.4 Call Price as a Function of Different Variables 140
B.5 Power Calls with $\alpha > 1$, and $\alpha < 1$ 161

Introduction

This book contains over 140 quantitative questions from actual job interviews in investment banking, investment management, and options trading. The interviewers use the same questions year-after-year, and here they are—with solutions! These questions come from all types of interviews (corporate finance, sales and trading, quant research, etc), but they are especially likely in quantitative capital markets job interviews. The questions come from all levels of interviews (undergrad, MBA, PhD), but they are especially likely if you have, or almost have, an MS or MBA.

This edition includes over 120 non-quantitative actual interview questions, and a new section on interview technique—based partly on my experiences interviewing candidates for the world's largest institutional asset manager.

The book is interspersed with boxed stories (including "horror stories") from job interviews. Some of these stories involve restatements of questions that appear elsewhere in the book.

All the quantitative questions are accompanied by detailed solutions. The questions are split into four categories: Purely Quantitative and Logic, Derivatives, Other Financial Economics, and Statistics (Chapters 1, 2, 3, and 4, respectively). The solutions appear in Appendices A, B, C, and D, respectively. Chapter 5 presents non-quantitative questions from actual interviews (with selected solutions in Appendix E).

I have removed two previous appendices, one on option pricing theory, and one with Hewlett Packard handheld calculator code, because I have written a new book that includes that material with far superior presentation (Crack [2004]). See the advertisement on the last page of this book for details, or go directly to `www.BasicBlackScholes.com`.

In the text, a name followed by a year (e.g., "Girsanov [1960]") refers to a work cited in the References for Further Research (following the appendices).

Questions in This Book

The questions in this book were collected by me from interviewees, interviewers, and others. I have taken the liberty of rewording them for maximum clarity because, unlike in an interview, you have no opportunity to ask me for clarification. Sometimes I give only part of a question that was asked; sometimes I combine related questions into a larger one. I remain faithful to the original problem statement wherever possible.

I often add a footnote to a question. The footnote contains a slight variation on the question. Unless otherwise indicated, these "footnote questions" are made up by me and are not actual job interview questions. All other questions come from actual job interviews (even the "condom question").

Many of the questions require a serious investment of your time. Knowing the answer is not enough in and of itself. If my answer is not clear after several readings, then consult the references I give.

Interviewees should attempt these questions without looking to the answers. Mastering the problem-solving process gets you the job. This may mean spending several days (or weeks) with a problem before you can figure it out exactly. Looking at the answer tells you how I did it; it does not tell you how to solve problems by yourself. The path of greatest resistance bears the highest rewards—do not look at the answers until you have done your very best with the questions!

Interviewers can use these questions as they stand. However, you should try to defeat interviewees who have already seen my questions and answers. One way is to make a "call" a "put," or "minimize" something that is "maximized" here, or use different numbers. I strongly encourage you to push students very hard for the underlying understanding. Ask them to explain the answer, not to simply solve the problem. This differentiates those who understand the problem from those who merely know the answer. The good ones meet the challenge; the bad ones do not. In my experience, many students can solve problems, but only a few genuinely understand what they are doing.

Will the questions in this book become obsolete or dated? The answer is no, for two reasons: First, many of the questions are "classics" that appear consistently year-after-year; and second, the body of quantitative skills required to solve these questions has remained unchanged for three decades. Even if some of the questions change, the skills required to solve them do not. It is these skills that my book promotes. For these reasons, it follows that these questions are genuinely timeless.[1]

Sometimes the classification of a question (and, therefore, the chapter it should appear in) is by no means clear. For example, some of the financial economics questions look like statistics questions, and I have placed questions on stochastic calculus in the derivatives chapter instead of the statistics chapter.

Some questions have more than one solution technique. The "right answer" is the wrong answer if you use a "brute-force" approach and completely miss an elegant alternative (I often give both techniques).

Some questions are more difficult than others. I have labelled difficult questions with two stars "(**)" and very difficult questions with three stars "(***)." By default, all other questions deserve one star. For the two-star or three-star questions, your approach, rather than your solution, may be of more importance. You should be able to set up a general framework for a solution. If you can solve such questions on the spot, you are doing well.

Some of the questions are at a low level, and you may think it beneath your dignity to answer them. I have, however, interviewed people who claim to have degrees in finance, economics, statistics or mathematics, who cannot answer basic finance, economics, statistics, or mathematics questions, respectively. If you think the basic questions are beneath you, then prove it by walking through them like a hot knife through butter. If you cannot answer the basic questions, however, either because you are rusty on the basics, or simply never understood them, then why should anyone hire you? No one will want to put you in front of their traders, other team members, or clients who will have basic questions-.

Knowledge of C (but probably not C++) is required for some types of quantitative Wall Street interviews (in such cases, C++ may be required for the job itself). I used to have a few C questions in my book without solutions, but I now think it is in your best interest to simply recommend Mongan and Suojanen (2000) to you.

You must have already heard all the ordinary interview advice (cover letters, appearance, comments on previous employers, use of bad language, chewing gum, researching people who will interview you, researching the firm, knowing your strengths and weaknesses, and so on); if not, then see Allen

[1]At first glance, some questions may seem dated (e.g., "Suppose that IBM is trading at $75 per share ..."). However, I could easily have made it "Stock XYZ" (contrary to the original wording), and you would not have noticed. Where possible, I retain the original wording for authenticity.

(2000). To answer the type of questions in this book, however, you may need the *extraordinary* advice in the next section.

The Interview

Will They Ask me These Questions?

Yes; you must assume that they will. You can hope for the best, but you must prepare for the worst. Some firms use the first round of interviews to get to know you with soft and non-quantitative questions. In this case, a second round typically follows with quantitative questions. Other firms use a quantitative first round to screen applicants up front. However, some firms ask no quantitative questions. There is thus a chance that you will see no quantitative questions. In this unlikely event, my quantitative questions will have increased your IQ, and my non-quantitative questions (Chapter 5) will have been of most assistance.

On the non-quantitative front, many interviewees have been asked, "Where did the Dow close yesterday?" or "Where did the Nikkei close?" or "Where is the long bond?" In addition to current knowledge, you should also know how these (and other) basic economic variables have changed over the recent past, and where they are relative to all-time highs and lows—see Chapter 5 for more examples. Even if you are very busy interviewing with many firms, you must not be found ignorant on such basic market knowledge.

ATQ!

"ATQ" stands for "Answer The Question!" Let us suppose that I am the interviewer and that I have little patience. I am busy. Damn busy! I have a deadline for my boss on a project that is due tonight (he is in an earlier time zone). I just walked away from the stack of work on my desk, and the computer simulation I desperately need to see the results of just so that I can talk to you. Spending 30 minutes with you means I get home at 10:30PM instead of 8:00PM, because I have to finish my project, and I will miss the last direct train. I earn \$250,000, \$500,000, \$1M, or more per annum. I got my job and kept it because I am efficent and I understand time management. I want to hire a good person, but if you waste my time then I will crucify you; perhaps not to your face, but to my colleagues, both at this firm, and at competing firms thinking of interviewing you.

I know from past experience that people with good resumes are not necessarily knowledgable in their claimed area of expertise. If you have a degree in finance, or mathematics, or whatever, and cannot answer a basic question in that area, then how the heck can I let you answer the phone when our traders call, or stick you in front of a client, or take you to a meeting with the portfolio managers, or have you join me in a conference call with my boss? That is, how can I hire you if you cannot answer basic questions? I know that there will come a time in this interview or the next when I have to push you to answer some quantitative questions, so that I can see what you understand and what you do not. Some of them will be basic, some of them not. I need to know the limits of your knowledge, and I cannot find them by asking soft wishy-washy questions about your resume.

If I ask you a question, then answer the damn question! If you know the answer, then tell me it. If you do not know the answer, but can work it out, then tell me that and outline the steps; I may be happy with that, and then not need to see the full derivation. If you have only a passing knowledge of the area, or no hope whatsoever of answering the question, then I need you to say so directly, and without wasting my time, so that I can you ask other questions. I need to know the boundaries of your abilities, and to find them I must ask you a mix of questions including ones that you cannot answer at all. Do not waste my time by floundering around and, in effect, drowning yourself in your own ignorance.

For example, suppose it is a bond trading job and I ask you whether the curvature in the plot of bond price versus yield to maturity is caused by changing Macaulay duration as yield changes. Let us suppose that you know the answer, but instead of giving it to me directly, you say:

"Well, that's an interesting question. We know that for a standard coupon-bearing bond with no embedded options, the plot of bond price versus yield to maturity is downward sloping and concave up. Let me draw that on the whiteboard here (draws picture). As yield rises, other things being equal, bond price falls, but the dollar rate at which the bond price falls actually decreases as yield to maturity rises. That is, the slope becomes less negative. Changing slope means that there is curvature, and sure enough the plot is concave up. Now, some people may think, naively, that the slope of the plot is just the Macaulay duration of the bond. Now, it is well known that as yield to maturity rises, other things being equal, the Macaulay duration of a standard coupon-bearing bond with no embedded options falls. So, these people would deduce, naively and incorrectly, that as yield to maturity rises, the changing slope is simply a reflection of changing Macaulay duration. However,

the simple fact that the Macaulay duration of a standard coupon-bearing bond with no embedded options is positive, and that the slope of our plot is negative, tells us that the slope is not the Macaulay duration. It is not the negative of the Macaulay duration either, and we can see that by looking at the case of a zero-coupon bond. Suppose we plot bond price versus yield to maturity for a ten-year zero with no embedded options. The plot is downward sloping and concave up as before, with slope becoming less negative as yield to maturity rises, but the duration is 10 years regardless of the yield—because it is a zero. That is, where the slope is of large magnitude, the Macaulay duration is ten; where the slope is of intermediate magnitude, the Macaulay duration is ten; where the slope is of small magnitude, the Macaulay duration is ten. Thus, slope does not equal Macaulay duration, or Macaulay duration, and the curvature of the plot cannot simply be a reflection of changing Macaulay duration. Now, the slope of the plot of the price of the standard coupon-bearing bond versus its yield to maturity is *a function of Macaulay duration, but it is also a function of bond price and yield to maturity. If we write down the slope explicitly, we see that it is* $-\frac{D}{(1+r)}P$, *where* D *is Macaulay duration,* P *is bond price, and* r *is yield. If we look at numerical examples, we can see however, that the duration does not change very much with changing yield. Indeed, as already mentioned, it does not change at all in the case of a zero, and low-coupon bonds are not that different from zeroes. Rather, it is the bond price that changes significantly with changing yield, and it is this that causes changes in the slope, thus producing curvature. Sure enough, in the case of a coupon-bearing bond, the changing Macaulay duration contributes to the change in slope, and thus to the curvature, but its contribution to curvature is much less important than the contribution of changing bond price. So,* no, *it is not changing Macaulay duration, but rather, changing price, that drives the change in slope, thus creating curvature."*

Well, you just spent two and one-half minutes of my valuable time saying that. That is ten percent of your interview time. In your favour, you got to the correct answer, which is "no," but in so doing you gave me so many words that I ceased caring whether you knew the answer or not. I did figure out, however, that if you were working on my team, I could not take you to a presentation to clients because you would take *for bloody ever* to answer their questions and bore the pants off them in the process. I also figured out that you really like hearing the sound of your own voice. You may well be someone who does not realise that time is money, that that money belongs to my clients, or to the firm, and that that money has a heck of a lot of zeros on the end of it.

You should have just answered "No, changing bond price drives changing slope and creates curvature." You could add that "Changing Macaulay duration contributes marginally to curvature for a coupon bearing bond, but not at all in the case of a zero." If the question has a "yes" or "no" answer, and you know the answer, then the first word out of your mouth should be "yes," or "no," respectively. Anything else means you are not getting to the point, and you are wasting my time and your golden opportunity! Obviously, you support your assertion immediately with more words, but answer the question first! ATQ!

I have had people talk a full ten minutes or more before coming anywhere near allowing me to detect whether they know the answer or not. After the first minute I have already decided that you are in the wrong building, and I am thinking about the stack of work on my desk. I stopped caring about your answer back in the first chapter of your saga. I am about to cancel the next person on your interview schedule because I value his or her time almost as much as I value my own. Unlike me, you get to go home early today.

To repeat, if the answer is "yes," and you know it, then say so! If the answer is "no," and you know it, then say so! You can add words after that, to support your answer, but for God's sake, get to the point! If you are on a date with a person you find exceptionally attractive, and you are dancing, and this person says "do you want to kiss me," are you going to talk to them for ten minutes about how you arrive at your decision or are you going to get to the point? Similarly, you must have had a professor at college who when asked a question from the audience mid-lecture would take five minutes giving his answer. When he got to the end of it you did not know what the answer was he had given, and you just wished he would shut up and move on. He invariably followed it up with "Is that clear now?" and no one dared say "no," for fear he would talk more about it. The bottom line is, answer the question! Remember ATQ, or even *ATFQ!*, if it helps hammer it into your skull.

If you do not know the answer, but know enough to try to work it out, than say something like "Hmmm. I do not know, but I think I can work it out. I know that the slope is given by $-\frac{D}{(1+r)}P$, where D is duration, P is price, and r is yield. I am not sure how much of the change in slope is explained by changes in each of D and P, but I do know that a zero has fixed D, so I suspect that changing P is more important than changing D." That is fine. You told me you did not know, and then you tried to work it out. That differs from knowing, but failing to tell me until the end of a saga.

If a question is not clear, be sure to ask for clarification. For example, "is it a straight bond with no embedded options?," "are there coupons?," etc.

If the interviewer tells you that your answer is incorrect when you know it is correct, and if you are dead sure of your answer, then defend youself to the hilt. Interviewers make mistakes, and you can earn their respect (and a job) by tactfully correcting them. Good people want you to do that in practice, though probably not in front of their colleagues.

Other Advice

You are not just interviewing for the job that was advertised. There are other openings in the firm that have not yet been advertised (and may never be advertised), and there are openings in other firms that your interviewer knows about because he or she knows people there. There will also be other openings at the interviewing firm in the future. If they like you and your CV, but do no think you are suited to that one job, they may recommend you strongly to another team leader within their firm or even at another firm. The implication of this is that if you discover quickly that you are not suited to the position advertised, or the firm, then you should steer the interview toward your strengths and ask the interviewer to keep you in mind for other positions. He or she may even tell you of another opening.

This works in reverse also. If your interview is awful, the interviewer will happily pass that information to other people when asked about you, or even without being asked if you really suck.

The finance community is relatively small and interwoven and corporate memory is long. If you interviewed at the firm before, your interviewer probably knows about it and will talk to the people you talked to. Indeed, if you worked/interviewed/studied anywhere in the world, the interviewer can find a former colleague, interviewer, adviser etc, of yours, who is known to them and who can comment on you. Your resume may have circulated widely within the firm, both in its local offices and overseas, before you set foot in the building or pick up the phone. Indeed, your resume might have circulated so widely that no one informs HR, and no one even remembers where your resume came from; that can explain why you never got any response, not even a rejection.

Your resume is a starting point. Do not inflate it. You *will* be asked about it. When a resume arrives on the desk of the interviewer, he or she looks at it and tries to figure out in advance some questions to ask. If you write on your resume that you took an option pricing class, and got an "A," then if the interviewer is an option pricing nut, you just guaranteed that the interview is going to get hot. If the area is a weakness for you, then do not make yourself a target. If you want to advertise that you took the class, then that is fine, but prepare yourself for incoming questions. ...and make sure your resume is proofread by people for whom English is their first language!

I received three cover letters that stand out in my mind. One from a young women applying for a junior quant position who stated that she had "a lot of love to give," one from a graduate of Rutgers who seemed to think I was sufficiently stupid not to have heard of Rutgers and felt the need to describe the school in some detail, and one from someone saying that they had always wanted to work in investment banking (when I was working for an asset management company). These were a source of amusement and ridicule, but nothing else. Why on earth would I interview you?

Be upbeat and enthusiastic. Even if the market is bad, and you are out of work. People like people who like them; it is that simple. If you tell me a tale of woe, all I can think is that "99% of your life is what you make it, and if your life sucks, you suck." Why would I want you sitting next to me at the office all day?

Do not smoke just before your interview. Get a stop smoking patch or something similar. The same goes for garlic for 24 hours before your interview. It stinks! Similarly, no one likes shaking hands with a limp dead fish. If your hands drip like a leaky fuacet, then put your hand in your pocket (warm and dry), or palm down on you lap right up until you get up to shake hands. It is simple but effective.

Intelligent or genuinely humourous small talk is fine, but do not make a fool of yourself. For example, one guy who came back a week later for a second round interview with me said "I remember you!" and "This is for the quant position, right?" Those were his first words!

Cover letters go in the garbage can, and e-mails are deleted. Make sure your e-mail address and phone number are on your resume. Similarly, buy an answering machine and check it often. If HR cannot find you quickly, then someone else can interview for your job before you.

Finally, the ex-post probability that you get the job is either zero or one. If you prepare as though it is zero, then it will be. If you prepare as though it can be one, then you can make it so.

Story: *A reader sent me the following e-mail: "...I bought your book... ...I just opened it to the first problem and was somewhat taken back by your solution. If you worked out the math, you would know that your answer is wrong. If you do not want to work out the math, then you could qualitatively grasp the mistake like so: ... ...Of course, to be quantitatively correct, you have to do the math. Always do the math... ...Hopefully, not too many interviewers have read this answer—or there will be lots of poor quants that will be turned away for being smart. Sincerely, YT.*

Like many an overconfident quant, YT jumped into the math without thinking *about the problem. Just as a pickpocket bumps you from the left while his accomplice takes money out of your right pocket, many of these questions are set up to distract you, and they are easier than they look.*

I shot YT back a tactful e-mail telling him why he was mistaken, and why no math was needed. I gave him a challenge question to solve to save face, but he crawled back under his rock, and has not been heard from since. How will he do in an interview?

Please feel free to send me e-mails with queries, corrections, alternative solutions, but especially with new interview questions. The errata can be found at the website below.

`www.InvestmentBankingJobInterviews.com`
`timcrack@alum.mit.edu`

Chapter 1

Purely Quantitative & Logic Questions

The only prerequisites for answering the questions in this chapter are elementary quantitative skills and common sense. Many questions in this chapter have two solution techniques: an elegant technique requiring little or no computation and a "hammer-and-tongs" brute-force approach. The technique you choose is revealing. Solutions for this chapter appear in Appendix A.

Question 1.1: You are given two glass jugs. Each contains the same volume, V, of liquid. One jug contains pure alcohol, and the other jug contains pure water. A modest quantity, Q, of water is poured from the water jug into the alcohol jug, which is then thoroughly mixed. The same modest quantity, Q, of (now diluted) alcohol is then poured back into the water jug to equalize the volumes of the jugs at their initial levels.

The initial concentration of alcohol in the alcohol jug equals the initial concentration of water in the water jug (at 100%). What is the relationship between the final concentrations of alcohol in the alcohol jug and water in the water jug?[1]

Question 1.2: What is the sum of the integers from 1 to 100?[2]

[1]This is not a chemistry problem. Please ignore the fact that mixing a volume V_1 of water with a volume V_2 of alcohol results in a total volume less than $V_1 + V_2$.

[2]More generally, what is the sum of the integers from 1 to n?

Question 1.3: (**) You are given a set of scales and 12 marbles. The scales are of the old balance variety. That is, a small dish hangs from each end of a rod that is balanced in the middle. The device enables you to conclude either that the contents of the dishes weigh the same or that the dish that falls lower has heavier contents than the other.

The 12 marbles appear to be identical. In fact, 11 of them are identical, and one is of a different weight. Your task is to identify the unusual marble and discard it. You are allowed to use the scales three times if you wish, but no more.

Note that the unusual marble may be heavier than the others, or it may be lighter. You are asked to both identify it and determine whether it is heavy or light.

Question 1.4: Interviewer: "You are a bug sitting in one corner of a *cubic* room. You wish to walk (no flying) to the extreme opposite corner (the one farthest from you). Describe the shortest path that you can walk. Be sure to mention direction, length, and so on."

Question 1.5: Picture a $10 \times 10 \times 10$ "macro-cube" floating in mid-air. The macro-cube is composed of $1 \times 1 \times 1$ "micro-cubes," all glued together. Weather damage causes the exposed (outermost) layer of micro-cubes to become loose. This outermost layer falls to the ground. How many micro-cubes are on the ground?

> **Story:** *One candidate for a futures trading position in Chicago was asked: "Would you rather be beaten up, beat someone up, or run around the block naked?" The last response did not get him the job. My wife was horrified to hear this story. Welcome to Chicago!*

Question 1.6: A mythical city contains 100,000 married couples but no children. Each family wishes to "continue the male line," but they do not wish to over-populate. So, each family has one baby per annum until the arrival of the first boy. For example, if (at some future date) a family has five children, then it must be either that they are all girls, and another child is planned, or that there are four girls and one boy, and no more children are planned. Assume that children are equally likely to be born male or female.

Let $p(t)$ be the percentage of *children* that are male at the end of year t. How is this percentage expected to evolve through time?

Question 1.7: How many degrees (if any) are there in the angle between the hour and minute hands of a clock when the time is a quarter past three?

Question 1.8: What is the first time after 3PM when the hour and minute hands of a clock are exactly on top of each other?

Question 1.9: There are 100 light bulbs lined up in a row in a long room. Each bulb has its own switch and is currently switched off. The room has an entry door and an exit door. There are 100 stockbrokers lined up outside the entry door. Each bulb is numbered consecutively from 1 to 100. Each stockbroker is numbered consecutively from 1 to 100.

Broker number 1 enters the room, switches on *every* bulb, and exits. Broker number 2 enters and flips the switch on every *second* bulb (turning off bulbs 2, 4, 6, ...). Broker number 3 enters and flips the switch on every *third* bulb (changing the state on bulbs 3, 6, 9, ...). This continues until all 100 brokers have passed through the room.

What is the final state of bulb number 64? Is it illuminated or dark?

Question 1.10: Exactly the same set-up as Question 1.9, with a different final question: How many of the light bulbs are illuminated after the 100th person has passed through the room, and which light bulbs are they?

Question 1.11: Your bedroom sock drawer contains eight red socks and 11 blue socks that are otherwise identical. The light is broken in your bedroom, and you must select your socks in the dark. What is the minimum number of socks you need to take out of your drawer and carry into your (well-lit) living room to guarantee that you have with you at least a matching pair to choose from?

> **Story:** *One of my students was asked to "Describe the best boss you have ever had." Watch out for the opposite question: "Describe the worst boss you have ever had." Your answer may indicate disloyalty to a (former) employer.*

Question 1.12: You and I are to play a competitive game. We shall take it in turns to call out integers. The first person to call out "50" wins. The rules are as follows:

1. The player who starts must call out an integer between one and 10, inclusive;
2. A new number called out must exceed the most recent number called by at least one and by no more than 10. For example, if the first player calls out "nine," then the range of valid numbers for the opponent is 10 to 19, inclusive.

Do you want to go first, and if so, what is your strategy?

Question 1.13: You are to open a safe without knowing the combination. Beginning with the dial set at zero, the dial must be turned counter-clockwise to the first combination number, (then clockwise back to zero), and clockwise to the second combination number, (then counter-clockwise back to zero), and counter-clockwise again to the third and final combination number, whereupon the door shall immediately spring open. There are 40 numbers on the dial, including the zero.

Without knowing the combination numbers, what is the maximum number of trials required to open the safe (one trial equals one attempt to dial a full three-number combination)?

Question 1.14: (**) You are given a set of scales and 90 coins (this question is similar to Question 1.3). The scales are of the old balance variety. That is, a small dish hangs from each end of a rod that is balanced in the middle. The device enables you to conclude either that the contents of the dishes weigh the same or that the dish that falls lower has heavier contents than the other. You must pay $100 every time you use the scales.

The 90 coins appear to be identical. In fact, 89 of them are identical, and one is of a different weight. Your task is to identify the unusual coin and to discard it while minimizing the maximum possible cost of weighing.[3] What is your algorithm to complete this task? What is the

[3] A slightly different task is to minimize the expected cost of weighing. Minimizing the expected cost of weighing does not necessarily minimize the maximum possible cost. This is a subtle distinction that you should not overlook.

most it can cost to identify the unusual coin (assuming your strategy minimizes the maximum possible cost)?

Note that the unusual coin may be heavier than the others, or it may be lighter. You are asked to both identify it and determine whether it is heavy or light.[4]

Question 1.15: **(***)** Suppose that the function $f(z)$ is complex valued in the complex plane. Suppose also that $f(z)$ is both bounded and entire. Prove that $f(z)$ must be a constant.[5]

Question 1.16: You start with a single lily pad sitting on an otherwise empty pond. You are told that the surface area of the lily pad doubles every day and that it will take 30 days for the single lily pad to cover the surface of the pond.[6]

If instead of one lily pad you start with eight lily pads (each identical in characteristics to the original single lily pad), how many days will it take for the surface of the pond to become covered?

Question 1.17: Another lily pad problem. There are 27 lily pads on a pond. The pond is 6,000 square feet in area. The lily pads are one square foot in area. Each lily pad doubles its size every day. How long until the pond is covered in lily pads?

> **Story:** *A student of mine was taken to a room and asked to choose a place to sit at a long oval-shaped table. He chose a place at random. Later the interviewer asked why he had chosen that spot. I think the intent was to see if he was a leader (sitting at the head) or a follower (sitting at the side).*

Question 1.18: Interviewer: "Alright, you're from MIT; you must be a quantitative type of person." Interviewee: (confidently, after a slight

[4]Does the answer change if you must identify the coin without saying whether it is heavy or light?

[5]Recall that an "entire" function is a function that is analytic in the entire finite complex plane. Thus, $f(z)$ may be represented by an everywhere-convergent power series: $f(z) = \sum_{n=0}^{\infty} a_n z^n$ (Holland [1973, p5]).

[6]The student who was asked this question says that his interviewer used the number 30. However, he suggested that I use the number 3,000 to make it more complicated. What is wrong with saying that it takes 3,000 days for the lily pad to cover the pond?

pause) "Yes indeed." Interviewer: "Give me the decimal equivalent of $\frac{13}{16}$ and of $\frac{9}{16}$."

Question 1.19: A snail is climbing up a 10-foot pole. It climbs up by three feet every day. Each night it sleeps. While sleeping, it slides down by one foot. When does it reach the top of the pole?

Question 1.20: **(**)** A windowless room contains three identical light fixtures, each containing an identical light bulb. Each light is connected to one of three switches outside of the room. Each bulb is switched off at present. You are outside the room, and the door is closed. You have one, and only one, opportunity to flip any of the external switches. After this, you can go into the room and look at the lights, but you may not touch the switches again. How can you tell which switch goes to which light?

Question 1.21: Inside of a dark closet are five hats: three blue and two red. Three smart men go into the closet, and each selects a hat in the dark and places it unseen upon his head. Each man knows both that the closet contains three blue hats and two red and that the other two men have the same knowledge.

Once outside the closet, no man can see his own hat. The first man looks at the other two, thinks, and says, "I cannot tell what colour my hat is." The second man hears this, looks at the other two, and says, "I cannot tell what colour my hat is either." The third man is blind. The blind man says, "Well, I know what colour my hat is." What colour is his hat, and how does he know?[7]

Question 1.22: **(**)** Find the smallest positive integer that leaves a remainder of 1 when divided by 2, a remainder of 2 when divided by 3, a remainder of 3 when divided by 4, ... and a remainder of 9 when divided by 10.

[7] In fact, the actual interview question (not provided to me) had exactly the same logical structure but had different "players."

Story: *1. During the interview, an alarm clock went off from the candidate's briefcase. He took it out, shut it off, apologized, and said he had to leave for another interview. 2. An applicant came in wearing only one shoe. She explained that the other shoe was stolen off her foot in the bus.*

Interview Horror Stories from Recruiters
Reprinted by kind permission of MBA Style Magazine
©*1996–2004 MBA Style Magazine,* www.mbastyle.com

Question 1.23: There are two motorcyclists on a single lane road. They are 25 miles apart. At a signal, they start moving toward each other with constant speeds. The first motorcyclist rides at 20 mph; the second rides at 30 mph. When the signal goes off, a fly on the helmet of the first motorcyclist is startled and starts flying toward the second motorcyclist at 40 mph. When the fly reaches the second motorcyclist (now moving toward the first), he immediately reverses course and flies back to the first motorcyclist. When the fly gets back to the first motorcyclist, he reverses course again. The fly continues to fly backwards and forwards between the two motorcyclists until they all collide. How many miles will the fly have travelled before his life is extinguished?

Question 1.24: A, B, C, D, E, F, G, H, and I, are the nine integers from one to nine (not necessarily in order). They satisfy the following constraints:

$$\begin{aligned} A+B+C+D &= 20, \\ B+C+D+E+F &= 20, \\ D+E+F+G+H &= 20, \quad \text{and} \\ F+G+H+I &= 20. \end{aligned}$$

What values are taken by each of A to I?

Story: *I recall reading in the WSJ about one young woman who was asked her greatest weakness (a common interview question). Without thinking, she blurted out the answer "chocolate!"*

Question 1.25: A small boat is floating in a swimming pool. The boat contains a very small but very heavy rock. If the rock is tossed out of the boat into the pool, what happens to the water level in the pool?

> **Story:** *Instead of being asked her greatest weakness, one of my students was asked: "Why shouldn't we hire you?" It is pretty difficult to manoeuvre your way out of that one!*

Question 1.26: Prove that the area of a triangle is given by

$$A = \sqrt{s(s-a)(s-b)(s-c)},$$

where a, b, and c are the side lengths, and $s \equiv \frac{a+b+c}{2}$ is half the perimeter.[8]

Question 1.27: A very large number, N, of people arrive at a convention. There are exactly N single rooms in the hotel where the convention takes place. Each guest is given a numbered key for a specific room. Before they even go upstairs, they are all invited to a large party in the banquet hall. To gain admittance to the hall, they have to give up their keys to a doorman. At the end of the evening, the guests are not sober enough to recall their room numbers, so the doorman simply hands out the keys randomly. Each guest ends up spending the night in a random room. What is the probability that at least one guest ends up in the room to which he or she was originally assigned?

Question 1.28: (**) In a certain matriarchal town, the women all believe in an old prophecy that says there will come a time when a stranger will visit the town and announce whether any of the menfolk are cheating on their wives. The stranger will simply say "yes" or "no," without announcing the number of men implicated or their identities. If the stranger arrives and makes his announcement, the women know that they must follow a particular rule: If on any day following the stranger's announcement a women deduces that her husband is not faithful to her, she must kick him out into the street at 10AM the next day. This action is immediately observable by every resident in the town. It is well known that each wife is already observant enough to know whether

[8] Mark Rubinstein kindly pointed out to me that this is "Heron's Formula."

any man (except her own husband) is cheating on his wife. However, no woman can reveal that information to any other. A cheating husband is also assumed to remain silent about his infidelity.

The time comes, and a stranger arrives. He announces that there are cheating men in the town. On the morning of the tenth day following the stranger's arrival, some unfaithful men are kicked out into the street for the first time. How many of them are there?

Question 1.29: In front of you are three poles. One pole is stacked with 64 rings ranging in weight from one ounce (at the top) to 64 ounces (at the bottom). Your task is to move all the rings to one of the other two poles so that they end up in the same order. The rules are that you can move only one ring at a time, you can move a ring only from one pole to another, and you cannot even temporarily place a ring on top of a lighter ring.

What is the minimum number of moves you need to make to achieve the task?

> **Story:** *Here are some common thinking questions from Section 5.5: "How many McDonald's fast food outlets are there in the US? How many gas stations are there in the US? How many elevators are there in the US?"*

Question 1.30: Solve the following ordinary differential equation (ODE):

$$u'' + u' + u = 1.$$

> **Story:** *One of my students went to an interview with a big-name Wall Street firm in New York. He was interviewed by five quantitative guys in a row. Each interview was one hour, and there were absolutely no breaks. He had to work through multiple quantitative problems on their blackboard. They gave him no lunch. He was exhausted and starving by the end. He was swearing black and blue about the "&@!#@$%'s" when he got back. He said "The Russian" was the worst.*

Question 1.31: Assume that the random variables X and Y are Normally distributed: $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$. The correlation between X and Y is ρ. How can you choose constants a and b such that you minimize the variance of the random variable sum $S = aX + bY$ under the constraints that $a + b = 1$, $0 \leq a \leq 1$, and $0 \leq b \leq 1$?[9]

Question 1.32: Suppose there is a straight coastline and a lighthouse that is $L = 3$ miles away from the coast. The light revolves at one revolution per minute. How fast is the beam of light travelling along the coastline? When the beam is $3L$ away from the coastal point closest to the light, how fast is the light travelling along the coast?

Question 1.33: I have a 20×20 chessboard and a very large box of identical cubes. Each square on the chessboard is the same size as the face of any cube. I am going to arrange piles of cubes on the chessboard in a special pattern. I align one edge of the board so it is running north–south. I start at the northwest corner by placing one cube on that square. Whenever I step to the south or the east, I place a pile of cubes containing one more cube than in the previous square. This produces the pattern in Figure 1.1. How many cubes in total are there on the chessboard?

1	2	3	4	$\cdots$	19	20
2	3	4	5	$\cdots$	20	21
3	4	5	6	$\cdots$	21	22
4	5	6	7	$\cdots$	22	23
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
19	20	21	22	$\cdots$	37	38
20	21	22	23	$\cdots$	38	39

Figure 1.1: Number of Cubes on Each Square of a 20×20 Chessboard (Q)

[9] Another version of this question asked in interviews is: "You are driving around with one wheel on the gravel and one wheel on the pavement. The variance of the gravel and pavement surfaces are described by σ_G^2 and σ_P^2. Whereabouts on the axle should you sit between $x = 0$ (right over the wheel on the gravel) and $x = 1$ (right over the wheel on the pavement) if you want the most comfortable ride?"

Question 1.34: You are standing at the centre of a circular field of radius R. The field has a low wire fence around it. Attached to the wire fence (and restricted to running around the perimeter) is a large, sharp-fanged, hungry dog who likes to eat any humans he can catch. You can run at speed v. Unfortunately, the dog can run four times as fast, at $4v$. The dog will do his best to catch you if you try to escape the field. What is your running strategy to escape the field without feeding yourself to the dog?

Question 1.35: Please prove that the following relationship holds:

$$\int_{-\infty}^{+\infty} e^{-x^2}dx = \sqrt{\pi}.$$

Question 1.36: What is $\int \sec\theta d\theta$ equal to?[10]

Question 1.37: Does the infinite sum $\sum_{n=1}^{\infty} e^{-\sqrt{n}}$ converge?

> **Story:** *One interviewee told me that the interviewers aim to put you under as much pressure as possible, and that "you never know when they are going to bring out the guy in the chicken suit."*

Question 1.38: What are $\sum_{k=1}^{n} k^2$, and $\sum_{k=1}^{n} k^3$?

Question 1.39: You are given eight balls. They appear identical, but one is heavier than the rest. As in the previous ball questions, you have a pair of scales. How do you find the heavy ball?

> **Story:** *She threw up on my desk and immediately started asking questions about the job, like nothing had happened.*
>
> *Interview Horror Stories from Recruiters*
> *Reprinted by kind permission of* MBA Style Magazine
> ©*1996–2004 MBA Style Magazine,* www.mbastyle.com

[10] Similarly, you could see questions on integrals (or derivatives) of $\sin\theta$, $\cos\theta$, $\tan\theta$, $\cot\theta$, and $\operatorname{cosec}\theta$.

Question 1.40: We are to play a game on a table in the next room. We each have an infinite bag of identical quarters (i.e., American 25-cent pieces). We will take it in turns to put one quarter on the table. Quarters may not overlap on the table. When there is no room left on the table to put another quarter, the winner is the last person to put a quarter on the table. Let me tell you that there does exist a strategy for winning and that this strategy is independent of the size of the table.

1. What is the shape of the table?
2. Do you start?
3. What is your strategy for winning?
4. Is there any case where this does not work?

Question 1.41: One analyst (John) is talking to another (Mary) while working on a deal book at 2AM Mary learns that John's sister has three children. "How old are the children?" asks Mary. "Well," replies John, "the product of their ages is 36." Mary thinks for a while and says, "I need more information." "Hmmm, the sum of their ages is the same as this figure right here," says John pointing at the spreadsheet. "Still not enough information," says Mary after thinking for a minute. "The eldest is dyslexic," says John. How old are the children?

Question 1.42: **(***)** You have 52 playing cards (26 red, 26 black). You draw cards one by one. A red card pays you a dollar. A black one fines you a dollar. You can stop any time you want. Cards are not returned to the deck after being drawn. What is the optimal stopping rule in terms of maximizing expected payoff? Also, what is the expected payoff following this optimal rule?[11]

Question 1.43: You have a chessboard (8×8) plus a big box of dominoes (each 2×1). I use a marker pen to put an "X" in the squares at coordinates (1,1) and (8,8)—a pair of diagonally opposing corners. Is it possible to cover the remaining 62 squares using the dominoes without any of them sticking out over the edge of the board and without any of them overlapping? You must not damage the board or the dominoes in

[11] Try the same question with four cards (two red, two black).

the process or do anything weird like standing them on their ends—just answer the question.[12]

Question 1.44: One of my students interviewed with some folks who "wanted to get an idea of his comfort with formulae and with explaining things to clients." They asked why it is that if p is a prime number greater than 3, then $p^2 - 1$ is always divisible by 24 with no remainder.

Question 1.45: You are bidding for a firm whose unknown true value is uniformly distributed between 0 and 1. Although you do not know the true value S of the firm, you do know that as soon as people learn that you have made a bid this news will cause the value to double to $2S$. Your bid, however, will be accepted only if it is at least as large as the original value of the firm. How do you bid so as to maximize your expected payoff?

Question 1.46: You have a string-like fuse that burns in exactly one minute. The fuse is inhomogeneous, and it may burn slowly at first, then quickly, then slowly, and so on. You have a match, and no watch. How do you measure exactly 30 seconds?

Question 1.47: How many places are there on the Earth where you can walk one mile south, one mile east, one mile north, and end up exactly where you started? Assume the Earth is a perfect sphere, that your compass bearing is constant on each leg of the walk, that all parts of the Earth are able to be walked upon, and that your feet are arbitrarily small.

Question 1.48: This is an absolute classic. A king demands a tax of 1,000 gold sovereigns from each of 10 regions of his nation. The tax collectors for each region bring him the requested bag of gold coins at year end. An informant tells the king that one tax collector is cheating and giving coins that are consistently 10% lighter than they should be, but he does not know which collector is cheating. The king knows that each coin should weigh exactly one ounce. How can the king identify the cheat by using a weighing device exactly once?

---

[12] Naoki Sato has suggested a follow up question. Place an "X" on two squares: one black, and one white. Can you cover the remaining squares with dominoes? See Answer 1.43 for the solution.

Question 1.49: Again, an absolute classic. You hire a man to work in your yard for seven days. You wish to pay him in gold. You have one gold bar with seven parts—like a chocolate bar. You wish to pay him one gold part per day, but you may snap the bar in only two places. Where do you snap the bar so that you may pay him at the end of each day, and so that on successive days he may use what you paid him previously to make change?

Question 1.50: How many consecutive zeroes are there at the end of 100! (100 factorial). For example, 12!=479,001,600 has two consecutive zeroes at the end.

Question 1.51: Why are images in a mirror flipped horizontally and not vertically? For example, although I wear my wristwatch on my left wrist, and my reflection wears his on his right wrist, my reflection is not standing on his head.

Question 1.52: I am told this is a genuine finance interview question. It had to be a trading interview, because no one but a trader would ask this in an interview. I considered transforming the question, but left it as is for authenticity. Avert your eyes if you are easily offended!

How can three men and one women have mutually safe heterosexual intercourse with just two condoms? Assume that no condom can break or leak, and that you cannot wash a used one.[13]

Question 1.53: Finally, can the mean of any two consecutive prime numbers ever be prime?

[13] With one man and three women, the answer is of similar type, but different.

Chapter 2

Derivatives Questions

A prerequisite for answering the questions in this chapter is knowledge of basic option pricing theory. I strongly recommend my book *Basic Black-Scholes: Option Pricing and Trading* (Crack [2004]) as the best resource. It provides a firm foundation in Black-Scholes option pricing, with practical advice about option trading. See the advertisement on the last page of this book, or go directly to `www.BasicBlackScholes.com`. Solutions for this chapter appear in Appendix B.

Question 2.1: All Black-Scholes assumptions hold. Assume no dividends. The stock price is \$100. The riskless interest rate is 5% per annum. Consider a one-year European call option struck at-the-money (i.e., strike equals current spot). If the volatility is zero (i.e., $\sigma = 0$), what is the call worth? After valuing the call, please tell me how to hedge the call.

Question 2.2: Two standard options have exactly the same features, except that one has long maturity, and the other has short maturity. Which one has the higher gamma?

Question 2.3: Suppose that IBM is trading at \$75 per share. What does it cost to construct a derivative security that pays exactly one dollar when IBM hits \$100 for the first time? Explain carefully the construction of the security. You may ignore IBM's dividends, assume a riskless interest rate of zero, assume all assets are infinitely divisible, ignore any short sale restrictions, and ignore any taxes or transactions costs.[1]

[1] It will obscure the question somewhat, but you may think of this derivative as a

Question 2.4: All Black-Scholes assumptions hold. Assume no dividends. Consider a standard European call and a standard European put on the same stock. Assume that each option has the same maturity, and is struck at-the-money (i.e., strike equals current spot). For the sake of simplicity, assume that the interest rate is zero. Draw the payoff diagrams for each option (i.e., terminal payoff to option versus level of underlying).

The put has limited downside potential and no upside; the call has unlimited upside and no downside. Given the random direction of the stock price movements between now and expiration, the disparity in potential payoffs seems to suggest that the call should be worth more than the put. However, put-call parity says that this is not so. Verify the put-call parity implications and reconcile them with the seemingly disparate potential payoffs.

Question 2.5: Assume a Black-Scholes world without dividends. Consider a standard European call struck at-the-money (i.e., strike equals current spot) with one year to maturity. If the interest rate is $r = 0.06$, is the option's delta greater or less than 0.5? What does it depend on?

Question 2.6: Assume a Black-Scholes world with continuous dividends. Consider a standard European call struck at-the-money (i.e., strike equals current spot) with one year to maturity. If the interest rate is $r = 0.06$, and dividends are at rate $\rho = 0.03$, can you tell whether the option's delta is greater or less than 0.5? What does it depend on?

> **Story:** *Some recent questions include "What do you think an investment banker does?" Not only that, but "Do you understand the hours investment bankers work and why?" Some of these folks look like Hell when you meet them. Are you sure about this career choice?*

Question 2.7: You are long a call option on MITCO stock. You have delta-hedged your position. You hear on the radio that the CEO of MITCO plunged from the roof of the MITCO skyscraper in New York

perpetual European up-and-out option with a knock-out barrier at \$100 and a rebate of \$1. You do not need to know what the strike is or whether it is a put or call, because this derivative is really just the rebate itself and nothing more.

(for personal reasons). The stock price plunges \$10 in sympathy. How do you adjust your hedge (qualitatively)? That is, do you borrow and buy stock or sell stock and lend? Explain carefully.

Question 2.8: How do you calculate an option's delta?[2]

Question 2.9: Explain very carefully the terms $N(d_1)$ and $N(d_2)$ that appear in the standard Black-Scholes European call option pricing formula without dividends.[3]

Question 2.10: Consider the European digital option (or "binary option") that pays a constant H if the stock price is above strike price X at expiration and zero otherwise. What is the price of this option, and how is it related to the price of the standard Black-Scholes European call option? Explain carefully.[4]

Question 2.11: Consider the European digital option (or "binary option") that pays H if the stock price is above strike price X at expiration and zero otherwise. How does the price of this option vary with volatility (that is, what is $\frac{\partial C}{\partial \sigma^2}$)? Intuitively? Rigorously? Explain carefully.

Question 2.12: Compare the "delta" of a standard European call option and the delta of a barrier option, for example a "down-and-out" call option.[5]

---

[2] Answer for a standard European call option (with and without dividends), and for an exotic option with no closed-form solution.

[3] Now use this explanation to deduce the standard Black-Scholes European put option pricing formula—if you can. Confirm that the pricing formulae verify the put-call parity relationship (with $D = 0$):

$$S(t) + p(t) = c(t) + Xe^{-r(T-t)} + D.$$

[4] This is the "cash-or-nothing" digital option. You should also be able to answer this question for the "asset-or-nothing" digital option (which gives you the asset if $S(T) > X$ and nothing otherwise).

[5] Is the answer different for an up-and-out call? Explain carefully. Incidentally, who would buy an up-and-out call? Well, suppose you expect only limited upside on a security. If you wish to participate in this upside without paying for what you consider to be very unlikely further price appreciation, then an up-and-out-call could be just what you want (see Derman and Kani [1993, pp3–4]).

Story: *One of my students was asked to "Describe the best party you have ever been to." She said this big-name Wall Street investment bank was looking for "fun loving" people.*

Question 2.13: (**) This is an applied theoretical option pricing problem taken from the telephone interview of an MIT student. You are given three time series of continuously compounded returns on an industry sector index: the ISI50. The time series are daily, weekly, and monthly over the same time period.[6] You are to price a standard European call option written on the level of the ISI50 with one month to expiration.

You decide to use the trusty Black-Scholes model. You observe all input variables except for the volatility term σ^2. Unsure of which of your three time series to use to estimate the volatility term, you calculate the sample volatility of each time series. You figure that the estimators $(\hat{\sigma}_d^2, \hat{\sigma}_w^2, \hat{\sigma}_m^2)$ should be related as $\hat{\sigma}_m^2 \approx 4\hat{\sigma}_w^2$, $\hat{\sigma}_m^2 \approx 20\hat{\sigma}_d^2$, and $\hat{\sigma}_w^2 \approx 5\hat{\sigma}_d^2$. You could thus get the monthly volatility either explicitly from the monthly estimate or implicitly from the weekly or daily estimates. You think the daily data are most reliable (they have the most observations).

You find, much to your horror, that $\hat{\sigma}_m^2 > 4\hat{\sigma}_w^2$, $\hat{\sigma}_m^2 > 20\hat{\sigma}_d^2$, and $\hat{\sigma}_w^2 > 5\hat{\sigma}_d^2$. Further investigation reveals that these differences are highly statistically significant. Your statistical observation is thus that the monthly volatility implicit in the daily and weekly time series is significantly smaller than the monthly volatility in the monthly time series.

How do you price the option? Explain your reasoning carefully.[7]

Question 2.14: For a standard European call option, draw the graph of the "delta" as a function of current stock price, $S(t)$.

Question 2.15: Consider a plain vanilla American call option on a non-dividend-paying stock. The price of the call is $C(t)$ at time t. The "intrinsic value" of the call is $\max[S(t) - X, 0]$ (where $S(t)$ is stock

[6] Feel free to assume that one week is exactly five days, one month is exactly 20 days, and that there are no missing observations or exchange holidays.

[7] Hint: Begin by explaining how and why your statistical observations could arise. What went wrong? Ask yourself whether Black-Scholes pricing is still applicable. If not, where do you turn?

price at time t, and X is exercise price). The excess of call value over intrinsic value is the "time value" of the option.[8]

Draw a graph of the time value, $C(t) - \max[S(t) - X, 0]$, versus $S(t)$. Explain carefully the different aspects of the plot.

> **Story:** *A student of mine was asked "How would you value yourself?" That is, put a dollar figure on your value using discounted cash flow analysis.*

Question 2.16: It is 10 months since you sold a one-year European call option to a customer. You have been delta-hedging your exposure to the written call since it was sold. The option is now well in-the-money, and the delta of your replicating portfolio is correspondingly high (at around 0.90, say).

Suppose that you watch the underlying stock price falling gently over the last two months of the life of the option. As the stock price falls over this time period, what happens to the delta of the replicating portfolio? That is, are you buying stocks or selling stocks as you watch the stock price fall? You may have to describe different possible scenarios—be clear on the assumptions you make.

Question 2.17: What do you know about jump processes and jump diffusion processes? Explain when the pricing formula for a call option written on an asset whose price level follows a jump process can and cannot be derived using the Black-Scholes/Merton no-arbitrage technique.[9]

---

[8]Perhaps a more natural definition of intrinsic value is $\max[S(t) - Xe^{-r(T-t)}, 0]$ (Merton [1973, p145]; Merton [1992, p260]; Smith [1976, p11]). What would the plot of time value versus $S(t)$ look like with this definition of intrinsic value?

[9] Describe the form of the pricing formula for a European call option written on a stock whose price level follows a jump diffusion process (using Merton's notation): $\frac{dS}{S} = (\alpha - \lambda k)dt + \sigma dZ + dq$, where $dq = 0$ if the "Poisson event" (i.e., the jump) does not occur, $dq = (Y-1)$ if the jump does occur, $(Y-1)$ is a spike producing a finite jump in stock price from S to SY, α is the instantaneous expected rate of return on the stock, σ^2 is the instantaneous variance or returns assuming no junp occurs, dZ is a standard Wiener process, λ is the number of arrivals that you expect per unit time, $k \equiv E(Y-1)$ where E is the expectation operator over the random variable Y, and dZ is assumed independent of the Poisson process dq (see Merton [1992, p313]).

Question 2.18: This question concerns the standard European call option on a non-dividend-paying stock. You are asked to draw three closely related graphs as follows:

1. Please draw the graph of call price at maturity (time T) versus terminal stock price, $S(T)$.
2. Please draw the graph of call price at time t versus the futures price $F(t, T)$. The futures price $F(t, T)$ is observed at time t, prior to maturity. The futures contract and the option both mature at the same date T.[10]
3. Now draw the graph of call price versus stock price at time t, prior to maturity.

Explain carefully the relationships between the three graphs.

Question 2.19: Consider two European call options on the same underlying stock. The options have the same strike price. Assume constant interest rates. One option matures in one year; the other option matures in four years. Suppose that you put $\sigma = 15\%$ into the Black-Scholes formula to value the one-year option. What value of σ do you put into the Black-Scholes formula to value the four-year option? Assume that you set $T - t = 1$ in the Black-Scholes formula in both cases (i.e., one unit of time equals four years in the second case but only one year in the first case).

Question 2.20: **(***)** The Black-Scholes formula is derived assuming the stock price process $S(t)$ follows a geometric Brownian motion: $dS(t) = \mu S(t)dt + \sigma S(t)dw(t)$, where $w(t)$ is a standard Brownian motion. Suppose instead that a stock price process $\mathcal{S}(t)$ follows an arithmetic Brownian motion: $d\mathcal{S}(t) = \mu dt + \sigma_A dw(t)$. Derive the pricing formula for a call option on $\mathcal{S}(t)$. Please assume that the option is at-the-money [i.e., $\mathcal{S}(t) = \mathcal{X}$], that the riskless interest rate $r = 0$, and that the stock pays no dividends.

Question 2.21: Interviewer: "You are fully familiar with Black-Scholes pricing aren't you?" Interviewee: (confidently, after a slight pause)

[10] It is only recently (November 2002) that futures on single stocks have been traded in the US. See `www.OneChicago.com` for details.

"Yes indeed." Interviewer: "What is the value of a three-month at-the-money (i.e., $S = X$) call option on a \$100 stock when the implied vol is 40? Please assume $r = 0$ (it is the least important ingredient anyway) and assume also that the stock pays no dividends. You have 10 seconds to perform the calculation in your head. Now tell me how your answer changes if it is instead a put."

Question 2.22: A customer calls up and wants a price on a European 100-day call option. You quote \$100. He calls back a minute later and wants a quote on the same option but with 200 days to maturity. How does the second price quote compare to the first price quote? Explain carefully.

Question 2.23: Assume a Black-Scholes world. You have a one-year European call option on a stock. There are no dividends, the interest rate is assumed to be zero, and the option is struck at-the-money (i.e., strike equals current spot). The current spot is \$100. The standard deviation of terminal stock price (conditional on current stock price) is \$10.[11] Is the call price closer to \$1, \$5, or \$10?[12]

Question 2.24: You hold a 100-day European call option on a stock with implied volatility 20. Suppose that you know right now that tomorrow the implied volatility will increase to 25, but that after that it will return to 20 for the remainder of its life. What extension to the life of the call would produce the same change in the present value of the call as the above-mentioned single-day increase in volatility (assuming a constant implied volatility at 20)? That is, other things being equal, what change in the term to maturity is equivalent to the quoted one-day change in the implied volatility? Explain carefully.

Question 2.25: You are long a straddle with a strike of \$25. The underlying is at \$25. The straddle costs you \$5 to enter. What price movement are you looking for in the underlying?

Question 2.26: You are considering two contracts: a Eurodollar *futures* contract, with six months to maturity, selling at 5%, settled on three-month LIBOR, marked-to-market every day; and a Eurodollar *forward*

[11] It follows that the standard deviation of continuously compounded returns is approximately 10% per annum.

[12] If the standard deviation is \$20 per annum, is the call price closer to \$5, \$10, or \$20?

contract, with six months to maturity, selling at 5%, settled on three-month LIBOR at maturity.

1. Which contract do you prefer (or are you indifferent)?
2. Do you think there is a mis-pricing?
3. If you go long one and short the other, which one should be long, and which one should be short (or are you indifferent)?

Question 2.27: You are to value a call option using Monte-Carlo simulation. Is it better to simulate the geometric Brownian motion (GBM) process for the call itself, or the GBM process for the underlying?

Question 2.28: Suppose that you hold a long position in mortgage-backed securities. If you are expecting a bond market rally, would you be better off with positive convexity or negative convexity?

> **Story:** *There is the old story of the candidate who flew to London for an interview. At the interview, the interviewer excused himself for a few minutes. However, before leaving he asked the interviewee to open a window. Once alone, the interviewee discovered that all the windows were sealed shut. Great! Michael Lewis (in his excellent book* Liar's Poker*) talks about this technique in use on Wall Street (Lewis, 1990, p27). He suggests that one desperate interviewee threw a chair through Lehman's 43rd floor window in Manhattan!*

Question 2.29: What is wrong with the following strategy for hedging a short call option: buy one share if the stock price exceeds the strike, and sell the share if the stock price falls below the strike?

Question 2.30: How fresh is your stochastic calculus? What can you tell me about $\int_0^T w(t)dt$, where $w(t)$ is a standard Brownian motion?

Question 2.31: What can you say about $\int_0^T w(t)dw(t)$, where $w(t)$ is a standard Brownian motion?

Question 2.32: **(**)** The payoff to a European "power call" is given by $\max(S^\alpha - X, 0)$. Derive the price of a European power call option using

Black-Scholes pricing.[13,14]

> **Story:** *Here is a question they can spring to see how keen you are about this particular job: "If we offer you a job right now, will you take it?"*

Question 2.33: Why do you get a "smile" effect when you plot implied volatilities of options against their strike prices?

Question 2.34: Is the price of a double-barrier, knock-out option (i.e. one with both up-and-out and down-and-out barriers) just the price of an up-and-out plus the price of a down-and-out?

Question 2.35: Consider an American-style double-barrier "out-in" call option. There is an out barrier above the current stock price (an "up barrier") and an in barrier below the current stock price (a "down barrier"). This option has a payoff only if all three of the following events happen: first, the stock price path includes a fall in price below the down barrier (i.e., the option is "knocked in"); second, the stock price path does not include a rise in price above the up barrier (i.e., the option is not "knocked out"); and third, the option is exercised when the stock price is above the strike (i.e., the option is in-the-money at exercise). This option is both path-dependent and American-style. Is there an easy technique for valuing the option?

Question 2.36: Suppose gold prices follow a Gaussian process.[15] The current price of gold is \$400. The riskless interest rate is zero. The volatility of gold in dollar terms is $\sigma = \$60$ per annum. What is the value today of a digital cash-or-nothing option that pays \$1 million in six months if the price of gold is at or above \$430?

---

[13]Try drawing the payoff diagrams for the cases $\alpha > 1$ and $\alpha < 1$. Add the current call value as a function of stock price to your diagrams.

[14]Jarrow and Turnbull (1996, p175) describe a "powered option" with payoff $[S(T)-X]^2$ if $S(T) \geq X$ and zero otherwise. I give the general result for the case $[S(T)-X]^\alpha$ in the solutions. Try to derive it before you peek at my solutions.

[15]This is an arithmetic Brownian motion. The future price of gold is thus assumed to be Normally distributed (not Lognormally as per Black-Scholes).

Question 2.37: Describe the analytical procedure for deriving (using calculus) the values of European digital asset-or-nothing and digital cash-or-nothing options.

> **Story:** *Here is a quirky thinking question from Section 5.5: "How many ping-pong balls can you fit in a jumbo jet (e.g., Boeing 747)?"*

Question 2.38: **(**)** What is the value of a perpetual (i.e., potentially infinitely lived) American put option?

Question 2.39: Let "L" denote the three-month US dollar LIBOR rate. Consider an interest rate swap arrangement where Party A pays L to Party B, and Party B pays $24\% - 2 \times L$ to Party A. Can you reverse engineer this deal and express it in simpler terms?

Question 2.40: If an option is at-the-money, about how many shares of stock should you hold to hedge the option?

Question 2.41: Compare the price of an option on a stock if the stock price follows mean reversion versus if the stock price does not.

Question 2.42: When can hedging an options position make you take on more risk?

Question 2.43: How do you hedge a written put on a stock if you can neither short any stock nor use options on any stock?

Question 2.44: You order a pizza for six people. The diameter of the pizza is 12 inches. What would the diameter have to be to feed eight people? Yes, this is a derivatives question.

Question 2.45: When do you want to be short a put option on IBM stock?

Question 2.46: You own two pieces of land—a huge field in Arizona and a tiny piece of beach in Florida. The field in Arizona is idle; you have no plans to develop the land in any way. The tiny beach in Florida is very popular. In fact, it is so popular that you charge a small entrance fee for beachgoers.

The government has offered to buy the Arizona field for $1 million. Your neighbour has offered to buy the Florida beach for $1 million as well. Other things being equal, which piece of land has the higher forward price?

Question 2.47: You have 30 days of "representative" stock price data. How do you calculate historical volatility $\hat{\sigma}^2$ to use in Black-Scholes?

> **Story:** *A student interviewing with Goldman, Sachs was asked how he would move Mount Fuji. One of my colleagues suggested the answer "Call Mohammed."*

Question 2.48: Suppose a "top issuer" (i.e., highest-rated financial institution used as a reference in setting the swap curve) issues a corporate bond for itself valued at 100. The issuer then reprices this bond using the swap curve. What price do they get (100, above 100, or below 100)? To clarify, they fix the coupon rate of the bond so that it is priced at par, and then they try pricing this same bond by discounting those previously set coupons using the swap curve. Is the answer par, above par, or below par?

Question 2.49: Suppose I don't know any mathematics. How do you explain to me why you use the riskless rate instead of the required return on the stock to derive the Black-Scholes formula?

Question 2.50: Are you better off using implied standard deviation or historical standard deviation to forecast volatility?

Question 2.51: According to Black-Scholes, which is more valuable: a European call option that is 10% out-of-the-money, or a European put option that is 10% out-of-the-money?

Question 2.52: Why are theta and gamma of opposite signs? Are they always of opposite signs?

Question 2.53: Suppose that the riskless rate is zero. Suppose that a stock is at $100, and one year from now will be at either $130, or $70, with probabilities 0.80 and 0.20 respectively. There are no dividends. What is the value of a one-year European call with strike $110?

Question 2.54: **(**)** Find a formula for the European-style "product call" with payoff $\max(S_1 \times S_2 - X, 0)$, where S_1 and S_2 are the prices of assets following geometric Brownian motions with correlated random increments. All other Black-Scholes assumptions apply.

> **Story:** *A friend in the City of London was interviewing a candidate for a position on a credit derivatives quant team. On asking the candidate why he moved out of theoretical physics, he replied: "Why does a bank robber rob a bank?" After asking him several probabilistic dice questions, the candidate replied: "I can't be bothered with this shit." On asking him why he left his previous job, he replied: "Because they were a bunch of wankers"—which translates to "jerk-offs" for those readers in the United States. This is a true story!*

Question 2.55: Are Asian options cheaper or more expensive than plain vanilla European-style options?

Question 2.56: When can a plain vanilla European-style call be treated as an American-style one? When can a plain vanilla European-style put be treated as an American-style one?

Chapter 3

Other Financial Economics Questions

As a prerequisite to answering the questions in this chapter, it is expected that you have completed an introductory course in financial economics (or equivalent independent study). You also need a good deal of common sense. Solutions for this chapter appear in Appendix C.

Question 3.1: Consider the following game: a player tosses a fair coin until a head appears; if the head occurs on the k^{th} toss, the player gets a payoff of $\$2^k$, and the game ends.[1]

1. What is the fair value of the game? That is, what is the expected payoff to a player?
2. A very important customer is on the line and wants you to quote him a bid-ask spread for exactly one play of the game. "Hurry up, I haven't got all day!" You have 15 seconds.

Question 3.2: If the standard deviation of continuously compounded annual stock returns is 10%, what is the standard deviation of continuously compounded four-year stock returns?

Question 3.3: From the term-structure of interest rates, you see that the five-year spot rate is 10% per annum and the 10-year spot rate is 15% per annum.

[1]This game is over 250 years old and is known as the "St. Petersburg Game." It is quoted by Daniel Bernoulli ([Latin version 1738]; [English translation 1954]).

What is the implied forward rate from year 5 to year 10?

Question 3.4: Explain carefully the difference between the "yield" on a bond and the "rate of return" on a bond.

Question 3.5: What is "chaos theory"? Can you use it to predict stock returns? If so, how?

> **Story:** *One of my students who got a job at a large mutual fund company described his firm's working environment as follows: "Dig a hole, fill the hole with water, fill the water with sharks, and promote anything that crawls out alive."*

Question 3.6: Draw the graph of bond price versus yield-to-maturity. Why is the curve convex?[2]

Question 3.7: The Capital Asset Pricing Model (CAPM) suggests that the plot of $E(r)$ versus β should be an upward sloping line through $(0,\ r_f)$ and $[1,\ E(r_M)]$—the Security Market Line (SML). Suppose, however, that when you plot average returns against estimated betas you find something else. Which of the following two scenarios is most likely?

1. An upward sloping curve beginning at $(0,\ r_f)$, wholly above the theoretical SML, initially more steep than the SML, but eventually roughly parallel to the SML
2. An upward sloping curve beginning at $(0,\ r_f)$, wholly below the theoretical SML, initially less steep than the SML, but eventually roughly parallel to the SML

Which CAPM assumptions (if any) are violated by the above two scenarios?

Question 3.8: From the term-structure of interest rates, you see that the two-year spot rate is 7.60% per annum, and the one-year spot rate is 7.15% per annum.

What is the implied forward rate for the second year?

[2]Can you give economic intuition for this convexity? What about mathematical intuition?

Question 3.9: Consider a six-month forward contract on a 10-year riskless discount (zero-coupon) bond.

1. Is the bond selling at a forward premium or a forward discount?
2. Does your answer change if the bond is a riskless *coupon* bond (assume the coupon rate exceeds the current risk-free rate)?

Question 3.10: You believe that the yield curve is going to steepen very soon. It may be a fall in short-term rates, a rise in long-term rates, or some combination of these. What strategy should you pursue in the bond market to position yourself to profit from your beliefs?

> **Story:** *Many people are asked: "Are you married? Do you have children? What does your spouse do? How is your family? Where were you born? How old are you?" These are all illegal interview questions in the US. Although illegal, you should be prepared to answer these questions.*

Question 3.11: Define "duration" and "convexity." Describe their properties and uses.

Question 3.12: Describe briefly the GARCH(1,1) model in qualitative terms. Now write down the formal GARCH(1,1) model and explain each term carefully.[3]

Question 3.13: You have a long position in a \$100 million 30-year bond. What can you do to limit your exposure to only \$50 million?

Question 3.14: **(**)** You hold an 8% coupon, 30-year, \$1,000 par, Mexican Brady bond. Interest rates in Mexico do not change. Interest rates in the US increase by 1%. What is the change in the price of your bond? Make any necessary assumptions.

Question 3.15: You construct a yield curve for (coupon-bearing) treasuries. A particular five-year corporate zero-coupon bond has a default risk premium of 1% over the level of your treasuries yield curve at the five-year mark. You believe that the yield curve is going to flatten in

[3] Note that GARCH is an acronym for Generalized AutoRegressive Conditional Heteroskedasticity. How do you estimate the model? Why was it introduced?

such a way that the default risk premium of the five-year corporate zero remains constant (short-term rates rise, long-term rates fall, and yield on the five-year coupon-bearing treasury and five-year corporate zero remain unchanged).

What strategy should you pursue using the five-year zero-coupon corporate bond and treasuries to position yourself to profit from your beliefs?

> **Story:** *One interviewee was asked, "If you are holding a dinner party, and you can invite any three dead people (presumably resurrected), who would you choose? Please do not choose any relatives." I suggested Charles Manson, Ted Bundy, and Jeffrey Dahmer as poor choices.*

Question 3.16: If the five-year interest rate is 10%, and the 10-year interest rate is 15%, then you can conclude that the forward rate from year 5 to year 10 is 20% to a first-order approximation. However, it is slightly higher than 20%. Explain, *using plain English*, why the forward rate has to be *higher* than the 20% approximate value mentioned above.

Question 3.17: Here is a simple game. You get to toss a fair coin now. If it is heads, you get seven dollars 18 months from now. If it is tails, you lose two dollars immediately. The one-year interest rate is 12% per annum. The two-year interest rate is 18% per annum.

How much are you prepared to pay to play this game?

Question 3.18: There are 20 traders in a room. They trade in 100 stocks. They trade for their own accounts and only amongst themselves—it is a "closed economy." Halfway through their morning trading session, a group of SEC officials arrives and announces that one of the traders has inside information on one stock and has been trading on it. The trader is not yet identified. The SEC officials seat themselves in the room to watch. What happens to trading volume after the SEC announcement? Explain carefully.

Chapter 4

Statistics Questions

The only prerequisites for answering the statistics questions in this chapter are elementary statistical skills. Solutions for this chapter appear in Appendix D.

Question 4.1: Consider the following game. The player tosses a die once only. The payoff is $1 for each "dot" on the upturned face. Assuming a fair die, at what level should you set the ticket price for this game?

Question 4.2: I will roll a single die no more than three times. You can stop me immediately after the first roll, or immediately after the second, or you can wait for the third. I will pay you the same number of dollars as there are dots on the single upturned face on my last roll (roll number three unless you stop me sooner). What is your playing strategy?[1]

Question 4.3: (**) Two sealed envelopes are handed out. You get one and your competitor gets the other. You understand that one envelope contains m dollars, and the other contains $2m$ dollars (where m is unstated).[2]

---

[1] If you were running this game, how much would you charge players for repeated plays of the game? Suppose instead an amended game is played: I roll a single die three times without pause, and the payoff to the player is the maximum of the three rolls. What is the expected payoff to the player? Can you tell up front whether the original or amended game has the higher expected payoff?

[2] This problem is over 40 years old and is known as the "Exchange Paradox."

1. *If* you peek into your envelope, you see \$X. However, you do not know whether your opponent has $\$2X$ or $\$\frac{1}{2}X$. *Without peeking*, what is your expected benefit to switching envelopes? What is your opponent's expected benefit to switching envelopes (assuming your opponent sees \$Y)? Should you switch? If you do, do you do it again for the same reason (assuming neither of you peeked)?
2. Suppose that you both peek into your envelopes initially. What is the payoff to switching? Should you switch? If you do, do you do it again for the same reason?

Question 4.4: They call this the "World Series" problem. Two football teams are to play each other until one has four wins and is declared the overall champion (they, therefore, play at most seven times). You are to wager on who the overall champion will be. If you bet \$100 on Team A and they are declared overall champion, you get back your \$100 plus \$100 profit. Team A has a 70% chance of winning any game; Team B has a 30% chance. You may adjust your wager (on either Team A or Team B) as each game is played, but the overall winner is not known until one team has four wins. How do you bet so that you are guaranteed to either win \$100 or lose \$100 (either way is alright)?

Question 4.5: You and I are to play a game. You roll a die until a number other than a one appears. When such a number appears for the first time, I pay you the same number of dollars as there are dots on the upturned face of the die, and the game ends. What is the expected payoff to this game?

Question 4.6: You are dealt exactly two playing cards from a well-shuffled standard 52-card deck. The standard deck contains exactly four Kings. What is the probability that both of your cards are Kings?

> **Story:** *I spoke recently with a contact at a big-name Wall Street investment bank. I told him that some of my students were about to fly out to New York to interview. He told me: "Make sure they have nice suits, good hair-cuts, and wear their wedding rings." I would not pass on this story if my contact were heterosexual (which he is not) instead of openly homosexual (which he is).*

Question 4.7: (**) This is one version of the famous "Let's Make a Deal" or "Monty Hall" game show question. There are three doors. You know that there is a prize behind one of them, and nothing behind the other two. The game show host tells you that you shall receive whatever is behind the door of your choice. However, before you choose, he promises that rather than immediately opening the door of your choice to reveal its contents, he will open one of the other two doors to reveal that it is empty. He will then give you the option to switch your choice. You may assume that the host is totally impartial—not malicious in any way.

You choose Door 3. He opens Door 2 and reveals that it is empty. You now know that the prize lies behind either Door 3 or Door 1. Should you switch your choice to Door 1?

I strongly recommend that you not look at the answer until you have done your best.

Question 4.8: You are presented with two empty jars and 100 marbles on a table. There are 50 white marbles and 50 black marbles. You are to put all 100 of the marbles into the two jars in any way you choose. I will then blindfold you. I will shake the jars up to ensure good mixing, and I will rearrange the placing of the jars on the table so that you do not know which one is which. You may then request either the "left-hand" or the "right-hand" jar. You get to choose exactly one jar, you are allowed to withdraw at most one marble from the jar, and you do not get a second chance if you are unhappy with your choice.

How many of each colour marble should you place in each jar to maximize the probability that your blindfolded random draw obtains a white marble?[3]

Question 4.9: (***) Your name is Mr. 10. You are standing in a field with two opponents: Mr. 30 and Mr. 60. Each of you has a gun and plenty of ammunition. Each of you is in clear sight of the others and well within firing range. The goal is to maximize the probability of survival. Unfortunately, you are not a very good shot. If you take a

[3] Can you answer the same question except that you are to *minimize* the probability of a white marble? Does minimizing the probability of a white marble maximize the probability of a black one?

shot at one of your opponents, you have only a 10% chance of killing him. Mr. 30 is a better shot; he has a 30% chance of killing whomever he shoots at. Mr. 60 is even better; he has a 60% chance of killing his target. You take turns shooting in a pre-arranged order: first you, then Mr. 30, then Mr. 60, and then through this cycle again and again until only one person remains.

You get to shoot first. At whom do you shoot?[4]

Question 4.10: Basketball! Your team is down two points, you are the best player, and you have the ball. There are only a few seconds left before the buzzer. You can take a shot from three-point land or move up and take one from two-point land. Historically, you have a 40% probability of getting the shot in from three-point land and a 70% probability of getting the shot in from two-point land.

Should you try for the three-point shot (a certain win if you make it), or should you try for the two-point shot? Note that a two-pointer produces a tie and puts you into overtime. We assume your team has a fifty-fifty chance of winning in overtime.

Question 4.11: I will spin a fair roulette wheel with only five sections. Four of the five sections pay \$1; the fifth pays \$5.

1. If the cost is \$1.50 per spin, and you may play as often as you want, should you play the game?
2. If the cost is \$1.50 per spin, and you may play exactly once, should you play the game?

> **Story:** *It is many years ago now, but I know of a well-qualified MIT student who got a job offer of \$X from a well known firm (a good offer at that time). He declined, telling them that they had misjudged him. They called him back a couple of days later and offered him \$X × 1.67 instead! Amazing! He took the job.*

Question 4.12: If you like gambling and you like betting on the outcome of sports matches, then you may like the "parlay card." A parlay card

[4] Does the answer change if the order is first you, then Mr. 60, then Mr. 30, then you, and so on?

lets you bet on the outcomes of more than one match. In order to win a parlay bet, you must be correct on each of the matches you bet upon. Parlay cards offer big payoffs if you are right on every match (some even offer a payoff for "almost wins").

Suppose that your bookie will give you 10-to-1 odds for a parlay bet that covers four sports matches (with no almost wins). Should you take the bet?[5]

Question 4.13: What is the standard deviation of $(1, 2, 3, 4, 5)$?

Question 4.14: Welcome to your interview. Sit in this chair. Excuse me while I tie your arms and legs to the chair. Thank you. Now we are going to play "Russian roulette." I have a revolver with six empty chambers. Watch me as I load the weapon with two contiguous rounds (i.e., two bullets side-by-side in the cylindrical barrel). Watch me as I spin the barrel. I am putting the gun against your head. Close your eyes while I pull the trigger. Click! This is your lucky day: you are still alive! Our game differs from regular Russian roulette because I am not going to add any bullets to the barrel before we continue, and I am not going to give you the gun.

My question for you: I am going to shoot at you once more before we talk about your resume. Do you want me to spin the barrel once more, or should I just shoot?

Question 4.15: You have a large jar containing 999 fair pennies and one two-headed penny. Suppose you pick one coin out of the jar and flip it 10 times and get all heads. What is the probability that the coin you chose is the two-headed one?

Question 4.16: Four cards are shuffled and placed face down in front of you. Their faces (hidden) display the four elements: water, earth, wind, and fire. You are to turn the cards over one at a time until you either win or lose. You win if you turn over water and earth. You lose if you turn over fire. What is the probability that you win?

Question 4.17: Two players A and B play a marble game. Each player has both a red and a blue marble. They present one marble to each

[5]Should you take the bet if the odds are 25-to-1?

other. If both present red, A wins \$3. If both present blue, A wins \$1. If the colours do not match, B wins \$2.

Is it better to be A, or B, or does it matter?

Question 4.18: A coin-making machine produces pennies. Each penny is manufactured to have a probability P of turning up heads. However, the machine draws P randomly from the uniform distribution on $[0, 1]$ so P can differ for each coin produced. A coin pops out of the machine. You flip it once, and it comes up heads. Given this information, what is the (conditional) distribution function $F_{P|H}(p)$ for the probability of a head for that coin (where "H" denotes conditioning on the head)?

What is the (conditional) distribution function for the probability of a head if you flip the coin 1,000 times and get 750 heads?

> **Story:** *One of my MIT students was exceptionally well qualified. He was also one of the nicest guys I have ever met. He was quiet and soft-spoken. He was very understated, the kind of guy you might not notice. His starting salary at a big-name Wall Street firm was \$300,000 per year (and that was a few years ago). The moral of the story: I don't care how hot you think you are—brains wins.*

Question 4.19: Two games are offered to you. In Game One, you roll a die once and you are paid \$1 million times the number of dots on the upturned face of the die. In Game Two, you roll a die one million times. For each roll you are paid \$1 times the number of dots on the upturned face of the die. You are risk averse. Which game do you prefer?

> **Story:** *1. Took a brush out of my purse, brushed his hair and left. 2. Pulled out a Polaroid camera and snapped a flash picture of me. Said he collected photos of everyone who interviewed him.*
> *Interview Horror Stories from Recruiters*
>
> *Reprinted by kind permission of* MBA Style Magazine
> ©*1996–2004 MBA Style Magazine,* `www.mbastyle.com`

Chapter 5

Non-Quantitative Questions

You probably picked this book up for the same reason I wrote it: you like solving quantitative problems. If you are anything like me, you probably hate those invasive, wishy-washy, touchy-feely, namby-pamby, non-quantitative interview questions that you cannot solve using mathematics. If you have prepared for the quantitative questions, but you are dreading those wishy-washy, non-quantitative ones, then you need to read this chapter.[1]

Some of these questions have a single correct answer (great!). Others are roughly what you might expect in some sort of Freudian couch session after having been arrested for machine-gunning all the bag boys in your local supermarket. Some of the questions depend upon knowledge of financial management; others depend upon how many drinks you had at that last party you went to (and you might not get the job if you didn't have any drinks at the last party or if you do not go to parties).

Questions are broken into five categories: Questions About You, Questions About Your Job Awareness, Questions About the Markets or The Economy, Questions About Financial Management, and "Thinking Questions." In the rare cases where I deem an answer necessary, the question is labelled with an "**(A)**," and its suggested answer appears in Appendix E.

[1]As mentioned in the introductory chapter, you can buy interview books containing general non-quantitative questions. However, these books do not cover finance-related non-quantitative questions or the quirky questions unique to investment banking interviews. Of these interview books, I can recommend Allen (2000) and Marler and Mattia (1995). The former is a comprehensive presentation of common questions and intelligent answers; the latter presents general advice and is positive, encouraging, and supportive (good for anyone who is nervous or lacking in confidence). Both books are inexpensive.

This chapter benefited very much from interview questions collected by second-year MBA students at MIT's Sloan School of Management.

5.1 Questions about You

Question 5.1.1: Tell me about yourself.

Question 5.1.2: Walk me through your resume.

Question 5.1.3: What are your career goals? How do you plan to achieve those goals?

Question 5.1.4: What do you see yourself doing in five years? Is this different from what you imagined when you entered the degree program at your college (if so, how so)?

Question 5.1.5: Describe your life experiences, explaining any major decisions you have made to date.

Question 5.1.6: What two or three accomplishments have given you particular satisfaction over your lifetime?

Question 5.1.7: Tell me in detail what you did while working for this company that appears on your resume.

Question 5.1.8: How would you value yourself? "Value" here means in financial terms.

Question 5.1.9: How do you evaluate your success or the success of others?

Question 5.1.10: How would you describe yourself? How would your friends describe you? How would a former supervisor describe you?

Question 5.1.11: What is your greatest strength?

Question 5.1.12: Describe a situation where you have successfully sold your ideas.

Question 5.1.13: What is your greatest weakness?

Question 5.1.14: What areas of your performance need improvement?

Question 5.1.15: Why shouldn't we hire you? The interviewee described this to me as a tough spin on the traditional "what is your greatest weakness" question.

Question 5.1.16: Tell me something you tried but ended up quitting on.

Question 5.1.17: What is the biggest risk you have taken in your life?

Question 5.1.18: Rate yourself on a scale of 1 to 10 on the type of risk taker you are. Tell me why and give examples to support your claims.

Question 5.1.19: Tell me about a goal you set for yourself in the past that turned out to be either too easy or too hard to achieve. What did you learn from the situation?

Question 5.1.20: What distinguishes you from other candidates we might hire?

Question 5.1.21: What do you do for fun?

Question 5.1.22: Describe the best party you have ever been to.

Question 5.1.23: What is the biggest investment mistake you have ever made?

Question 5.1.24: Tell me about a time when you had to deal with a highly ambiguous situation. What did you do? How did you deal with it?

Question 5.1.25: Please describe an ethical dilemma you have faced at work, and tell me how you handled the situation.

Question 5.1.26: How good are your writing skills? Please give me some convincing evidence.

Question 5.1.27: If you could go on a cross-country car trip with any three people, who would you choose?

Question 5.1.28: If you were holding a dinner party, and you could invite any three dead people (presumably resurrected), who would you choose? Please do not choose any relatives.

Question 5.1.29: Why did you decide to apply to your MBA college? Did you apply to other MBA programs (if so, which ones and why)?

Question 5.1.30: What do you do if the "picture-in-picture" does not work on your television? Yes, one of my students was asked this in a banking interview!

Question 5.1.31: How would you evaluate your experiences at your MBA college?

Question 5.1.32: What are the strengths and weaknesses of your MBA program?

Question 5.1.33: Describe a situation in which you had to make a decision based on very little information.

Question 5.1.34: Tell me about a situation when you were chosen as a leader by the members of your group.

Question 5.1.35: Repeat the conversation that you had with your teammates when things did not go well in your group.

Question 5.1.36: What have you enjoyed most about your experiences at your MBA college? What would you change?

Question 5.1.37: What is your GPA at your college? What about your GMAT score?

Question 5.1.38: Which courses did you enjoy most at your MBA college (and why)?

Question 5.1.39: How has your course work at your MBA college helped you to develop skills relevant to this job?

Question 5.1.40: What has been most difficult for you at your MBA college, and how have you dealt with it?

Question 5.1.41: How much of your education did you personally fund?

Question 5.1.42: How do you spend your time outside of school and work? How do you balance your life?

Question 5.1.43: Are you married? Do you have children? What does your spouse do? How is your family? Where were you born? How old are you? These are all illegal questions in Canada and the US. Most questions about family background, religion, marital status, and so on are illegal. Although illegal, you should be prepared to answer these questions.

Question 5.1.44: Describe your typical day.

Question 5.1.45: Are you innately intelligent, or do you have to work really hard?

> **Story:** *A student of mine interviewing at Goldman, Sachs was asked to draw a picture of himself! They gave him pencil and paper, and he drew a picture (into his picture he also drew books, friends, and other goofy stuff to indicate that he was not retarded).*

Question 5.1.46: At interview end: "Is there anything important you have not had a chance to tell me?"

Question 5.1.47: At interview end: "Do you have any questions you would like to ask me?"

5.2 Questions about Your Job Awareness

Question 5.2.1: How does this position in this company fit into your career development plans? What other career options are you considering?

Question 5.2.2: Why do you want to work for this employer?

Question 5.2.3: Sell yourself to me. Prove to me that you are someone I should seriously consider for our firm.

Question 5.2.4: What motivates you to put forth your best effort? What type of work environment brings out your best effort?

Question 5.2.5: What rewards do you seek from work? What rewards do you seek from this particular job (or company)?

Question 5.2.6: Why are you not better matched with Firm X (our competitor)?

Question 5.2.7: Do you have any geographical preferences? What are your thoughts about travel or relocation?

Question 5.2.8: What do you see yourself contributing to our organization, both in the short term and in the long term?

Question 5.2.9: What other companies are you interviewing with, and how do we compare?

Question 5.2.10: **(A)** What do you think of our tombstone in today's *Wall Street Journal*?

Question 5.2.11: Why fixed income rather than equities?

Question 5.2.12: What do you think it takes to be successful in this position (or this organization)?

Question 5.2.13: Why do you want to work as a trader?

Question 5.2.14: What do you think traders do?

Question 5.2.15: If you were in my position, interviewing candidates for this position, what qualities would you seek? How would you evaluate candidates?

Question 5.2.16: Describe the best boss you have ever had. How would you define the qualities of a good manager?

Question 5.2.17: What do you think an investment banker does?

Question 5.2.18: Do you understand the hours investment bankers work and why?

Question 5.2.19: Describe how you build relationships in a new job.

Question 5.2.20: Imagine you have received three job offers. How will you decide which one to accept?

Question 5.2.21: If you were to get two other job offers in addition to one from us, from which firms would they most likely come, would you take them, and why?

Question 5.2.22: Some people say investment banking is not value adding. How do you refute that?

Question 5.2.23: If we offer you a job right now, will you take it?

5.3 Questions about the Markets or the Economy

Question 5.3.1: Where is the DOW, or S&P500, or NIKKEI, or FTSE, or Hang Seng, or....? How does it compare now to where it has been over the last two years? Where do you see it two weeks from now (or six months from now)?

Question 5.3.2: Where is the JPY, or GBP, or DEM, or FFR, or....? How does it compare now to where it has been over the last two years? Where do you see it two weeks from now (or six months from now)?

Question 5.3.3: What is LIBOR, and what is today's LIBOR rate?[2]

Question 5.3.4: Why invest in a particular market (e.g., Korea, Russia, Germany)?

Question 5.3.5: How would the following affect interest rates: Saddam Hussein starts making trouble in the Middle East; there is another Asian currency crisis; Monica Lewinsky (alleged mistress of the US president Clinton) hits the headlines again as the alleged mistress of the current US president.

[2]They probably mean the benchmark three- or six-month US dollar LIBOR rates, but they might not say that. There are several different dimensions here: you should understand the distinction between US dollar LIBOR and British pound LIBOR, between three-month LIBOR and six-month LIBOR, and between LIBOR (London InterBank Offer Rate) and PIBOR (Paris InterBank Offer Rate). If you do not, look in your favourite investments book, or use a WWW search engine.

Question 5.3.6: **(A)** When inflationary fears arise, the government has two forms of macroeconomic policy to try to slow the economy down. Name these and explain them in a few words.

Question 5.3.7: What stock do you recommend and why?

Question 5.3.8: Tell me about a stock you like or hate and why.

Question 5.3.9: What should be the (CAPM) beta for Intel Corp.?

Question 5.3.10: Where do you think the US economy will go over the next year?

Question 5.3.11: **(A)** What are the "Dow Jones Dogs"?

Question 5.3.12: Do you think the stock market is efficient (in an EMH sense)?

Question 5.3.13: Suppose you are actively investing to beat the market. Are there more opportunities (i.e., inefficiencies) in the S&P500 or in the 500 largest stocks in Europe?

5.4 Financial Management Questions

Question 5.4.1: How would you value a company? (This is a very popular question indeed.)

Question 5.4.2: Suppose that the S&P500 index has a P/E ratio of 20. How would you value a manufacturing company with earnings of one million dollars?

Question 5.4.3: What key financial ratios do you look at when trying to determine a firm's financial health from its balance sheet?

Question 5.4.4: Why do pharmaceutical companies increase drug prices when they come off patent protection?

Question 5.4.5: Describe the CAPM.

Question 5.4.6: Can a company function without working capital?

Question 5.4.7: What happens to a company's balance sheet if the company buys an asset? Walk me through the steps.

Question 5.4.8: **(A)** When are motor vehicles owned by the company not recorded on the balance sheet as PPE (physical plant and equipment)?

Question 5.4.9: How would you market this financial product (e.g., a structured note)?

Question 5.4.10: Imagine you are giving a presentation to a client and they tell you your numbers are wrong. What would you do?

Question 5.4.11: How do you use DCF to value a skyscraper in order to sell it? You need to come up with current revenue, costs, net income, estimates of future cash flows, and a discount rate.

Question 5.4.12: Kirk Kekorian attempted to force Chrysler to rid itself of what he called "excess cash"—through higher dividends and a stock buy back. What do you think of this?

Question 5.4.13: How would you market this company to our clients?

Question 5.4.14: Have you ever had to fire someone? If so, how did you handle this situation?

Question 5.4.15: Forecast the income statement for Duracell for this year.

Question 5.4.16: **(A)** In the calculation of free cash flow (i.e., FCF), does the level of long-term debt matter?

Question 5.4.17: How do you calculate VaR (i.e., Value at Risk)?

Question 5.4.18: Have you heard of LTCM?

Question 5.4.19: What is the difference between default risk and prepayment risk?

Question 5.4.20: What is kurtosis?

5.5 Thinking Questions

These "thinking questions" lie between quantitative and non-quantitative. Most of these questions have in common that they have some sort of precise solution. However, if you know exactly how many McDonald's outlets there are in the US (one of the questions) and say so directly, then you have missed the point. The interviewer wants you to work the answer out and describe your reasoning. These are thus "thinking questions," not calculation ones. The answer per se is not important; it is your reasoning that counts.

Question 5.5.1: If a cannonball is dropped in the deepest part of the Earth's oceans, how long will it take to reach the ocean floor?

Question 5.5.2: **(A)** How many McDonald's fast food outlets are there in the US?

Question 5.5.3: How many gas stations are there in the US?

Question 5.5.4: **(A)** You are in a jail cell alone stripped of your possessions. It is Friday afternoon, and you desperately need a cigarette. How do you force the guard to give you one?

Question 5.5.5: How many elevators (i.e., "lifts" if you are British) are there in the US?

Question 5.5.6: How would you value an option on (famous basketball player) Michael Jordan?

Question 5.5.7: **(A)** I toss a coin 100 times and get 100 heads in a row. What is the probability that the next outcome will be a head?

Question 5.5.8: How many ping-pong balls can you fit in a jumbo jet (e.g., Boeing 747)?

Question 5.5.9: How would you move Mount Fuji?

Question 5.5.10: How do you weigh a jet aeroplane without using scales?

Question 5.5.11: You have a five-gallon jar and a three-gallon jar. You can have as much water as you want. How do you put exactly four gallons into the five-gallon container? This is too easy for me to supply an answer.

Question 5.5.12: Estimate the annual demand for car batteries.

Question 5.5.13: What would you estimate to be the size of the racquetball market in the US?

Question 5.5.14: You are to build a plant for Coors to serve all beer customers in the state of Ohio. How large would you build it? That is, specifically how many cans (of the new wide-mouth variety) do you anticipate being demanded for the year?

Question 5.5.15: Why do beer cans have tapered tops and bottoms?

Question 5.5.16: **(A)** Explain why aeroplanes can fly.

Question 5.5.17: How many fish are there in the Earth's oceans?

Question 5.5.18: How many barbers are there in Chicago?

Question 5.5.19: **(A)** Finally, why are manhole covers round?

Story: *I was telephone-interviewing a candidate for an active equity research job in London. His job would be the creation, testing, and implementation of strategies for beating the market. I asked him if he could draw upon his considerable experience in the markets to suggest to me a strategy he had heard of for beating the market. There was a very long pause, after which he answered simply "no." What did he think I was going to ask him about?!*

Appendix A

Purely Quantitative & Logic Answers

This appendix contains answers to the questions posed in Chapter 1.

Answer 1.1: This question has appeared over and over again. Although simple, it is rarely answered well. No calculation is required to determine the answer. If you used *any* algebra whatsoever, stop now, go back, reread the question, and try again.

When the quantity Q of water is poured into the alcohol jug, the concentration of alcohol in the alcohol jug becomes $\frac{V}{V+Q}$. After mixing and pouring some back, the concentration of alcohol in the alcohol jug does not change again (because no new water is added). However, when the diluted alcohol is poured back into the water jug, the concentration of water in the water jug changes from 100% to $\frac{V}{V+Q}$. That is, the final concentrations are identical.

How do you see that the final concentrations must be identical? Remember, you do not need any calculations at all. In fact, the only reason for any calculation is if you also want to find out what the final concentrations are (you were not asked this, but if you wish to work it out, your calculations need not go beyond those of the previous paragraph).

Here is how it works. At the end of the process, both jugs contain the same volume of fluid as they did at the start. The only way for the concentration of alcohol (for example) to have changed from 100%

is if some alcohol was displaced by water. Similarly, the only way for the concentration of water to have changed from 100% is if some water was displaced by alcohol. Volume is conserved (both total volume and volume in each jug), so all that has happened is that identical quantities of water and alcohol have traded places (and these identical quantities are slightly less than Q). By symmetry, the concentrations of alcohol in the alcohol jug and water in the water jug must be identical.

Answer 1.2: This is a very common question, and a very simple one. You need to figure out the sum: $1+2+3+\cdots+99+100$. There are several ways to do this.

FIRST SOLUTION

A simple technique is to note that the first and last terms add to 101. The second and second-to-last terms also add to 101. The same is true of the third and third-to-last terms. Continuing in this fashion, you soon find yourself with 50 pairs of numbers adding to 101; 50 times 101 is 5,050.

SECOND SOLUTION[1]

A simple technique you can picture easily is the following:

$$\begin{array}{cccccc} 1 & 2 & 3 & \ldots & n-1 & n \\ \underline{n} & \underline{n-1} & \underline{n-2} & \ldots & \underline{2} & \underline{1} \\ n+1 & n+1 & n+1 & \ldots & n+1 & n+1. \end{array}$$

There are n terms each equal to $n+1$. The required sum is half the grand total: $\frac{n(n+1)}{2}$.

THIRD SOLUTION

I read somewhere many years ago that the high school drop-out Albert Einstein devised the following alternative solution technique at age 15. Think of each summand, i, in the sum $\sum_{i=1}^{100} i$ as a group of i marbles in a row from $i=1$ to $i=100$ (see the array following). Stacking each row of marbles on top of each previous row, you get the array including both the diagonal and the lower-triangular off-diagonal. Were the array full, it would contain $100 \times 100 = 100^2$ marbles. So, your answer must be roughly half this (roughly 50×100). This is not exact because although the array contains two triangular-shaped off-diagonals (upper

[1] I thank Tom Arnold for this solution technique. I am responsible for any errors.

and lower), there is only one diagonal. If you add another diagonal, and *then* split the total in two, you get the right answer. The diagonal contains 100 marbles, so the right answer must be $\frac{100^2+100}{2} = 5,000+50$, as before.

$$
\begin{array}{c|cccccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 & \cdots & 100 \\
\hline
1 & \bullet & & & & & & & \\
2 & \bullet & \bullet & & & & & & \\
3 & \bullet & \bullet & \bullet & & & & & \\
4 & \bullet & \bullet & \bullet & \bullet & & & & \\
5 & \bullet & \bullet & \bullet & \bullet & \bullet & & & \\
6 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \\
100 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \cdots & \bullet
\end{array}
$$

More generally, the sum from 1 to n may be written down as $\frac{n^2+n}{2} = \frac{n(n+1)}{2}$. Just picture the square array of side length n, add another diagonal, and split the total in half.

To calculate $\frac{n(n+1)}{2}$ quickly in your head, note that one of n or $n+1$ must be even and thus divisible by two. You should divide the even number by two and multiply the odd number remaining by the result. In our case,

$$\frac{100 \times 101}{2} = \frac{100}{2} \times 101 = 50 \times 101 = 5,050.$$

Finally, note that two more solutions appear in the answers to Question 1.38, starting on page 102.

Answer 1.3: This question has been very popular indeed. Sometimes it is golf balls, sometimes marbles, sometimes coins. Most students find it very challenging.[2]

[2] I heard about one guy who got home, took 12 golf balls, and tried to solve this by physically manipulating them. I understand that he was still unsuccessful. This particular solution technique combines independent contributions of Juan Tenorio, Bingjian Ni, Yi Shen, and Jinpeng Chang. I am responsible for any errors.

The first step is to split the 12 marbles into three groups of four. Each group of four has two subgroups, a singleton and a triplet: $\{\{1\}_A, \{3\}_A\}$, $\{\{1\}_B, \{3\}_B\}$, and $\{\{1\}_C, \{3\}_C\}$.

Compare $\{\{1\}_A, \{3\}_A\}$ to $\{\{1\}_B, \{3\}_B\}$. If they balance, then the odd ball is in group C. In this case, compare $\{3\}_C$ to $\{3\}_B$. If $\{3\}_C$ is heavier (or lighter), then comparing any two marbles from within $\{3\}_C$ immediately locates the odd one; if $\{3\}_C$ balances $\{3\}_B$, then compare $\{1\}_C$ to $\{1\}_B$ to see whether $\{1\}_C$ is heavier or lighter.

If the initial comparison is unbalanced, say $\{\{1\}_A, \{3\}_A\}$ is heavier than $\{\{1\}_B, \{3\}_B\}$, then rotate groups $\{3\}_A$, $\{3\}_B$, and $\{3\}_C$ and compare $\{\{1\}_A, \{3\}_B\}$ to $\{\{1\}_B, \{3\}_C\}$ (while holding out $\{\{1\}_C, \{3\}_A\}$). If they balance, then a heavy marble is in $\{3\}_A$ and comparing any two marbles from within $\{3\}_A$ immediately locates the odd one. Suppose they do not balance. If $\{\{1\}_A, \{3\}_B\}$ is heavy, then either $\{1\}_A$ is heavy, or $\{1\}_B$ is light. Compare $\{1\}_A$ to $\{1\}_C$ to finish. If $\{\{1\}_A, \{3\}_B\}$ is light, then $\{3\}_B$ is light and comparing and two marbles within $\{3\}_B$ immediately locates the light one.

In each case, only three weighings are needed. This technique is generalized in Answer 1.14 (the "90-coin problem").

Answer 1.4: This is cute. You (the bug) cannot fly; you have to walk. You must find the shortest path from corner to corner.

In any world, the shortest path between two points is called a "geodesic." On a spherical world (e.g., the Earth's surface), a geodesic is an arc of a "great circle." A great circle is a circle on the surface of the sphere with diameter equal to the diameter of the sphere. For example, aeroplanes typically follow great circles above the Earth (because it is the shortest path and, therefore, the most fuel-efficient path).

Like a spherical world, the cubic-room world has a two dimensional surface. However, the lack of curvature in the cubic-room world means that the shortest distance between two points must be a straight line rather than an arc of a great circle (in a world without curvature, geodesics are straight lines).

If the cubic room is opened up and flattened out it can be seen that the shortest path is a straight line from one corner to the other. In the unflattened room, this straight line corresponds to two line segments that meet exactly halfway up one wall-floor or wall-wall boundary. Direct computation using Pythagoras' Theorem[3] reveals that the total path length is $\sqrt{5}$ units.

Answer 1.5: The 10×10 macro-cube question has been very popular. The most common mistake is for students to *count* the number of 1×1 micro-cubes on each face and add them up. Even if you do the mathematics correctly (and most people do not), you miss the whole point.

If you focus on the 1×1 micro-cubes on the faces and how to count them directly (e.g., How many faces? How many on each face? How many edges?), you miss the point. Go back now and figure out a better way. As I stated before, the path of greatest resistance bears the highest rewards, so read no further unless you did it a better way.

You must look for structure in a problem that leads you to a simple and speedy solution. The most structure here is to be found in the macro-cube you start with and the (now slightly smaller) macro-cube that remains. The difference between their volumes is how many micro-cubes fell.

The volume of a cube of side length n is n cubed; that is, n^3. The answer is, therefore, $10^3 - 8^3$.

How do you calculate this without a calculator? You should know that 10^k is a 1 with k zeroes attached, so $10^3 = 1,000$. You should know that $8 = 2^3$ and, therefore, that $8^3 = 2^{3\times3} = 2^9$. You should definitely know that 2^{10} is 1,024. Thus, 2^9 is half of 2^{10} and, therefore, equal to 512. The answer is $1,000 - 512 = 488$. A common mistake is for students to think the answer is $10^3 - 9^3 = 271$, because only "one layer" fell off (you should of course know what 9^3 is also without having to work it out).

Answer 1.6: This is a good question. Students tend to overlook the brilliantly simple situation described. If you did any mathematics whatsoever, you probably missed the point.

[3]Recall Pythagoras' Theorem. Consider a triangle with side lengths X, Y, and Z. If the angle between the sides of length X and Y is 90°, then it is a "right-angle" triangle. The side of length Z (the "hypotenuse") must be the longest side, and it must be that $X^2 + Y^2 = Z^2$. In this case, the path is the hypotenuse of a triangle of side lengths 2 and 1 in the flattened-out room or two hypotenuses of triangles each of side lengths 1 and $\frac{1}{2}$ in the un-flattened room. In either case, the path is of total length $\sqrt{5}$.

No calculation is needed to see that at each stage an equal number of male babies and female babies are expected to be born. The proportions of male and female children are, therefore, expected to remain equal at 50%.

Still stuck? Here are the details (assuming equal numbers of boys and girls are born): by the end of the first year, the 100,000 families have 50,000 boys and 50,000 girls. The proportion of male children stands at 50%. By the end of the second year, half of the 100,000 families (the ones without a son) have another child. This adds 25,000 boys and 25,000 girls. There are now 75,000 boys and 75,000 girls. The proportion of male children still stands at 50%. There are still 25,000 families without a son. They add another 12,500 boys and an equal number of girls, and so on.

Some people are tempted to suppose that because all large families have many daughters and a single son, there must be more girls than boys. However, there are not many large families.[4]

Answer 1.7: I like this one. Students have given me answers to this one ranging from 0° to 75° (and many answers in between). The big hand is on the three; the little hand is one-quarter of the way between the three and the four. The answer must be one-quarter of one-twelfth of 360°. That is, one-quarter of 30°. That is, 7.5° (or $\frac{\pi}{24}$ radians if you like measuring angles in radians).[5]

You should focus on what you know (the angle is non-zero, the big hand is on the three, one hour is one-twelfth of the full circle, and 15 minutes is one-quarter of one hour) and make sure that your answer accords with intuition. For example, if you get 75°, then something is wrong with your reasoning (or you have never owned an analogue wristwatch).

[4]In fact, the average number of children per family is only $\sum_{k=1}^{\infty}\frac{k}{2^k}=2$ (obtained using basic probability theory and the following algebraic result derived by me: $\sum_{k=1}^{\infty}\frac{k}{x^k}=\frac{x}{(x-1)^2}$, for $|x|>1$).

[5]There are 2π radians in a full circle. Thus, $360^\circ=2\pi$ radians; $180^\circ=\pi$ radians; $90^\circ=\frac{\pi}{2}$ radians; and so on. It is just another way of measuring angles.

Story: *1. He whistled when the interviewer was talking. 2. Asked who the lovely babe was, pointing to the picture on my desk. When I said it was my wife, he asked if she was home now and wanted my phone number. I called security.*

Interview Horror Stories from Recruiters
Reprinted by kind permission of MBA Style Magazine
©*1996–2004 MBA Style Magazine,* www.mbastyle.com

Answer 1.8: I present both a brute-force approach (with some algebra) and an elegant approach looking at the bigger picture.

FIRST SOLUTION

By the time the minute hand gets to the three, the hour hand will have moved on a little. There is then a "catch-up time" of a minute or two after 3:15, so the hands must be coincident around 3:16 or 3:17. As time passes from three o'clock to four o'clock, the minute hand whips 60 minutes around the entire face, and the hour hand moves slowly from pointing at the 3 to pointing at the 4 (marked off as five increments of one minute on my watch). At all times, the proportion of the full 60 minutes traversed by the minute hand equals the proportion of the five increments of one minute traversed by the hour hand. Let M minutes after 3PM be the first time at which the hands are coincident, then our argument above implies directly that the proportion of the full 60 minutes covered by the minute hand equals the aforementioned catch-up time as a proportion of the five increments of one minute:

$$\begin{aligned} \frac{M}{60} &= \frac{M-15}{5} \\ &= \frac{M}{5} - 3 \\ \Rightarrow \frac{M}{5} - \frac{M}{60} &= 3 \\ \Rightarrow \frac{11M}{60} &= 3 \\ \Rightarrow M &= \frac{180}{11} = 16\frac{4}{11}. \end{aligned}$$

Thus, the hands are coincident at 3:16 and $\frac{4}{11}^{th}$'s of a minute. That is 3:16:21.82.

SECOND SOLUTION
At high noon, the hands are coincident. At midnight, the hands are coincident. The time between noon and midnight is cut into 11 equally-spaced time intervals of one-eleventh of 12 hours (1:05:27.27). At the end of each of these intervals, the hands are coincident. For the question at hand, the answer is three times one-eleventh of 12 hours: 3:16:21.82. The full set of "coincident times" are as follows:

$$\left(\begin{array}{r} 1:05:27.27 \\ 2:10:54.55 \\ 3:16:21.82 \\ 4:21:49.09 \\ 5:27:16.36 \\ 6:32:43.64 \\ 7:38:10.91 \\ 8:43:38.18 \\ 9:49:05.45 \\ 10:54:32.73 \\ 12:00:00.00 \end{array}\right).$$

Answer 1.9: Well now, this looks pretty complicated the first time you see it. However, there is a simple way to figure it out. If you think about it, you see that the only brokers who touch the switch for light bulb number 64 are those whose numbers are divisors of 64.

That is, light bulb number 64 has its state of illumination changed by brokers whose numbers are factors of 64. That is, brokers 1, 2, 4, 8, 16, 32, and 64 flip the switch. Because light bulb 64 is originally *off*, it must be that after this odd number of switches it is *on*. See Answer 1.10 for a closely related but more general solution technique.

Answer 1.10: If you now know the answer to Question 1.9, you should be able to figure this one out swiftly. If you have not yet figured out Question 1.9, then read no further—solve that one first.

The only way for a light bulb to be illuminated after the 100th person has passed through is if its switch was flipped an odd number of times. The switch for light bulb number K gets flipped only by people whose numbers are factors of K. Thus, the only light bulbs illuminated at the conclusion are those with a number that has an odd number of factors.

However, factors for numbers go in pairs. For example, 32 has factors $(1, 32)$, $(2, 16)$, and $(4, 8)$. This means that 32 has an even number of factors, and bulb 32 is not illuminated at the conclusion. In fact, at first glance, all numbers have an even number of factors.

However, you do get an odd number of factors if two factors (one pair) are identical. For example, 64 has $(8, 8)$ as one pair. If one pair of factors are identical, then the original number must be a "perfect square." Therefore, the only numbers with an odd number of factors are the perfect squares.

There are exactly 10 perfect squares between 1 and 100, and they are 1, 4, 9, 16, 25, 36, 49, 64, 81, and 100 (1^2, 2^2, 3^2, ... 10^2). These are the numbers of the 10 bulbs that are illuminated after the 100th person has passed through the room.

Answer 1.11: This is an old favourite. I have tried this out on students and have received almost all imaginable responses. The answer is three, and it cannot be anything else. Two socks can be different, but a third must match one of the first two—giving a matching pair.

Answer 1.12: This has appeared several times. You get the answer by working backwards. If I am your opponent, and I am able to call out "39," then you cannot reach 50, but I can after you say whatever you say. So, my goal is to call out "39." However, if I am able to call out "28," then you cannot get to 39, but I can after you say whatever you say. So, my goal is to call out "28." To get to 28, I need only to be able to call out "17," and to do this, I need only to be able to call out "6."

So, my strategy, as your opponent, is to get onto the series 6, 17, 28, 39, 50 at whichever point I can. If you get to go first, you should call out "6." As long as you know the winning numbers and stick to them, you cannot lose. If you start with anything other than 6, I cannot lose.

Answer 1.13: Safe-cracking in a finance interview? Yes indeed. The naive answer is that there are 40 possible numbers for the first combination, 40 for the second, and 40 for the third. It would then take at most 40^3 possible trials to get the safe open. That is 64,000 trials. There are two factors that reduce this number considerably. The first you should have figured out; the second you are excused.

The first factor is that although three numbers are required to open the safe, you need only find the first two of them. If you dial the first two numbers correctly, then you need only turn the dial until the safe pops open. You do not need to know the last number. This gives 40^2 possible combinations. That is only 1,600 trials. For extra safe-cracking advice along these lines see Gleick (1993, pp189–190).[6]

The interviewer in this case suggested a second factor, as follows (and I think it is a little unfair to any interviewee who is inexperienced in safe-cracking). The safe is a mechanical device designed with a particular tolerance for inaccuracy. If the first combination number is 14, then dialing either 13, 14, or 15 suffices. This tolerance for inaccuracy brings you down to roughly $\left(\frac{40}{3}\right)^2$ trials. This is less than 200 trials.

Answer 1.14: To minimize the maximum possible cost of weighing, your strategy must use the scales as few times as possible, wherever the location of the "bad" coin. From Answer 1.3, you know that you may need as many as three weighings to find a bad coin in a group of 12. You have 90 coins, so it must take at least four weighings. However, by the same argument, if you had 144 coins (12 groups of 12), you could identify a bad group of 12 in three weighings and then the bad coin in another three. So, (because 90 is less than 144) it should take no more than six weighings—either four, five, or six weighings.

In fact, it takes only five weighings (and at most \$500) to both find the bad coin and identify it as heavy or light. I present two quite different solution techniques plus a third quasi-solution: the first is an ingenious "hammer-and-tongs" technique, the second is slightly more structured, and the third generalizes the second but applies only in special cases. In each case, it takes only five weighings to both find the bad coin and identify it as heavy or light.

FIRST SOLUTION[7]

The technique here is similar to the solution technique to Question

[6] I went to a presentation at MIT at which Jim Gleick (pronounced "Glick") talked about his then soon-to-be-published book "Genius." He talked about genius in general and Richard P. Feynman in particular. Feynman was an interesting guy, and this is a good book about him.

[7] I thank Eva Porro (then at the Universidad Complutense de Madrid) for this solution technique. I am responsible for any errors.

1.3—be sure to answer that question before answering this one. The first move is to divide the 90 coins into three groups of 30. Weigh two of the groups of 30 coins. Either the scales balance, or they do not. If the scales balance, then you are left with one group of 30 containing the bad coin. You may "hold out" 10 of these 30 and compare the remaining 20—with 10 on each side. If the scales balance, you get one group of 10 containing the bad coin. If they do not balance, you have one group of 10 coins potentially heavy and one group of 10 coins potentially light. Stop here if you just wanted to know how to start the solution process. This should be enough for you to finish.

Return for a moment to the case in which the two groups of 30 do not balance. Place 10 potentially heavy coins and 10 potentially light coins on each side of the scales, while holding out 10 potentially heavy coins and 10 potentially light coins. Whether they balance or not, you can immediately identify one group of 10 coins that is potentially heavy and one group of 10 coins that is potentially light—the other 40 are "good" coins.

Thus, after two weighings, the problem reduces either to one group of 10 coins containing the bad coin (no further information) or two groups of 10 coins (where one group potentially contains a heavy coin, and the other potentially contains a light one). I need only demonstrate the solution technique for each case.

Suppose you have 10 coins, and one of them is bad. You can find the bad one in three weighings simply by adding two good coins and following Answer 1.3 for the 12-ball case. This finds you the bad coin in a total of five weighings.

Suppose instead that you have the two groups of 10 coins (where one group potentially contains a heavy coin, and the other potentially contains a light one). Use the notation "3$\uparrow$" to denote three potentially light coins, "3$\downarrow$" to denote three potentially heavy coins, and "1good" to denote one coin known to be "good." In this case, you begin at the end of the second weighing with $\{10\uparrow\}$ and $\{10\downarrow\}$ on the scales. Hold out 3$\uparrow$ and 3$\downarrow$ coins and place the following on the scales for weighing number three: $\{3\uparrow, 4\downarrow\}$ versus $\{3\downarrow, 4\uparrow\}$.

Suppose the scales balance with $\{3\uparrow, 4\downarrow\}$ versus $\{3\downarrow, 4\uparrow\}$. Then you have 3$\uparrow$ and 3$\downarrow$ coins left. Hold out 1$\uparrow$ and 1$\downarrow$ and weigh $\{1\uparrow, 1\downarrow\}$

versus $\{1\downarrow,1\uparrow\}$. If these balance, weigh the hold out $1\uparrow$ against 1 good coin to find the bad one; if they do not balance, you get $1\uparrow$ and $1\downarrow$ from the light and heavy sides, respectively; and you need only compare one of them to a good coin to find the bad one. This gives a total of five weighings in either case.

Suppose the scales do not balance with $\{3\uparrow,4\downarrow\}$ versus $\{3\downarrow,4\uparrow\}$. If the first group appears lighter, then you get $3\uparrow$ and $3\downarrow$ coins as in the previous paragraph—able to be solved in a total of five weighings. If the second group appears lighter, then you get $4\downarrow$ and $4\uparrow$ coins. This is just like the first weighing of two groups of four in the 12-ball problem in Question 1.3, and you know the bad coin can be identified in only two more weighings by rotating "triplets." In each case, the bad coin is both located and identified as heavy or light in only five weighings.

> **Story:** *One of my students was told "Take your jacket off—it's going to get hot in here."*

SECOND SOLUTION[8]

Begin by noting that if you have a group of 3^k coins that is known to contain a heavy coin, it takes only k weighings to identify it. You can see this as follows: split the group of 3^k coins into three subgroups each of size 3^{k-1}; now compare any two subgroups on the scales. Whether the scales balance or not, you know immediately which of the three subgroups contains the heavy coin. It thus takes only one weighing to go from a group of 3^k coins known to contain a heavy coin to a group of 3^{k-1} coins known to contain a heavy coin. Proceeding in this fashion, it takes k weighings to go from a group of 3^k coins known to contain a heavy coin to a single coin known to be heavy. The same result applies if the initial sample is known to contain a light coin.

Therefore, if you know that the bad coin in your sample is heavy (or if you know that it is light), Table A.1 gives the correspondence between sample size and number of weighings required to locate the bad coin. I now use Table A.1 to answer the question. Begin by splitting the sample into as few groups of form 3^k as possible.[9] In this case, $90 = 81 + 9$,

[8] I thank Bingjian Ni for this solution technique. I am responsible for any errors.

[9] Can you make a conjecture about the sample size, its ternary (i.e., base three) representation, and the number of weighings needed to find the bad coin/marble?

Sample Size	Weighings (if bad coin is known heavy)
3	1
9	2
27	3
81	4
243	5
⋮	⋮

Table A.1: Weighings Needed to Find Bad Coin
If you have a sample of coins and you know that there is a bad coin in your sample and that it is heavy (or if you know that it is light), then the table gives the number of weighings required to locate the bad coin.

so you choose one group of 81 and one group of nine. Split the group of 81 into three subgroups of 27. Call these groups $\{27\}_A$, $\{27\}_B$, and $\{27\}_C$. Now use the scales to compare groups $\{27\}_A$ and $\{27\}_B$. Now use the scales again to compare groups $\{27\}_A$ and $\{27\}_C$. If the bad coin is in the group of 81, then these two weighings are sufficient to identify which subgroup of 27 the bad coin falls into and whether it is heavy or light. Consulting Table A.1, you can see that in this case it takes only three additional weighings to find the bad coin.

If both the initial weighings balance (i.e., $\{27\}_A$ versus $\{27\}_B$ and $\{27\}_A$ versus $\{27\}_C$ both balance), then the bad coin is in the group of nine. Compare the group of nine to nine good coins taken from the group of 81. This tells you whether the bad coin is heavy or light. Consulting Table A.1, you can see that in this case, it takes only two more weighings to find the bad coin. Alternatively, you could have split the group of nine into three groups of three and weighed two pairs of them. This identifies the group of three containing the bad coin and tells you whether it is heavy or light. Consulting Table A.1, you can see that in this case, it takes only one more weighing to find the bad coin. In each case, the bad coin is both located and identified as heavy or light in only five weighings (at a maximum cost of \$500).

THIRD SOLUTION[10]

Suppose you are given $N = \frac{3^n-3}{2}$ balls for some positive integer n. The balls appear identical, but one ball is odd—either heavy or light; you do not know which. Then it takes n weighings to both find the odd ball and identify it as heavy or light (see Table A.2).

Balls Supplied $N = \frac{3^n-3}{2}$	Weighings Needed n
3	2
12	3
39	4
120	5
363	6
$\vdots$	$\vdots$

Table A.2: Weighings Needed to Find Bad Coin

If you have a sample of coins and you know that there is a bad coin in your sample but not whether it is heavy or light, then the table gives the number of weighings required to locate the bad coin.

It is no coincidence that the first column in Table A.2 is the partial sums of the first column in Table A.1—this technique generalizes the second. I prove the particular case $N = 120$ (i.e., $n = 5$), but the proof generalizes directly to any $N(n) = \frac{3^n-3}{2}$.

Put the 120 balls into three groups of 40. Each group of 40 is a cohort of subgroups of size 3^k for $k = 0$ to $k = n-2$: $\{\{1\}_A, \{3\}_A, \{9\}_A, \{27\}_A\}$; $\{\{1\}_B, \{3\}_B, \{9\}_B, \{27\}_B\}$; and $\{\{1\}_C, \{3\}_C, \{9\}_C, \{27\}_C\}$.

Compare cohorts A and B. If they balance, then you have 80 good balls, and cohort C contains the bad ball. In this case, compare $\{27\}_C$ to the good balls $\{27\}_A$. If $\{27\}_C$ contains the bad ball, then Table A.1 says you need three more weighings. If $\{27\}_C$ is good, then compare $\{9\}_C$ to the good balls $\{9\}_A$. If $\{9\}_C$ contains the bad ball, then Table A.1 says you need two more weighings. If $\{9\}_C$ is good, then compare $\{3\}_C$ to

[10] I thank Yi Shen for this solution technique. I am responsible for any errors.

the good balls $\{3\}_A$. You need only one more weighing—either because $\{3\}_C$ is bad (see Table A.1), or because $\{3\}_C$ is good (thus $\{1\}_C$ is bad).

If the initial comparison of A and B does not balance, then rotate (like changing car tyres) groups $\{27\}_A$, $\{27\}_B$, and $\{27\}_C$ and compare $\{\{1\}_A, \{3\}_A, \{9\}_A, \{27\}_B\}$ to $\{\{1\}_B, \{3\}_B, \{9\}_B, \{27\}_C\}$ while holding out $\{\{1\}_C, \{3\}_C, \{9\}_C, \{27\}_A\}$. If they balance, $\{27\}_A$ contains the bad ball, and it is known to be heavy or light, and Table A.1 says three more weighings are needed. Otherwise, the scales tilt the same way, you can discard the $\{27\}$'s, rotate the $\{9\}$'s, and compare $\{\{1\}_A, \{3\}_A, \{9\}_B\}$ to $\{\{1\}_B, \{3\}_B, \{9\}_C\}$ while holding out $\{\{1\}_C, \{3\}_C, \{9\}_A\}$. You either find $\{9\}_A$ is bad, known heavy or light, and apply Table A.1, or you discard the $\{9\}$'s and rotate the $\{3\}$'s. Continue reducing the problem until convergence—in at most five weighings.

One problem with this method is what to do if given only 90 balls (more than 39 but fewer than 120). I guess you ask for 30 extra ones and then follow the procedure for 120.

Answer 1.15: This question is unusually esoteric, but I like it. The result is known as Liouville's Theorem. It can be proved directly using Picard's Theorem,[11] or with very slightly more work, using a Cauchy integral.[12] I have chosen, however, to prove it from first principles.[13]

Begin by proving a lemma (a "helping theorem" to be used in a later proof). The lemma is used in the proof of a theorem that answers the interview question. If you have a mathematical background but cannot answer the question, you should read the statement of the lemma, and stop reading there. You should then try to complete the remainder of the proof on your own. This is more satisfying than seeing the full proof.

[11]Picard: A non-constant entire function assumes every complex value, with at most one possible exception. Thus, a bounded function must be a constant.

[12]Cauchy integral: let $C(z_0, r)$ be a circle or radius r about arbitrary $z_0 \in \mathbb{C}$, then $f'(z_0) = \frac{1}{2\pi i}\int_{C(z_0,r)} \frac{f(z)}{(z-z_0)^2}dz$ (a contour integral), so $|f'(z_0)| \leq \frac{1}{r^2}\sup_{z\in\mathbb{C}}|f(z)|$. Bounded f implies the RHS tends to zero as $r \to \infty$. Thus, $f'(z_0) \equiv 0$, for arbitrary $z_0 \in \mathbb{C}$, and f must be a constant.

[13]I thank Naoki Sato and Thomas C. Watson for suggesting the Picard Theorem and Cauch integral approaches, respectively. Any errors are mine.

LEMMA: Maximum Modulus of Coefficients of a Power Series[14]
Suppose that $f(z)$ is analytic in the disc $|z| \leq r < \infty$. Let $M(r) \equiv \max\{|f(z)| : |z| \leq r\}$. Then the coefficients a_p in the power series expansion $f(z) = a_0 + a_1 z + \cdots + a_p z^p + \ldots$ satisfy the following bound:

$$|a_p| \leq \frac{M(r)}{r^{\,p}} \quad \text{for} \;\; p = 1, 2, 3 \ldots .$$

PROOF OF LEMMA
With $f(z) = a_0 + a_1 z + \cdots + a_p z^p + \ldots$ in $|z| \leq r < \infty$, divide by z^p to get

$$\frac{f(z)}{z^p} = a_0 z^{-p} + a_1 z^{-p+1} + \cdots + a_p + \ldots .$$

Now change to polar coordinates. Hold $r = |z|$ constant, and integrate $\frac{f(z)}{z^p}$ from $\phi = 0$ to $\phi = 2\pi$ [where $\phi \equiv arg(z)$]:

$$\int_0^{2\pi} \frac{f(z)}{z^p} d\phi = \int_0^{2\pi} a_0 z^{-p} d\phi + \cdots + \int_0^{2\pi} a_p d\phi + \ldots .$$

The integral is convergent because, for fixed r and varying ϕ, the original series converges uniformly. Each term in the series expansion contributes an integral of form

$$a_{k+p} \int_0^{2\pi} z^k d\phi.$$

However, for $k \neq 0$ this integral contributes zero:

$$\begin{aligned} a_{k+p} \int_0^{2\pi} z^k d\phi &= a_{k+p} \int_0^{2\pi} r^k [\cos(k\phi) + i \;\sin(k\phi)] \; d\phi \\ &= a_{k+p} \; r^k \left[\frac{\sin(k\phi)}{k} - i \; \frac{\cos(k\phi)}{k} \right]_0^{2\pi} \\ &= 0. \end{aligned}$$

[14]This lemma and its proof are adapted from Holland (1973, pp9–10), with copyright permission from Academic Press.

Thus, the only term that contributes anything to $\int_0^{2\pi} \frac{f(z)}{z^p} d\phi$ is $\int_0^{2\pi} a_p d\phi = 2\pi a_p$, (when $k = 0$). It follows that

$$a_p = \frac{1}{2\pi} \int_0^{2\pi} \frac{f(z)}{z^p} d\phi \quad \text{for} \quad p = 0, 1, 2, \ldots .$$

Now, with $M(r) = \max\{|f(z)| : |z| \leq r\}$, and $|z| = r$, it follows that for integer $p > 0$, you get

$$\begin{aligned}
|a_p| &\leq \frac{1}{2\pi} \int_0^{2\pi} \frac{|f(z)|}{|z|^p} d\phi \\
&= \frac{1}{2\pi r^p} \int_0^{2\pi} |f(z)| d\phi \\
&\leq \frac{1}{2\pi r^p} \int_0^{2\pi} M(r) d\phi \\
&= \frac{M(r)}{2\pi r^p} \int_0^{2\pi} d\phi \\
&= \frac{M(r)}{r^p} \quad \square
\end{aligned}$$

I now present the interview question as a theorem and use the previous lemma in its proof.

THEOREM: Bounded Entire Function[15]
If $f(z)$ is entire and bounded, then $f(z)$ is a constant.

PROOF OF THEOREM
Assuming that $f(z)$ is entire implies that $f(z)$ is analytic in the entire finite complex plane. Thus, the Taylor series $f(z) = \sum_{n=0}^{\infty} a_n z^n$ converges for all $|z| < \infty$. If the stated bound (of the theorem) is $|f(z)| \leq M$, say, then from the lemma it follows that

$$0 \leq |a_n| \leq \frac{M(r)}{r^n} \leq \frac{M}{r^n} \quad \text{for all } n > 0 \text{ and all } r.$$

If you let $r \to \infty$, then $0 \leq |a_n| \leq 0$ for all $n > 0$. Thus, $a_n = 0$ for all $n > 0$, and $f(z) = a_0$, a constant. $\square$

[15]This theorem and its proof are adapted from Holland (1973, p10), with copyright permission from Academic Press.

Answer 1.16: This is a nice question. If you can figure out the correct relationship between the eight lily pads and the single one, you get the answer. If you do not have it yet, or you think it is 3.75 days, then you should stop reading now, and go back and try again. I am serious; this is a nice question, and you lose a great deal by peeking at the answers to help you out.

The naive answer is that it takes $\frac{30}{8} = 3.75$ days. This is, of course, incorrect. The lily pads in the question all grow at the same rate. This means that you may think of the eight lily pads as being equivalent to one big lily pad. Indeed, when the single lily pad is three days old, it has the same area as the eight lily pads do at time zero. This means that you may think of the eight lily pads as a single lily pad that is three days old. It takes another 27 days for a three-day-old single lily pad to cover the pond, so it also takes 27 days for the eight lily pads to cover the pond.

The interviewee suggested that I use 3,000 days instead of 30 days as the time it takes for the single lily pad to cover the pond. The idea was to make the question more confusing. The problem with this is that, no matter how small the initial lily pad (assuming it is visible to the naked eye), it will cover the surface of the Earth within 100 days and the entire solar system not long after. Within 3,000 days, the universe will be blotted out—such is the power of compound growth.

Answer 1.17: You use the same idea as in the previous lily pad question. Each pad needs to cover $\frac{6{,}000}{27}$ square feet to choke the pond. The size of each pad is 2^N after N days, so you need to solve: $\frac{6{,}000}{27} = 2^N$. The solution is $N = \frac{\log\left(\frac{2{,}000}{9}\right)}{\log(2)} \approx 7.8$ days.[16]

Without a calculator, you can still do it in your head. You calculate $\frac{6{,}000}{27}$ as approximately 200. You know that 2^8 is 256, which exceeds 200; whereas 2^7 is 128, which falls short, so eight days should do the trick.

Answer 1.18: Decimal pricing was introduced on the New York Stock Exchange in 2001. I have left this question in because there are many

[16] An interesting aside here is that it does not matter which logarithm function you use. The result is the same regardless of the base.

people who lived with eighths and sixteenths for most of their working life, and they may be tempted to ask you about it.

Most people stumble a little. Do not memorize all the possible sixteenths before your interview—you have worse things to worry about (much worse). Add or subtract one-sixteenth to get the requested fraction into quarters or eighths and then compensate for your adjustment.

You should remember that $\frac{1}{8}$ is 0.1250 and deduce from that that $\frac{1}{16}$ is 0.0625 (you should be able to give *any* eighth in decimal form). The fraction $\frac{13}{16}$ is only one-sixteenth away from $\frac{12}{16}$ which is exactly three quarters (0.7500). You need only add 0.0625 to 0.7500 to get 0.8125. Similarly, $\frac{9}{16}$ is only one-sixteenth over one half, and is, therefore, 0.5625.

Answer 1.19: A common question. The naive answer is that the snail climbs a net of two feet per day, so it reaches the 10-foot mark at the end of the fifth day. However, on the morning of the fifth day, the snail starts out at the eight-foot mark (having slid down from the nine-foot mark overnight). Two-thirds of the way through the fifth day, the snail reaches the 10-foot mark and stops because there is no pole left to climb.

Answer 1.20: This is not so much a quantitative question as it is a pure logic question.[17] Here are a couple of answers (the first is a slight modification of the actual solution given to the interviewee):

1. Turn Switch #1 on. Wait a while. Then turn it off while simultaneously turning Switch #2 on. Go into the room. The illuminated light corresponds to Switch #2. The warm non-illuminated bulb corresponds to Switch #1. The cold non-illuminated bulb corresponds to Switch #3 (in the interviewer's version, Switch #1 is not turned off again).
2. Guess. You have a one-in-six chance if they are random. However, light switches are not usually random. If you assume the switches are physically located in an order that relates to the physical placement of the bulbs (as they usually are), then you have a fifty-fifty chance!

[17] I thank Dahn Tamir for assistance on this question. I am responsible for any errors.

Answer 1.21: The only way the first man can know the colour of his own hat is if he sees the other two wearing red hats—of which there are only two. However, the first man does *not* know his hat colour, so the other two must be wearing either both blue or one red and one blue. The second man, upon hearing the first, knows then that he and the third man are either both wearing blue hats, or one wears a red hat, and one a blue. If he still does not know what colour hat he is wearing, it must be because the third man is wearing a blue hat. Why? Well, if the third man wears red, then that pinpoints his own hat as blue since this is the only option left from the choice of either both blue or one red and one blue. Since the second man does not know his hat colour, then the third man must be wearing blue. The third man, upon hearing the first two, deduces that his own hat is blue via the same reasoning.

Answer 1.22: This is a very nice question indeed. You may be looking to the solutions for a hint. My first hint is that, if you are using linear algebra (i.e., solving systems of equations by substitution) then stop that right now. There are nine equations in 10 unknowns, so this will get you nowhere. In fact, there are infinitely many integers that solve the problem statement; we are searching for the smallest such number. My second hint is that you might like to try drawing a picture.

FIRST SOLUTION

My first solution technique begins with the simultaneous equation approach and quickly abandons it. Let X denote the solution. Then I know there exist positive integers X_2, X_3, X_4, ... X_{10}, such that

$$\begin{aligned} X &= 2 \times X_2 + 1, \\ X &= 3 \times X_3 + 2, \\ X &= 4 \times X_4 + 3, \\ X &= 5 \times X_5 + 4, \\ &\vdots \\ X &= 10 \times X_{10} + 9. \end{aligned}$$

Looking at the equations, it is clear that the coefficients on the right-hand side differ from the remainders by only one. If we simply add one to both sides of each equation, then the coefficients and the remainders will be identical, and we can collect terms to obtain the following:

$$\begin{aligned}
X+1 &= 2\times X'_2, \\
X+1 &= 3\times X'_3, \\
X+1 &= 4\times X'_4, \\
X+1 &= 5\times X'_5, \\
&\vdots \\
X+1 &= 10\times X'_{10}.
\end{aligned}$$

Where $X'_n \equiv X_n + 1$ for each n between 2 and 10. With this simple restatement, the problem now requires that we find the smallest number X, such that $X+1$ is perfectly divisible by 2, 3, 4, 5, 6, 7, 8, 9, and 10. That is, find a number X, such that $X+1 = LCM(2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10)$, where $LCM(\cdot)$ is the lowest common multiple operator. Given various redundancies, we conclude that $X = LCM(6,\ 7,\ 8,\ 9,\ 10) - 1 = 2520 - 1 = 2519$ is the solution.

Looking at my restatement of the problem, it should be clear that if X solves $X + 1 = K \times LCM(6,\ 7,\ 8,\ 9,\ 10)$, for any positive integer K, then X is also a solution to the problem (but not the smallest unless $K = 1$). That is, $X = K \times LCM(6,\ 7,\ 8,\ 9,\ 10) - 1 = K \times 2520 - 1$ is also a solution for any positive integer K.

I think the interviewer would have been perfectly happy to hear that $X = LCM(6,\ 7,\ 8,\ 9,\ 10) - 1$, without your having to find the LCM. However, this does leave one question unanswered: What is the most efficient way to find the lowest common multiple of a group of numbers?[18]

SECOND SOLUTION

I did not discover the first solution by looking at the simultaneous equations and using algebra; I discovered it by drawing a picture. It is somewhat difficult to reproduce my picture, but here is an attempt using a sporting analogy (see Figure A.1).

[18] In this case, if you factor each of 6, 7, 8, 9, and 10, you get 2×3, 7, 2×4, 3×3, and 2×5. The LCM, when factorized, must include each of these expressions. Indeed, $2\times 3\times 3\times 4\times 5\times 7 = 2,520$—the LCM.

I am searching for a number X with the divisor/remainder properties described. I have nine runners to help me: Mr. 2, Mr. 3, ... Mr. 9, and Mr. 10. They are assembled at the start of an arbitrarily long, dead-straight, sand-covered, nine-lane racetrack that has distances measured in metres, beginning at "0" at the starting line (bear with me on this). Like many race tracks made for people, the people do not all start in the same place; their positions are staggered (which makes no sense for a straight track in the real world). Mr. 2 starts 1 metre from the zero line. Mr. 3 starts 2 metres from the zero line. Mr. 4 starts 3 metres from the zero line, ... and Mr. 10 starts 9 metres from the zero line.

The gun fires, and the race begins. Each Mr. n runs taking steps of length n metres (for n between 2 and 10). The runners each leave footprints in the sand on the track. Let them run for a very long time and then look at the footprints (we do not care who wins). Starting at the zero line, the divisor/remainder properties of X imply that the first time you find a row of nine footprints adjacent to each other must be after X metres. The first time the footprints (the solid bullets in Figure A.1) are aligned vertically, is when they have reached the solution, X metres from the zero starting line. We can see that the number of metres they step out before beginning is one less than their step size when they run. If you look one pace backwards from the start line (back to position -1 on the race track), then it should be clear that the distance from the -1 position to X (i.e., $X + 1$) is just the LCM of the step sizes taken by the runners (how else could the footprints be adjacent?). It should also be clear that if you look beyond X, you will find another set of adjacent footprints after you travel another LCM metres. This solution is identical to the first.

Answer 1.23: This is easier than it sounds. You do not need any infinite sums, and, if you used them, go back and try again before reading on. For every two miles covered by the first motorcyclist, the second covers three miles. Two plus three is five, and there are five multiples of five between them. This means they will meet after the first has travelled 10 miles and the second 15. We know that the fly moves at twice the speed of the first motorcyclist, so it must cover 20 miles before its miserable life ends.

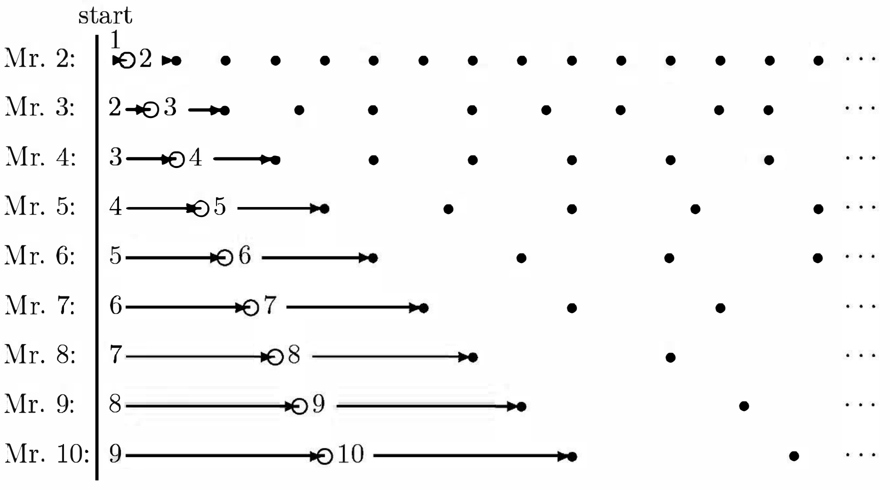


Figure A.1: A Road Race Analogy for the LCM Problem
In the picture, the nine people run side by side. Mr. 2 steps out one metre to start (the hollow bullet) and then take steps of length two metres (his footsteps are solid bullets). Mr. 3 steps out two metres to start and then take steps of length three, and so on up until Mr. 10, who steps out nine metres to start and then takes steps of length 10.

Answer 1.24: Let me begin by repeating the constraints:[19]

$$
\begin{aligned}
A+B+C+D &= 20, \\
B+C+D+E+F &= 20, \\
D+E+F+G+H &= 20, \quad \text{and} \\
F+G+H+I &= 20.
\end{aligned}
$$

We have four equations in nine unknowns. The additional information (A to I are some permutation of the integers 1 to 9) restricts the solution space, but there can be no unique solution.

If the first four (A to D) and the last four (F to I), each add to 20, then because $\sum_{i=1}^{i=9} i = 45$, it follows immediately that $E = 5$. If we

[19] I thank Dahn Tamir for suggesting this solution technique. I am responsible for any errors.

subtract the second constraint from the first and use $E = 5$, we get $A = F + 5$. If we subtract the fourth constraint from the third, we get $I = D + 5$.

The derived restrictions $A = F + 5$, and $I = D + 5$ imply that F and D must be in the set $\{1, 2, 3, 4\}$. Once F and D are chosen, A and I are determined within the set $\{6, 7, 8, 9\}$. There are thus $\begin{pmatrix} 4 \\ 2 \end{pmatrix} = \left(\frac{4!}{4!(4-2)!}\right) = 12$ (i.e., "four-choose-two") permutations of F and D. This leaves B, C, G, and H floating in the remaining four spaces. However, subtracting the second constraint from the third implies that $B + C = G + H$. There are four choices for B, but, once B is chosen, C is determined. With B and C chosen, there are two ways to allocate G and H to the remaining two slots. There are thus $12 \times 4 \times 2 = 96$ different solutions. Here are several solutions (reverse the orderings to get several more):

$$\begin{array}{c} 6\ 8\ 4\ 2\ 5\ 1\ 3\ 9\ 7, \\ 6\ 8\ 4\ 2\ 5\ 1\ 9\ 3\ 7, \\ 6\ 4\ 8\ 2\ 5\ 1\ 3\ 9\ 7, \\ 6\ 4\ 8\ 2\ 5\ 1\ 9\ 3\ 7. \end{array}$$

Story: *He took off his right shoe and sock, removed a medicated foot powder and dusted it on the foot and in the shoe. While he was putting back the shoe and sock, he mentioned that he had to use the powder four times a day, and this was the time.*

Interview Horror Stories from Recruiters
Reprinted by kind permission of MBA Style Magazine
©*1996–2004 MBA Style Magazine,* www.mbastyle.com

Answer 1.25: First of all, "very small" is classic physics slang for very, very small (i.e., so small that it is a pinpoint mass). If the rock is tossed overboard, the water level falls as though water equal in mass to the mass of the rock is being sucked out of the pool. The rock forces the boat to displace the rock's mass of water. After the rock is gone, the

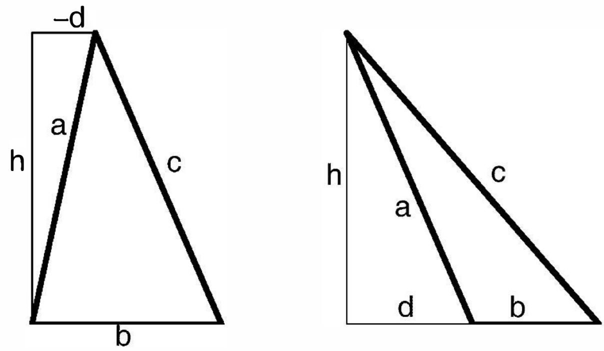


Figure A.2: Two Possible Triangle Configurations
For both triangles configurations, the sides are a, b, and c, and the height is h. The variable d is defined so that $b+d$ measures the distance from the lower right corner to the point where a vertical dropped from the peak touches the base. It follows that $d<0$ in the first case, and $d>0$ in the second.

boat rises up, and the water level falls down (Archimedes' Principle).[20]

The next time you are washing dishes, try this experiment. With the sink half-full of water, float a drinking glass. Now drop a steel ball bearing gently into the glass. The glass sinks down, displacing a mass of water equal in mass to the mass of the ball bearing, and the water level rises. Now pluck the ball bearing from the glass, using a magnet. The reverse happens, the glass rises, and the water level falls as though water equal in mass to the mass of the ball bearing is being sucked out of the sink.

Answer 1.26: Let A denote the area of a triangle of sides a, b, and c.[21] We may make several statements that apply to any triangle and which are clearly visible in Figure A.2:

1. The area A is given by $A=\frac{1}{2}bh$.

[20] Archimedes said simply that an object in a fluid experiences an upwards force equal to the weight of the fluid that is displaced by the object.

[21] I thank Thomas C. Watson for comments on an earlier version of this proof. Any errors are mine.

2. Pythagoras' Theorem implies that $a^2 = h^2 + d^2$ and $c^2 = h^2 + (b + d)^2$.
3. The first Pythagorean result implies $h^2 = a^2 - d^2$. When the two Pythagorean results are subtracted from each other, they imply $d = \frac{c^2 - a^2 - b^2}{2b}$.

If we combine the above results, we get:

$$\begin{aligned} A^2 &= \frac{1}{4}b^2h^2 \\ &= \frac{1}{4}b^2(a^2 - d^2) \\ &= \frac{1}{4}b^2\left[a^2 - \frac{(c^2 - a^2 - b^2)^2}{4b^2}\right] \\ &= \frac{1}{4}b^2a^2 - \frac{1}{16}(c^2 - a^2 - b^2)^2. \end{aligned}$$

Now, s equals half the perimeter, so $s = \frac{a+b+c}{2}$. It follows that $a = (2s - b - c)$, $b = (2s - a - c)$, and $c = (2s - a - b)$. If we plug each of these into A^2, above, and perform considerable tedious algebra, we arrive at the polynomial[22]

$$A^2(s) = 3s^4 - 4(a+b+c)s^2 + [a^2+b^2+c^2+5(ab+ac+bc)]s^2 + (a+b+c)abc.$$

This may be factored into

$$A^2(s) = [3s - (a + b + c)](s - a)(s - b)(s - c).$$

We have $3s = s + (a+b+c)$, by definition of s, so, $A^2(s) = s(s-a)(s-b)(s-c)$, and thus $A(s) = \sqrt{s(s-a)(s-b)(s-c)}$.

[22]Every term in a polynomial involves positive integral powers of literal numbers (i.e., the letters that represent numbers) pre-multiplied by a factor that does not contain the literals. So, $2x^2y^2 + 5z^3 + 2$ is a polynomial, but $4\sqrt{y} + 2$ is not. The "degree" of a polynomial is the degree of the term having the highest degree and non-zero coefficient. The polynomial $2x^2y^2 + 5z^3 + 2$ is of degree four ($2 + 2 = 4$). The polynomial $4x^2 + 2x + 1$ is of degree two. The Fundamental Theorem of Algebra says that every polynomial equation of form $f(x) = 0$ (i.e., only one literal) has at least one root (or "zero"). If the polynomial is of degree n, then it has n roots (or zeroes), where repeated roots are counted as often as their multiplicity. The Unique Factorization Theorem says that a factorization of the polynomial $f(x)$ into products of terms of form $(x - \text{root}_i)$ is unique up to trivial sign changes and ordering of terms. See Spiegel (1956) and Spiegel (1981) for more details—both part of the excellent Schaum Outline Series of books.

Answer 1.27: I present two solution techniques: a "hammer-and-tongs" brute-force approach and an elegant alternative.

FIRST SOLUTION[23]

Consider first the simple cases in which there are two or three guests. It soon becomes clear that you need to consider many different overlapping events and that you need to account for intersections of events. That is, you need basic set theory.

Let A_k, for $k = 1, \ldots, N$, denote the event that the kth guest sleeps in the room to which he or she was originally assigned (i.e., his or her "own room"). What we need to find is the probability that *at least one* of the guests ends up in his or her own room. This event is the union of the individual events and occurs with probability: $P(\bigcup_{k=1}^{N} A_k)$.

If you draw the familiar case of three intersecting circles—each representing an event—it is relatively straightforward to deduce the following "inclusion-exclusion" formula:

$$\begin{aligned} P\left(\bigcup_{k=1}^{N} A_k\right) = \sum_i P(A_i) &- \sum_{1\leq i<j\leq N} P(A_i \cap A_j) \\ &+ \sum_{1\leq i<j<k\leq N} P(A_i \cap A_j \cap A_k) - \cdots \\ &+ (-1)^{N+1} P(A_1 \cap \cdots \cap A_N). \end{aligned}$$

All you are doing here is adding the original event probabilities, then taking out the intersections where you double counted, then adjusting for the fact that you over-compensated, and so on—all of which is very easily seen if you draw intersecting sets for the case $N = 3$. To figure this out, we need to find the probability of the intersection of any group of events. Given the symmetry here, we can, without loss of generality, look at the events in the following order: $1, 2, \ldots, N$.

Given the random allocation of keys, each guest is equally likely to end up in his or her own room. That is, $P(A_i) = \frac{1}{N}$ for any $i \in \{1, 2, \ldots, N\}$. If guest 1 is given his own key, then guest 2 has a chance of only $P(A_2|A_1) = \frac{1}{N-1}$ of getting her own key back. So, $P(A_1 \cap A_2) = P(A_2|A_1)P(A_1) = \frac{1}{N(N-1)}$. In fact, this result is more general:

[23] I thank Taras Klymchuk for suggesting a very similar solution technique. I am responsible for any errors.

$$
\begin{aligned}
P\left(\bigcap_{k=1}^{m} A_i\right) &= \frac{1}{N(N-1)\cdots(N-m+1)} \\
&= \frac{(N-m)!}{N!} \\
&= \frac{1}{m!\begin{pmatrix} N \\ m \end{pmatrix}}, \quad \text{for any } m \in 1,\ 2,\ \ldots,\ N.
\end{aligned}
$$

We may now plug this formula for probabilities of intersecting events back into the original inclusion-exclusion formula:

$$
\begin{aligned}
P\left(\bigcup_{k=1}^{N} A_k\right) &= \sum_{m=1}^{N}(-1)^{m+1}\begin{pmatrix} N \\ m \end{pmatrix} P\left(\bigcap_{k=1}^{m} A_i\right) \\
&= \sum_{m=1}^{N}\frac{(-1)^{m+1}}{m!} \\
&= -1 \times \sum_{m=1}^{N}\frac{(-1)^{m}}{m!} \\
&= 1 - \sum_{m=0}^{N}\frac{(-1)^{m}}{m!} \\
&\rightarrow 1 - e^{-1} = \frac{e-1}{e} \quad \text{as} \quad N \rightarrow \infty.
\end{aligned}
$$

The final result is about 63%, but a guess of $\frac{2}{3}$ is close enough.

SECOND SOLUTION[24]

Simplify the problem by assuming that the "very large number" of people is almost an infinite number. In this case, it is as though each person is first in line to be allocated a key because the previous finite number of people are negligible compared to the almost infinite number of people waiting to receive keys. It follows that each person has the same probability, $\frac{1}{N}$, of being allocated his or her key. Let X be the

[24] I thank Jason Roth for supplying this technique. I am responsible for any errors.

number of people who end up sleeping in their own rooms, then

$$\begin{aligned} P(X \geq 1) = 1 - P(X = 0) &= 1 - \left(1 - \frac{1}{N}\right)^N \\ &= 1 - \left(1 + \frac{-1}{N}\right)^N \\ &\rightarrow 1 - e^{-1} = \frac{e-1}{e} \quad \text{as} \quad N \rightarrow \infty. \end{aligned}$$

Answer 1.28: The answer involves both mathematical induction and game theoretic arguments.

If there is *exactly one* cheating man in the town, Mr. C, say, then every wife except Mrs. C knows who he is. Not only that, but Mrs. C is unaware of any cheats—the stranger's announcement comes as a shock to her. Immediately after the stranger's announcement, Mrs. C asks: "Who can be cheating if I have seen no cheats?" The only possible answer is it is Mr. C. Come the next morning, his happy days are over, and out he goes.

Suppose instead that there are *exactly two* cheating men in town: Mr. C1 and Mr. C2. In this case, Mrs. C1 knows there is one cheat in town (Mr. C2), and Mrs. C2 knows there is one cheat in town (Mr. C1)—the stranger's announcement comes as no shock to either woman. Each thinks there is only one cheat in town and fully expects him to be kicked out the next morning (each wife thinks the other poor woman is in the position of Mrs. C mentioned above). The first morning after the announcement comes, and the streets are bare. Mrs. C1 concludes that Mrs. C2 did not kick her husband out because she did not think he was a cheat. How could Mrs. C2 be so foolish? Mrs. C1 knows that Mrs. C2 believes the prophecy, so the only possible reason for Mrs. C2 not to have reacted is if Mrs. C2 saw a cheat herself. Mrs. C1 asks herself: "Who did Mrs. C2 see cheating, when the only cheat I can see is Mr. C2?" The only possible answer is that it is her own man, Mr. C1. Come the second morning after the announcement, Mr. C1 and Mr. C2 are both kicked out (the latter because Mrs. C2 went through the same thought process).

Suppose now that there are *exactly three* cheating men in town: Mr. C1, Mr. C2, and Mr. C3. In this case, each of Mrs. C1, Mrs. C2, and Mrs. C3 thinks that there are two cheats in town and believes in the innocence of her own man. However, come the second morning, they are each very surprised to find the streets empty. Had there been exactly two cheats, as each of the wives had surmised, then the cheats should have been kicked into the street two mornings after the announcement—as per the argument above. The empty street means that a third cheater exists—one previously assumed innocent! So, three mornings after the announcement, all three cheating men are bounced out into the street.

More generally, let me assert that if there are exactly n cheats, then they will all be kicked out into the street on the nth morning after the stranger's announcement. If my assertion is true for n cheats, and a wife sees n other cheats but finds the streets bare on the nth morning, then she is shocked to conclude that her own man must be unfaithful to her. She (and each of the other n wives) will kick her man out the next morning. That is, if there are $n+1$ cheats, then they will be kicked out on the $(n+1)$st morning. That is, if my assertion is true for n cheats, then it is also true for $n+1$ cheats.

I proved my assertion to be true for n equal to each of one, two, and three. It follows my mathematical induction that it is true for all n (in fact, I needed only to prove it for $n=1$ for the proof to go through).

It follows that if cheating men are kicked into the street for the first time on the tenth morning after the stranger's announcement, then there must be exactly 10 of them.

Answer 1.29: This is one of the easier questions in the book. If you are peeking here for a solution, then go back and think about mathematical induction.

Let $V(n)$ denote the minimum number of moves needed for n rings. I *assert* that $V(n) = 2^n - 1$, for all positive integers n. The proof uses mathematical induction.

<u>Case $n = 1$:</u> With one ring, it certainly takes exactly one move. My assertion is thus true for the case $n = 1$.

<u>Case $n = N$:</u> Suppose that my assertion is true for $n = N$, and consider the case $n = N + 1$. By assumption, it takes $V(N) = 2^N - 1$ moves to get the first N rings to pole #2. Use one additional move to get ring #$(N+1)$ to pole #3. Now use $V(N) = 2^N - 1$ moves to move

the N rings on pole #2 to pole #3. The total number of moves used is $2V(N)+1 = 2(2^N - 1)+1 = 2^{(N+1)} - 1$. However, this is just $V(N+1)$. That is, if my assertion is true for $n = N$, it is also true for $n = (N+1)$.

The result follows, because I showed my assertion is true for $n = 1$, and I showed that if my assertion is true for $n = N$, then it is also true for $n = N + 1$. In particular, because I showed the assertion to be true for $n = 1$, it follows that it must be true for $n = 2$. It then follows that because the assertion is true for $n = 2$, it is also true for $n = 3$, ... and so on, up to ∞.[25]

In the particular case of $n = 64$ rings, it takes $V(64) = 2^{64} - 1$ moves. This number is large:

$$V(64) = 2^{64} - 1 = 18,446,744,073,709,551,616.$$

At one move per second, it would take you 584.5 billion (i.e., 584.5 thousand million) years to complete this task. The Earth will fall into the Sun in less than one-hundredth of this time period.

Answer 1.30: The ordinary differential equation (ODE) $u'' + u' + u = 1$ has a simple solution. This is a second-order linear ODE with constant coefficients, so we need only search for solutions to the homogeneous form $u'' + u' + u = 0$, and then tag on a solution to the specific nonhomogeneous equation given.

Solutions to a second-order linear homogeneous ODE of form $Au'' + Bu' + Cu = 0$ are of form[26]

$$u(x) = ae^{\lambda_1 x} + be^{\lambda_2 x},$$

where λ_1 and λ_2 are the roots of the characteristic equation:

$$A\lambda^2 + B\lambda + C = 0.$$

It follows (using the quadratic formula) that

$$\lambda_1, \lambda_2 = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} = \frac{-1 \pm \sqrt{1-4}}{2} = \frac{-1}{2} \pm \frac{\sqrt{3}}{2} i,$$

[25] A natural question to ask is how I guessed that $V(n) = 2^n - 1$ to begin with. I got this because I figured that $V(n+1) = 2V(n) + 1$ had to hold, and $V(1) = 1$ is obvious. These together are sufficient to deduce the functional form of $V(n)$.

[26] Unless $\lambda_1 = \lambda_2 = \lambda$, say (i.e., a repeated root), in which case solutions are of form $u(x) = axe^{\lambda x} + be^{\lambda x}$.

where $i \equiv \sqrt{-1}$. In our case, $u = 1$ is a solution to the specific nonhomogeneous ODE, so the general solution must be of form

$$u(x) = ae^{\left(\frac{-1}{2}+\frac{\sqrt{3}}{2}i\right)x} + be^{\left(\frac{-1}{2}-\frac{\sqrt{3}}{2}i\right)x} + 1,$$

for arbitrary constants a and b. To pinpoint a and b, you need two initial conditions (not supplied here) in addition to the ODE.

Answer 1.31: The obvious application is to proportions of a portfolio invested in risky assets. Make the substitution $b = 1 - a$. Then the variance of the sum is

$$V(S) = a^2\sigma_X^2 + 2a(1-a)\rho\sigma_X\sigma_Y + (1-a)^2\sigma_Y^2.$$

The first-order condition is $\frac{\partial V(S)}{\partial a} = 0$. The partial derivative is:

$$\begin{aligned}\frac{\partial V(S)}{\partial a} &= 2a\sigma_X^2 + 2\rho\sigma_X\sigma_Y - 4a\rho\sigma_X\sigma_Y + 2(1-a)(-1)\sigma_Y^2\\ &= 2\left[a(\sigma_X^2 - 2\rho\sigma_X\sigma_Y + \sigma_Y^2) + \rho\sigma_X\sigma_Y - \sigma_Y^2\right].\end{aligned}$$

Thus, the particular a that satisfies the first-order condition is

$$a^* = \frac{\sigma_Y^2 - \rho\sigma_X\sigma_Y}{(\sigma_X^2 - 2\rho\sigma_X\sigma_Y + \sigma_Y^2)}.$$

We should check the second-order condition:

$$\left.\frac{\partial^2 V(S)}{\partial a^2}\right|_{a=a^*} > 0,$$

to make sure this is a minimum, not a maximum. This is straightforward here because:

$$\begin{aligned}\frac{1}{2}\frac{\partial^2 V(S)}{\partial a^2} &= \sigma_X^2 - 2\rho\sigma_X\sigma_Y + \sigma_Y^2\\ &\geq \sigma_X^2 - 2(+1)\sigma_X\sigma_Y + \sigma_Y^2\\ &= (\sigma_X - \sigma_Y)^2\\ &\geq 0,\end{aligned}$$

and the first inequality is strict unless $\rho = +1$.

In fact, I have solved the unconstrained problem—ignoring the constraint $0 \leq a \leq 1$. If a^* breaches the constraints, the constrained solution for a is either 1 or 0, depending upon whether σ_X or σ_Y is the smaller respectively.[27]

In the special case where $\rho = -1$ (perfect negative correlation), the solution for a^* is given by

$$\begin{aligned} a^* &= \frac{\sigma_Y^2 - \rho\sigma_X\sigma_Y}{(\sigma_X^2 - 2\rho\sigma_X\sigma_Y + \sigma_Y^2)} \\ &= \frac{\sigma_Y^2 + \sigma_X\sigma_Y}{(\sigma_X^2 + 2\sigma_X\sigma_Y + \sigma_Y^2)} \\ &= \frac{\sigma_Y(\sigma_X + \sigma_Y)}{(\sigma_X + \sigma_Y)^2} \\ &= \frac{\sigma_Y}{(\sigma_X + \sigma_Y)}, \end{aligned}$$

and this particular a^* gives variance of $aX + bY$ equal to zero.

Answer 1.32: The lighthouse question is an old favourite.[28] The lighthouse is a distance L from the coast. The beam of light casts a "spot" a distance R across the sea from the lighthouse (see Figure A.3).

FIRST SOLUTION

Using the coordinates in Figure A.3, the coastline is the line $x = L$. The spot hits the coastline at the coordinate pair $(x, y) = (L, L\tan(\theta))$, where θ is the angle between the beam and the x-axis. Suppose $\theta = 0$ when $t = 0$, then $\theta = \omega t = \frac{2\pi}{60}t$ where $\omega = \frac{2\pi}{60}$ is the angular velocity in radians per second (the beam makes one revolution per minute and t is measured in seconds). The speed V of the spot along the coastline is the partial derivative of $y = L\tan(\theta) = L\tan(\omega t)$ with respect to t.

$$V = \frac{\partial}{\partial t}[L\tan(\theta)] = \frac{\partial}{\partial t}[L\tan(\omega t)] = \omega L \sec^2(\omega t) = \frac{\omega L}{\cos^2(\omega t)}$$

From Figure A.3 we see that $\cos(\omega t) = \frac{L}{R}$, so we conclude that $V = \frac{\omega L}{(L/R)^2} = \frac{\omega R^2}{L}$. In our particular case, with $\omega = \frac{2\pi}{60}$ radians per second

[27]The a^* will breach the constraints if the correlation ρ is large enough or the standard deviations are disparate enough that either $\frac{\sigma_X}{\sigma_Y} < \rho$ or $\frac{\sigma_Y}{\sigma_X} < \rho$.

[28]I thank Valeri Smelyansky for advice. Any errors are mine.

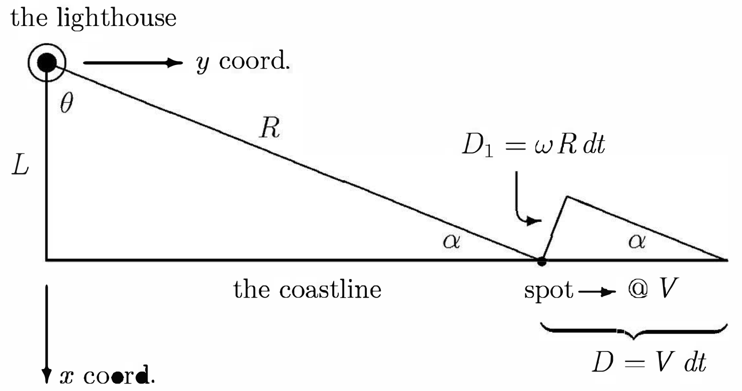


Figure A.3: The Lighthouse Problem
Refer to this figure for both solutions to the lighthouse problem. The first solution uses the x-y coordinates and θ; the second solution uses α, D, and D_1; both solutions use L, R, and V.

and $L = 3$, the speed of the beam is $\frac{\pi R^2}{90}$ miles per second. When the beam is $3L = 9$ miles down the coastline, $R^2 = 10L^2 = 90$, and the speed is just π miles per second.

More than once, students have suggested to me that the velocity V is a constant (i.e., V is the same regardless of how far along the coast the spot is cast)—this is clearly incorrect.

SECOND SOLUTION

Follow the beam's course for a small time interval dt. In Figure A.3, we see that the beam's spot covers a distance $D = V dt$ along the coast, while the beam's "perpendicular motion" covers a distance $D_1 = \omega R dt$ (where V is the spot's speed along the coast, and $\omega = \frac{2\pi}{60}$ radians per second is the beam's angular velocity). For small dt, the distance triangle is a right-angle triangle, so $\sin(\alpha) = \frac{D_1}{D} = \frac{\omega R dt}{V dt} = \frac{\omega R}{V}$. Looking at the larger triangle, we see $\sin(\alpha) = \frac{L}{R}$. Thus, $V = \frac{\omega R^2}{L}$, as before.

Answer 1.33: There are many different ways to solve this problem. Let me begin with a "hammer-and-tongs" approach using algebra. When you see how neat the solution is, try to come up with an argument that uses no mathematics whatsoever (I present such an argument after the hammer-and-tongs approach).

FIRST SOLUTION

Identify the squares using horizontal and vertical indices, counting from the northwest corner. Let i count down and j count across. Then it is readily seen that the square with coordinates (i, j) has $(i + j - 1)$ cubes on it. It follows that the total number of cubes is given by (in the general case of an $n \times n$ chessboard)[29]

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n}(i+j-1) &= \sum_{i=1}^{n}\sum_{j=1}^{n}[(i-1)+j] \\
&= \sum_{i=1}^{n}\left[n(i-1)+\frac{n(n+1)}{2}\right] \\
&= \left[n\left(\frac{n(n+1)}{2}-n\right)+n\left(\frac{n(n+1)}{2}\right)\right] \\
&= n\left(\frac{n^2+n-2n+n^2+n}{2}\right) \\
&= n^3.
\end{aligned}$$

The answer n^3 is extremely neat and tidy. In the special case where $n = 20$, there are $20^3 = 8,000$ cubes on the board.

SECOND SOLUTION

With an answer this neat, there must be a non-algebraic solution. Imagine the 20×20 chessboard in front of you, with the stacks of cubes on it as in Figure A.4. Now slice through the cubes horizontally at height 20 units. The cubes above the slice all lie in the southeast lower-triangular section below the non-leading diagonal. Now flip the above-the-slice cubes across the diagonal from southeast to northwest. They will fill the lower stacks to a height of 20 units. You now have a solid cube, and the result follows immediately.

Answer 1.34: The naive strategy is to run directly away from the dog toward the edge of the field. However, at speed v, it takes you $\frac{R}{v}$ units of time to get to the perimeter, while it takes the dog only $\frac{1}{2}\frac{2\pi R}{4v} = \frac{\pi R}{4v} \approx \frac{3}{4} \times \frac{R}{v}$ units of time to get there—so he will meet you and eat

[29] I make use of the property that $\sum_{j=1}^{n} j = \frac{n(n+1)}{2}$ (see Question 1.2).

1	2	3	4	$\cdots$	19	20
2	3	4	5	$\cdots$	20	21
3	4	5	6	$\cdots$	21	22
4	5	6	7	$\cdots$	22	23
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
19	20	21	22	$\cdots$	37	38
20	21	22	23	$\cdots$	38	39

Figure A.4: Number of Cubes on Each Square of a 20×20 Chessboard (A) The figure shows the number of cubes on each square of a chessboard, starting with one in the northwest corner and stepping up one each time you step south or east.

you. You somehow need to get further from him and closer to the fence before you make a run for it.

Suppose you behave somewhat like the dog. Step away from the centre of the circle until you are at a radius of $\frac{R}{4}$. Now constrain yourself to running circuits around that radius. It takes you $\frac{\pi R}{4v}$ units of time to run half-way around this circle. The dog can also run half-way around the field in the same time. That is, at this radius, you and the dog are perfectly matched in your abilities to run around in circles.

Now suppose you step slightly closer to the centre of the circle. Let us say you now start running around in circles of radius $\frac{R}{4} - \epsilon$, for some small ϵ. In this case, you have a slight advantage over the dog: you can run around your circle in slightly less time than he can run around his. As you run, the dog tries to track you. However, you are gaining a little on the dog with every circuit. Eventually, you will be at the "top" of your circle, when he is at the "bottom" of his. Now it is time to make a run for it. You only have to travel a distance of $R - \left(\frac{R}{4} - \epsilon\right) = \frac{3}{4}R + \epsilon$. The dog has to travel a distance πR to meet you. You can outrun him as long as the time it takes you at speed v is less than the time it takes him at $4v$, that is:

$$\begin{aligned}
\frac{\frac{3}{4}R+\epsilon}{v} &< \frac{\pi R}{4v} \\
\Leftrightarrow \frac{3}{4}R+\epsilon &< \frac{\pi R}{4} \\
\Leftrightarrow 3R+4\epsilon &< \pi R \\
\Leftrightarrow \qquad \epsilon &< \frac{(\pi-3)R}{4}.
\end{aligned}$$

It follows that if you choose an ϵ such that $0<\epsilon<\frac{(\pi-3)R}{4}$, then you can run in a circle of radius $\frac{R}{4}-\epsilon$ until you are as far from the dog as possible and then you can escape by running away from him.

Answer 1.35: There are two methods. The first method assumes a known probability result (this may be sufficient for you); the second method subsumes the first by proving the aforementioned probability result before proceeding.

FIRST SOLUTION

The integral is immediately recognized as a simple transformation of an integral over the entire domain of a Normally distributed random variable.

The standard Normal distribution has probability density function $f(u) \equiv \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}$, for $-\infty<u<+\infty$. Integrating over the entire domain must produce total probability mass of unity:

$$\int_{-\infty}^{+\infty} f(u)du = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}du = 1.$$

If we substitute in $x=\frac{1}{\sqrt{2}}u$ (to make the integral look like the one we seek), then $dx=\frac{1}{\sqrt{2}}du$, and we get:

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{\pi}}e^{-x^2}dx = 1.$$

Multiply both sides by $\sqrt{\pi}$, and the result follows immediately.

SECOND SOLUTION

Let $I = \int_{-\infty}^{+\infty} e^{-x^2}dx$. Then squaring gives the following:

$$\begin{aligned}
I^2 &= \left(\int_{-\infty}^{+\infty} e^{-x^2}dx\right)\cdot\left(\int_{-\infty}^{+\infty} e^{-y^2}dy\right)\\
&= \int_{x=-\infty}^{+\infty}\int_{y=-\infty}^{+\infty} e^{-\left(x^2+y^2\right)}dydx\\
&\overset{\text{see below}}{=} \int_{\theta=0}^{2\pi}\int_{r=0}^{+\infty} e^{-r^2}rdrd\theta\\
&= \int_{\theta=0}^{2\pi} -\frac{1}{2}e^{-r^2}\Big]_{r=0}^{\infty} d\theta\\
&= \frac{1}{2}\int_{\theta=0}^{2\pi} d\theta = \pi.
\end{aligned}$$

Thus, $I = \sqrt{\pi}$, as required.

The above basis change from Cartesian coordinates to polar coordinates uses the transformation $x = r\cos\theta$, and $y = r\sin\theta$. This implies that $x^2 + y^2 = r^2$. However, there is more to it than this. You also need to know the general result:

$$\int_x\int_y f(x,y)dxdy = \int_\theta\int_r f(x(r,\theta),y(r,\theta))rdrd\theta.$$

The "r" in the integrand on the right-hand side is the "Jacobian" of the transformation from Cartesian to polar coordinates. The Jacobian, J, is just the determinant of the matrix of partial derivatives of the transformation. That is:

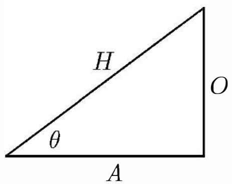
$$\begin{aligned}
\int_x\int_y f(x,y)dxdy &= \int_\theta\int_r f(x(r,\theta),y(r,\theta))Jdrd\theta, \text{ where}\\
J \equiv \frac{\partial(x,y)}{\partial(r,\theta)} &\equiv \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta}\end{vmatrix}\\
&= \begin{vmatrix} \frac{\partial[r\cos\theta]}{\partial r} & \frac{\partial[r\cos\theta]}{\partial \theta}\\ \frac{\partial[r\sin\theta]}{\partial r} & \frac{\partial[r\sin\theta]}{\partial \theta}\end{vmatrix}\\
&= \begin{vmatrix} \cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{vmatrix}\\
&= r\cos^2\theta + r\sin^2\theta\\
&= r(\cos^2\theta + \sin^2\theta) = r.
\end{aligned}$$

For more on Jacobians, determinants, and transformations, consult DeGroot (1989, pp162–166) and Anton (1988, pp1068–1069).

Answer 1.36: Of all the simple trigonometric functions that you might have been asked to integrate, $\int \sec\theta d\theta$ has arguably the most complicated answer.

Perhaps it is useful to review quickly the definitions of these trigonometry functions. Consider a right-angle triangle. Let θ be one of the acute angles (i.e., one of the two angles of less than 90 degrees). Let the lengths of the sides of the triangle be denoted by "A" (the side adjacent to the angle θ), "O" (the side opposite to the angle θ), and "H" (the hypotenuse—opposite the right angle), then the elementary trigonometric functions may be defined as in Table A.3.[30]

$\sin\theta = \frac{O}{H}$	$\operatorname{cosec}\theta = \frac{H}{O} = \frac{1}{\sin\theta}$
$\cos\theta = \frac{A}{H}$	$\sec\theta = \frac{H}{A} = \frac{1}{\cos\theta}$
$\tan\theta = \frac{O}{A} = \frac{\sin\theta}{\cos\theta}$	$\cot\theta = \frac{A}{O} = \frac{1}{\tan\theta} = \frac{\cos\theta}{\sin\theta}$



Mnemonic: "SOH-CAH-TOA"

Table A.3: Trigonometrical Functions: Definitions
These definitions are for the triangle illustrated. The sides are of length A (adjacent to the angle θ), O (opposite to the angle θ), and H (the hypotenuse).

For the particular problem given, $\int \sec\theta d\theta$, we see in Table A.4 that the answer is $\ln|\sec\theta + \tan\theta|$ (up to an arbitrary constant of integration), which is readily verified via differentiation.

[30]The trigonometric functions' names are short for sine, cosine, tangent, cotangent, secant, and cosecant.

$\int f(x)dx$	$f(x)$	$f'(x)$
$-\cos\theta$	$\sin\theta$	$\cos\theta$
$\sin\theta$	$\cos\theta$	$-\sin\theta$
$\ln\lvert\sec\theta\rvert$	$\tan\theta$	$\sec^2\theta$
$\ln\lvert\sin\theta\rvert$	$\cot\theta$	$-\mathrm{cosec}^2\,\theta$
$\ln\lvert\sec\theta+\tan\theta\rvert$	$\sec\theta$	$\sec\theta\tan\theta$
$\ln\lvert\tan\frac{1}{2}\theta\rvert$	$\mathrm{cosec}\,\theta$	$-\mathrm{cosec}\,\theta\cot\theta$

Table A.4: Trigonometrical Functions: Calculus

The middle column gives a trigonometrical function. The columns to the left and right give the integral of the function (ignoring arbitrary constant), and the derivative of the function, respectively.

Answer 1.37: The sum $\sum_{n=1}^{\infty} e^{-\sqrt{n}}$ takes the form $\sum_{n=1}^{\infty} a_n$, where $a_n \equiv e^{-\sqrt{n}}$. There is a whole host of tests for the convergence of such sums. Before we look at these, a short review of the terminology is in order.

A "sequence" is a set of numbers a_1, a_2, a_3, ... indexed in a particular order corresponding to the natural numbers. We may denote the sequence as "$\{a_n\}$." Each number, a_n, in the sequence is a "term." The "limit of a sequence" exists and is equal to $l < \infty$ if the numbers a_n get closer and closer to l as n gets larger. That is, $\lim_{n\to\infty} a_n = l$. If such a limit exists, then the sequence is said to "converge" to that limit, and the limit is unique. If a sequence does not converge, then it "diverges." There is no mention of additivity here: A sequence is a succession, not a sum.

A "series" is formed from a sequence via partial sums. Let $S_1 = a_1$, $S_2 = a_1 + a_2$, $S_3 = a_1 + a_2 + a_3$, and so on, so that $S_n = \sum_{i=1}^{n} a_i$ is the nth "partial sum" of the sequence $\{a_n\}$. Then the sum $\sum_{n=1}^{\infty} a_n$ is referred to as an "infinite series." The infinite series $\sum_{n=1}^{\infty} a_n$ is said to be "convergent" if the sequence of its partial sums $\{S_n\}$ is convergent.

A necessary (but not sufficient) condition for convergence of an infinite series $\{a_n\}$ is that $a_n \to 0$ as $n \to \infty$. In our case, $a_n = e^{-\sqrt{n}} \to 0$, so we cannot reject convergence.

The first (of several) formal tests that comes to mind is **The Ratio Test** (for series with positive terms only):

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}\begin{cases} <1, & \Rightarrow \sum a_n \ \text{ converges,}\\ >1, & \Rightarrow \sum a_n \ \text{ diverges,}\\ =1, & \Rightarrow \ \text{the test fails.}\end{cases}$$

In our case, the test fails because

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\lim_{n\to\infty}\frac{e^{\sqrt{n}}}{e^{\sqrt{n+1}}}=1.$$

The next formal test for convergence that comes to mind is **The nth Root Test** (for series with positive terms only):

$$\lim_{n\to\infty}\sqrt[n]{a_n}\begin{cases} <1, & \Rightarrow \sum a_n \ \text{ converges,}\\ >1, & \Rightarrow \sum a_n \ \text{ diverges,}\\ =1, & \Rightarrow \ \text{the test fails.}\end{cases}$$

In our case, the test fails because

$$\begin{aligned}\lim_{n\to\infty}\sqrt[n]{a_n} &= \lim_{n\to\infty}\left(e^{-\sqrt{n}}\right)^{\frac{1}{n}}\\ &= \lim_{n\to\infty} e^{-n^{\frac{1}{2}}\times n^{-1}}\\ &= \lim_{n\to\infty} e^{-n^{-\frac{1}{2}}}\\ &= \lim_{n\to\infty}\frac{1}{e^{\frac{1}{\sqrt{n}}}}\\ &= 1.\end{aligned}$$

When the two tests above fail, we head for **Raabe's Test** (for series with positive terms only):

$$\lim_{n\to\infty}\left[n\left(1-\frac{a_{n+1}}{a_n}\right)\right]\begin{cases} >1, & \Rightarrow \sum a_n \ \text{ converges,}\\ <1, & \Rightarrow \sum a_n \ \text{ diverges,}\\ =1, & \Rightarrow \ \text{the test fails.}\end{cases}$$

In our case, the test indicates convergence because

$$\lim_{n\to\infty}\left[n\left(1-\frac{a_{n+1}}{a_n}\right)\right] = \lim_{n\to\infty}\left[n\left(1-e^{\sqrt{n}-\sqrt{n+1}}\right)\right] > 1.$$

However, the algebraic proof that that last limit exceeds one is by no means elegant. Instead of proving it, I present another test of convergence that is both elegant and conclusive.

> **Story:** *1. She wore a Walkman and said she could listen to me and the music at the same time. 2. Balding candidate abruptly excused himself. Returned to office a few minutes later, wearing a hairpiece.*
>
> *Interview Horror Stories from Recruiters*
>
> *Reprinted by kind permission of* MBA Style Magazine
>
> ©*1996–2004 MBA Style Magazine,* www.mbastyle.com

The Integral Test applies to a series $\sum a_n$ of positive terms. Let $A(x)$ denote the function of x obtained by replacing n in a_n by x. Then if $A(x)$ is decreasing and continuous for $x \geq 1$,

$$\sum_{n=1}^{\infty} a_n \quad \text{and} \quad \int_{x=1}^{+\infty} A(x)dx$$

either both converge, or both diverge (Anton [1988, p623]).

In our case, $a_n = e^{-\sqrt{n}}$. To test for convergence of $\sum_{n=1}^{\infty} a_n$, we may look at convergence of $\int_{x=1}^{+\infty} A(x)$, where $A(x) \equiv e^{-\sqrt{x}}$ is seen to be both decreasing and continuous for $x \geq 1$.

We need $\int e^{-\sqrt{x}}dx$. My first guess for this integral is

$$\int e^{-\sqrt{x}}dx = e^{-\sqrt{x}} + x^{\frac{1}{2}}e^{-\sqrt{x}} = (1+\sqrt{x})e^{-\sqrt{x}}.$$

However, differentiation shows that I am out by a factor of -2. If you cannot guess this directly, you need some practice with integration by parts. We get the following integral:

$$\begin{aligned}
\int_{x=1}^{+\infty} e^{-\sqrt{x}}dx &= -2(1+\sqrt{x})e^{-\sqrt{x}}\Big|_{1}^{\infty} \\
&= 2(1+\sqrt{x})e^{-\sqrt{x}}\Big|_{\infty}^{1} \\
&= 2(1+\sqrt{x})e^{-\sqrt{x}}\Big|_{1} - 2(1+\sqrt{x})e^{-\sqrt{x}}\Big|_{\infty} \\
&= \frac{4}{e} - 2\lim_{x\to\infty}\frac{(1+\sqrt{x})}{e^{\sqrt{x}}} \\
&= \frac{4}{e} - 2\lim_{u\to\infty}\frac{(1+u)}{e^{u}} = \frac{4}{e},
\end{aligned}$$

because $\sqrt{x}\to\infty$ if and only if $x\to\infty$, and $\lim_{u\to\infty}\frac{(1+u)}{e^u}=0$ is well known. It follows that the series is convergent! Incidentally, the limit of the series $\sum_{n=1}^{+\infty} e^{-\sqrt{n}}$ is only slightly below $\frac{4}{e}$.

The **Comparison Test** says that if there exists $N<\infty$ such that $0\leq a_n\leq b_n$ for $n\geq N$, and if $\sum_{n=1}^{+\infty} b_n$ is convergent, then so too is $\sum_{n=1}^{+\infty} a_n$. This also works in reverse: If $\sum_{n=1}^{+\infty} a_n$ is divergent, so too is $\sum_{n=1}^{+\infty} b_n$. This test requires that you construct b_n. In our case, if $n\geq 3$, then $a_n=e^{-\sqrt{n}}<\frac{1}{n^2}=b_n$, and $\sum_{n=1}^{+\infty}\frac{1}{n^2}$ is known to converge! In fact, $\sum_{n=1}^{+\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$ (Spiegel [1968, p108]).

It is worth noting that if $\sum|a_n|$ converges, then $\sum a_n$ converges also. The former is referred to as "absolute convergence." Thus, absolute convergence of an infinite series is sufficient for convergence. Absolute convergence is not, however, a necessary condition for convergence. A series that is convergent, but not absolutely convergent, is said to be "conditionally convergent."

A final convergence test we might have tried is **Gauss' Test** (for series with positive terms only): If $\frac{a_{n+1}}{a_n}=1-\frac{\mathcal{L}}{n}+\frac{b_n}{n^2}$, where there exists an $M>0$, and an N such that $|b_n|<M$ for all $n>N$, then the series $\sum_{n=1}^{\infty} a_n$ is

1. convergent if $\mathcal{L}>1$, and
2. divergent or conditionally convergent if $\mathcal{L}\leq 1$.

For more information on tests of convergence of series, look to your favourite calculus book. Most of the above-mentioned tests should appear if the book is worthwhile.

Answer 1.38: The sums in Table A.5 are well known (the first is discussed extensively in the answer to Question 1.2).[31] You should certainly know

$\sum_{k=1}^{n} k$	$\frac{n(n+1)}{2}$
$\sum_{k=1}^{n} k^2$	$\frac{n(n+1)(2n+1)}{6}$
$\sum_{k=1}^{n} k^3$	$\frac{n^2(n+1)^2}{4}$

Table A.5: Sums of k, k^2, and k^3

the first sum by heart, and you should note that the third is the first squared (Grahame Bennett has given me a very elegant geometrical proof of this). My first solution uses a sensible guess plus induction. My second solution is similar, but requires that you notice or already know a special result.

FIRST SOLUTION[32]

Given that $\sum_{k=1}^{n} k = n(n+1)/2$, let us assume that $\sum_{k=1}^{n} k^i$ equals an $(i+1)^{th}$-order polynomial $f^{(i)}(n)$. In the case $i = 2$ (i.e., trying to find $\sum_{k=1}^{n} k^2$), let this polynomial be $f^{(2)}(n) = an^3 + bn^2 + cn + d$. We can calculate $f^{(2)}(n)$ for $n = 1$–4, and set up the system Equations A.1.

$$\begin{array}{rcrcrcrcrcr} f^{(2)}(1) & = & a & + & b & + & c & + & d & = & 1 \\ f^{(2)}(2) & = & 8a & + & 4b & + & 2c & + & d & = & 5 \\ f^{(2)}(3) & = & 27a & + & 9b & + & 3c & + & d & = & 14 \\ f^{(2)}(4) & = & 64a & + & 16b & + & 4c & + & d & = & 30 \end{array} \tag{A.1}$$

Standard row-reduction techniques soon yield $a = 1/3$, $b = 1/2$, $c = 1/6$, and $d = 0$. Thus, $f^{(2)}(n) = \frac{2n^3+3n^2+n}{6} = \frac{n(n+1)(2n+1)}{6}$. Having obtained a formula that works for $n = 1$–4, we must now prove, by induction, that if it works for n, then it works for $n+1$. That is, we

[31] The fourth-order result is not well known: $\sum_{k=1}^{n} k^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}$ (Spiegel [1968, p108]).

[32] I thank Vince Moshkevich for suggesting this technique. Any errors are mine.

show that $f(n) + (n+1)^2 = f(n+1)$:

$$
\begin{aligned}
f(n) + (n+1)^2 &= \frac{2n^3 + 3n^2 + n}{6} + (n^2 + 2n + 1) \\
&= \frac{2n^3 + 9n^2 + 13n + 6}{6} \\
&= \frac{(n+1)(n+2)(2n+3)}{6} \\
&= \frac{[(n+1)][(n+1)+1][2(n+1)+1]}{6} \\
&= f(n+1), \text{ as required.}
\end{aligned} \tag{A.2}
$$

The case $f^{(3)}(n) = \sum_{k=1}^{n} k^3$ may be proved in exactly the same fashion, and I leave it as an exercise.

SECOND SOLUTION

If you notice that $\sum_{k=1}^{n} 1 = n$ and $\sum_{k=1}^{n} k = n(n+1)/2$, you might deduce the following pattern:[33]

$$
\begin{aligned}
\sum_{k=1}^{n} 1 &= n \\
\sum_{k=1}^{n} k &= \frac{n(n+1)}{2} \\
\sum_{k=1}^{n} k(k+1) &= \frac{n(n+1)(n+2)}{3} \\
\sum_{k=1}^{n} k(k+1)(k+2) &= \frac{n(n+1)(n+2)(n+3)}{4} \\
&\vdots
\end{aligned}
$$

These can each be proved easily using mathematical induction. For example, the third equality immediately above is easily proved true when $n = 1$ (both sides equal 2). To prove this third equality in general, we now need only show that increasing n by one on each side of the equality has the same incremental effect on both sides. That is,

[33] I thank David Maslen for suggesting this technique. Any errors are mine.

we need only show that the right-hand side evaluated at $n = (N+1)$ less the right-hand side evaluated at $n = N$ gives what would be the $(N+1)^{st}$ term in the summation on the left-hand side:

$$\begin{aligned}
& \frac{(N+1)(N+2)(N+3)}{3} - \frac{N(N+1)(N+2)}{3} \\
= & \frac{(N+1)(N+2)}{3}((N+3)-N) \\
= & \frac{(N+1)(N+2)}{3}(3) \\
= & (N+1)(N+2) \\
= & k(k+1)\Big|_{k=N+1} \quad \text{QED.}
\end{aligned}$$

The results we are interested in follow quite easily now because, for example, $\sum_{k=1}^{n} k^2 = \sum_{k=1}^{n} k(k+1) - \sum_{k=1}^{n} k$, and we have expressions for both the latter summations.

Answer 1.39: You know one of the eight balls is heavy. Compare one group of three to another group of three. You need only one more weighing—for a total of two weighings.

Answer 1.40: You win if you can place the last coin on the table and leave no space for me to place a further coin. A necessary condition is that the table be radially symmetric. That is, there must exist a central point on the table (at its centre of mass if the table is of uniform density and thickness) such that any line drawn upon the table passing through this central point is evenly bisected at this central point. Simple examples are a square, an ellipse, a rectangle, a circular disc, etcetera.

You should play first and place your first quarter at the centre of the (round, square, ...) table. You should make subsequent moves by imitating me: Place your quarter in the mirror image of my position when viewed looking through the central point. This ensures victory because if I can still place a coin on the table, then so can you.[34]

Although radial symmetry is necessary, it is not sufficient. The strategy does not necessarily work if the table is a regular shape but not a

[34] I thank Tim Hoel and Victor H. Lin for this elegant solution technique.

simply-connected one; for example, an annulus.[35] If the table is an annulus and the hole in the middle is bigger than a quarter, then the only change to your winning strategy is that you should let me go first.

Answer 1.41: This one has been popular since mid-1999. The numbers used and the situation described differ from question to question, but the general solution technique is always the same. Factor the product into all possible triplets: (1,2,18), (1,3,12), (1,4,9), (1,6,6), (2,2,9), (2,3,6), and (3,3,4). Which one is it? Well, Mary knows the sum, and these potential triplets sum to 21,16,14,13,13,11, and 10, respectively. Knowing the sum was not sufficient for Mary to pin down the triplet, so it must be a triplet with a non-unique sum: 13 in this case. This cuts down the candidates to (1,6,6) and (2,2,9). John says the eldest is dyslexic, so there must be an eldest (ignoring rubbish answers like one twin is 20 minutes older than the other). That just leaves (2,2,9).

Answer 1.42: Deciphering the optimal strategy is analogous to locating an optimal stock price exercise boundary for an American-style option. Calculating the expected payoff to the game, assuming the optimal strategy, is analogous to valuing an American-style option. Like valuing an American option, you have to work backward through a decision tree, calculating the expected payoffs to proceeding versus stopping at each node. For two, four, six, and eight cards, the expected payoff to the game is $\$\frac{1}{2}$, $\$\frac{2}{3}$, $\$\frac{17}{20}$, and \$1, respectively, when following the optimal strategy. The two-card game decision tree is a subtree of the four-card game decision tree, so later results can be appended to earlier ones. Stop reading here and try to replicate these numbers.

Let R and B denote the number of red and black cards, respectively, when you begin play ($R = B = 26$ in our case). Let r and b denote the number of red and black cards remaining in the deck at some intermediate stage of the game when you are trying to decide whether to take another card. You get $+1$ for each red card drawn and -1 for each black card drawn, so your current accumulated score is the

[35] An "annulus" is a disc with a hole in the centre—like a musical compact disc, for example. An annulus is path connected (any two points may be joined by a line), and is therefore connected (it cannot be split into two non-empty non-intersecting open sets), but it is not simply connected (which requires path-connectivity and that any loop may be shrunk continuously within the set).

number of reds drawn so far less the number of blacks drawn so far: $(R-r)-(B-b)$. The expected value of the game $V(r,b)$ is the current accumulated score plus the additional expected value, if any, remaining in the deck, assuming optimal play. With r and b cards remaining, denote this additional expected value as $E(r,b)$. Thus, the value of the game is $V(r,b)=(R-r)-(B-b)+E(r,b)$. Simple logic dictates that $E(r,b)$ is defined recursively as follows:[36]

$$E(r,b)=\begin{cases} 0, \text{ if } r=0 \\ r, \text{ if } b=0 \\ \max\left\{0, \frac{r}{r+b}[1+E(r-1,b)]+\frac{b}{r+b}[-1+E(r,b-1)]\right\}, \\ \text{otherwise.} \end{cases}$$

Table A.6 gives all the necessary information for you to solve the easier game in which there are four cards of each colour. It is worth noting that the recursive definition of $E(r,b)$, when seen in action in Table A.6, produces a complicated Pascal's Triangle type of calculation when working from the lower right to the upper left.[37] That is, $E(r,b)$ in each cell depends on $E(r,b)$ in the cells immediately to the right and below. As mentioned previously, in the two-, four-, six-, and eight-card games, the expected payoffs are $\$\frac{1}{2}$, $\$\frac{2}{3}$, $\$\frac{17}{20}$, and \$1, respectively, and these are visible on the leading diagonal in Table A.6.

If the additional remaining value in the deck when playing optimally is zero [i.e., $E(r,b)=0$], you are not indifferent about continuing. Rather, you want to quit immediately because in every case except one, $E(r,b)=0$ implies $\frac{r}{r+b}[1+E(r-1,b)]+\frac{b}{r+b}[-1+E(r,b-1)]$ is negative, and that the "max" function is being used in the recursive definition of $E(r,b)$. The only exception is when $(r,b)=(1,2)$, and even then $\frac{r}{r+b}[1+E(r-1,b)]+\frac{b}{r+b}[-1+E(r,b-1)]$ is zero and a risk-averse player would quit. When playing optimally, the last card drawn is always red. That is, you never pick a black card and then quit. The optimal score to quit at cannot be negative because you always have the safety net of a zero payoff for sure if you draw every card.

[36] I thank Paul Turner for solving this problem when it was posted as a challenge question on my web site. Any errors are mine.

[37] Pascal's Triangle has the following rows: [1], [1 1],[1 2 1], [1 3 3 1], [1 4 6 4 1], and so on. Apart from the 1's, each item is the sum of the two items above. The $(n+1)^{st}$ row gives the coefficients in the polynomial expansion of $(a+b)^n$.

0	(4,4)	1	(3,4)	2	(2,4)	3	(1,4)	4	(0,4)
1	$(\frac{4}{8}, \frac{4}{8})$	$\frac{12}{35}$	$(\frac{3}{7}, \frac{4}{7})$	0	$(\frac{2}{6}, \frac{4}{6})$	0	$(\frac{1}{5}, \frac{4}{5})$	0	$(\frac{0}{4}, \frac{4}{4})$
1	Y	$1\frac{12}{35}$	Y	2	$N\bullet$	3	NN	4	NN
-1	(4,3)	0	(3,3)	1	(2,3)	2	(1,3)	3	(0,3)
$1\frac{23}{35}$	$(\frac{4}{7}, \frac{3}{7})$	$\frac{17}{20}$	$(\frac{3}{6}, \frac{3}{6})$	$\frac{1}{5}$	$(\frac{2}{5}, \frac{3}{5})$	0	$(\frac{1}{4}, \frac{3}{4})$	0	$(\frac{0}{3}, \frac{3}{3})$
$\frac{23}{35}$	Y	$\frac{17}{20}$	Y	$\frac{6}{5}$	Y	2	$N\bullet$	3	NN
-2	(4,2)	-1	(3,2)	0	(2,2)	1	(1,2)	2	(0,2)
$2\frac{2}{5}$	$(\frac{4}{6}, \frac{2}{6})$	$1\frac{1}{2}$	$(\frac{3}{5}, \frac{2}{5})$	$\frac{2}{3}$	$(\frac{2}{4}, \frac{2}{4})$	0	$(\frac{1}{3}, \frac{2}{3})$	0	$(\frac{0}{2}, \frac{2}{2})$
$\frac{2}{5}$	Y	$\frac{1}{2}$	Y	$\frac{2}{3}$	Y	1	$N\bullet$	2	NN
-3	(4,1)	-2	(3,1)	-1	(2,1)	0	(1,1)	1	(0,1)
$3\frac{1}{5}$	$(\frac{4}{5}, \frac{1}{5})$	$2\frac{1}{4}$	$(\frac{3}{4}, \frac{1}{4})$	$1\frac{1}{3}$	$(\frac{2}{3}, \frac{1}{3})$	$\frac{1}{2}$	$(\frac{1}{2}, \frac{1}{2})$	0	$(\frac{0}{1}, \frac{1}{1})$
$\frac{1}{5}$	Y	$\frac{1}{4}$	Y	$\frac{1}{3}$	Y	$\frac{1}{2}$	Y	1	$N\bullet$
-4	(4,0)	-3	(3,0)	-2	(2,0)	-1	(1,0)	0	(0,0)
4	$(\frac{4}{4}, \frac{0}{4})$	3	$(\frac{3}{3}, \frac{0}{3})$	2	$(\frac{2}{2}, \frac{0}{2})$	1	$(\frac{1}{1}, \frac{0}{1})$	0	(0,0)
0	Y	0	Y	0	Y	0	Y	0	$N\bullet$

Table A.6: The Red/Black Card Game

Each cell below is laid out as $\boxed{\begin{array}{cc} \text{Accum. Score} & (r,b) \\ E(r,b) & (p_{red}, p_{black}) \\ V(r,b) & Y, N, \text{ or } NN \end{array}}$, where r and b are the number of red and black cards remaining in the deck, "Accum. Score" is the accumulated score so far (i.e., $(R-r)-(B-b)$, where $R=B=4$ in this case); p_{red} and p_{black} are the probabilities that the next card drawn will be red or black, respectively; $E(r,b)$ is the expected additional value remaining in the deck assuming optimal play; $V(r,b)$ is the expected payoff of the game, assuming you start in the top left cell (it is the sum of the two items above it); "Y" means *yes* you should continue playing; "$N\bullet$" means *no* you should halt (the bullet helps your eye see the boundary), and "NN" means *no* you should halt, but you should also note that it is *not* possible to get to this cell if you start with an even number of each colour card and play optimally.

The optimal score to quit at is a non-increasing step function of the number of black cards drawn (drawing red cards has no effect on the optimal score to quit at). Drawing black cards can lower the optimal score to quit at. In the eight-card game, the optimal stopping rule is: If you have turned over zero or one black card, then quit if you can get to a score of 2 without seeing another black card; if you have turned over two or three black cards, then quit if you can get to a score of 1 without seeing another black card; if you have turned over four black cards, then the best you can do is draw every card and get a payoff of 0. In the $2n$ card game with n red cards and n black cards, the expected payoffs are shown in Table A.7.[38]

$2n$	$r = b = n$	$V(r,b)$ (ratio)	$V(r,b)$ (decimal)
2	1	$\frac{1}{2}$	0.500000000000
4	2	$\frac{2}{3}$	0.666666666667
6	3	$\frac{17}{20}$	0.850000000000
8	4	$\frac{1}{1}$	1.000000000000
10	5	$\frac{47}{42}$	1.119047619048
12	6	$\frac{284}{231}$	1.229437229437
14	7	$\frac{4{,}583}{3{,}432}$	1.434110334110
$\vdots$	$\vdots$	$\vdots$	$\vdots$
52	26	$\frac{41{,}984{,}711{,}742{,}427}{15{,}997{,}372{,}030{,}584}$	2.624475548994

Table A.7: E(Payoff) in Red/Black Card Games ($2n$ cards, n red, n black) These expected payoffs are derived using the same rules used in the eight-card game. I have included the ratio form of the expected payoff in case anyone spots a simple pattern.

Answer 1.43: No, definitely not. You cannot tile the 62 squares with the dominoes. If you cannot see why, then go back and think again. This one is too good to waste by peeking at the answers—stop reading here and try again.

Each domino covers two squares that are side-by-side on the board. Each of these pairs of squares consists of a black and a white. As you place the dominoes, you cover the same number of black squares as white ones. However, the two squares that are off limits are the same colour (opposing corners on a chessboard must be the same colour). Thus, the number of white squares to be covered is not the same as the number of black, and the dominoes cannot cover them all.

[38] I thank David Maslen for the final ratio in the table. Any errors are mine.

Naoki Sato has supplied an answer to his follow up question. Imagine a closed path on the chessboard that passes through every square exactly once (moving horizontally and vertically, eventually returning to the original square). The two "X"s, unless adjacent, divide this path into two sections. Since one "X" is on black, and one is on white, the two sections each cover an even number of squares. They may thus be tiled using the dominoes. If the two "X"s are adjacent, the solution is obvious.

Answer 1.44: A prime number has no factors other than itself and 1. Thus, 4 is not prime because it has factors: (1,4), and (2,2). Drawing a number line might be a good way to explain this to an interviewer, but I will just use words.

1. A prime p bigger than 2 cannot be an integer multiple of 2, else it would not be prime. Thus, a prime bigger than 2 must be odd. Thus, $p-1$ is even. Thus, $p-1=2n$ for some positive integer n. Thus, $p=2n+1$.
2. A prime p bigger than 3 cannot be an integer multiple of 3, else it would not be prime. However, draw a number line and it must be that either $p-1$ or $p+1$ (but not both) is a multiple of 3. That is, p is 1 away from a multiple of 3, but we do not know in which direction. Thus, $p \pm 1 = 3m$ for some positive integer m, where $\pm$ means exactly one of $+$ or $-$, but not both. Thus, $p = 3m \pm 1$.
3. The question asks about p^2-1. From #1 we see that $p^2-1 = 4n^2+4n = 4n(n+1)$. One of n or $n+1$ must be even, and with that 4 there, we see that p^2-1 contains a factor of 8 (i.e., $2\times2\times2$).
4. From #2, we see that $p^2-1 = 9m^2 \pm 6m = 3m(m\pm 2)$. Thus, p^2-1 contains 3 as a factor.
5. If we picture p^2-1 factored out into all possible numbers of smallest possible size, then the results from #3 and #4 cannot overlap. That is, p^2-1 contains factors of $2\times2\times2$ and 3; thus, p^2-1 is an integer multiple of 24.

Answer 1.45: Let B be your bid. Let S be the true value of the firm. The density function of S equals unity for $0 \leq S \leq 1$, and zero otherwise.

Your payoff P is

$$P(S) = \begin{cases} 2S - B, & \text{if } B > S \\ 0, & \text{otherwise.} \end{cases}$$

The maximum post-bid firm value is 2, so you should bid no more than 2. You want to maximize $E[P(S)]$ with respect to choice of B in the interval $[0, 2]$. Your expected payoff is

$$\begin{aligned} E[P(S)] &= \int_{S=0}^{S=1} P(S) \cdot 1 \cdot dS \\ &= \int_{S=0}^{S=\min(B,1)} (2S - B)dS \\ &= (S^2 - BS)\Big|_{S=0}^{S=\min(B,1)} \\ &= \begin{cases} 0, & \text{if } B \leq 1 \\ 1 - B, & \text{if } B > 1, \end{cases} \end{aligned}$$

so you should bid less than or equal to 1 and expect to break even.

Answer 1.46: What is going to happen if you light both ends simultaneously? The two fizzing sparking flames are going to burn toward each other and meet. When they meet 60 seconds worth of fuse has been burnt in two sections that each took the same amount of time. How much time? It has to be exactly 30 seconds because they both took the same time, and these times add to 60 seconds. Of course, you have to bend the fuse so that you can light both ends simultaneously and when they meet it probably won't be in the centre of the fuse.

Answer 1.47: I first heard the S-E-N problem in 2002. If your answer is "none" or "one," then go back and think again. There are, in fact, an uncountably infinite number of starting points that solve this problem.

First of all, you could start at the north pole. On the middle leg of your walk you would always be one mile south of the north pole, so the final leg would put you back where you started. Second, if you start at a point close to the south pole but one mile north of a line of latitude having circumference one mile, then the middle leg of your walk begins and ends in the same spot; the final leg takes you back to

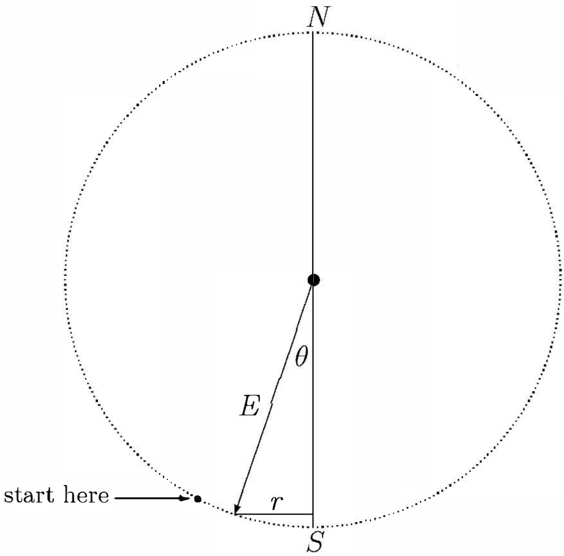


Figure A.5: S-E-N Problem: The Earth

The Earth is a perfect sphere with radius E. You start your trek one mile north of a line of latitude having circumference $1/n$ miles, and radius r miles (so $2\pi r = 1/n$). You must start a distance of $1 + E \cdot \arcsin \frac{1}{2\pi n E}$ miles from the south pole—see Answer 1.47.

your starting point. There are infinitely many such starting points on the line of latitude that is one mile north of the line of latitude having circumference one mile.

Similarly, if you start slightly further south, at a point one mile north of a line of latitude having circumference one-half mile, then the middle leg of your walk begins and ends in the same spot, and the final leg returns you to your starting point.

More generally, if you begin on a line of latitude one mile north of a line of latitude having circumference $1/n$ miles, then you will walk one mile south, circle the line of latitude n times, and return to your starting point.

In the latter case, how far is your starting point from the south pole? Well, assume the Earth is perfectly spherical, and let E be its radius. Let r be the radius of the line of latitude having circumference $1/n$ miles, so, $2\pi r = 1/n$. A simple sketch shows that the angle θ between the axis of the Earth, and a line drawn from the centre of the Earth to any point on the line of latitude having circumference $1/n$ miles, satisfies $\sin\theta = r/E$ (see Figure A.5 and the trig' review on page 97). Thus, the arc length from the pole to this line of latitude is the fraction $\frac{\arcsin\frac{r}{E}}{2\pi}$ of the circumference of the Earth, $2\pi E$. That is, the arc length is $E \cdot \arcsin\frac{1}{2\pi n E}$ (using $r = 1/(2\pi n)$). You start one mile north of this, at a distance of $1 + E \cdot \arcsin\frac{1}{2\pi n E}$ miles from the south pole.

Answer 1.48: The king should take one coin from bag one, two coins from bag two, three coins from bag three, and so on, finishing with ten coins from bag ten. Place this collection on the weighing device, and look for the discrepancy from $\sum_{i=1}^{10} i$ ounces. If the actual wieght is 0.40 ounces short, for example, then bag four is light, and collector four is the cheat.

Answer 1.49: Snap the bar into pieces that are one, two, and four parts long, respectively. On day one, give him one part. On day two, exchange your two parts for his one. On day three, give him back the one part. On day four, exchange four parts for his three. On day five, give him one more part. On day six, exchange your two parts for his one. One day seven, give him back the one part.

Answer 1.50: $100! = 100 \times 99 \times 98 \times \ldots 3 \times 2 \times 1$. Factor each number and count how many supply a 5. Combine the 5's with all the 2's going spare to get the 10's that give 0's at the end of 100!. The following supply a 5 (or two, as indicated): 5, 10, 15, 20, 25(2), 30, 35, 40, 45, 50(2), 55, 60, 65, 70, 75(2), 80, 85, 90, 95, 100(2). This gives the 24 zeroes at the end of 100!:[39]

			933	26215
44394	41526	81699	23885	62667
00490	71596	82643	81621	46859
29638	95217	59999	32299	15608
94146	39761	56518	28625	36979
20827	22375	82511	85210	91686
40000	00000	00000	00000	00000

[39] Type `vpa factorial(100) 158` in MATLAB; vpa is variable precision arithmetic.

Answer 1.51: Let us attack the mirror problem in stages.

Your Perspective, No Rotations: Put your wristwatch on your left wrist and stand facing a mirror with your arms held out as though you are being crucified (it is a tough interview remember). Your reflected self's wristwatch-bearing arm is pointing the *same* direction as yours. Your wristwatch is to the left of your head, and your reflected self's wristwatch is also to *your* left of his or her head. *There has been no flipping of left for right.* Similarly, if your head is pointing up, then your reflected self's head is also pointing up, and *there has been no flipping of up for down.*

Perhaps this is clearer if you write a sentence on a transparent plastic sheet, and hold the sheet in front of your body, as though there is no mirror at all and you are simply reading what you have just written. Now look in the mirror. The reflection of your sheet in the mirror is *not* reversed. That is, the left-most word is still left most, the right-most word is still right most, and you can still read the reflected image from left to right.

Viewed from your perspective, everything about you that is left, right, up, or down is still left, right, up, or down, respectively, in your reflected image. There is thus *no* flipping of left for right or up for down. What *has* flipped is that if you are facing east, then your reflection is facing west. It does not matter for the sentence written on the transparent sheet, because it has no depth. It does matter for you, because your reflected nose is facing the opposite direction.

Your Perspective, Rotation of Yourself: If your interviewer suggests that there really is a flipping left for right of your reflected self, but not up for down, then this requires an implicit rotation of your perspective about a vertical axis, to place your right-handed self into the imagined boots of your reflected self who is left handed and standing on the other side of the mirror. To get a one-to-one mapping (so the boots fit), however, you still need to flip yourself left for right (without changing the direction in which you are facing) because your wristwatch is on your left wrist—the opposite of your reflected self. Had you instead rotated yourself about a horizontal axis, and then attempted to place yourself into the imagined boots of your reflected self, you would find your noses pointing the same direction and your wristwatches on the same sides, but your head would be between the feet of your reflected self, and to get into those imagined boots, you would still need to flip yourself up for down (without changing the direction in which you are facing).

The fact that neither a rotation about a horizontal nor a vertical axis suffices to place you into the imagined boots of your reflected self, serves to confirm my earlier assertion: There has not been a flipping of left for right, or up for down, but rather, a flipping in the direction of the depth. If your interviewer firmly believes that a mirror does flip left for right, then he or she is predisposed toward rotation about a vertical axis (something many of us do every day), and has not thought through the consequences of the attempted one-to-one mapping.

Answer 1.52: Yes, it can be done, in theory if not in practice. If you are stuck and looking for a hint, think about inverting a condom and covering it with another.

Let us label the condoms $C1$, and $C2$, and the men $M1$, $M2$, and $M3$. $M1$ wears $C1$ with $C2$ placed over it. $M2$ then uses $C2$, which is still clean inside. $M3$ then wears $C1$ inverted ($C1$'s outside, you will recall, was kept clean by $C2$), and places the twice-used $C2$ over it. Don't try this at home.

Answer 1.53: No, of course not. Replace the word "prime" with any other word, and the answer is still no. If they are consecutive, then by definition there are none of them in between!

Appendix B

Derivatives Answers

This appendix contains answers to the questions posed in Chapter 2.

Answer 2.1: Most students incorrectly deduce that the call option is worthless. If this is your conclusion, stop looking at the answers and go back and think again. You missed the point. Many students think that zero volatility means the stock price is going nowhere. However, volatility of returns is, by definition, the average deviation from expected returns. It follows that zero volatility means the stock price drifts up at the expected return on the stock with no deviations from this path.

With no volatility, the stock is riskless. In the absence of arbitrage opportunities, the stock must offer an expected return equal to the riskless rate. This is true in both the real world and the theoretical risk-neutral world. This result (expected return equals r) is very strange in the real world—stocks normally offer higher returns. Do note, however, that *all* stocks in the risk-neutral world have expected return equal to the riskless rate. Although I discuss option pricing in the risk-neutral world, the same arguments apply in the real world in the no-volatility case.[1]

The required rate of return on the stock is the riskless interest rate. It

[1] If option pricing is done using real world probabilities rather than risk-neutral ones, then the discount rate on the option is a path-dependent random variable that changes as the stock price changes (Arnold and Crack [2003]). Such a model allows inference of real world probabilities of a real option project being successful, a financial option finishing in-the-money, or a corporate bond defaulting.

follows that with no volatility, the stock price rises to about \$105 for sure.[2] That is, the option finishes in-the-money for sure and is thus riskless. The discounted expected payoff is thus roughly $\frac{(\$105-\$100)}{1.05} = \frac{5}{1.05}$. At 5%, you lose about five cents on every dollar when you discount over a year. The discounted expected payoff is, therefore, about \$4.75, and this is the call value.

This is a good place to mention an often overlooked connection between options and forwards. Suppose that $S(t)$ is the price today of a stock that pays no dividends. Let r denote the continuously compounded interest rate per annum. Then a fair price for delivery of the stock at time T is: $F = S(t)e^{r(T-t)}$. In the absence of volatility, the expected time-T stock price is just the forward price. Once volatility is introduced into the picture, the distribution of terminal stock price $S(T)$ becomes spread out. However, the mean of the (risk-neutral) distribution of $S(T)$ is unchanged, and this mean equals the forward price, which is also unchanged: $S(t)e^{r(T-t)}$. That is, the expected time-T stock price in the risk-neutral world is just the forward price.

Back to the option at hand: With no volatility, the value of the option at time t is just the discounted expected time-T payoff in a risk-neutral world:

$$\begin{aligned} c(t) &= e^{-r(T-t)}\max(S(T)-X,0) \\ &= e^{-r(T-t)}\max(S(t)e^{r(T-t)}-X,0) \\ &= e^{-r(T-t)}\max(F-X,0), \end{aligned}$$

where $F = S(t)e^{r(T-t)}$ is the forward price for the stock, and X is the strike price. If follows that the option has value if and only if the forward price exceeds the option's strike.

How do you hedge this? If $F > X$, the option will finish in-the-money for sure, so you need a delta of $+1$. If $F \leq X$, the option will die worthless for sure, so you need a delta of 0 (who would buy the option in this case anyway?).

Answer 2.2: The gamma of an option is the rate of change of its delta, Δ, with respect to stock price—denoted Γ. Option gamma is also called

[2]If the interest rate is an effective (i.e., simple) rate, then this is exact; if it is continuously compounded, this is an approximation.

"curvature," or "convexity." Gamma is non-negative for standard puts and calls (their deltas rise with increasing S). From put-call parity, it should be clear that the gamma of a call is the same as the gamma of a put.

Option value "decays" toward kinked final payoff as expiration approaches (see Figure B.1—first panel). This time decay is called "theta." Although theta is negative for plain vanilla European calls, and for American puts and calls, a deep in-the-money European put decays *upward* in value (i.e., it has positive theta).

Theta is large and negative for at-the-money options, and it increases in magnitude as maturity approaches. Theta and gamma are typically of opposing signs, so large negative theta goes hand-in-hand with large positive gamma. That is, shortening maturity accelerates at-the-money option prices towards the kink and also gives more curvature (i.e., gamma) in the plot of option value as a function of stock price (see Figure B.1—third panel).

The maturity/gamma relationship is reversed away from the strike price. If a call is deep in-the-money, then $\Delta \rightarrow 1$, as expiration approaches (for a deep in-the-money *put*, $\Delta \rightarrow -1$ as expiration approaches). Thus, short maturity calls or puts that are deep in-the-money have deltas that do not vary much as S changes. With little variation in delta, the gamma is close to zero. If an option is instead deep *out*-of-the-money, then its gamma is also close to zero because its delta is close to zero with little variation across S. It follows that for away-from-the-money standard options, shorter maturity implies lower gamma for both puts and calls (see Figure B.1—third panel).

The gamma (i.e., convexity) for a standard European call on a stock that pays a continuous dividend at rate ρ is given as follows:

$$\Gamma(t) \equiv \frac{\partial^2 c(t)}{\partial S(t)^2} = \frac{e^{-\rho(T-t)-\frac{1}{2}d_1^2}}{S(t)\sigma\sqrt{2\pi(T-t)}},$$

where

$$d_1 = \frac{\ln\left(\frac{S(t)}{X}\right) + (r - \rho + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

With $(T-t) > 0$, the formula for Γ shows that as $S(t) \rightarrow \infty$, the numerator goes to zero (because $d_1 \rightarrow \infty$), and the denominator goes to infinity. Both limits have the same effect on Γ, pushing it to zero. Similarly, if $(T-t)>0$, then as $S(t)\rightarrow 0$, $d_1^2\rightarrow\infty$ so the numerator goes to zero again. However, having S in the denominator pushes Γ in the opposite direction as $S\rightarrow 0$. The exponentiation of d_1^2 in the numerator is much more powerful than the linearity of S in the denominator, so the ratio, Γ, is forced to zero as $S\rightarrow 0$.

If the option is exactly at-the-money [i.e., $S(t)=X$], then as maturity approaches, you have a knife-edge singularity. You get $d_1\rightarrow 0$, so the numerator of Γ goes to **1**. However, the denominator tends to zero, so the ratio, Γ, blows up. That is, you get "infinite gamma" at the kink as maturity approaches.

Infinite gamma means the sensitivity of delta to small changes in price of the underlying is infinite. This means that the delta can jump from one-half up to one, or down to zero with just a hair's breadth move in the stock price. In this knife-edge scenario, any delta-hedge that you establish is extremely sensitive to a move in the underlying—you are not hedged.

If you try gamma-hedging (adding traded options to your delta-hedge to replicate the convexity of the derivative), you will need many traded options in your hedge portfolio, and it may become difficult to manage the position.[3] Your problems are similar (but much worse) if hedging barrier options (i.e., "knock-outs") as the price of the underlying approaches the knock-out barrier. The problem is worse near a knock-out's barrier than near a standard call's kink. This is because the knock-out's delta can jump from one to zero whereas the standard call's delta jumps only from one-half to zero, or one-half to one.

For American-style options (or more complicated Europeans), you have no closed form formulae. You will probably have to calculate Γ using numerical techniques.

Answer 2.3: Most students incorrectly deduce that the derivative security is worth \$1. If you got this answer, go back right now and think some more. I present two solution techniques: the first uses standard no-

[3]In practice, even with a day left to maturity, although the gamma can be quite large, you might need only 10 three-month calls to replicate the convexity of a standard call with one day to maturity—we are not talking infinity here.

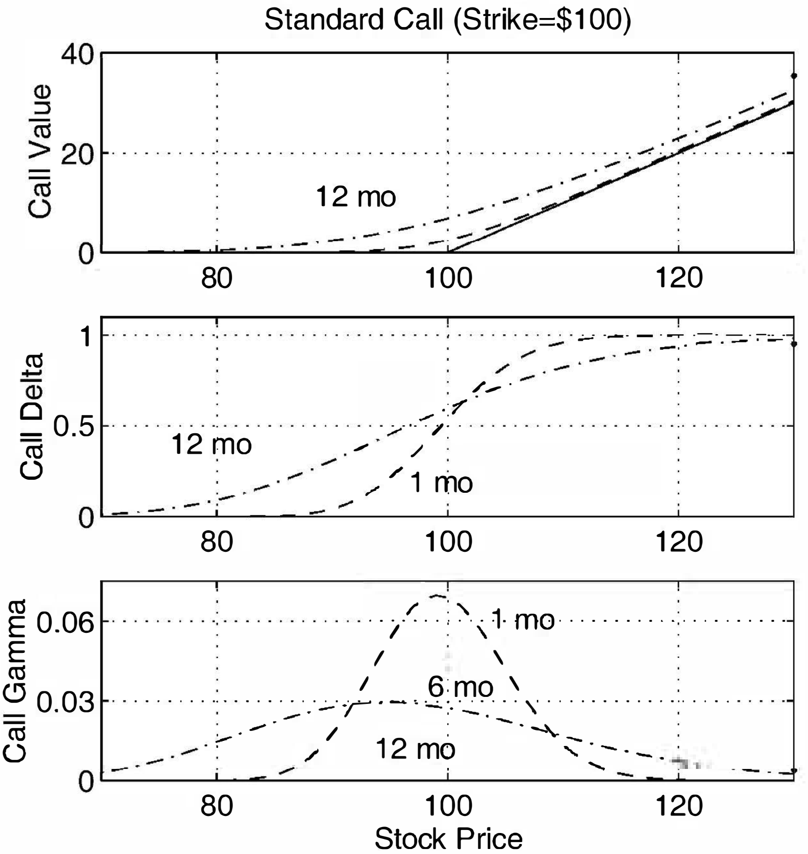


Figure B.1: Standard Call: Price, Delta, and Gamma.
For maturities of 12 months "····", six months "·—·—", and one month "————", the call price, delta, and gamma are plotted as a function of price of underlying (see Answer 2.2).

arbitrage arguments; the second is more advanced and uses partial differential equations (PDE's).[4]

FIRST SOLUTION

You are an investment banker. Assume there exists a derivative security that promises one dollar when IBM hits \$100 for the first time. If this security is marketable at *more* than \$0.75, then you should issue 100 of them and use \$75 of the proceeds to buy one share of IBM. If IBM ever hits \$100, sell the stock and pay \$1 to each security holder as contracted. You sell the securities, perfectly hedge them, and still have money in your pocket. By no-arbitrage, the security cannot sell for more than \$0.75.

The converse is that if this security costs *less* than \$0.75, you should buy 100 of these securities financed by a short position in one IBM share. For this to establish \$0.75 as a lower bound on the security price (and, therefore, to pinpoint the price at \$0.75—the solution given to the interviewee by the Wall Street firm), you need to assume that you can roll over a short position *indefinitely*. This assumption seems reasonable for moderate amounts of capital. However, it is not clear to me that this is a reasonable interpretation of "ignore any short sale restrictions" when larger quantities of capital are involved. As one colleague said to me: "If it were possible to short forever, I'd short stocks with face value of a billion dollars, consume the billion, and roll over my short position forever." This seems to be an arbitrage opportunity.

We conclude that \$0.75 is a clear upper bound by no-arbitrage, and thus \$1 cannot be the correct answer. Whether or not \$0.75 is also a lower bound is arguable (but it seems to make sense for moderate amounts of capital). The second solution technique also establishes \$0.75 as the value of the security.

SECOND SOLUTION

This technique is more advanced and may be beyond the average student.[5] The derivative value V must satisfy the Black-Scholes PDE

[4] If you want a good introductory book on PDE's, I recommend Farlow (1993). I loved this book when I was a student. I still find it a breath of fresh air compared to my other math books.

[5] I thank Alan J. Marcus for suggesting this type of solution technique. I am responsible

(Wilmott et al. [1993]):

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

The boundary conditions that make sense for $V(S,t)$ are:

$$\begin{aligned} V(S=100,\ \text{any } s>t) &= \$1, \quad \text{and} \\ V(S=0,\ \text{any } s>t) &= \$0. \end{aligned}$$

Let us simplify our lives by searching initially for a solution that is affine in S: $V(S,t) = kS(t) + l$, for some constants k and l.[6]

The two boundary conditions imply that

$$\begin{aligned} k \times \$100 + l &= \$1, \quad \text{and} \\ k \times \$0 + l &= \$0. \end{aligned}$$

From these we deduce that $k = \frac{1}{100}$, and $l = 0$. The functional form $V(S,t) = \frac{1}{100}S(t)$ satisfies the Black-Scholes PDE and the two boundary conditions and is thus the derivative value. In the special case where $S(t) = \$75$, we get $V = \$0.75$, as for the first technique.

Answer 2.4: The key here is the shape of the risk-neutral distribution of final stock price, $S(T)$, conditional on current stock price, $S(t)$. Many students mistakenly assume the distribution of final stock price to be both symmetric and Normal. The distribution is in fact Lognormal.

The Lognormal distribution is "right skewed," also known as "positively skewed." It looks as though its top has been shoved from the right while keeping its base fixed.

If we start with $S(t) = X$, and $r = 0$, then the skewness in the distribution of $S(T)$ means that the final stock price is more likely to end up below the strike than above it.[7] The call has bigger potential

for any errors.

[6] An affine function involves both a linear portion, kS, and a constant, l. On a two-dimensional plot, a linear function goes through the origin; whereas an affine function may have a non-zero intercept.

[7] With $r = 0$, the median of the risk-neutral distribution of $S(T)$ conditional on $S(t)$ is $S(t)\ e^{(r-\frac{1}{2}\sigma^2)(T-t)} = S(t)\ e^{-\frac{1}{2}\sigma^2(T-t)} < S(t)$. The option is struck at-the-money [i.e., $S(t) = X$], so the median is below the strike.

payoffs than the put but (because of skewness) lower probabilities of achieving them. The put has smaller potential payoffs than the call but (because of skewness) higher probabilities of achieving them. The bigger payoffs and lower probabilities for the call exactly match the smaller payoffs and higher probabilities for the put. It follows that the put and call have the same risk-neutral expected payoff and, therefore, have the same value. It is straightforward to confirm this equality of values using put-call parity.

Answer 2.5: Without dividends, the standard Black and Scholes (1973) pricing formula for the European call option is given by

$$\begin{aligned} c(t) &= S(t)N(d_1) - e^{-r(T-t)}XN(d_2), \quad \text{where} \\ d_1 &= \frac{\ln\left(\frac{S(t)}{X}\right) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \quad \text{and} \\ d_2 &= d_1 - \sigma\sqrt{T-t}. \end{aligned}$$

The option's "delta" is given by $\frac{\partial c(t)}{\partial S(t)} = N(d_1)$. With the option struck at-the-money, $S(t) = X$, and thus, $\ln\left(\frac{S(t)}{X}\right) = 0$ [remember that $\ln(1) = 0$]. All other terms in d_1 are positive. Therefore, $d_1 > 0$, and $N(d_1) > 0.5$ (remember that $N(0) = 0.5$ and $N(\cdot)$ is an increasing function of its argument). Thus, an at-the-money option on a non-dividend-paying stock always has a delta slightly greater than one-half.

Answer 2.6: With continuous dividends at rate ρ, the standard Black-Scholes pricing formula for the European call option is given by[8]

$$\begin{aligned} c(t) &= S(t)e^{-\rho(T-t)}N(d_1) - e^{-r(T-t)}XN(d_2), \quad \text{where} \\ d_1 &= \frac{\ln\left(\frac{S(t)}{X}\right) + (r - \rho + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \quad \text{and} \\ d_2 &= d_1 - \sigma\sqrt{T-t}. \end{aligned}$$

The option's "delta" is given by $\frac{\partial c(t)}{\partial S(t)} = e^{-\rho(T-t)}N(d_1)$. With the option struck at-the-money, $S(t) = X$, and thus, $\ln\left(\frac{S(t)}{X}\right) = 0$ [remember that

[8]This extension of Black-Scholes is due originally to Merton (1973, Footnote 62). Note, however, that his original formula is not entirely correct (he omits the dependence of d_1 and d_2 on ρ).

$\ln(1) = 0]$. This, combined with $r > \rho$ yields $d_1 > 0$, and thus $N(d_1) > 0.5$. The naive answer is that $N(d_1) > 0.5$ and that this is the delta—forgetting that $e^{-\rho(T-t)}$ premultiplies $N(d_1)$ in the continuous-dividend case. In general, you cannot tell whether the delta, $e^{-\rho(T-t)}N(d_1)$, is larger or smaller than 0.5: it depends upon the size of σ^2. However, in this particular case, $\rho = 0.03$ is so small that $\Delta > 0.5$ for any σ.

Answer 2.7: Almost every student I have asked has got the answer to this one backwards at first. This is unfortunate, because it is a commonly asked question. Think it through carefully before answering, and do not get caught out. The delta is the number of units of stock in the replicating portfolio. Other things being equal, the delta falls with a fall in stock price. However, you are long the call and short the replicating portfolio. This means that the number of units of stock you are short has to fall. So, you must borrow more money and buy back some stock.

If you got it wrong, think about it as follows. Ask yourself how the replicating portfolio changes (e.g., delta falls, so less stock is needed in the replicating portfolio). Then ask yourself whether you are long or short the replicating portfolio (you are short here). If you are short, be sure to reverse the implications (less stock shorted means you must borrow to buy some back).

Answer 2.8: With the standard European call, you have a simple closed-form expression for the option's delta. For example, (under the Black and Scholes [1973] assumptions) the delta of a standard European call on a non-dividend-paying asset is equal to $N(d_1)$ where

$$d_1 = \frac{\ln\left(\frac{S(t)}{X}\right) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

See Answer 2.6 for the delta in the case where there are continuous dividends at rate ρ.

Unfortunately, only a few known options have closed-form pricing formulae. For exotic options with no closed-form pricing formula, you need a pricing algorithm. This may be a Monte-Carlo simulation,[9] a

[9]As an introduction to exotic options and Monte-Carlo techniques, I recommend the Monte-Carlo chapter of my book *Basic Black-Scholes* (Crack [2004]). The earliest Monte-

binomial tree, or a numerical PDE solution routine. By varying the input value of the current level of underlying, you can use the pricing algorithm to calculate a numerical derivative of price with respect to level of underlying; i.e., the delta. All you are doing is using the computer rather than the calculus to tell you how the option price changes with a change in stock price.

Answer 2.9: The pricing formula for the standard Black-Scholes European call option on a non-dividend-paying stock is:

$$c(t) = S(t)N(d_1) - e^{-r(T-t)}XN(d_2),$$

where d_1 and d_2 are as previously defined. $N(d_1)$ is the option's "delta," sometimes denoted "Δ." $\Delta = N(d_1)$ is the same thing as the partial derivative of call price with respect to underlying: $\frac{\partial c(t)}{\partial S(t)}$. It measures how the call price changes per unit change in the price of the underlying.

Another interpretation of the terms involves a replicating portfolio. $\Delta = N(d_1)$ is the number of units of stock you must hold in a continuously rebalanced portfolio that replicates the payoff to the call. The term $e^{-r(T-t)}XN(d_2)$ is the value of the borrowing (or a short position in bonds) required in a continuously rebalanced portfolio that replicates the payoff to the call. The value of the borrowing in the replicating portfolio is always less than or equal to the value of the replicating portfolio's long position in the stock. This is equivalent to stating that the call has non-negative value.

Another interpretation of the terms involves expected benefits and expected costs to owning the call. The term $S(t)N(d_1)$ is the discounted value of the expected *benefit* of owning the option (expectations taken under a risk-neutral probability measure). Why is the

Carlo reference I know of in option pricing is Boyle (1977). Boyle also gives techniques for accelerating the convergence of Monte-Carlo estimation and some references to the mathematics literature (see Hull [1997, pp365–368] for other techniques). For background information on the development of exotics and the players in the market, see Fraser (1993); for a slightly higher-level than Hunter and Stowe (1992), see Ritchken, Sankarasubramanian, and Vijh (1993) or Hull (1997); at a slightly higher level still, see Goldman, Sosin, and Gatto (1979) or Conze and Viswanathan (1991). Note that the value of a look-back option to buy at the minimum or sell at the maximum might arguably be considered an upper bound on the value of market timing skills—see Goldman, Sosin, and Shepp (1979) for more details.

$N(d_1)$ there? Well, $N(\cdot)$ is a cumulative density function, so it must be that $N(d_1) \leq 1$. This in turn implies that $S(t)N(d_1) \leq S(t)$. This is because the future benefit of owning the option is $S(T)$ if the option finishes in-the-money and zero if it finishes out-of-the-money (or "under water"). This benefit is strictly dominated by a long position in the stock (a position that returns $S(T)$ regardless of whether the option is in- or out-of-the-money and costs $S(t)$ now). It follows that you value the benefit from the call at less than the long position in the stock, $S(t)N(d_1) \leq S(t)$. It is for this reason that the $N(d_1)$ term multiplies the $S(t)$ term.[10]

The term $e^{-r(T-t)}XN(d_2)$ is the discounted value of the expected *cost* of owning the option (with expectations taken under a risk-neutral probability measure). You can see all the components of the discounted expected value as follows: $N(d_2)$ is the (risk-neutral) probability that the call option finishes in-the-money (see extended discussion in Crack [2004]); X is your cost if it does; and $e^{-r(T-t)}$ is the discounting factor.

Here is a summary of the foregoing paragraphs (where "$P(in)$" denotes the risk-neutral probability that the call finishes in-the-money):

$$c(t) = \overbrace{S(t)\underbrace{N(d_1)}_{\Delta}}^{\text{stock position \& benefit}} \overbrace{\underbrace{-e^{-r(T-t)}X\underbrace{N(d_2)}_{P(in)}}_{\text{borrowing \& cost}}}^{\text{bond position}}$$

The value of the standard European put on a non-dividend-paying stock may now be deduced. The present value of the *benefit* of owning the put is $e^{-r(T-t)}X[1-N(d_2)]$, where $[1-N(d_2)]$ is the (risk-neutral) probability that the put option finishes in-the-money (i.e., the call finishes out-of-the-money), X is your payoff if it does, and $e^{-r(T-t)}$ is the discounting factor.

[10]The first term is $S(t)N(d_1) = e^{-r(T-t)}E^*[S(T)\mathcal{I}_{S(T)>X}|S(t)]$, where E^* denotes expectation taken with respect to the risk-neutral probability measure, and $\mathcal{I}_{S(T)>X}$ is as given in Equation B.1.

$$\mathcal{I}_{S(T)>X} = \begin{cases} 1 \text{ if } S(T) > X, \\ 0 \text{ if } S(T) \leq X. \end{cases} \tag{B.1}$$

The present value of the *cost* of owning the put option is $S(t)[1-N(d_1)]$. There are two probabilistic interpretations of $N(d_1)$, each under a competing martingale measure (see Crack [2004]).

Using the property that $[1 - N(z)] = N(-z)$, the value of the put option must be

$$p(t) = e^{-r(T-t)}XN(-d_2) - S(t)N(-d_1),$$

where d_1 and d_2 are as already defined for the call.

Put-call parity says that

$$S(t) + p(t) = c(t) + Xe^{-r(T-t)} + D.$$

If you plug in $c(t) = S(t)N(d_1) - e^{-r(T-t)}XN(d_2)$, and $D = 0$, you do indeed get that $p(t) = e^{-r(T-t)}XN(-d_2) - S(t)N(-d_1)$, as deduced above.

Answer 2.10: Questions about a "digital option" or "binary option" are quite common. The digital "cash-or-nothing" option that pays H if $S(T) > X$ has a value of $He^{-r(T-t)}N(d_2)$. This is simply the discounted (risk-neutral) expected payoff to the option: $N(d_2)$ is the (risk-neutral) probability that the option finishes in-the-money; H is the payoff if it does; and $e^{-r(T-t)}$ is the discounting factor. H is sometimes called the "bet." If H is chosen to equal the strike of the standard Black-Scholes option, then the cash-or-nothing option has the same value as the second term in the Black-Scholes formula: $e^{-r(T-t)}XN(d_2)$.

The first term in the Black-Scholes formula, $S(t)N(d_1)$, is the value of a long position in a digital "asset-or-nothing" option. A long position in the asset-or-nothing option, combined with a short position in the cash-or-nothing option, replicates the payoff to the European call—and, therefore, has the same value (you should draw the payoff diagrams to verify this).[11]

Be sure to see Question 2.11 and Answer 2.11 for more details on the binary option.

[11] As an aside, you might like to note that the payoff to the European call may also be replicated by using barrier options: you need a "knock-out" call option plus a "knock-in" call option.

Answer 2.11: I look at this intuitively first and then more rigorously. Intuitively, if the digital "cash-or-nothing" option is deep in-the-money, you are just waiting for your fixed cash payoff, and increases in volatility can only decrease your payoff. If you are deep out-of-the-money, you are expecting nothing, and increases in volatility can only increase your payoff. If $c(t)$ is the price of the digital cash-or-nothing option, then somewhere around the at-the-money position, the sign of $\frac{\partial c(t)}{\partial \sigma^2}$ must change.

Rigorously, if $c(t)$ is the price of the digital cash-or-nothing option, then direct calculation (under Black-Scholes assumptions) shows that

$$\frac{\partial c(t)}{\partial \sigma^2} > 0 \text{ if and only if } S(t) < Xe^{-(r+\frac{\sigma^2}{2})(T-t)}.$$

Another (equivalent) way of looking at this is that $\frac{\partial c(t)}{\partial \sigma^2} > 0$ if and only if the probability of finishing in-the-money increases with an increase in σ^2, and this is so if and only if $S(t) < Xe^{-(r+\frac{\sigma^2}{2})(T-t)}$.

Figure B.2 (on page 129) shows $\frac{\partial \text{ CALL PRICE}}{\partial \sigma^2}$ for the asset-or-nothing digital option, the cash-or-nothing digital option, and the standard call (all options are European). The price of the standard call is just the difference between the prices of the asset-or-nothing digital option and the cash-or-nothing digital option. Differentiation is a linear operation, so the sensitivity of the standard call to volatility is just the difference between the sensitivity of the asset-or-nothing digital option and the cash-or-nothing digital option.

It is clear from Figure B.2 that the price of the standard call is increasing in volatility. This should come as no surprise. A call option is an insurance policy. It puts a floor on your losses. When there is more risk about, the premium (i.e., call price) should be higher. In the same way, you should be happy to pay more for fire insurance if you find out that your next-door neighbour is an arsonist. See Chance (1994) for more details on the sensitivity of option value to the various input parameters.

For the cash-or-nothing, the boundary on the sign of $\frac{\partial c(t)}{\partial \sigma^2}$ is always slightly less than X (see Figure B.2 for a clear illustration). Thus, if you are in-the-money, or at-the-money, more volatility is always bad;

if you are very slightly out-of-the-money, more volatility is still bad [when $Xe^{-(r+\frac{\sigma^2}{2})(T-t)} < S(t) \leq X$]; if you are well out-of-the-money, more volatility is always good [when $S(t) < Xe^{-(r+\frac{\sigma^2}{2})(T-t)}$]. This differs from the standard European call option for which $\frac{\partial c(t)}{\partial \sigma^2}$ is always non-negative (see Figure B.2).

You might ask why the boundary on the sign of $\frac{\partial c(t)}{\partial \sigma^2}$ is always slightly less than X, rather than exactly at X. The relationship between volatility and skewness in the (Lognormal) distribution of final stock price is where the explanation lies. There are two forces at work: First, an increase in σ^2 tends to spread out the distribution of $S(T)$, putting more probability weight into the tails; second, increasing σ^2 drags down the median of the distribution, tending to pull probability weight leftward and out of the right tail, thus increasing the skewness.[12] If the strike price is at or below the median of $S(T)$ (so the option is in-the-money, or very slightly out-of-the-money), then both forces push probability mass leftward, increasing the likelihood of finishing out-of-the-money. However, if the strike price is far above the median of $S(T)$ (so the option is far out-of-the-money), then the increasing spread of the distribution dominates the leftward move of the median, and the probability of finishing in-the-money increases with increasing σ^2. For the forces to be balanced, the option must be struck above the median of the distribution of $S(T)$. The strike price that just balances the influence of both forces is $X = S(t)e^{(r+\frac{\sigma^2}{2})(T-t)}$. At this strike, the option is insensitive to instantaneous changes in σ^2 and it is slightly out-of-the-money: $S(t) = Xe^{-(r+\frac{\sigma^2}{2})(T-t)}$.

Answer 2.12: The naive and time-consuming way to find the delta for the knock-out option (or "barrier option") is to differentiate the closed-form pricing formula for the down-and-out, find $\frac{\partial C}{\partial S}$, and compare it to the same quantity for the standard call.[13] It is more elegant to use com-

[12]The mean of the risk-neutral distribution of $S(T)$ conditional on $S(t)$ is $S(t)e^{r(T-t)}$; the median is $S(t)e^{\left(r-\frac{1}{2}\sigma^2\right)(T-t)}$.

[13]The closed-form valuation formula for the down-and-out, together with discussion, is in Merton ([1973, pp175–76]; [1992, p302]). It takes around 15 minutes to differentiate it by hand carefully and about the same time to program the numerical derivative in MATLAB. The down-and-out option was introduced by Gerard Snyder (1969). See his paper for a butchers at the operations of the options markets in the late sixties.

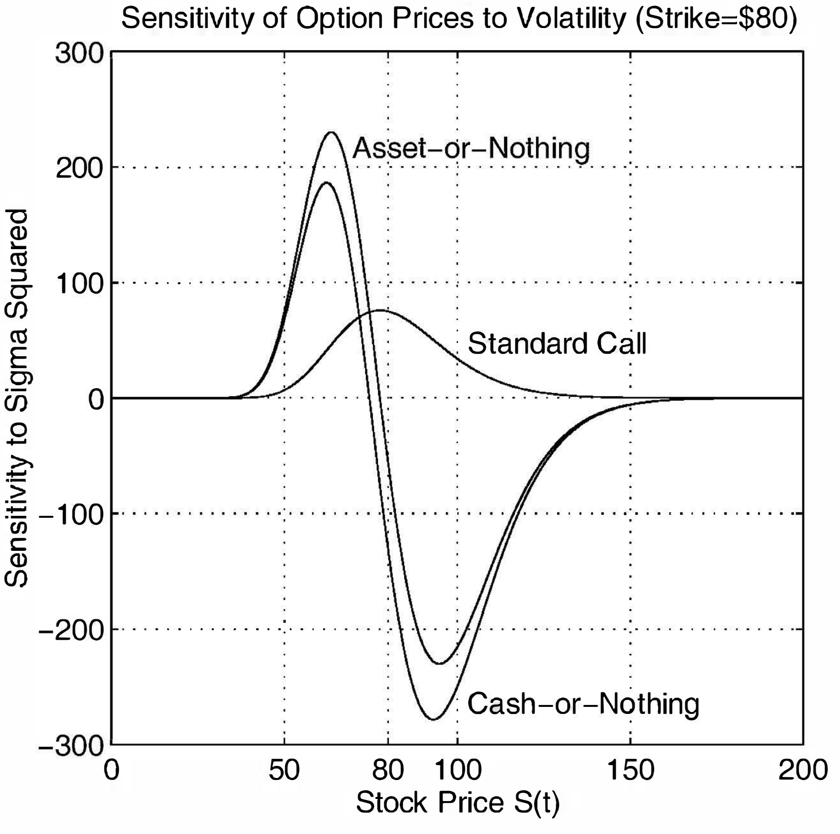


Figure B.2: Sensitivity of Option Prices to Volatility

The figure plots $\frac{\partial \text{ CALL PRICE}}{\partial \sigma^2}$ (i.e., "vega") using parameters $X = 80$, $r = 0.05$, $T - t = 1$, and $\sigma = 0.20$. The asset-or-nothing call price is always more sensitive to σ^2 than the cash-or-nothing call price. The difference between the sensitivities of each digital option is thus non-negative. The standard European call is equivalent to a long position in the asset-or-nothing and a short position in the cash-or-nothing. The response of the standard call price to increases in σ^2 is thus non-negative.

mon sense and some limiting relationships to deduce the relationships between the deltas of the knock-outs and the standard call.

The delta is the sensitivity of call price to underlying. This means that the option's delta is just the slope when you plot call value, $c(t)$, against underlying value, $S(t)$. Do not get this plot mixed up with the *payoff* diagram (the one with a "kink" for a standard call). See Figure B.1 (on page 119).

Now, everything you can do with a down-and-out option, you can also do with a standard option. On top of that, you still have a standard option in your hands in cases where the down-and-out gets "knocked out." It follows that the standard call is more versatile than the down-and-out call and must be more expensive. Thus, the value of the standard call must plot above the value of the down-and-out call for any value of the underlying. However, the two calls have the same value for very large $S(t)$—because the down-and-out option is unlikely to get knocked out. Both valuation curves are smooth, so the down-and-out call's valuation curve must be steeper [it starts lower than the standard call and "finishes" in the same place for high $S(t)$]. A steeper valuation curve when plotted against the level of underlying means precisely that the delta is higher for the down-and-out call than for the standard call.

For the up-and-out, you get a different answer. As before, the up-and-out is a knock-out option and is cheaper than the standard call. However, the standard call option and the up-and-out call option have the same value for very small $S(t)$—because the up-and-out option is unlikely to get knocked out. Both valuation curves are smooth, so the up-and-out call's valuation curve must be flatter (it starts in the same place as the standard call and finishes lower). A flatter valuation curve means precisely that the delta is lower for the up-and-out call than for the standard call.

Thus, the following relationships hold for the deltas of the different options:

$$\Delta_{up-and-out\ call} \leq \Delta_{standard\ call} \leq \Delta_{down-and-out\ call}$$

To hedge a short position in a down-and-out call, you need to buy more units of stock than you do to hedge a short position in a standard call. The value of the down-and-out call is more sensitive than the standard call to changes in the value of the underlying stock.

Note that increasing the term to maturity or increasing the knock-out price both increase the likelihood that a down-and-out call will be knocked out. This makes the down-and-out call even cheaper relative to the standard call. In fact, if the down-and-out call is very likely to be knocked out, the plot of down-and-out call price against stock price can become concave. Conversely, if the term to maturity is very short and the knock-out price is very low, the standard call and the down-and-out call have virtually identical prices (because the knock-out is very unlikely to be knocked out).

Answer 2.13: Your observation is that the sample variances are not linear in time and that the differences are statistically significant. This is equivalent to rejecting the null hypothesis of a random walk using a "variance ratio" test (Lo and MacKinlay [1988]).[14] This is contrary to the random walk assumptions of the Black-Scholes model.

The observations are consistent with the empirical findings that some financial stock indices are positively autocorrelated at weekly return intervals (Lo and MacKinlay [1988]).[15] This predictability influences the theoretical value and the empirical estimate of the diffusion coefficient σ (Lo and Wang [1995]). An adjustment can be made to the Black-Scholes formula to account for the predictability that is not part of the original Black-Scholes model. A new diffusion process that captures the predictability can be defined (Lo and Wang [1995]).

With the new specification, the autocorrelation is described using a more complicated drift in the diffusion. The drift is now important for pricing the option. In the old specification, drift was not important (Black and Scholes [1973]; Merton [1973]).

The final pricing formula takes the same form as the original Black and Scholes (1973) formula. However, the way in which the volatility term σ is estimated changes. An increase in autocorrelation may either increase or decrease the value of σ—it depends upon the specification of the drift (Lo and Wang [1995, p105]).

[14] See also Peterson et al. (1992) for related variance ratio testing in the commodities market; their findings lead them to a brief discussion of option pricing in the presence of autocorrelation.

[15] Autocorrelation in a time series is correlation between observations and themselves lagged. It is also known as "serial correlation." Its presence neither implies, nor is implied by, the presence of a drift. Consult your favourite statistics book for more information.

The presence of autocorrelation in stock returns is only one example of a real world divergence from the Black and Scholes (1973) assumptions. For example, Thorp (1973) discusses the effect of restrictions on short sales proceeds. See Hammer (1989) for a discussion of other deviations.

Answer 2.14: This is a common question. Stock price, $S(t)$, ranges from \$0 to \$$\infty$; the "delta" varies from 0 to +1. When $S(t)$ is very low (well out-of-the-money), delta is close to zero; when $S(t)$ is very high (well in-the-money), delta is close to one; when $S(t) = X$ (at-the-money), delta is very slightly higher than one-half (assuming no dividends). The curve is smooth and looks very much like a cdf (cumulative distribution function). This is not surprising, given that delta $= N(d_1) = N(d_1(S))$, and $N(\cdot)$ is a cdf, and $d_1(S)$ is an increasing function of S. The delta is illustrated in the second panel in Figure B.1 (on page 119).

How about the intuition? The delta is how many units of stock you need to hold to hedge a short call option. If your call option is deep in-the-money, you need one unit of stock because the option will be exercised and the stock will be called; if your option is deep out-of-the-money, you need no stock because the option will expire worthless and the stock will not be called; if your option is at-the-money, you are not too sure, and you have about one-half a unit of stock just in case.

Answer 2.15: Wirh no dividends, it is never optimal to exercise the plain vanilla American call option prior to maturity because the option is worth more "alive" than "dead." If you never exercise early, then the "American" feature of the call is not valuable. Thus, the standard American call option and European call option (on a non-dividend-paying stock) have equal values. See Crack (2004) for extensive discussi●n.

Figure B.3 (on page 134) plots the time value of the call option, $c(t) - \max[S(t) - X, 0]$, against $S(t)$ for the parameter values $X = 80$, $r = 0.05$, $T - t = 1$, and $\sigma = 0.20$. I have replaced the American call value $C(t)$ with the European value $c(t)$ because they are the same thing for plain vanilla call options in the absence of dividends.

The time value (the height in Figure B.3) tends to zero as expiration approaches—regardless of stock price. The existence of positive time value (i.e., value over and above exercise value) means that there is value in waiting to exercise. It is this value that makes the American call more valuable alive than dead. However, this does not mean that you should continue to hold the option. Rather, it means that if you wish to exit the call position, you should sell it, not exercise it. The time value is easily seen by looking at the excess of option price over intrinsic value in the "Listed Options Quotations" (i.e., options on individual stocks) in the *Wall Street Journal*.

How does time value arise? There are costs to exercising a call prior to maturity: you lose the interest you would have earned on the strike price and you lose the ability to exercise later. These costs are both intimately linked with the time to maturity, and thus they decline to zero as maturity approaches. There is a benefit to early exercise of a call: you capture any dividend payment on the underlying. In the presence of dividends, you gain the benefits with least cost by waiting until just prior to the ex-dividend day to exercise. In this case, you would exercise only if the benefit outweighs the costs. In practice, these costs of early exercise typically outweigh the benefit until the last ex-dividend date during the life of the option (Cox and Rubinstein [1985, p144]). By this time the costs of early exercise have depreciated substantially. A very large expected dividend might also trigger early exercise.

Answer 2.16: The naive answer is that as stock price falls, so too does the delta. However, this ignores the influence of the passage of time on your hedge. This is a good question, because you must think in both dimensions.

Two opposing forces are at work here: First, other things being equal, the delta of a call option that is in-the-money rises toward +1 as the option gets closer to expiration;[16] second, other things being equal, as stock price falls, the delta of a call option falls.

If stock price is observed to fall gently over the final two months, and the option remains in-the-money, the approach of the expiration date

[16] If the option you sold finishes in-the-money, you need to be long the stock because it will be called away. Of course, if the option is out-of-the-money, the approach of the expiration pushes the delta down to zero. If the option is at-the-money, then (assuming a non-dividend-paying stock) the delta tends to a number slightly greater than one-half as the expiration date approaches (Cox and Rubinstein [1985, Figure 5-13, p223]).

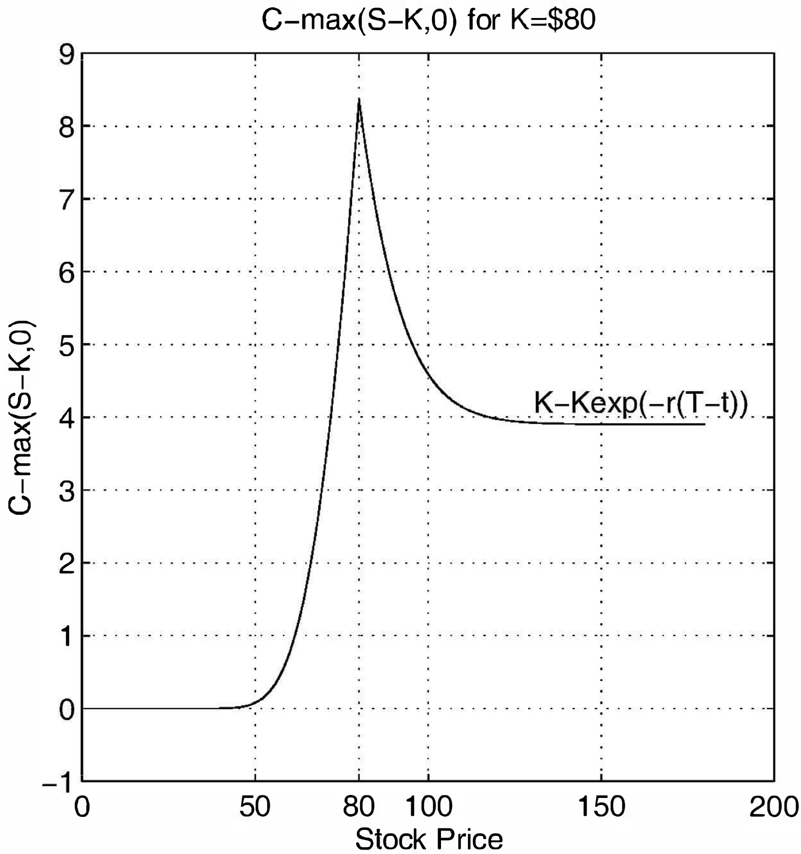


Figure B.3: Time Value of a European Call ●ption
The difference $c(t) - \max[S(t) - X, 0]$ is the value of not exercising. When the option is deep out-of-the-money, $\max[S(t) - X, 0] = 0$, and $c(t)$ is approximately zero. When the option is deep in-the-money, you save X by not exercising now, but it costs you the present value of exercising at maturity: $X \times e^{-r(T-t)}$. The left-hand limit of $X - X \times e^{-r(T-t)} = X \times \left(1 - e^{-r(T-t)}\right)$ is always non-negative. The "kink" in $\max[S(t)-X, 0]$ puts the "cusp" in the plot of $c(t) - \max[S(t)-X, 0]$ versus $S(t)$ at $S(t) = X$.

pushes the delta up to +1. If the fall in stock price is a little stronger, you may see the delta fall somewhat initially (and you will sell stock in your hedge portfolio). However, if the option finishes in-the-money, then the delta rises to +1 at the end of the life of the option (and you will buy stock in your hedge portfolio).

Answer 2.17: This is a good question. Introductory courses typically do not say much about jump processes.

The Black and Scholes (1973) model naively assumes that stock prices are continuous. That is, they assume that you can draw the price history without lifting your pencil from the paper. You need only stand on the floor of an exchange,[17] watch a real-time feed (on a Bloomberg terminal, Quantex box, ...), or read the WSJ headlines after an "event" to see that prices do not move smoothly. Indeed, the fact that stock prices are typically quoted with a minimum tick size (either exchange-imposed or effective) means that stock prices *cannot* move continuously. You can think of big stock price jumps as being stock price responses to the arrival of information in the market; small stock price jumps might just be due to the random ebb and flow of non-information-based (i.e., liquidity-related) transactions.

A "jump" price process is a price process that has infrequent jumps (i.e., discontinuities) in it. If the jump process is a very simple one, the Black-Scholes/Merton no-arbitrage technique can still be used to hedge and price options on an asset whose price follows the process. If the jump process is more complicated, the no-arbitrage technique breaks down. See the following discussion, and go to the references if you need more details. I have included some lengthy comments and references. This is because I think it is relevant, and it is often not covered in introductory courses.

[17] I have been on the floors of the New York Stock Exchange (NYSE), Chicago Board of Trade (CBOT), Boston Stock Exchange (BSE), and Dunedin Stock Exchange—long since replaced by screen trading—during trading hours. I have also visited the Chicago Board Options Exchange (CBOE), the Chicago Stock Exchange (CSE), and the old Paris Bourse. The new (at 1997) financial futures floor at the CBOT is big enough to land a 747 jumbo jet with space to spare—and it is noisy as hell. Conversely, the BSE is small and quieter than your typical MBA computer lab. I forecast that by 2005, all the Chicago futures exchange floors will be deserted—replaced by electronic trading. The NYSE may take a little longer, but I think it will suffer the same fate.

A simple jump process example (that is *not* a diffusion) has $\frac{dS}{S} = \mu dt + (J-1)d\pi$ (Cox and Ross [1976, p147]). In this example, $J-1$ is the jump amplitude (where $J \geq 0$), $d\pi$ takes the value $+1$ with probability λdt and 0 with probability[18] $1-\lambda dt$. The percentage stock price change $\frac{dS}{S}$ can thus jump suddenly to $J-1$ (which may itself be random); such a jump pushes S to SJ.

In this simple example, if J is fixed (i.e., non-random), a riskless hedge portfolio *can* be formed, and options on an asset whose price follows this simple jump process *can* be valued using the Black-Scholes/Merton no-arbitrage technique. This should come as no big surprise. The only real difference between this "pure Poisson process" case, and the simple binomial option pricing situation (Sharpe [1978]; Cox, Ross, and Rubinstein [1979]; Rendleman and Bartter [1979]; Cox and Rubinstein [1985]; Crack [2004]) is that the arrival time of the jump up or jump down is a random variable. You do not need to know *when* the stock price will jump to hedge the risk in a binomial setting. This "pure Poisson process" is a special case of a more general jump diffusion process discussed next.

Consider the jump diffusion process $\frac{dS}{S} = (\alpha - \lambda k)dt + \sigma dZ + dq$ (described in detail in my Footnote 9 to Question 2.17 on page 29). When $\sigma = 0$ and $Y \equiv dq + 1$ is non-random, you get Cox and Ross's simple jump process above, and the no-arbitrage technique can be used to hedge and price options on the jump process.[19] Otherwise, when $\sigma > 0$ and $\text{var}(Y) \geq 0$ it is not possible to form a riskless hedge portfolio or use the no-arbitrage technique (Cox and Ross [1976, p147]; Cox and Rubinstein [1985, pp361–371]; Merton [1992, p316]). Both the (non-jump) diffusion process and the (non-diffusion) simple jump process are the continuous limits of discrete binomial models. However, the jump-diffusion is not. It is for this reason that a riskless hedge cannot be formed in the jump-diffusion case (Cox and Rubinstein [1985, pp361–371]).

The fundamental reason that the no-arbitrage technique can be used to hedge and price options in the standard Black-Scholes world is linear-

[18] In this example, π is a continuous time "Poisson process." The term λ is the "intensity" of the process.

[19] I thank John Cox for explaining to me why such jump processes can be perfectly hedged (personal communication [February 17, 1994]).

ity. In continuous time, the Black-Scholes option price is an instantaneously linear function of the stock price. Portfolio building is a linear operation, and it follows that payoffs to the option can be perfectly replicated by building and continuously rebalancing a portfolio of the stock and the bond. Linearity breaks down when the jump term has positive variance—the call price becomes a nonlinear function of the stock price and perfect hedging is not possible (Merton [1992, p316]).

Although the no-arbitrage technique fails to price the option on the jump diffusion process, you can price the option using an *equilibrium* argument. An instantaneous CAPM (capital asset pricing model) approach may be used—as it was in the original Black and Scholes (1973) paper. The information that causes jumps may be assumed to be firm-specific (i.e., unsystematic and diversifiable).[20] You can hedge out the non-jump part of the option and deduce that the remainder (the jump) must have zero beta and, therefore, a riskless rate of return. This yields a partial differential equation that can be solved to give the call option price as an infinite summation:

$$C(S(t), T-t) = \sum_{n=0}^{\infty} \left\{ \frac{\exp[-\lambda(T-t)][\lambda(T-t)]^n}{n!} \times \right.$$
$$\left. E_n\{W[S(t)X_n \exp(-\lambda k(T-t)), (T-t); X, \sigma^2, r]\} \right\}.$$

Here X_n is a random variable with the same distribution as the product of n independent and identically distributed random variables each identically distributed to the random variable Y (recall that $Y-1$ is the random percentage change in stock price when a jump occurs), $X_0 \equiv 1$, E_n is the expectation operator over the distribution of X_n, and $W[S, (T-t); X, \sigma^2, r]$ is the standard Black-Scholes pricing formula (see Merton (1992, pp318–320) for a full discussion of the foregoing).

You cannot perfectly hedge the call when the underlying follows the general jump diffusion [$\sigma > 0$, var$(Y) \geq 0$]. However, you can hedge out the continuous parts of the stock and option price movements. This leaves you with a risky hedge portfolio following a pure jump process (with stochastic jump size). If you follow the Black-Scholes hedge when

[20] Note that in situations where the size of the jump is assumed to be systematic, the risk-neutral pricing technique cannot be used to value options. Hull (1997, p449, Footnote 14) directs the reader to Naik and Lee (1990) for a discussion of this point.

you are short the option, then most of the time you earn more than the expected rate of return on the risky hedge portfolio (an "excess return"). However, if one of those occasional jumps occurs (i.e., news arrives), you suffer a reasonably large loss. The jumps occur just infrequently enough that, on average, they balance the excess returns on the Black-Scholes hedge; and, on average, the hedge returns zero. In general, there is no way to adjust the parameters of the hedge technique (σ^2, for example) to get a better hedge (see Merton [1992, pp316–317] for a full discussion of the issues).[21]

Finally, if the underlying asset price is modelled as a jump process, the standard Black-Scholes call option formula mis-prices the option. Both the magnitude and the direction of the mis-pricing of the Black-Scholes model relative to the jump model vary with the distributional assumption for the size of the jump component (Trippi et al. [1992]).

Answer 2.18: Most students upon whom I have tested this one make several mistakes. The most common mistake is in the plot of call price at time t (i.e., prior to maturity) against a range of values for the underlying stock. Most students are under the impression that the call price is asymptotic to the 45° line rising from the strike price;[22] Merton (1973) demonstrates that this is not true. If you made this mistake, stop reading here and go back and try again. The next most common mistake is in drawing the plot of call price versus futures price—the answers I have seen vary dramatically and have nearly all been incorrect.

The correct plots appear in Figure B.4 (on page 140). The parameters used are $X = 80$, $r = 0.05$, $T - t = 1$, and $\sigma = 0.20$. The plot of call value against terminal stock price is the classic "kinked" call option payoff (the top plot in Figure B.4). Call value (terminal payoff) rises at 45° from the point $S(T) = X$.

The plot of call price versus futures price is a smooth curve that is asymptotic to the line $C = 0$ when the futures price, $F(t, T)$, is very

[21] For theoretical and empirical comparisons of the Merton (1976) jump process call option pricing and the standard Black-Scholes pricing, see Ball and Torous (1985).

[22] A curve is "asymptotic" to a line (i.e., an asymptote) if the curve gets closer and closer to the line. For example, $y = \frac{1}{x}$, for $x > 0$ is asymptotic to the line $y = 0$ as $x \to \infty$ and asymptotic to the line $x = 0$ as $y \to \infty$.

small, and asymptotic to the line that rises at 45° from the point $F(t,T) = X$ when the futures price is very large. See the middle plot in Figure B.4.

The plot of call price versus stock price, $S(t)$, is a smooth curve that is asymptotic to the line $C = 0$ when the stock price, $S(t)$, is very small, and asymptotic to the line that rises at 45° from the point $S(t) = Xe^{-r(T-t)}$($= \$76.10$ here) when the futures price is very large.[23] See the bottom plot in Figure B.4. The last two results are tied together by the fact that $F(t,T) = X \Leftrightarrow S(t) = Xe^{-r(T-t)}$.

At time t prior to maturity, the call price is lower if the futures price is equal to \$10 than it is if the stock price is equal to \$10. This is because the futures price represents expected future value in some sense, and this is not worth as much as current value (\$10 today is worth more than \$10 tomorrow).

Answer 2.19: This is a fundamental question. If it takes you more than five seconds to answer this, you are in trouble. Black and Scholes (1973) assume an arithmetic Brownian motion in log price. This assumption yields a geometric Brownian motion in price and an arithmetic Brownian motion in continuously compounded returns. Volatility of continuously compounded stock returns, σ^2, grows linearly with time for an arithmetic Brownian motion. The four-year σ^2 is four times the one-year σ^2. It follows that the four-year σ is two times the one-year σ. The answer is, therefore, 30%.

If $r > 0$, you also need to adjust the value of r that you use—r is four times as large when one period is four years as compared to when one period is one year.

Answer 2.20: Suppose that the process $\mathcal{S}(t)$ is an arithmetic Brownian motion of form

$$d\mathcal{S}(t) = \mu dt + \sigma_A dw(t),$$

where μ is the instantaneous drift per unit time, σ_A is the instantaneous volatility of $\mathcal{S}(t)$, and $w(t)$ is a standard Brownian motion (see Crack [2004] for introductory discussion of Brownian motions). Under

[23] Note that this implies that the time value $\{c(t) - \max[S(t) - X, 0]\} \to \{X - Xe^{-r(T-t)}\}$ as $S(t) \to \infty$. See Figure B.3 (on page 134) for a plot of time value versus stock price.

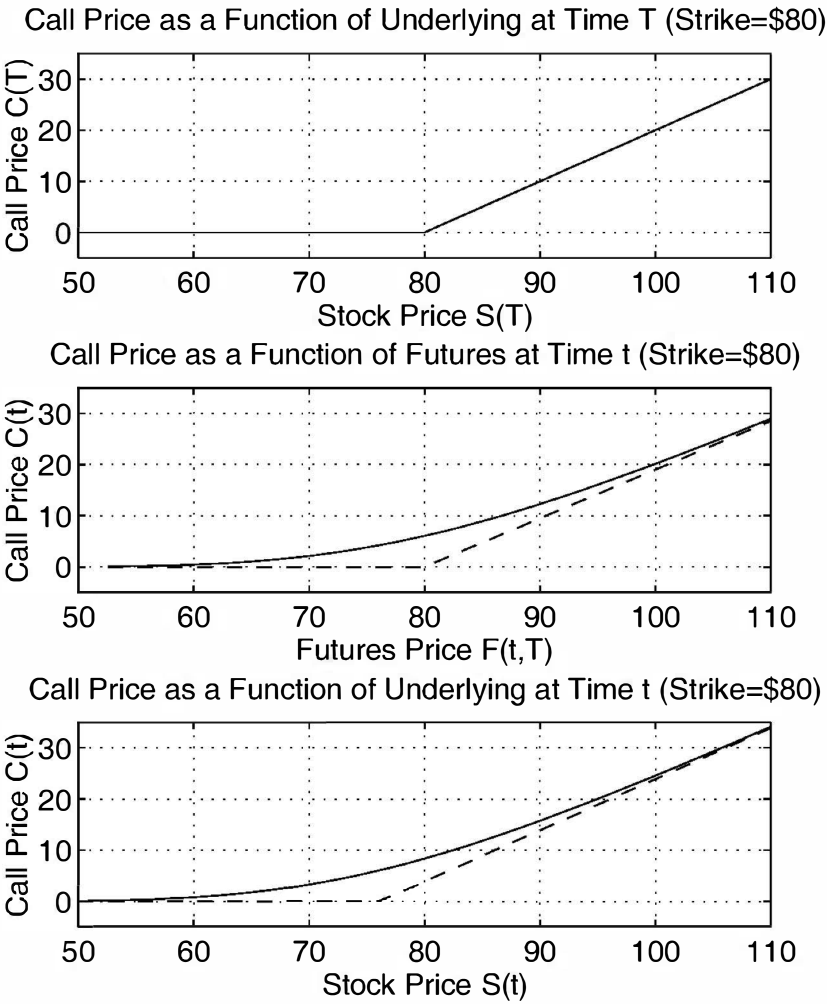


Figure B.4: Call Price as a Function of Different Variables
Call price is plotted as a function of price of underlying and of futures price (see Answer 2.18).

our assumptions, and under the risk-neutral probability measure, the process is $d\mathcal{S}(t) = \sigma_A dw^*(t)$.

Assume a strike of $\mathcal{X}$. Note that under the risk-neutral probability measure with $r = 0$ the process $\mathcal{S}(t)$ is given by Equation B.2.

$$\mathcal{S}(t) = \sigma_A w^*(t). \tag{B.2}$$

The call price is the discounted expected payoff under the risk-neutral probability measure, as follows:

$$\begin{aligned} c(t) &= e^{-r(T-t)}\ E^*[\max(\mathcal{S}(T) - \mathcal{X}, 0) \mid \mathcal{S}(t)] \\ &= E^*[\max(\mathcal{S}(T) - \mathcal{X}, 0) \mid \mathcal{S}(t)] \end{aligned}$$

From Equation B.2, it follows that

$$\begin{aligned} \mathcal{S}(T) &= \mathcal{S}(t) + \sigma_A(w^*(T) - w^*(t)) \\ &= \mathcal{S}(t) + \sigma_A \mathcal{W}^*, \end{aligned}$$

where $\mathcal{W}^* \equiv w^*(T) - w^*(t)$ is Normal $\mathcal{N}(0, T-t)$ under the risk-neutral probability measure. Now let "v" play the part of $\mathcal{W}^*$ distributed as $\mathcal{N}(0, T-t)$. Then the call price is given by the following integration over the Normal density:

$$c(t) = \int_{v_0}^{+\infty} (\mathcal{S}(t) + \sigma_A v - \mathcal{X})\ f_V(v)\ dv,$$

where

$$f_V(v) = \frac{1}{\sqrt{2\pi}\ \sqrt{T-t}} e^{-\frac{1}{2}\left(\frac{v}{\sqrt{T-t}}\right)^2}$$

is the pdf of $v \sim \mathcal{N}(0, T-t)$, and

$$v_0 \equiv \frac{\mathcal{X} - \mathcal{S}(t)}{\sigma_A}$$

is the boundary value of v at which $(\mathcal{S}(t) + \sigma_A v - \mathcal{X})$ changes sign.

The remainder of the proof is left to the reader. The final result is

$$c(t) = \sigma_A \sqrt{T-t}\ \left\{ \frac{e^{-\frac{1}{2}d^2}}{\sqrt{2\ \pi}} + N(d)\ d \right\}, \tag{B.3}$$

$$\text{where } d = \frac{\mathcal{S}(t) - \mathcal{X}}{\sigma_A \sqrt{T-t}}.$$

The arithmetic Brownian motion pricing formula (Equation B.3, above) is not well known. This is because an arithmetic Brownian motion is not a reasonable assumption for a price process: arithmetic Brownian motions can assume negative values. However, the geometric Brownian motion assumed by Black and Scholes (1973) is always non-negative, just as a price process should be. The importance of pricing options on stock catapulted the Black-Scholes formula (and the geometric Brownian motion beneath it) to super-stardom, while the pricing formula for the arithmetic Brownian motion languishes in relative obscurity.

Let's have a little history. Louis Jean Baptiste Alphonse Bachelier finished his mathematics PhD thesis at the Sorbonne in Paris in January 1900.[24] The topic of his thesis was the pricing of options contracts traded on the Paris Bourse.[25] Bachelier (1900) assumes that stock prices are Normally distributed and follow an arithmetic Brownian motion. He also assumes that expected returns on stocks (and on investments in general) are zero. Bachelier was the first to publish payoff diagrams for a European call option. Bachelier was also the first mathematician to use the "reflection principle."[26] Bachelier's derivation of the mathematical properties of Brownian motion predates by five years Albert Einstein's 1905 work on Brownian motion (Einstein [1905]). Bachelier even tested the predictions of his model using actual option prices on the Paris Bourse and found them not too far wrong.[27]

---

[24]The Sorbonne was the prestigious University of Paris founded by Robert de Sorbon in 1253. The Sorbonne was split into 13 units during the period 1968–1970. Nowadays, the name "Sorbonne" refers to the original university or to three of the 13 units that retain the title as part of their name. I had the pleasure of visiting the Sorbonne as a tourist in both 1998 and 1999. It is on the Left Bank, not far from Notre Dame.

[25]A very brief look at Bachelier's model is in Appendix A of Smith (1976); a full translation appears in Cootner (1964). Note that my option pricing formula, Equation B.3, is mathematically equivalent to Equation A.5 in Smith (1976). On an historical point of some coincidence, note that as I write this it is Louis Bachelier's 125th birthday. Bachelier was born in Le Havre, France, on March 11th, 1870.

[26]It is also known as the method of "reflected images." If you do not yet know what the reflection principle is, you probably do not need to know. If you are curious, see Harrison (1985, p7) for details.

[27]Samuelson (1973, Footnote 2, p6) compares Bachelier (1900) and Einstein (1905). He declares Bachelier dominant "in every element of the vector." See Samuelson (1973) for further discussion (and criticism) of Bachelier and other topics in the mathematics of speculative prices.

Unfortunately, Bachelier's assumptions violate some basic economic principles. In particular, he violates limited liability, time preference, and risk aversion (see Samuelson [1965, p13] for discussion). However, the significant contributions of Bachelier's thesis mean that he is rightfully considered the "father of modern option pricing theory" (Sullivan and Weithers [1991]).

In the special case when $\mathcal{S}(t) = \mathcal{X}$ (the option is struck at-the-money), Equation B.3 reduces to[28]

$$c_A(t) = \sigma_A \sqrt{\frac{T-t}{2\pi}}, \tag{B.4}$$

where my "A" indicates that the underlying process, $\mathcal{S}(t)$, is an *arithmetic* Brownian motion, and σ_A is the standard deviation of the level of $\mathcal{S}(t)$.

Equation B.4 was derived assuming $r = 0$, and $\mathcal{S}(t) = \mathcal{X}$, plus the horrible assumption of arithmetic Brownian motion. You might reasonably ask how does Equation B.4 compare to Black-Scholes for an at-the-money call option when $r = 0$?

The Black-Scholes formula for pricing a standard European call on a non-dividend-paying stock reduces to Equation B.5 in the special case when $r = 0$ and $S(t) = X$ (i.e., the option is struck at-the-money):

$$c_{BS}(t) = S(t)\left[N\left(+\frac{\sigma}{2}\sqrt{T-t}\right) - N\left(-\frac{\sigma}{2}\sqrt{T-t}\right)\right], \tag{B.5}$$

where $S(t)$ follows a *geometric* Brownian Motion, and σ is the standard deviation of continuously compounded returns on the stock price, $S(t)$.

When σ is small, Equation B.5 may be approximated as[29]

$$c_{BS}(t) \approx S(t)\sigma\sqrt{\frac{T-t}{2\pi}}. \tag{B.6}$$

[28] This is Equation 4.7 in Samuelson (1973).

[29] This approximation appears in Brenner and Subrahmanyam (1988). They use a Taylor series derivation, but less formally it follows because $[N(z)-N(-z)]$ is just the area under the standard Normal pdf from $-z$ to z. With σ small, you can approximate the area by length times height. The length is $\sigma\sqrt{T-t}$; for small σ, the height is close to the height of the standard Normal pdf at its peak: $\frac{1}{\sqrt{2\pi}}$ (recall that $\frac{1}{\sqrt{2\pi}} \approx 0.4$).

Compare Equation B.6 with Equation B.4. In the arithmetic Brownian motion case, σ_A is the standard deviation of the level of the price process $\mathcal{S}(t)$; in the geometric Brownian motion case, σ is the standard deviation of continuously compounded returns. Standard deviation of price is, however, approximately equal to price times the standard deviation of continuously compounded returns. It follows that the pricing in Equations B.4 and B.6 is consistent, even though the first uses an arithmetic Brownian motion (supposedly incorrect), and the second uses a geometric Brownian motion. Thus, the Black-Scholes formula reduces to the century-old Bachelier formula.

In my other book (Crack [2004]) I demonstrate the general ABM case where we assume neither that the option is at the money, nor that $r = 0$.[30] The formula for the call option price in this case is given by Equations B.7–B.9. See Crack (2004) for full details of the derivation.

$$
\begin{aligned}
c(t) &= e^{-r(T-t)}\sigma_A\sqrt{\frac{e^{2r(T-t)}-1}{2r}}\left[N'(d)+N(d)\cdot d\right] && \text{(B.7)}\\
&= e^{-r(T-t)}\sigma_A\sqrt{\frac{e^{2r(T-t)}-1}{2r}}\left[\frac{e^{-\frac{1}{2}d^2}}{\sqrt{2\,\pi}}+N(d)\cdot d\right] && \text{(B.8)}\\
\text{where } d &= \frac{S(t)e^{r(T-t)}-X}{\sigma_A\sqrt{\frac{e^{2r(T-t)}-1}{2r}}}. && \text{(B.9)}
\end{aligned}
$$

Answer 2.21: Black-Scholes in your head!? This technique is so well known that some interviewers just ask if you can do it, and if you say yes they move on. It's not worth the gamble if you don't know it.[31]

Traders use the arithmetic Brownian motion approximation (or Black-Scholes reduced formula) from Answer 2.20 as a rough but fundamental call pricing relationship:

$$
c(t) \approx \sigma S\sqrt{\frac{T-t}{2\pi}}, \qquad \text{(B.10)}
$$

where σ is the standard deviation of returns or where σS is replaced by the standard deviation of prices. You should also note that this

[30] I thank Mikhail Voropaev for contributing this idea. Any errors are mine.

[31] See Haug (2001) for related material.

versatile little formula prices *both* puts and calls. Why is this? Well, if interest rates are low, and the option is struck at-the-money, then in the absence of dividends a call and put have the same value—just use put-call parity.

Many times interviewees are asked to price an option in their head where the interest rate is zero and the option is struck at-the-money. You should, therefore, know that the option pricing formulae of both Black-Scholes and Bachelier reduce to Equation B.10 and that it works for both for puts and calls. I expect you to be able to evaluate Equation B.10 in your head in less than 10 seconds if asked to in an interview. How can you do this so quickly? Well, $\frac{1}{\sqrt{2\pi}} \approx 0.4$, and for three months, six months, or one year to maturity, you have $\sqrt{0.25} = 0.50$, $\sqrt{0.50} \approx 0.70$, and $\sqrt{1} = 1$, respectively. Of course, it helps that they usually give you easy numbers. For example, if $S = \$100$, $\sigma = 0.40$, and $(T-t) = 0.25$, the formula gives \$8 $(0.4 \times 0.4 \times 100 \times 0.5)$ whereas Black-Scholes proper gives \$7.97—not bad at all! The approximation is usually accurate to within a couple of percentage points.

Answer 2.22: This is a common type of question requiring fundamental knowledge. The only thing that changes between the two options is the time until expiration. The important knowledge here is how the value of a call changes with time to expiration.

You should remember that in the special case where $r = 0$, and the option is struck at-the-money [so that $S(t) = X$], the Black-Scholes European call option pricing formula may be approximated by the following (see discussion on page 144):

$$c(t) = \sigma\sqrt{\frac{T-t}{2\pi}},$$

where σ is not standard deviation of continuously compounded returns, but standard deviation of *price*. From this approximation, you can see that if the call is at-the-money, the call value increases at something like the square root of the term to maturity (if you double term to maturity, value increases by 40% to 50%). You must be very comfortable with this approximation.

The above approximation is a good place to start. However, a full answer recognizes that the response of call value to term to maturity depends heavily upon whether the call is in-the-money, at-the-money, or out-of-the money. Sensitivity to term to maturity decreases as you move into-the-money, down to zero in the limit if you are very deep in-the-money. Sensitivity to term to maturity increases as you move out-of-the-money. Doubling the term to maturity can easily double, triple, or quadruple the value of the call if it is well out-of-the-money. The effect is greater the further out-of-the-money the call is—this is why deep out-of-the-money options are sometimes called "lottery tickets."

You can see this effect clearly if you compare the prices of actively traded equity options and LEAPS.[32] Go to the listed options quotations in the third section of the *Wall Street Journal*. Choose a stock for which both equity options and LEAPS are traded. Compare the prices of call options on your chosen stock that have the same strike, but different terms to maturity (e.g., three months, six months, one year, and two years). If you do this comparison for different strike prices, you should see that an extension in term to maturity has the most impact on call option prices when the options are out-of-the-money. You should see that for call options that are deep out-of-the-money, doubling the term to maturity can easily quadruple the value of the call option.

In simple terms, if you extend the term to maturity, the option has more opportunities to finish in-the-money, the present value of the cost of exercising decreases, and the call value increases. The increase in the call value depends upon the initial likelihood that the call will finish in-the-money. This likelihood is small if the option is well out-of-the-money. Thus an increase in term to maturity produces a proportionately greater increase in the value of an option that is out-of-the-money.[33]

A caveat. The approximation formula $c(t) = \sigma\sqrt{\frac{T-t}{2\pi}}$ prices at-the-

[32]LEAPS are "Long-term Equity AnticiPation Securities." That is, LEAPS are long-term options. LEAPS have terms to maturity of up to three years. The term to maturity of standard exchange-traded equity options does not exceed eight months. LEAPS are not exotic options, but exchange-traded standardized options contracts. Like standard equity options, all LEAPS are American-style options. Unlike standard equity options, equity LEAPS all expire in January; index LEAPS all expire in December (Options Clearing Corporation [1993]).

[33]For a very helpful practitioner view on the interpretation of partial derivatives of call price with respect to each option pricing parameter, see Chance (1994).

money European-style puts and calls when $r = 0$. However, it has its limitations. For example, if $r \neq 0$, then the value of a deep in-the-money European put *decreases* as time to maturity extends. If the put is deep in-the-money, then life is already as good as it gets (the put has limited upside). You want to exercise now and take the money. Extending the life of the option pushes the expected benefit further away and decreases the put's value.

Answer 2.23: I give two different methods for answering this question. If the standard deviation is \$20, not \$10, then double the answers given.

FIRST SOLUTION

As a loose rule of thumb, the standard deviation of price per period (\$10 here) is a rough measure of the average possible upside move or downside move in stock price over the next period. You have approximately half-a-chance of finishing in-the-money, and half-a-chance of finishing out-of-the-money (or "under water," as it is sometimes called). The expected payoff is, therefore, roughly $\left(\frac{1}{2} \times \$0\right) + \left(\frac{1}{2} \times \$10\right) = \$5$. In fact, the shape of the Lognormal distribution of final stock price means that the expected payoff is slightly less than \$5 (it is around \$4).

SECOND SOLUTION

In the case where $r = 0$, and the option is struck at-the-money [so that $S(t) = X$], the Black-Scholes option pricing formula may be approximated by the following (see discussion on page 144):

$$c(t) = \sigma\sqrt{\frac{T-t}{2\pi}},$$

where σ is not standard deviation of continuously compounded returns, but standard deviation of *price*. With $T - t = 1$, and $\frac{1}{\sqrt{2\pi}} \approx 0.40$ (memorize that one), the standard deviation of price of \$10 implies a call price of around \$4. Note that this technique is more accurate than the first, giving \$4 instead of \$5.

Answer 2.24: The answer cannot be found exactly in the Black-Scholes framework, but you can get a good estimate.[34] Increasing the implied

[34]Francis Longstaff suggested to me that an important option pricing problem is the handling of "event risk" (personal communication [September 25, 1998]). For example, how do you price a 14-day option on a stock whose CEO is scheduled to make an important

volatility σ by 25% (from 0.20 to 0.25) on one day out of 100 in the option's life is the same (to a first-order approximation) as increasing σ^2 by 50% on one day out of 100 in the option's life.[35] This averages out to something like increasing σ^2 by 0.5% for every day remaining in the option's life (i.e., multiplying the average σ^2 by a factor 1.005).[36] Using the approximation (see page 144) $c(t) \approx \sigma S(t)\sqrt{\frac{T-t}{2\pi}}$, we see that multiplying σ^2 by M has the same effect on $c(t)$ as multiplying $T-t$ by M. This is because each of σ^2 and $(T-t)$ appear in the option formula under a square root sign—either implicitly or explicitly. It follows that multiplying σ^2 by a factor 1.005 is equivalent to increasing the term to maturity by something like 0.5% (i.e., one-half of a day for a 100-day option). That is, increasing σ^2 by 50% on one day is equivalent to increasing the length of one day by 50%.

Either of the adjustments mentioned increases the value of an at-the-money option by a factor of about $\sqrt{M}$—a quarter of a percent here. Note that the equivalence of the 50% increase in σ^2 on one day and the extension of option life by half a day is a general result—because variance is linear in time. However, the conclusion that either of these adjustments increases option value by about a quarter of a percent applies only to at-the-money options. If an option is deep in-the-money, the adjustments mentioned may have little or no effect on option value; if an option is deep out-of-the-money, the adjustments mentioned may increase the option value by substantially more than a quarter of a percent.

Answer 2.25: A give-away question! A long straddle is a long call plus a long put with the same strike. If you hold the straddle until maturity, then you need a price change of more than \$5 either way in the underlying to profit. A smaller price change, however, can lead to profits if it happens before maturity. For example, using Black-Scholes

announcement in seven days. I think this question is a loose attempt at this issue.

[35]The "implied volatility" is the volatility figure implicit within an option price, assuming that market participants value options using the Black-Scholes formula. The "implied vol" appears first in the literature in Latané and Rendleman (1976).

[36]As an aside, note this for the Black-Scholes formula: If we increase the calendar term to maturity, but still call it "one period," then we need to increase σ. However, if we increase $(T-t)$ ("term to maturity" or "the number of periods") without changing the length of one period in the model, we do not need to change σ.

(ignoring that CBOE equity options are American-style), if $\sigma = 0.357$, $T - t = 0.5$, $S = \$25$, and $r = 0.02$, then a straddle struck at \$25 costs \$5. If the price of the underlying suddenly jumps to \$27, then the straddle is suddenly worth \$5.50 and you have an immediate 10% gain. See Table B.1 for details.

	Stock Price = \$25.00	Stock Price = \$27.00
Price of the Call ($X = \$25$)	\$2.625	\$3.875
Price of the Put ($X = \$25$)	\$2.376	\$1.626
Price of the Straddle (sum)	\$5.001	\$5.502

Table B.1: Straddle Prices when the Stock Price Jumps

The option prices in the table are calculated using volatility of $\sigma = 0.357$ per annum, time to maturity of $T - t = 0.5$ years, a riskless rate of $r = 0.02$ per annum, and the Black-Scholes formula. A long straddle is a long call plus a long put with the same strike. A straddle struck at \$25 costs \$5 when stock price is $S = \$25$, but if the stock price jumps immediately to \$27, the straddle is worth \$5.50, giving an immediate 10% gain, ignoring transactions costs.

Answer 2.26: The Eurodollar futures contract is the most popular short-term interest rate futures contract. The contract value used for marking-to-market at the end of the day is $\$10,000 \times [100 - \frac{90}{360}\delta]$, where δ is the settlement discount rate. Between settlements, the market participants determine, through supply and demand, what is considered a fair discount. At maturity, the discount δ must converge to the three-month LIBOR US dollar rate. Note that if the discount is 5%, then $\delta = 5.0$ in the above calculation, not 0.05.

The three-month LIBOR rate is typically about 40 to 50 basis points (0.40 to 0.50 percentage points) higher than the yield on three-month treasury bills (this compensates for default risk of London banks).[37] The discount rate δ is thus highly correlated with US interest rates. The contract value is highly negatively correlated with δ, and thus highly negatively correlated with US interest rates.

[37]This is an estimate only. There is tremendous variation in the spread. The "Ted spread" is the Eurodollar futures less T-bill futures index point spread (with same delivery month). It can also vary tremendously.

Suppose that you are long a Eurodollar future. If US interest rates rise, the contract value declines, and you finance your loss at a relatively high rate. If US interest rates fall, then the contract value rises, but you invest the marked-to-market gains at relatively low rates. If you hold the forward, rather than the future, you do not have day-to-day gains and losses, so you are not hurt in the same way by these opportunity costs. Other things being equal, you would rather have the Eurodollar forward contract than the Eurodollar futures contract. If the discounts are the same (as stated), then there is a mis-pricing.[38] With the current mis-pricing, I would choose to go long the Eurodollar forward and short the Eurodollar future.

Answer 2.27: It makes much more sense to simulate the underlying and find the payoffs to the call, than it does to simulate the process for the call itself. It is difficult to model the call, because the instantaneous volatility of the call changes whenever the leverage of the call changes (assuming the underlying is of constant volatility). The leverage of the call changes whenever the stock price moves (and it even changes if the stock price does not move—simply because of time decay).

Answer 2.28: I think a quick review of "mortgage-backs" is in order before addressing the question. Mortgage-backed securities are shares in portfolios of mortgages. The value of all mortgage-backed securities outstanding was around \$1.5 trillion as of mid-1997—that is, \$1,500,000,000,000.

Owners of mortgage-backs are exposed to "prepayment risk," and "extension risk." Prepayment risk is the risk that interest rates will fall, and borrowers will exercise their right to refinance at lower rates (they exercise their call option on the mortgage). The problem is that the holders of the mortgage-backs, therefore, get repaid when interest rates are low—the worst possible time to receive the money. Conversely, extension risk is the risk that interest rates will rise, and borrowers will slow down their rate of repayment—meaning that holders of mortgage-backs get fewer dollars to invest at precisely the best time for them to be

[38] If the underlying were strongly *positively* correlated with US interest rates, then the futures contract would be more attractive than the forward. This is because daily gains can be invested at relatively high rates, while daily losses are financed at relatively low rates (see Hull [1997, pp55–56] for more details).

investing. Mortgage-back investors thrive when interest rate volatility is low.

The simplest mortgage-back is a "pass-through"—each share in the mortgage pool provides a prorata share in the cash flows to the pool, and thus each share has identical risk and return characteristics.

Collateralized mortgage obligations (CMO's) are a type of mortgage-back that splits the mortgage pool up into "tranches" (the French word for "slice"). Unlike a pass-through, which gives equal shares to all holders, the tranches are unequal shares. Take a simple example with only four tranches: "A," "B," "C," and "Z." The A, B, and C shares all receive regular coupons. The A shares are retired (i.e., the principal is repaid) ahead of the other tranches by using the earliest prepayments by borrowers. The B shares are retired, through prepayments, only after the A shares are gone. The C shares are retired, through further prepayments by borrowers, only after the A and B shares are gone. The Z shares receive no payouts whatsoever until all of the A, B, and C shares are gone. You may think of the Z shares as being like zero-coupon bonds with a life equal to the life of the longest-lived mortgages in the pool. CMO tranches thus provide *different* risk-return profiles—in contrast to pass-throughs.

With borrowers long a call on the mortgage (i.e., the right to buy back the mortgage by prepayment), holders of mortgage-backs are short a share of each of these calls from the mortgage pool. You will recall, of course, that long calls have positive convexity and that short calls have negative convexity.

In the absence of the call feature of a mortgage-back (the fact that borrowers have the right to prepay early), the mortgage back has positive convexity as a function of interest rates—just like an ordinary non-callable bond (Sundaresan [1997, p393]). However, when interest rates are low, the call feature becomes important to borrowers. If interest rates fall, all borrowers will refinance by the time rates have fallen to some critical value. At this stage, the mortgage-back is worth par. When interest rates are low, the importance of the call feature (a short call position to the mortgage-back holder) means that the mortgage-back can acquire negative convexity.[39] Negative convexity is also called

[39] In fact, this is a feature of any callable bond—if interest rates fall far enough, the call

"compression to par" because of the convergence of the security's value to par as interest rates fall (Sundaresan [1997, p394]).

Note that although the mortgage-back value may have negative convexity, it is still downward sloping as a function of interest rates. However, if interest rates are low and close to the coupon rate of the mortgage back, then an increase in volatility of interest rates can decrease the value of the security (Sundaresan [1997, p394]). This result follows because the holder of the mortgage back is short the calls that the borrowers are long—and calls increase in value with volatility.

Now to the interview question. If you are long mortgage-backs, and you expect a bond market rally, then you expect bond yields to fall and bond prices to rise. Thus, your position will gain in value. The question is, which sign on convexity would maximize the gain (+ or −)? Positive convexity provides a steeper downward sloping plot of security value as a function of interest rates, and this in turn implies a larger gain if rates fall—thus, you prefer positive convexity.[40]

A full answer notes that we have assumed a *parallel* shift in the yield curve. If the yield curve steepens or flattens, the answer could change. Whether additional convexity helps or hurts you depends upon the type of yield-curve shift and the particular bonds under consideration. It needs to be evaluated on a case-by-case basis. See the related discussion beginning on page 198.

Answer 2.29: The hedging strategy is naive. This is called a "stop-loss strategy" (Hull [1997, p310]). At first glance, it replicates the payoff to the call. However, purchases and sales cannot be made at the strike price. When the stock is near the strike, you cannot know whether it will cross over the strike price or not. You have to wait until the stock price crosses the strike price. This means you end up making purchases at a price slightly higher than the strike and sales at a price slightly lower than the strike. The closer to the strike you try to time your trades, the more frequently you can expect to have to trade. You can

feature kicks in and imparts negative convexity to the security.

[40]Convexity is not such an issue if you expect a bond market rout. When prices fall and rates rise, prepayment becomes less attractive, and the call option in the hands of the borrowers assumes less importance—and so does the negative convexity the call is able to impart to your mortgage-back security value.

get eaten alive by transactions costs (see Hull [1997, p310]).

A second criticism is that the timing of the cash flows to the option and the hedge are different—it is not a hedge (see Hull [1997, p310]).

> **Story:** *Late one winter's evening at MIT (1994 I think), I was helping Franco Modigliani operate our photocopier. We somehow got onto the topic of the Crash of 1987 and he said "Yes, that is when I made all my money." He said he had been watching the market and, thinking it overvalued, he had bought out-of-the-money index puts (presumably S&P500 index options at that time). He made a bundle. He said he had tried it several times since then without success. At my office doorway another time, he told me that when pronouncing his name I should "drop the 'g'—it's the mark of a true Italian"—and that is how he pronounces it.*

Answer 2.30: This question and the next are the most popular stochastic calculus interview questions. Although this is ostensibly a stochastic calculus question, the answer relies only upon Riemann calculus. If you were stuck and looking for a hint, then maybe this is enough to get you going.[41]

Let $I_T(\omega)$ denote the integral $\int_0^T w(t,\omega)dt$. In this integral, t measures time along sample paths, and ω is an element of the sample space Ω (i.e., ω corresponds to a particular possible sample path). Since $w(t)$ has continuous paths with probability one (i.e., for almost every $\omega \in \Omega$ the path is continuous), this integral is a Riemann integral evaluated pathwise for any fixed $\omega \in \Omega$. The Riemann integral is just (in its simplest form):

$$I_T = \lim_{\Delta t \to 0} S_n, \text{ where } S_n \equiv \sum_{i=1}^{n} (t_i - t_{i-1})\, w(t_{i-1}), \text{ and}$$

$$\Delta t \equiv \max_i (t_i - t_{i-1}),\ 0 = t_0 < t_1 < \ldots < t_n = T.$$

[41] I thank Taras Klymchuk for suggesting a related solution technique. I am responsible for any errors.

We may rearrange terms as follows:

$$\begin{aligned}
S_n &= \sum_{i=1}^{n} (t_i - t_{i-1})\, w(t_{i-1}) \\
&= (t_1 - t_0)w(t_0) + (t_2 - t_1)w(t_1) + \ldots + (t_n - t_{n-1})w(t_{n-1}) \\
&= -t_0 w(t_0) + t_1\left[w(t_0) - w(t_1)\right] + t_2\left[w(t_1) - w(t_2)\right] \\
&\qquad\qquad + \ldots + t_{n-1}\left[w(t_{n-2}) - w(t_{n-1})\right] + t_n w(t_{n-1}) \\
&= -t_0 w(t_0) + \sum_{i=1}^{n-1} t_i\left[w(t_{i-1}) - w(t_i)\right] \\
&\quad + t_n \sum_{i=i}^{n-1}\left[w(t_i) - w(t_{i-1})\right] + t_n w(t_0) \quad \text{(a telescoping series)} \\
&= \sum_{i=i}^{n-1} (t_n - t_i)\left[w(t_i) - w(t_{i-1})\right], \quad \text{a.e.}
\end{aligned}$$

The last line follows because $w(t_0) \equiv w(0) = 0$ a.e. (i.e. almost everywhere) by definition. So, S_n is just a weighted sum of increments of a standard Brownian motion. It is well known that such increments are independently Normally distributed and that a finite sum of constant-weighted independent Normals is also Normal. Thus, S_n is Normal for each n. It can be shown that in the limit as $\Delta t \to 0$ (or as $n \to \infty$), the integral I_T is also Normal.

A Normal distribution is completely determined by its first two moments: the mean and the variance. We need only calculate the mean and variance of I_T to pinpoint the distribution. The mean is just $E(I_T) = \int_0^T E(w(t))dt = 0$. With a mean of zero, the variance is just the second non-central moment:

$$\begin{aligned}
V(I_T) &= E(I_T^2) \\
&= E\left\{\left(\int_0^T w(t)dt\right)\left(\int_0^T w(s)ds\right)\right\} \\
&= \int_0^T \int_0^T E\left[w(t)w(s)\right] dtds.
\end{aligned}$$

Now, you will recall that $w(t)$ is a process with independent increments. Let us assume, for the moment, without loss of generality, that $s < t$.

Then, $w(t) = w(s) + (w(t) - w(s))$, and $w(t)w(s) = w^2(s) + (w(t) - w(s))w(s)$. It follows that

$$E[w(t)w(s)] = E\left[w^2(s)\right] = s,$$

using independent increments and the fact that $w(s)$ has a variance of s. More generally, $E[w(t)w(s)] = \min(t, s)$. Thus, we may write the variance of the integral as follows:

$$\begin{aligned}
V(I_T) &= \int_0^T \int_0^T E[w(t)w(s)]\,dtds \\
&= \int_0^T \int_0^T \min(t, s)dtds \\
&= \int_0^T \left(\int_0^s tdt + \int_s^T sdt \right) ds \\
&= \int_0^T \left(\frac{s^2}{2} + s(T - s) \right) ds \\
&= \int_0^T \left(sT - \frac{s^2}{2} \right) ds = \left(\frac{T^3}{2} - \frac{T^3}{6} \right) = \frac{T^3}{3}.
\end{aligned}$$

Therefore, conditional on time 0 information, $I_T(\omega)$ is distributed as $\mathcal{N}(0, T^3/3)$.

Answer 2.31: Do you need a hint? This problem requires Itô's Lemma and not much else. Now go back to the problem and stop peeking at the solutions.[42]

If we apply Itô's Lemma to $F(t, w) \equiv \frac{w^2(t)}{2}$, we find[43]

$$dF = F_t dt + F_w dw + \frac{1}{2} F_{ww} (dw)^2 = w(t)dw(t) + \frac{1}{2} dt.$$

[42] I had the pleasure of attending a conference in honour of Norbert Wiener at MIT in October 1994. Two seats to my left sat Kiyoshi Itô—of Itô's Lemma fame. Professor Itô was not old, but neither was he young. He was of small build and very distinguished looking. He spoke clearly in somewhat halting English, and his good-natured humour was infectious. Paul Samuelson and Robert Merton also spoke, and it seems that Itô's Lemma was in fact a footnote in a paper of Itô's. They joked that it should be called "Itô's Footnote" instead—but that does not have the same ring to it.

[43] I thank Taras Klymchuk for suggesting this solution technique. I am responsible for any errors.

This notation means precisely

$$F(T) - F(0) = \int_0^T w(t)dw(t) + \frac{1}{2}\int_0^T dt = \int_0^T w(t)dw(t) + \frac{T}{2}.$$

Given the definition of $F(t)$, it follows immediately that

$$\int_0^T w(t)dw(t) = \frac{w^2(T) - T}{2} \quad \text{a.e.}$$

It should be noted that the expected value of the right-hand side of the equality is zero. This is consistent with the expected value of the left-hand side of the equality being zero also.

Answer 2.32: There are two ways to proceed: the first way is to work out the pricing formula from first principles; the second way is to use Black-Scholes option pricing as it stands and make some ad hoc adjustments to it to account for the power payoff.

FIRST SOLUTION

I was unable to find a published pricing formula for the power call (with payoff $\max[S^\alpha - X, 0]$) or for the power put (with payoff $\max[X - S^\alpha, 0]$), so I followed a straight discounted expected payoff approach under risk-neutral probabilities. It is relatively straightforward to show that the value at time t of European power call and put options maturing at time T is given as follows:

$$\begin{aligned}
c(t) &= S^\alpha(t)e^{m(T-t)}N(d_1') - e^{-r(T-t)}XN(d_2'), \quad \text{and} \\
p(t) &= e^{-r(T-t)}XN(-d_2') - S^\alpha(t)e^{m(T-t)}N(-d_1'), \qquad \text{where} \\
d_1' &= \frac{\ln\left(\frac{S(t)}{K}\right) + (r + (\alpha - \frac{1}{2})\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \\
d_2' &= \frac{\ln\left(\frac{S(t)}{K}\right) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} = d_1' - \alpha\sigma\sqrt{T-t}, \\
K &\equiv X^{\frac{1}{\alpha}}, \quad \text{and} \quad m \equiv \left(r + \frac{1}{2}\sigma^2\right)(\alpha - 1).
\end{aligned}$$

In the case $\alpha = 1$, the power option pricing formulae reduce to the standard Black-Scholes call and put pricing formulae.

The "delta" of the power call can be found by differentiating the power call pricing formula with respect to $S(t)$. The delta for the power call is given by:

$$\begin{aligned}\Delta_{power\ call} &\equiv \frac{\partial c(t)}{\partial S(t)} \\ &= \alpha S^{(\alpha-1)} e^{m(T-t)} N(d_1') \\ &\quad + \frac{X^{(1-\frac{1}{\alpha})} n(d_2' + \sigma\sqrt{T-t})[e^{-\frac{1}{2}(T-t)\sigma^2(\alpha-1)^2} - 1]}{\sigma\sqrt{T-t}},\end{aligned}$$

where $n(\cdot)$ is the Normal pdf function $n(x) \equiv \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, and m, d_1', and d_2' are as defined above.

It is interesting to note that because d_2' has the same functional form as the original Black-Scholes d_2, then the term $d_2' + \sigma\sqrt{T-t}$ appearing in the delta has the same functional form as the original Black-Scholes d_1. However, this differs from the power call's d_1' which contains an α term.

How does the power call's delta behave as $S(t)$ gets large? Well, as $S(t)$ gets large, both $d_1', d_2' \to \infty$. Thus, $N(d_1') \to 1$, and $n(d_2' + \sigma\sqrt{T-t}) \to 0$. It follows that

$$\Delta_{power\ call} \approx \alpha S^{(\alpha-1)} e^{m(T-t)}, \quad \text{for large} \quad S(t).$$

It follows that if $S(t)$ is large, then as $(T-t) \to 0$, we get that

$$\Delta_{power\ call} \approx \alpha S^{(\alpha-1)}.$$

This should come as no surprise: If the power call is deep in-the-money, and there is little time to maturity, then its sensitivity to changes in $S(t)$ will be about the same as the sensitivity of $S^{\alpha}(t)$ to changes in $S(t)$. The latter sensitivity is just

$$\frac{\partial S^{\alpha}(t)}{\partial S(t)} = \alpha S^{(\alpha-1)}.$$

One implication of this is that the delta of a power call continues to change as $S(t)$ increases.

What may come as a surprise is the shape of the power call option pricing function (see Figure B.5). If $\alpha > 1$, the plot of $c(t)$ versus $S(t)$ is *steeper* than and above the plot of $\max(S^{\alpha} - X, 0)$ for large $S(t)$ (it decays down toward the payoff as maturity approaches). If $\alpha < 1$, the plot of $c(t)$ versus $S(t)$ is less steep than and *below* the plot of $\max(S^{\alpha} - X, 0)$ for large $S(t)$ (it decays up toward the payoff as maturity approaches). Only in the case $\alpha = 1$ do the results agree with those for the standard call: The plot of call value as a function of stock price is less steep than and above the plot of $\max(S - X, 0)$. In all cases, the plot of $c(t)$ as a function of $S(t)$ is above the plot of $\max(S^{\alpha} - X, 0)$ for small $S(t)$.

Mathematically, the approximation $\Delta_{power\ call} \approx \alpha S^{(\alpha-1)} e^{m(T-t)}$ drives the results for large $S(t)$ (together with the fact that m is positive if $\alpha > 1$ and negative if $\alpha < 1$). Economically, the time value of the option drives the results. When $\alpha > 1$, the power of S is so high that the option value grows more quickly with increasing S than does the intrinsic value. When $\alpha < 1$, the option value grows less quickly than does the intrinsic value, and the European nature of the option means that there is "negative time value" for having to wait for such a low payout.

The payoff diagram for the power call is a little strange because the "kink" does not occur at $S = X$, but at $S = X^{\frac{1}{\alpha}}$—see Figure B.5. For example, if $\alpha = 2$, the payoff diagram is flat until $S(T) = \sqrt{X}$ and then is an upward sloping portion of the parabola $S^2(T) - X$. If $\alpha > 1$, the delta of the power call will be higher than the delta of a standard call with strike $X^{\frac{1}{\alpha}}$ because the payoff diagram is steeper. Conversely, if $\alpha < 1$, the delta of the power call will be lower than the delta of a standard call with strike $X^{\frac{1}{\alpha}}$ because the payoff diagram is less steep.

In the power option pricing formulae, d_2' has the same functional form as the d_2 in the regular Black-Scholes. The only difference is that you have $\ln\left(\frac{S(t)}{K}\right)$, where $K = X^{\frac{1}{\alpha}}$, in place of $\ln\left(\frac{S(t)}{X}\right)$. The reasoning follows a Z-score argument (see details in Crack [2004]). In the standard Black-Scholes formula, $N(d_2)$ is the (risk-neutral) probability that the call finishes in-the-money; it is the probability that $S(T) > X$. In the power call option formula, $N(d_2')$ is the (risk-neutral) probability that the power call finishes in-the-money. For the power call, this is the probability that $S^\alpha(T) > X$. This is the same as the probability that $S > X^{\frac{1}{\alpha}}$. This probability is in turn just the standard $N(d_2)$, in the case where the strike is given by $K \equiv X^{\frac{1}{\alpha}}$.

To extend the formulae to the case of continuous dividends at rate ρ, replace $S(t)$ by $S(t)e^{-\rho(T-t)}$ throughout the power option pricing formulae to yield

$$
\begin{aligned}
c(t) &= S^\alpha(t)e^{(m-\alpha\rho)(T-t)}N(d_1') - e^{-r(T-t)}XN(d_2'), \quad \text{and} \\
p(t) &= e^{-r(T-t)}XN(-d_2') - S^\alpha(t)e^{(m-\alpha\rho)(T-t)}N(-d_1'), \qquad \text{where} \\
d_1' &= \frac{\ln\left(\frac{S(t)}{K}\right) + (r - \rho + (\alpha - \frac{1}{2})\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \\
d_2' &= \frac{\ln\left(\frac{S(t)}{K}\right) + (r - \rho - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} = d_1' - \alpha\sigma\sqrt{T-t}, \\
K &\equiv X^{\frac{1}{\alpha}}, \quad \text{and} \quad m \equiv \left(r + \frac{1}{2}\sigma^2\right)(\alpha - 1).
\end{aligned}
$$

SECOND SOLUTION

An alternative to the full and formal pricing formulae given is an approximation using the standard Black-Scholes formula. Simply use d_2' exactly as above (with $K = X^{\frac{1}{\alpha}}$ for the reasons given), use $d_1' = d_2' + \sigma\sqrt{T-t}$, and replace $S(t)$ by $S^\alpha(t)$ in each of $c(t)$ and $p(t)$. However, be warned, this is an approximation only. If α is far from one (say above 1.2 or below 0.8), or time to maturity is longer than about six months, or implied volatility is bigger than about 0.40, the approximation is poor.

JARROW AND TURNBULL'S POWERED CALL

Jarrow and Turnbull ask their readers to value a call with payoff $[S(T)-K]^2$ if $S(T) \geq K$ and zero otherwise (Jarrow and Turnbull [1996, p175]). Assuming a Black-Scholes world, it is easy to show that the value of this call at time t prior to maturity is as follows:

$$\begin{aligned} c(t) &= S^2(t)e^{(r+\sigma^2)(T-t)}N(d_0) - 2KS(t)N(d_1) \\ &\qquad\qquad + e^{-r(T-t)}K^2N(d_2), \quad \text{where} \\ d_l &= \frac{\ln\left(\frac{S(t)}{K}\right) + \left(r + \left[\frac{3}{2} - l\right]\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \quad \text{for } l = 0, 1, 2. \end{aligned}$$

More generally, the following result (derived by me)

$$E^*[S^\alpha(T)|S(T) \geq K] = S^\alpha(t)e^{\alpha\left(r+\left[\frac{\alpha-1}{2}\right]\sigma^2\right)(T-t)}N(d_{2-\alpha}),$$

where E^* is expectation with respect to the risk-neutral probability measure and d_l is as above, allows you to value the powered call with general payoff:

$$c(T) = \begin{cases} [S(T) - K]^\alpha & ; S(T) \geq K \\ 0 & ; S(T) < K, \end{cases}$$

for non-negative integer α. The general pricing formula is

$$\begin{aligned} c(t) &= \sum_{j=0}^{\alpha}(-K)^{\alpha-j}\binom{\alpha}{j}S^j(t)e^{\left[(j-1)\left(r+j\frac{\sigma^2}{2}\right)(T-t)\right]}N(d_{2-j}), \quad \text{where} \\ d_l &= \frac{\ln\left(\frac{S(t)}{K}\right) + \left(r + \left[\frac{3}{2} - l\right]\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \quad l = 2, 1, \ldots, 2-\alpha, \end{aligned}$$

and $\binom{\alpha}{j} \equiv \frac{\alpha!}{j!(\alpha-j)!}$ is the usual binomial coefficient.

The reader should check that in the special case $\alpha = 2$, the general formula reduces to that previously given, and that in the special case $\alpha = 1$, the general formula reduces to standard Black-Scholes.

Answer 2.33: If the Black-Scholes assumptions are correct, then the implied volatilities of options (those backed out of the Black-Scholes pricing formula given the other pricing parameters) should fall on a horizontal line when plotted against strike prices of the options used. However, the patterns that result include smiles and skewed lines depending upon the underlying asset and the time period (Hammer [1989]; Sullivan [1993]; Murphy [1994]; Derman and Kani [1994]). Fifteen years

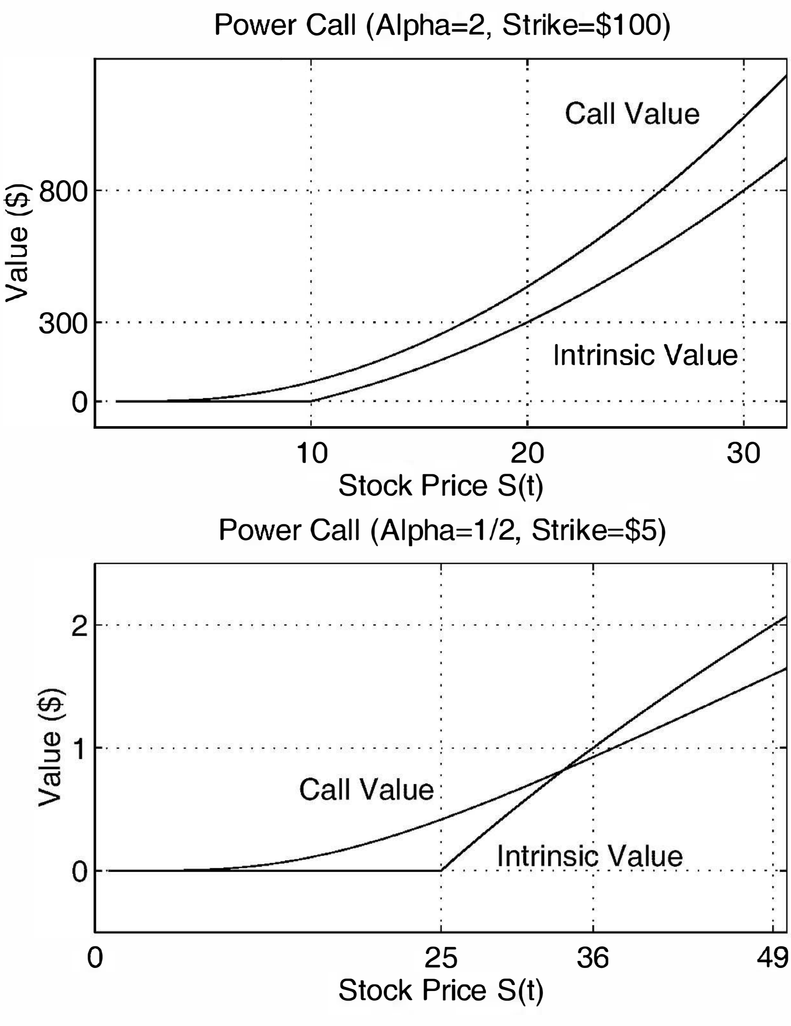


Figure B.5: Power Calls with $\alpha > 1$, and $\alpha < 1$

The power call prices are plotted as a function of price of underlying. Note that the "kink" in the payoff diagram does not occur at the strike K, but rather at $K^{\frac{1}{\alpha}}$ (see Question 2.32).

ago, you typically got smiles when you plotted the implied volatilities against strikes. Nowadays you are more likely to get skews, or smirks.[44]

What is happening may be viewed in some different and related ways. Option prices are determined by supply and demand, not by theoretical formulae. The traders who are determining the option prices are implicitly modifying the Black-Scholes assumptions to account for volatility that changes both with time and with stock price level. This is contrary to the Black and Scholes (1973) assumption of constant volatility irrespective of stock price or time to maturity. That is, traders assume $\sigma = \sigma(S(t), t)$, whereas Black and Scholes assume σ is just a constant.[45]

If volatility is changing with both level of the underlying and time to maturity, then the distribution of future stock price is no longer Lognormal. The distribution must be something different. Black-Scholes option pricing takes discounted expected payoffs relative to a Lognormal distribution. As volatility changes through time, you are likely to get periods of little activity and periods of intense activity. These periods produce peakedness and fat tails respectively (together called "leptokurtosis"), in stock returns distributions. Fat tails are likely to lead to some sort of smile effect, because they increase the chance of payoffs away-from-the-money.[46]

These irregularities have led to "stochastic volatility" models that account for volatility changing as a function of both time and stock price level (Hull and White [1987]; Scott[1987]; Wiggins [1987]; Hull [1997]). Applications to FOREX options include Chesney and Scott (1989) and Melino and Turnbull (1990). The effect of stochastic volatility on options values is similar to the effect of a jump component: both increase the probability that out-of-the-money options will finish in-the-money and increase the probability that in-the-money options will finish out-of-the-money (Wiggins [1987, pp360–361]). Whether the smile

[44]Another related deviation from Black-Scholes pricing is that implied volatilities when plotted against term to maturity produce a "term-structure of volatility." That is, traders use different volatilities to value long-maturity and short-maturity options (Derman and Kani [1994, pp2–3]; Hull [1997, pp503–504]).

[45]Black (1976) is the earliest paper I know of that acknowledges that $\sigma \uparrow$ as $S \downarrow$, and vice versa.

[46]The interaction of skewness and kurtosis of returns gives rise to many different possible smile effects (Hull [1997, Section 19.3]; Krause [1998, pp145-148]).

is skewed left, skewed right, or symmetric in a stochastic volatility model depends upon the sign of the correlation between changes in volatility and changes in stock price (Hull [1997, Section 19.3]).

Answer 2.34: This question is probably supposed to invoke misleading memories of the barrier option parity relationship: Other things being equal, a down-and-out call plus a down-and-in call is the same as a standard call. However, a double-barrier knock-out is not the same as an up-and-out together with a down-and-out. The latter pair of options is more valuable than the double-barrier knock-out. The most obvious reason is that if the underlying moves one way, then the double knock-out is knocked out, but a portfolio of a down-and-out plus a down-and-in still contains one option. That is, the pair of knock-outs is more versatile—and thus more valuable.

A double-barrier knock-out can be priced using a lattice (e.g., binomial) method. It may also be priced using the Kunitomo-Ikeda formula (Kunitomo and Ikeda [1992]; Musiela and Rutkowski [1997, p211]; or the user-friendly Haug [1997, p72]). The Kunitomo-Ikeda formula is an infinite series. Typically, only the leading few terms are needed for practical purposes (Kunitomo and Ikeda [1992, p286]). More terms may be needed if volatility is high, term to maturity is long, or the distance between the barriers is small (in each case this increases the likelihood of knockout and the pricing is more difficult).

Answer 2.35: A path-dependent option is one where the final payoff depends upon the stock price path followed. If the stock price ends up between the barriers, the option has different values, depending upon whether it was knocked in or knocked out (or both). Path-dependent options can typically be priced using Monte-Carlo methods. However, Monte-Carlo does not work for American-style options. Standard lattice techniques (e.g., binomial option pricing) do not usually work for path-dependent options.[47] However, you can price the "out-in" derivative using standard lattice methods, as follows. The parity relationship for knock-outs says that a down-and-out plus a down-and-in is a standard option. We can generalize this to conclude that an out-in plus

[47]Standard lattice techniques can be modified to allow pricing of path-dependent options. However, a couple of conditions involving complexity of the payoffs need to be satisfied (Hull and White [1993]; Hull [1997]).

a double-barrier knock-out is the same as an up-and-out (other things being equal). It follows that the out-in is worth the excess of the value of the up-and-out over the double-barrier knock-out. Both these knock-outs can be priced using standard lattice techniques.

Answer 2.36: Let $G(\cdot)$ denote the gold price. Now is time t, and time T is six months from now. The naive (and incorrect) step is to conclude that a volatility of $\sigma = \$60$ per annum translates to a six-month volatility of \$30. In fact, volatility grows with the square root of the term. Thus, \$60 per year translates to about $\sqrt{\frac{1}{2}} \times \$60 \approx \$42$ per half-year.

How do we find the probability that the option finishes in-the-money, $P(G(T) > 430)$? With $r = 0$ there is no drift in the risk-neutral world, so the distribution of $G(T)$ is centred on $G(t) = \$400$, with standard deviation roughly \$42. Thus:

$$
\begin{aligned}
P(G(T) > 430) &= P(G(T) - G(t) > 30) \\
&= P\left(\frac{G(T) - G(t)}{42} > \frac{30}{42}\right) \\
&\approx 1 - N\left(\frac{3}{4}\right),
\end{aligned}
$$

where $N(\cdot)$ is the cumulative standard Normal function. The last step follows because $\frac{G(T)-G(t)}{42}$ is roughly standard Normal. We know that $N(0) = 0.50$, and $N(1) = 0.84$, so $N\left(\frac{3}{4}\right) \approx 0.75$.

We conclude that there is roughly a 25% chance that the digital option finishes in-the-money. With a bet size of \$1 million, and a riskless interest rate of zero, the discounted expected payoff (in a risk-neutral world) is roughly \$250,000. The erroneous $\sigma = \$30$ gives an incorrect value of only about \$160,000.

Answer 2.37: The prices of the digital asset-or-nothing, $da(t)$, and the digital cash-or-nothing (with a "bet" size of \$1), $dc(t, \$1)$, are just the two parts of the Black-Scholes formula:

$$\begin{aligned} c(t) &= da(t) - X dc(t, \$1), \text{ where} \\ da(t) &= S(t)N(d_1), \\ dc(t, \$1) &= e^{-r(T-t)}N(d_2), \\ d_1 &= \frac{\ln\left(\frac{S(t)}{X}\right) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \text{ and,} \\ d_2 &= d_1 - \sigma\sqrt{T-t}. \end{aligned}$$

A Black-Scholes derivation using discounted expected payoffs under risk-neutral probabilities (Crack [2004]) contains implicit derivations of both digital option values. These may be identified if the initial step is re-expressed in the following form:

$$\begin{aligned} da(t) &= e^{-r(T-t)}\ E^*[S(T)\mathcal{I}_{S(T)>X} \mid S(t)], \quad \text{and,} \\ dc(t, \$1) &= e^{-r(T-t)}\ E^*[\mathcal{I}_{S(T)>X} \mid S(t)], \end{aligned}$$

where $\mathcal{I}_{S(T)>X}$ is the indicator function:

$$\mathcal{I}_{S(T)>X} = \begin{cases} 1 \text{ if } S(T) > X, \\ 0 \text{ if } S(T) \leq X. \end{cases}$$

Answer 2.38: Let us review quickly standard American options before looking at the perpetual option. American options are harder to price than European ones. Puts are harder to price than calls. An American put is hardest of all to price because early exercise can in general be optimal at any time for an American-style put. This differs from an American-style call, for which early exercise is optimal only at a few dates during the option's life (just prior to ex-dividend days). In fact, the problem is so hard that no exact pricing formula exists for standard American put options.

Black and Scholes (1973) value European-style puts and calls. If a stock does not pay dividends, then a European call and an American call have the same value (there is no incentive to exercise early). Thus, American calls on non-dividend-paying stocks can be valued using Black-Scholes. The introduction of dividends complicates matters.

However, an approximate pricing formula (Black [1975]) and an exact pricing formula (Roll [1977b]; Geske [1979]; Whaley [1981]) for American calls on dividend-paying stocks are known (see Hull [1997, Chapter 11]). American puts are more complicated. The dividend issue is not as important for puts as for calls because it is the receipt of the strike, not the dividends, that encourages early exercise of a put. Although no exact American put pricing formula exists, there are approximations (Parkinson [1977]; MacMillan [1986]; Barone-Adesi and Whaley [1987]). See the summaries in Tables B.2 and B.3.

Now to the perpetual American put. Extending the life of an option in perpetuity eases the pricing burden (removing the dependence on time turns a PDE into an ODE). Pricing the perpetual American put was a question on a problem set I had as a student in Robert C. Merton's derivatives course at Harvard in 1991. I reproduce my solution here.

Let V denote the value of a perpetual American put option on a stock. Let S denote the stock price. Let X denote the strike price. Assume that the stock pays continuous dividends at rate ρ. Let σ and r denote the volatility of stock returns and the riskless interest rate respectively. The Black-Scholes PDE is given by (Wilmott et al. [1993]):

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - \rho) S \frac{\partial V}{\partial S} - rV = 0.$$

However, for a *perpetual* put, time decay must be zero (it cannot age if it can live forever). Thus, the PDE becomes an ODE:

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - \rho) S \frac{\partial V}{\partial S} - rV = 0.$$

Let $\underline{S}$ denote the lower exercise boundary (this is how low the stock has to go before exercise of the put becomes optimal—it has to be determined). Then, we have the boundary conditions

$$\begin{aligned} V(S = \underline{S}) &= X - \underline{S}, \\ \left.\frac{\partial V}{\partial S}\right|_{S=\underline{S}} &= -1, \\ V(S) &\leq X. \end{aligned}$$

The second condition is the "high contact" condition.

All of this ODE's solutions may be represented as a linear combination of any two linearly independent solutions. It follows that

$$V(S) = A_1 V^1(S) + A_2 V^2(S),$$

where A_1 and A_2 are constants, and V^1 and V^2 are linearly independent solutions of the ODE. My guess is that $V^1 = S^{\lambda_1}$, and $V^2 = S^{\lambda_2}$ for some constants λ_1, and λ_2.[48] Substitution of V^i into the ODE yields (for $i = 1,\ 2$, and for $\underline{S} \leq S$):

$$\left[\frac{1}{2}\sigma^2\lambda_i(\lambda_i - 1) + (r - \rho)\lambda_i - r\right] S^{\lambda_i} = 0.$$

Rearranging and collecting terms in λ_i, we get for $i = 1,\ 2$:

$$\frac{1}{2}\sigma^2\lambda_i^2 + \left(r - \rho - \frac{1}{2}\sigma^2\right)\lambda_i - r = 0.$$

This is a quadratic formula, with solutions for λ_i:

$$\begin{aligned}
\lambda_1 &= \frac{-\left(r - \rho - \frac{1}{2}\sigma^2\right) + \sqrt{\left(r - \rho - \frac{1}{2}\sigma^2\right)^2 + 2\sigma^2 r}}{\sigma^2}, \quad \text{and} \\
\lambda_2 &= \frac{-\left(r - \rho - \frac{1}{2}\sigma^2\right) - \sqrt{\left(r - \rho - \frac{1}{2}\sigma^2\right)^2 + 2\sigma^2 r}}{\sigma^2}.
\end{aligned}$$

The solutions for λ_i can be seen to satisfy $\lambda_1 > 0$ if $r > 0$, and $\lambda_2 < 0$ if $r > 0$. Let us now consider the behaviour of the general solution we have derived: $V(S) = A_1 S^{\lambda_1} + A_2 S^{\lambda_2}$. First of all, with $\lambda_1 > 0$, and $\lambda_2 < 0$, then

$$\lim_{S \to +\infty} \left(A_1 S^{\lambda_1} + A_2 S^{\lambda_2}\right) = \pm\infty, \quad \text{if } |A_1| > 0.$$

However, the boundary conditions put both upper and lower finite bounds on the value of the put. Therefore, $A_1 = 0$, and $V(S) = A_2 S^{\lambda_2}$. Now, the first boundary condition tells us that

$$V(\underline{S}) = A_2 \underline{S}^{\lambda_2} = X - \underline{S},$$

[48] Why make this guess? Look at the ODE: the degree of the derivatives of V and the degree of S in the coefficients move together (both two, then both one, then both zero). This suggests solutions that are powers of S.

so it follows that $A_2 = \frac{X-\underline{S}}{\underline{S}^{\lambda_2}}$, which yields:

$$V(S) = \left(\frac{X-\underline{S}}{\underline{S}^{\lambda_2}}\right) S^{\lambda_2} \;=\; (X-\underline{S})\left(\frac{S}{\underline{S}}\right)^{\lambda_2}.$$

To pinpoint the solution, we must determine the value of the lower exercise boundary $\underline{S}$. The second of our boundary conditions says $\frac{\partial V}{\partial S}\Big|_{S=\underline{S}} = -1$. We can solve for $\underline{S}$ using this.

$$\begin{aligned}
\frac{\partial V}{\partial S} &= \lambda_2(X-\underline{S})\left(\frac{S^{\lambda_2-1}}{\underline{S}^{\lambda_2}}\right)\\
\Rightarrow \quad \frac{\partial V}{\partial S}\bigg|_{S=\underline{S}} &= \lambda_2\frac{(X-\underline{S})}{\underline{S}} = -1\\
\Rightarrow \lambda_2(X-\underline{S}) &= -\underline{S}\\
\Rightarrow \quad \underline{S} &= \frac{\lambda_2 X}{\lambda_2 - 1}.
\end{aligned}$$

Thus, for $S \geq \underline{S} \equiv \frac{\lambda_2 X}{\lambda_2 - 1}$, the perpetual American put is worth

$$\begin{aligned}
V(S) &= (X-\underline{S})\left(\frac{S}{\underline{S}}\right)^{\lambda_2} = \left(\frac{X}{1-\lambda_2}\right)\left[\frac{(\lambda_2-1)S}{\lambda_2 X}\right]^{\lambda_2}, \quad \text{where}\\
\lambda_2 &= \frac{-\left(r-\rho-\frac{1}{2}\sigma^2\right) - \sqrt{\left(r-\rho-\frac{1}{2}\sigma^2\right)^2 + 2\sigma^2 r}}{\sigma^2}.
\end{aligned}$$

For $0 \leq S \leq \frac{\lambda_1 X}{(\lambda_1 - 1)} \equiv \overline{S}$, it may be shown using similar techniques that a perpetual American call is worth

$$\begin{aligned}
V(S) &= \left(\frac{X}{\lambda_1 - 1}\right)\left[\frac{(\lambda_1-1)S}{\lambda_1 X}\right]^{\lambda_1}, \quad \text{where}\\
\lambda_1 &= \frac{-\left(r-\rho-\frac{1}{2}\sigma^2\right) + \sqrt{\left(r-\rho-\frac{1}{2}\sigma^2\right)^2 + 2\sigma^2 r}}{\sigma^2}.
\end{aligned}$$

It is worth noting that (theoretically at least) a perpetual European call is worth the same as the stock price, whereas a perpetual European put is worth zero (look at the limiting behaviour of the Black-Scholes formula).[49]

	European-Style		American-Style	
	Put	Call	Put	Call
No Div's	Black-Scholes Put Formula	Black-Scholes Call Formula	No Exact Formula (use approx. formula, tree, or finite differences)	Black-Scholes Call Formula (early exercise is never optimal)
Lump Sum Div' D	Use $S^*=S-PV(D)$ in Black-Scholes	Use $S^*=S-PV(D)$ in Black-Scholes	No Exact Formula (use approx. formula, tree, or finite differences)	Roll, Geske, Whaley Formula, or Black's Pseudo Formula
Cont. Div's at rate ρ	Use $S^*=Se^{-\rho(T-t)}$ in Black-Scholes (Merton's Formula)	Use $S^*=Se^{-\rho(T-t)}$ in Black-Scholes (Merton's Formula)	No Exact Formula (use approx. formula, tree, or finite differences)	Adjust Roll, Geske, Whaley formula
$S=(\frac{USD}{FX})$	Use $\rho = r_{FX}$ in Merton (Garman-Kohlhagen)	Use $\rho = r_{FX}$ in Merton (Garman-Kohlhagen)	Use $\rho = r_{FX}$	Use $\rho = r_{FX}$
All Cases	Monte-Carlo, Tree, or Finite Differences		Tree or Finite Differences	

Table B.2: Pricing Methods Summary: Plain Vanilla Options
Pricing methods for European- or American-style plain vanilla puts or calls where the underlying pays no dividends, pays a lump sum dividend, pays continuous dividends, or is a foreign currency.

Answer 2.39: If you subtract LIBOR, denoted "L," from both payments, it seems that Party B is paying $24\% - 3 \times L$. This is three times $8\% - L$. The quoted swap is, therefore, equivalent to three swaps, each of which is a swap of LIBOR for 8% fixed (where Party A pays LIBOR, and Party B pays 8%).

Answer 2.40: If you sold the option, you should hold about one-half a

[49]Note that in the case of the perpetual American call, $\lim_{\rho \to 0} \lambda_1 = 1$, and $\lim_{\lambda_1 \to 1} V(S) = S$. That is, with no dividends, the perpetual American call has the same value as the stock—just like the perpetual European call.

European-Style		American-Style	
Path-Independent	Path-Dependent	Path-Independent	Path-Dependent
Tree, Monte-Carlo, or Finite Difference	Monte-Carlo, Finite Difference, Tree (difficult)	Trees or Finite Differences	Finite Differences or Trees (difficult)
... or a Formula if You Can Derive It			

Table B.3: Pricing Methods Summary: Exotic Options
Summary of pricing methods for exotic options that are European- or American-style, path-independent or path-dependent.

share to hedge. If you bought the option, you should short about one-half a share to hedge. If you are at-the-money, there is about a fifty-fifty chance the option finishes in-the-money; and with this expectation, you need about one-half a share to hedge.

Answer 2.41: Mean reversion is the tendency for a variable to return to some sort of long-run mean. Interest rates are generally considered to be mean-reverting: they go up, they go down, but they eventually return to some sort of long-term average. In the case of a mean-reverting stock price, the stock price would tend to be pulled back to the average if the price rises or falls very far. This may reduce volatility and make the option cheaper.

A model of mean reversion makes sense for interest rates, and for stock returns, but it is by no means clear to me that it makes sense for stock *prices*. Bates argues that strong mean reversion in stock prices is implausible because of speculative opportunities available from buying when $S < \bar{S}$ and selling when $S > \bar{S}$ (Bates [1995, pp7–8]). Lo and Wang say that autocorrelation in asset *returns* can increase or decrease σ (and the option price) and that it depends upon the specification of the drift in the model (Lo and Wang [1995, p105]).

Mean reversion is really just negative autocorrelation at some horizon. At short horizons (e.g., daily or weekly), stock index returns are positively autocorrelated (Lo and MacKinlay [1988]). At longer horizons (e.g., three or four years), Fama and French (1988) and Poterba and Summers (1988) say that stock returns are negatively autocorrelated (i.e., mean reverting). However, evidence for this is weak (Richardson [1993]). Lo and MacKinlay (1988, p61) say that longer-term positive autocorrelation is not inconsistent with shorter-term negative autocorrelation (i.e., mean reversion). Peterson et al. (1992) and Lo and Wang (1995) discuss option pricing when asset returns are autocorrelated. Crack and Ledoit (1998) discuss hypothesis testing when asset returns are autocorrelated.

Answer 2.42: Hedging can increase your risk if you are forced to both buy short-dated options and hedge them. In this case, to hedge, you need to short the stock. If the stock price rises up to the strike, and the options (be they puts or calls) expire worthless, then you lose on both the options and the short stock position. By hedging, you end up worse off than if you had not hedged.

Answer 2.43: This is a common question. You can hedge the written put by shorting an asset whose returns are correlated with returns on the underlying stock. Ideally, this would just be the stock itself. However, it is not always possible to short stock. Shorting some index futures would give you an (imperfect) hedge. You need to know either the beta or the correlation of the stock relative to the index to apportion the hedge correctly.

Answer 2.44: People are fed by the area, A, of the pizza. $A = \pi r^2 = \pi \left(\frac{d}{2}\right)^2 = \frac{\pi}{4}d^2$, where d is the diameter. Thus, $d = \sqrt{\frac{4}{\pi}}\sqrt{A}$. Multiplying A by $\frac{8}{6}$ requires a multiplicative change of $\sqrt{\frac{8}{6}}$ in d. That is, $d' = \sqrt{\frac{8}{6}}d = 13.86$ inches. Without a calculator, the square root of $(1+X)$ is roughly $(1+\frac{X}{2})$, so $\sqrt{\frac{8}{6}} \approx \sqrt{1.33} \approx 1.15$. Fifteen percent of 12 is 1.8, so the answer is roughly 13.8 inches.

Why is this a derivatives question? Using the approximation $c = S\sigma\sqrt{\frac{T-t}{2\pi}}$, a question with the same answer is: a six-month at-the-money call has price \$12; what is the price of the eight-month call?

Answer 2.45: You want to be short a put if you expect a price rise. In this case, you expect to keep the option premium when the option expires worthless.

Answer 2.46: A fair price for future delivery of an asset depends upon the spot price and the cost of carry. The cost of carry includes the cost of money (i.e., an interest rate), dividend income, storage costs, and the convenience yield. The only difference between the two pieces of land is the entrance fee to the beach. This is a dividend that lowers the forward price of the beach relative to the field.

Answer 2.47: There are two important points: use of logarithms, and division by $T-1$. Begin by calculating continuously compounded returns (as used in Black-Scholes):

$$\begin{aligned} X_t &\equiv \ln(1+R_t) \\ &= \ln\left(1+\frac{P_t - P_{t-1}}{P_{t-1}}\right) \\ &= \ln\left(\frac{P_t}{P_{t-1}}\right). \end{aligned}$$

With 30 stock prices, you get $T = 29$ returns. Now calculate the standard sample mean and variance. Remember to divide by $T-1 = 28$ in the variance estimator to get an unbiased small sample estimator of historical volatility (DeGroot [1989, p413]).

$$\begin{aligned} \hat{\mu} &= \frac{1}{T}\sum_{t=1}^{T} X_t \\ \hat{\sigma}^2 &= \frac{1}{T-1}\sum_{t=1}^{T}(X_t - \hat{\mu})^2. \end{aligned}$$

Some people may even leave off the "$-\hat{\mu}$" in the $\hat{\sigma}^2$ calculation because mean daily stock returns are typically so tiny compared to volatility, but I prefer to leave it in.

Answer 2.48: The key is default risk, but let's start with a quick swap curve review. Swap rates are fixed rates quoted by dealers against the floating leg (e.g., six-month USD LIBOR) of an interest rate swap. The "swap buyer" is the fixed-rate payer and is said to be "long the swap" (although I have also heard the reverse). The swap curve is inferred from quoted swap rates for different maturities in the same manner that a zero-coupon yield curve (i.e., a "spot curve") is bootstrapped from the yields on coupon bearing bonds of different maturities. Swaps dealers can do customized deals offering different quoted swap rates to companies of different credit rating; however, dealers tend to quote the same swap rate to companies of different credit rating but ask for different amounts of collateral based on the rating (personal communication with a NY dealer [April, 1999]).[50] The collateral and subsequent margin calls essentially resolve the credit issues.

The settlement features of an interest rate swap mean that default risk in a swap is higher than in a eurodollar futures contract but lower than in a bond (Minton [1997, p253]). The reasoning is as follows. The settlement rate for the futures contract is reset daily by market forces, but the swap typically resets only every six months. Both the futures contract and the swap are marked-to-market and use margins, but the futures contract is backed by the triple-A-rated futures clearing corporation as a counterparty of last resort and so the futures contract is less credit-risky than the swap. The swap differs from the bond because no principal changes hands.[51] At initiation, the value of a swap contract is zero; but during the life of the swap, as interest rates rise and fall, the value of the contract can become positive or negative, respectively, to the swap buyer. Although a bondholder is always worried about default risk, the swap buyer worries about default risk only when the swap has positive value. Default on a swap is thus less likely than default on a bond because default on a bond requires only that the company be in financial distress, whereas default on a swap requires both that the company be in financial distress and that the remaining value of the swap be positive. The joint probability of both events needed for swap default is less than the single probability needed for bond default (Minton [1997, p262–263, p267]).

It follows that the coupon rate on a bond will be higher than the quoted swap rate for a swap of the same maturity. This is true for all maturities, so bootstrapping the swap curve from swap rates of swaps of different maturities and bootstrapping the zero-coupon yield curve (i.e., the spot curve) from coupon rates of bonds of different maturities

[50]Minton (1997, p252) confirms that swap quotes assume no credit enhancement (e.g., margins or marking to market).

[51]Although true for an interest rate swap, this is not so in a forex swap, where principal changes hands at initiation and conclusion of the life of the swap.

produces a swap curve strictly below the zero-coupon yield curve.[52] It follows that when you discount the cash flows to the bond using the swap curve, you get a number above that which you would get when you discount the cash flows to the bond using the zero-coupon yield curve (i.e., above par).

Answer 2.49: Try a simple economics argument. The option must cost the same as a replicating portfolio—else there is money to be made. This result is driven by no-arbitrage and is thus independent of risk preferences. I can ease my calculations by assuming risk-neutrality for everyone in the economy. In such an economy in equilibrium, the required return (and thus the expected growth rate and also the discount rate) for all traded securities is the riskless rate. I price the option as if we are in this economy (and the option pricing is immune to this assumption).

Answer 2.50: Options live in the future, not the past: Today is the first day of the rest of the life of a traded option. Setting aside problems with volatility smiles and skews, the implied volatility (or "implied standard deviation") is a market-consensus forecast of volatility over the remaining life of the option. It would be logical, therefore, that implied volatility is a better predictor of future volatility than is historical volatility. Indeed, this is found empirically for both FOREX (Xu and Taylor [1995]) and for equity indices (Fleming [1998]).

Answer 2.51: Assume that $S = \$100$ so that the call has strike $X_c = 110$, and the put has strike $X_p = 90$. The distribution of future stock prices is Lognormal in the Black-Scholes model and is thus skewed with its mean higher than the median, which is in turn higher than the mode (the peak). The median of the distribution of future stock prices is $Se^{(r-\frac{1}{2}\sigma^2)}$. The term $(r - \frac{1}{2}\sigma^2)$ tends to be close to zero, so the median is approximately S. It follows that $X_c = 110$ and $X_p = 90$ are roughly equidistant from the median. Half the distribution is above the median and half is below. The probability that the call finishes in-the-money is $P(S(T) > X_c) = P(S(T) > 110) \approx \frac{1}{2} - P(100 < S(T) < 110)$.

[52] In early 1999, two-year swap rates were about 40 basis points higher than US treasuries, about 5 basis points lower than the yields on AAA-rated debt, about 30 basis points lower than the yields on AA-rated debt and about 40 basis points lower than the yields on BBB-rated debt.

The probability that the put finishes in-the-money is $P(S(T) < X_p) = P(S(T) < 90) \approx \frac{1}{2} - P(90 < S(T) < 100)$. However, the distribution is so skewed that $P(90 < S(T) < 100) \gg P(100 < S(T) < 110)$ (at the median the probability density function is downward sloping). Thus, $P(S(T) > X_c) \gg P(S(T) < X_p)$, and the call is more likely to finish in-the-money than the put, and this typically makes the call more valuable.

Answer 2.52: The derivative value V must satisfy the Black-Scholes PDE (Wilmott et al. [1993]):

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

These derivatives are just the theta (Θ), gamma (Γ), and delta (Δ), respectively, so we rewrite the PDE as

$$\Theta + \frac{1}{2}\sigma^2 S^2 \Gamma + rS\Delta - rV = 0.$$

The last two terms may be written as $r(S\Delta - V)$, and they offset to some extent. The entire PDE adds to zero, so that leaves $\Theta + \frac{1}{2}\sigma^2 S^2 \Gamma$ taking a value close to zero. This means that Θ and Γ are typically going to be of opposite signs. Not only that, but their magnitudes are going to be correlated. For example, if Θ is large and negative then Γ is probably large and positive (e.g., an at-the-money call close to maturity has these properties).

There is one exception amongst plain vanilla puts and calls. A deep in-the-money European put has positive Θ, and as long as it is not extremely deep in-the-money, it also has positive Γ.

> **Story:** *1. Man wore jogging suit to interview for position as financial vice president. 2. Interrupted to phone his therapist for advice on answering specific interview questions.*
>
> *Interview Horror Stories from Recruiters*
>
> *Reprinted by kind permission of* MBA Style Magazine
>
> ©*1996–2004 MBA Style Magazine,* www.mbastyle.com

Answer 2.53: If you said you have an 80% chance of getting \$20, and a 20% chance of getting nothing, giving an expected payoff of \$16, which you then discount at zero to get an answer of \$16 for the call value, you are wrong! Sure enough, the call does have an expected payoff of \$16 in the real world, but the discount rate is not zero. The discount rate is some leveraged version of the discount rate on the stock, and you do not have that information. Try again, then come back here for the answer below.

We do not know the discount rate on the stock. We do not know the discount rate on the option. We must use risk-neutral valuation. The risk-neutral probability π^* of an up move in the stock satisfies:

$$\begin{aligned} S &= e^{-r(\Delta t)}\left[\pi^* Su + (1-\pi^*)Sd\right], \\ \text{that is, } \$100 &= \pi^*\$130 + (1-\pi^*)\$70, \end{aligned}$$

where r is the riskless rate (zero here), u is the multiplicative "up" growth factor in the stock (1.30 here), and d is the multiplicative "down" growth factor in the stock (0.70 here). See Crack (2004) or you favourite option pricing book for deeper details of binomial/lattice pricing. Simple algebra yields $\pi^* = 0.50$. The value of the call is then

$$\begin{aligned} c &= e^{-r(\Delta t)}\left[\pi^* \max(0, Su - X) + (1-\pi^*)\max(0, Sd - X)\right] \\ &= 1 \cdot [0.50 \cdot (\$130 - \$110) + 0.50 \cdot (\$0)] \\ &= \$10. \end{aligned}$$

Answer 2.54: The product call pricing formula is so simple that you could simply say "here is the answer, it looks like regular dividend-adjusted Black-Scholes but you replace $S(t)$ by the product $S_1(t) \times S_2(t)$, and you replace σ by $\sigma' \equiv \sqrt{\sigma_1^2 + \sigma_2^2 + \rho\sigma_1\sigma_2}$ where ρ is the instantaneous correlation between the Wiener processes (i.e., Brownian motions) driving S_1 and S_2, and of course the answer is symmetric in S_1, S_2, and their associated 'dividend yields'." However, the derivation is very instructive in risk-neutral pricing, PDE's, and similarity solutions, and I cannot find it in my books so I think it belongs here.

One application of the product call is to the pricing of foreign equity options struck in a domestic currency (Haug [1997, pp102–103]). For example, a US investor has the right to buy one share of NTT corporation stock (trading in Tokyo at JPY price S_2), but the call option strike price is in USD.[53] In this case, the payoff is $\max[S_1(T) \times S_2(T) - X, 0]$, where S_1 is the $\frac{USD}{JPY}$ exchange rate, S_2 is the JPY price of NTT, and T is the expiration date.

Make the following definitions:

$$
\begin{aligned}
S_1(t) &= \frac{\text{USD}}{\text{JPY}}(t) \\
S_2(t) &= \frac{\text{JPY}}{\text{Share of NTT}}(t) \\
r_{US} &= \text{US riskless interest rate} \\
r_{JP} &= \text{Japanese riskless interest rate} \\
q &= \text{NTT's continuous dividend yield} \\
dS_1 &= r_1 S_1 dt + \sigma_1 S_1 dw_1 \\
dS_2 &= r_2 S_2 dt + \sigma_2 S_2 dw_2 \\
\sigma_1 &= \text{Volatility of } dS_1/S_1 \text{ process} \\
\sigma_2 &= \text{Volatility of } dS_2/S_2 \text{ process} \\
r_1 &= \text{Drift of } dS_1/S_1 \text{ process} \\
r_2 &= \text{Drift of } dS_2/S_2 \text{ process} \\
\rho dt &= E[(dw_1) \cdot (dw_2)] \;=\; \text{instantaneous correlation} \\
X &= \text{USD-denominated strike price.}
\end{aligned}
$$

So, what exactly are r_1 and r_2 in a risk-neutral world? The answer depends upon whether we look from a US or a Japanese perspective (Hull [1997, p301]). We shall use the US perspective. For S_1 from the US perspective, the risk-neutral process has $r_1 = r_{US} - r_{JP}$. For S_2 from the Japanese perspective, $r_2 = r_{JP} - q$, but from the US perspective, $r_2 = r_{JP} - q + (-\rho) \cdot \sigma_1 \sigma_2$, where $-\rho$ is the instantaneous correlation between the Wiener processes driving the two JPY-denominated processes $S_2(t)$ and $\frac{JPY}{USD}(t)$. This correlation is the negative of that between the Wiener processes driving $S_2(t)$ and $S_1(t) = \frac{1}{\frac{JPY}{USD}(t)}$ (Hull

[53] Please note that this is *not* a quanto option. Quantos are currency translated options, and so is this, but a quanto takes the JPY price of the foreign security and simply replaces the JPY symbol with a USD symbol when calculating the payoff (Haug [1997, p104]; Hull [1997, p298]; Wilmott [1998, p155]). The JPY security payoff is said to be "quantoed" into USD.

[1997, p301]). Thus, the risk-neutral drifts from the US perspective are

$$r_1 = r_{US} - r_{JP}, \text{ and } r_2 = r_{JP} - q - \rho\sigma_1\sigma_2,$$

but we shall continue to work with r_1, and r_2, and then plug these in at the end. From our stochastic calculus training we know that as long as dynamic replication is possible, then de-trended prices of traded assets are martingales in the risk-neutral economy (Huang [1992]; Crack [1999, Section 5.2]). A bullet-point review is called for before proceeding.

• RISK – NEUTRAL PRICING REVIEW •

- The technical requirement for dynamic replication to be possible is described nicely in Jarrow and Rudd (1983). Essentially, it requires that for very small time horizons the value of the derivative and the value of the underlying(s) be perfectly linearly correlated. A diffusion or a simple jump process satisfies this, but if the underlying stock price follows a jump-diffusion process (regardless of whether the jump size is deterministic, stochastic, diversifiable, or non-diversifiable), then a replicating portfolio cannot be formed, and the no-arbitrage pricing method fails (Cox and Rubinstein [1985, Chapter 7]; Merton [1992]).

- If dynamic replication is possible, then by no-arbitrage the value of the derivative equals the start-up cost of a replicating portfolio.

- If the replication recipe is known (perhaps via an equilibrium CAPM pricing approach as in the original Black and Scholes [1973] paper), then no two economic agents can disagree on the correct arbitrage-free price of the derivative. It follows that regardless of what we assume about the preferences of the agents in the economy, the pricing of the derivative will be the same.

- We ease our calculations substantially by proceeding *as if* the agents in the economy are risk-neutral.[54] That is, although they see the risk, they ignore it completely.

[54] Important: We are not assuming anyone is really risk-neutral. It is simply that options prices are immune to assumptions about risk preferences, and this proves to be a very helpful assumption.

- In a risk-neutral economy people care only about expected return, so in equilibrium all traded assets must offer the same expected return (else investors would still be shorting low-yield securities to invest in high-yield ones and we would not yet be in equilibrium). The existence of a government-backed fixed-rate riskless asset means that the riskless rate is the equilibrium required return on all securities in this hypothetical world.

- If risk is not priced by agents in the economy, then traded security prices (including derivatives) are simply discounted expected payoffs where discounting uses the riskless rate, and all traded security prices are assumed to drift upwards at the riskless rate (less any dividend yield, of course—so that total yield is the riskless rate). If risk were priced, then discount rates would need to be risk adjusted, perhaps via the CAPM (Arnold and Crack [2003]).

- Let $B(t) \equiv e^{rt}$ denote the price of a riskless money market instrument (i.e., you invest \$1 at time 0, and it grows at riskless rate r). Then $B(t)$ drifts upward at the riskless rate. The money market account serves as a benchmark for performance in both the real and risk-neutral worlds. It seems natural to express other asset prices in terms of units of this asset.[55] That is, instead of looking at security price $P(t)$, look at $\frac{P(t)}{B(t)}$.

- With $B(t)$ drifting upward at the riskless rate, and $P(t)$ expected to drift upward at the same rate in equilibrium in the risk-neutral world, it follows that $\frac{P(t)}{B(t)}$ is expected to have no drift. Another way to say this is that for any $\Delta t > 0$,

$$E^* \left[\frac{P(t+\Delta t)}{b(t+\Delta t)} \bigg| \frac{P(t)}{B(t)} \right] = \frac{P(t)}{B(t)},$$

  where E^* denotes expectation in the risk-neutral world.

- Let $P^\dagger(t) \equiv \frac{P(t)}{B(t)}$, then the previous result says that for any $\Delta t > 0$,

$$E^* \left[P^\dagger(t+\Delta t) \,\middle|\, P^\dagger(t) \right] = P^\dagger(t).$$

[55] This is referred to as a change of "numeraire." A numeraire is a base unit of measurement. This is similar to changing units of measurement from USD to GBP, say, except that here we choose a USD-denominated money market account instead of GBP.

That is, the best guess of where $P^\dagger$ will be in the future (in the risk-neutral world) is where it is today. This is akin to the efficient markets hypothesis. A random variable with this property is called a "martingale."[56]

- When we assume that traded securities prices have required returns equal to the riskless rate in the risk-neutral world, we are really just redistributing the probabilities we associate with possible final security price outcomes.[57] However, some things stay the same. For example, if a stock price outcome occurs with probability **0** in the real world, then it still occurs with probability **0** in the risk-neutral world (thus, the range of possible outcomes does not change, only their probability of occurrence; and the transformation of probabilities moves the expected return on IBM, say, from 12% per annum to whatever the T-bill yield happens to be). Similarly, if a stock price outcome occurs with probability **1** in the real world, then it still occurs with probability **1** in the risk-neutral world.

- In probability theory, the mathematical function that allocates probability weight to outcomes in the sample space is called a "measure." Two probability measures that reassign probabilities to outcomes without changing the range of possible outcomes (as above) are called "equivalent measures."[58]

- Thus, in the risk-neutral world, we reallocate probabilities in an equivalent manner (i.e., same range of possible outcomes), and the price of any traded asset—when "de-trended" by the money market account—follows a martingale. The probability measure (i.e., allocation of probabilities to outcomes) in the risk-neutral world is thus called an "equivalent martingale measure." You see

[56] Note that there is a competing stock-numeraire world where if the bond de-trended by the stock follows a martingale, then $N(d_1)$ in the Black-Scholes formula is the probability that the call finishes in the money (see Crack [2004] for details). It is a Z-score argument similar to the one that establishes $N(d_2)$ as the probability that the call finishes in the money in a world in which the stock de-trended by the bond follows a martingale.

[57] Note the word "traded" here. A futures price, for example, is not the price of a traded asset, so its drift need not be r.

[58] The relationship between the two measures is captured by the Radon-Nikodym derivative. See Baxter and Rennie (1998, p65) for simple intuition and Musiela and Rutkowski (1992, pp114, 121) for the advanced mathematics.

this expression in the more advanced literature.

- Two natural derivative pricing methods fall out of all of the above. The first uses discounted expected payoffs, the second uses PDE's.
- First Method (Cox and Ross [1976]): Let V be the derivative price we seek, then the martingale property applied to de-trended V (i.e., $V^{\dagger} = V/B = Ve^{-rt}$) implies

$$\begin{aligned} V^{\dagger}(t) &= E^{*}\left[V^{\dagger}(T)\left|V^{\dagger}(t)\right.\right] \\ \Rightarrow V(t)e^{-rt} &= E^{*}\left[V(T)e^{-rT}\left|V(t)\right.\right] \\ \Rightarrow V(t) &= e^{-r(T-t)}E^{*}\left[V(T)\left|V(t)\right.\right]. \end{aligned}$$

  I derive Black-Scholes in Crack [2004] using precisely this approach: discounted expected payoff in a risk-neutral world.

- Second Method (Harrison and Kreps [1979]): Let V be the derivative price we seek, then the martingale property applied to detrended V (i.e., $V^{\dagger} = V/B = Ve^{-rt}$) implies that $dV^{\dagger}$ has no time trend; that is, no drift. We can apply Itô's Lemma to $V^{\dagger}$ to calculate

$$dV^{\dagger} = [\text{time trend}]dt + \sum_{i}[\text{diffusion coefficients}]_{i}dw_{i},$$

  where dw_i is the i^{th} Brownian motion driving the underlyings. If V is a function of $S(t)$ and t only, and $dS(t) = rSdt + \sigma Sdw$ then

$$\begin{aligned} dV^{\dagger}(S(t),t) &= d[V(S(t),t)e^{-rt}] \\ &\overset{\text{Itô}}{=} \left(V_S dS + V_t dt + \frac{1}{2}V_{SS}(dS)^2\right)e^{-rt} \\ &\qquad -rVe^{-rt}dt \\ &= \left(\frac{1}{2}V_{SS}\sigma^2 S^2 + V_S rS + V_t - rV\right)e^{-rt}dt \\ &\qquad +V_S\sigma Se^{-rt}dw, \end{aligned}$$

  where we used $(dw \cdot dw) = dt$, and $(dt \cdot dw) = 0$ (Merton [1992, p122–123]).

- However, $V^{\dagger} = Ve^{-rt}$ is a martingale in the risk-neutral world by construction, so it must be that there is no drift term. Thus, we deduce that

$$\frac{1}{2}V_{SS}\sigma^2 S^2 + V_S r S + V_t - rV = 0.$$

If we know the boundary conditions, we may now solve this (Black-Scholes) PDE to find the option value $V(S(t), t)$. A different initial process for dS will yield a different PDE. We will now do exactly this for the product call.

• $\mathcal{END\ OF\ RISK-NEUTRAL\ PRICING\ REVIEW}$ •

The time-t value of the European-style product call expiring at time-T is simply its discounted expected payoff in a risk-neutral world:

$$V(S_1(t), S_2(t), t) = e^{-r_{US}(T-t)}\ E^*\{\max[S_1(t)S_2(t) - X|\Omega_t]\},$$

where E^* denotes expectation taken with respect to the risk-neutral probability measure from the US perspective, and Ω_t is the time-t information set. We could work this out directly (it would be a double integral with respect to the two Brownian motions), but let us instead use the PDE approach.

Given the nature of the product call, I am going to guess that the solution is a function of only two variables, not three: $V(S_1, S_2, t) = \kappa H(\eta, t)$ for some constant κ and $\eta = S_1 \cdot S_2$ (see analogous guess in Wilmott [1998, p155]). I will need to use Itô's Lemma soon so I will now work out all the partial derivatives for the change of variables.

$$\begin{aligned}
\frac{\partial}{\partial S_1} &= \frac{\partial \eta}{\partial S_1}\frac{\partial}{\partial \eta} = S_2 \frac{\partial}{\partial \eta} \\
\frac{\partial}{\partial S_2} &= \frac{\partial \eta}{\partial S_2}\frac{\partial}{\partial \eta} = S_1 \frac{\partial}{\partial \eta} \\
\frac{\partial^2}{\partial S_1^2} &= S_2 \frac{\partial \eta}{\partial S_1}\frac{\partial^2}{\partial \eta^2} = S_2^2 \frac{\partial^2}{\partial \eta^2} \\
\frac{\partial^2}{\partial S_2^2} &= S_1 \frac{\partial \eta}{\partial S_2}\frac{\partial^2}{\partial \eta^2} = S_1^2 \frac{\partial^2}{\partial \eta^2} \\
\frac{\partial^2}{\partial S_1 \partial S_2} &= \frac{\partial}{\partial \eta} + S_2 \frac{\partial \eta}{\partial S_2}\frac{\partial^2}{\partial \eta^2} = \frac{\partial}{\partial \eta} + S_1 S_2 \frac{\partial^2}{\partial \eta^2},
\end{aligned}$$

and $\frac{\partial}{\partial t}$ is unchanged.

From our risk-neutral pricing review, we know $Ve^{-r_{US}t}$ is a martingale in the risk-neutral world, so it has no time trend. We need only find the coefficient of dt in $d[Ve^{-r_{US}t}]$ and equate it to zero. There are two Brownian motions, so we need the two dimensional Itô's Lemma (Merton [1992, p122]; Hull [1997, p304]), and $d[Ve^{-r_{US}t}]$ is itself a geometric Brownian motion (GBM):

$$\begin{aligned}
&d[Ve^{-r_{US}t}] \\
&= -r_{US}Ve^{-r_{US}t}dt + e^{-r_{US}t}dV \\
&\overset{\text{Itô}}{=} -r_{US}Ve^{-r_{US}t}dt + e^{-r_{US}t} \times \left(\frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S_1}dS_1 + \frac{\partial V}{\partial S_2}dS_2\right. \\
&+ \left.\frac{1}{2}\frac{\partial^2 V}{\partial S_1^2}(dS_1)^2 + \frac{1}{2}\frac{\partial^2 V}{\partial S_2^2}(dS_2)^2 + \frac{\partial^2 V}{\partial S_1 \partial S_2}(dS_1 \cdot dS_2)\right) \\
&= e^{-r_{US}t} \times \left\{\left[-r_{US}V + \frac{\partial V}{\partial t} + \frac{\partial V}{\partial S_1}r_1 S_1 + \frac{\partial V}{\partial S_2}r_2 S_2\right.\right. \\
&+ \left.\frac{1}{2}\frac{\partial^2 V}{\partial S_1^2}\sigma_1^2 S_1^2 + \frac{1}{2}\frac{\partial^2 V}{\partial S_2^2}\sigma_2^2 S_2^2 + \frac{\partial^2 V}{\partial S_1 \partial S_2}\rho\sigma_1\sigma_2 S_1 S_2\right] dt \\
&+ \left.\left[\frac{\partial V}{\partial S_1}\sigma_1 S_1 dw_1 + \frac{\partial V}{\partial S_2}\sigma_2 S_2 dw_2\right]\right\}, \text{ which is a GBM,}
\end{aligned}$$

where we used the earlier definitions of dS_1, dS_2, and so on.

We now take the time trend coefficient of dt, equate it to zero, use the change of variables $V(S_1, S_2, t) = \kappa H(\eta, t)$, where $\eta = S_1 S_2$, and drop the common terms $e^{-r_{US}t}\kappa$:

$$\begin{aligned}
-r_{US}H + H_t + r_1 S_1 S_2 H_\eta + r_2 S_1 S_2 H_\eta + \frac{1}{2}\sigma_1^2 S_1^2 S_2^2 H_{\eta\eta} \\
+\frac{1}{2}\sigma_2^2 S_1^2 S_2^2 H_{\eta\eta} + S_1 S_2 \sigma_1 \sigma_2 \rho [H_\eta + S_1 S_2 H_{\eta\eta}] = 0
\end{aligned}$$

Now collect terms and use $\eta = S_1 S_2$:

$$\begin{aligned}
H_t + \eta H_\eta(r_1 + r_2 + \rho\sigma_1\sigma_2) + \frac{1}{2}(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)\eta^2 H_{\eta\eta} \\
-r_{US}H = 0.
\end{aligned}$$

Now plug in $r_1 = r_{US} - r_{JP}$ and $r_2 = r_{JP} - q - \rho\sigma_1\sigma_2$, and let $\sigma' \equiv \sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}$ to deduce:

$$H_t + \eta H_\eta(r_{US} - q) + \frac{1}{2}\sigma'^2 \eta^2 H_{\eta\eta} - r_{US}H = 0.$$

The latter PDE is just the regular Black-Scholes PDE with continuous dividends and with special volatility σ'. Recalling our definition of η, we get a "similarity solution" by using what we already know about the Black-Scholes solution to this PDE:

$$\begin{aligned}
c(t) &= S_1(t)S_2(t)e^{-q(T-t)}N(d_1) - e^{-r_{US}(T-t)}XN(d_2), \text{ where} \\
d_1 &= \frac{\ln\left(\frac{S_1(t)S_2(t)}{X}\right) + (r_{US} - q + \frac{1}{2}\sigma'^2)(T-t)}{\sigma'\sqrt{T-t}}, \\
d_2 &= d_1 - \sigma'\sqrt{T-t}, \text{ and} \\
\sigma' &= \sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}.
\end{aligned}$$

Reassuringly, this is identical to Equation (2.65) in Haug (1997, p103).

> **Story:** *1. Said he wasn't interested because the position paid too much. 2. While I was on a long-distance phone call, the applicant took out a copy of* Penthouse, *and looked through the photos only, stopping longest at the centerfold.*
>
> *Interview Horror Stories from Recruiters*
> *Reprinted by kind permission of* MBA Style Magazine
> ©*1996–2004 MBA Style Magazine,* www.mbastyle.com

Answer 2.55: An Asian option is an average rate option. The underlying is a time series average of prices. Changes in average prices are much less volatile than changes in consecutive prices. Other things being equal, this lower volatility makes Asian options less expensive than plain vanilla options.

Answer 2.56: If the riskless rate is positive, and there are no dividends, early exercise is not optimal for an American-style call, and the European and American call have the same value. If the riskless rate is zero, then there is no incentive for early exercise of an American-style put. In this case, the European and American put have the same value.

Of course, that's little consolation to you if you are short an American-style option, a retail investor decides to exercise non-optimally, and you are assigned.

Appendix C

Other Financial Economics Answers

This appendix contains answers to the questions posed in Chapter 3.

Answer 3.1: This is a very old problem, and a common interview question. The probability that the first head occurs on toss k is $\left(\frac{1}{2}\right)^k$; this event carries with it a payoff of $\$2^k$. The contribution of toss k to the expected payoff is thus $\left(\frac{1}{2}\right)^k \times \$2^k = \$1$. This is the same for each k. The expected payoff to the game as a whole is the summation over all k of these payoffs. This is $\$1 + \$1 + \$1 + \cdots = \$\infty$. The expected payoff to the game is infinite!

This is called the "St. Petersburg Game." The fact that the expected payoff to the game is infinite, and that no one in his or her right mind would pay more than a few hundred dollars to play, is why it is sometimes called the "St. Petersburg Paradox." There are several ways that you can think about this sensibly.

One way is to note that "value" is not the same thing as "expected payoff";[1] value equals *utility* of expected payoff. Most people cannot distinguish between very large amounts of money.[2] This means that $\$2^{50}$ is not worth twice $\$2^{49}$. However, these very large amounts are

[1]It is important to note that the Weak Law of Large Numbers fails if the expectation is not finite (Feller [1968, pp251]).

[2]Bernoulli ([1738]; [1954]) suggests that utility of payoffs should depend upon how wealthy you are. For a practitioner's view of utility, see Kritzman (1992a).

counted in exactly this way when calculating expected payoff to the game as a whole. If you think that $\$2^k = \$2^{50}$ (essentially) for all $k \geq 50$, then the expected payoff to the game is finite:

$$\$50 + \$2^{50} \times \left(\frac{1}{2^{51}} + \frac{1}{2^{52}} + \frac{1}{2^{53}} + \ldots \right) = \$51.$$

A spread could be quoted around this, maybe (\$10, \$200). How much would you pay your customer to play? How much would you charge your customer?

A second way to think about this is in terms of default risk.[3] We need to quote the bid (what we pay) and the ask (what the customer pays). For the bid, it is the customer's default risk we need to worry about. Let us assume a wealthy customer who defaults above one million dollars. In this case, the customer defaults after (about) 20 tosses. Assuming the investment bank is of large scale, a payoff from the customer between two dollars and one million dollars is of relatively small size. The investment bank takes such bets every day, and this one is uncorrelated with all the others. At this level, we could argue that the investment bank is risk-neutral and so the bid is exactly \$20 with no risk premium.

For the ask, it is the company that risks bankruptcy and default. Let us assume that the company files for bankruptcy after losing one billion dollars (on the order of magnitude of Barings, and Metallgesellschaft)—approximately $\$2^{30}$. The expected value of the game to the customer is thus about \$30—the bank defaults after 30 tosses. However, your career and the holdings of all the shareholders can be destroyed by this bet, so you had better add a considerable risk premium. You might want to go all the way up to \$200 and quote a bid-ask of (\$20, \$200)—it depends upon your degree of risk aversion.

Each of these two solutions uses a truncation method. Another related way to think about this is in terms of feasibility. If it does take more than 50 tosses to get a head, then the payoff is not feasible because $\$2^{50}$ is more dollars than there are atoms in the universe, and whoever sold the ticket to the game is—by the laws of physics—unable to pay. See also Feller (1968, pp251–253).

[3] I thank Olivier Ledoit for suggesting this solution technique. I am responsible for any errors.

Answer 3.2: This is a frequent question. Assuming continuously compounded returns follow an arithmetic Brownian motion (see Crack [2004]), variance of returns grows linearly with the compounding period. This is because consecutive returns in a random walk are independent, and the variance of a sum of independent random variables is just the sum of the variances. This means that the four-year σ^2 equals four times the one-year σ^2. It follows that the four-year σ is two times the one-year σ. The answer is, therefore, 20%.

Answer 3.3: This is a very common term-structure question. You should be able to do this in your head almost instantly. Think of it this way: the rate over the first five years and the rate over the second five years must average out to give the rate over the full 10 years. That is, the average of 10% and the unknown forward rate must give 15%. The unknown must be 20%. To work it out quickly, note that the unknown (20%) is as far above the average (15%) as the known (10%) is below it.

In fact, if you work it out exactly, the forward rate is

$$\left[\frac{(1.15)^{10}}{(1.10)^5}\right]^{\frac{1}{5}} - 1 = 20.227\%.$$

You are making a "first-order" approximation when you do the simple averaging, but you end up quite close. For a practitioner's viewpoint on the term-structure of interest rates, take a look at Kritzman (1993b).

Answer 3.4: The "yield" on a bond is the "internal rate of return" or "yield-to-maturity" or "promised yield"; it is what you earn if you hold the bond to maturity. The "rate of return" on a bond is the internal rate of return of the realized cash flows to the bond-holder. If the bond is sold before maturity, the (realized) rate of return can be positive or negative.

Suppose you buy a bond promising 5%, but interest rates rise dramatically soon after your purchase. If you then sell the bond, you record a capital loss and a negative rate of return. However, if you hold the bond to maturity, you get your promised 5%.

Answer 3.5: Chaos theory came out of MIT in the early 1960's. Professor Edward N. Lorenz (now Professor Emeritus in the department of

Earth, Atmospheric and Planetary Science) discovered that computer-simulated nonlinear mathematical equations describing the evolution of weather patterns are very sensitive to the starting values of the variables (Lorenz [1963]).[4]

This "sensitive dependence on initial conditions" is the first of three characteristics most often associated with chaos theory. The second characteristic is that the nonlinear systems describing chaotic systems are nonrandom. That is, they are "deterministic," not "stochastic." However, the output of the (often very simple nonlinear) systems can appear quite random. The third characteristic is "self-similarity": the physical system looks similar at different levels of magnification. It is self-similarity that gives rise to the "fractals" that you may have seen elsewhere. Fractals are often associated with the mathematician Benoit Mandelbrot (now at Yale).

There are several different definitions of "chaos" in the literature. These definitions are beyond the scope of this book. See Brock et al. (1991, pp8–17) for further details. For a low-level broad introduction to chaos, see Gleick (1987).

Can you use chaos theory in finance? This was a hot topic in the late 1980's and early 1990's. Many academic economics and finance papers were written on the subject. The few that made any sense found nothing reliable. The others were written by ignorant people who jumped on the bandwagon; they should never have published their empty papers.

After reading more than 150 journal articles and half a dozen books on chaos theory and writing a 100-page graduate level thesis on chaos theory applied to financial economics and publishing one paper in the *Journal of Finance*, I am quite pessimistic. My co-author and I hypothesized considerable discreteness-induced bias in the popular "BDS test" for chaos in equity data (Crack and Ledoit [1996]). It seems that our hypothesis has now been confirmed (Krämer and Runde [1997]).

If you want to predict stock returns, I recommend that you use neural

[4]I had the pleasure of attending some Independent Activities Period (IAP) classes taught by Prof. Lorenz at MIT in 1994/1995. He reminded me a little of a slim Dave Thomas (you know, the Wendy's guy). He was younger than I expected, and very good-natured.

nets or some other nonlinear technique. In my opinion, any predictability that you can discern with chaos theory (e.g., "nearest neighbour" prediction techniques) is better investigated using the other nonlinear techniques available to you. Give it up—chaos theory is great in the physical sciences, but it is a lost cause in finance.

Answer 3.6: Look at Table C.1 on page 199. The slope of the price-yield curve is $-\frac{D}{(1+r)}P$, where D is Macaulay duration, P is bond price, and r is yield. Changing slope (i.e., curvature) is driven almost entirely by changing P, because Macaulay duration, D, changes very little with changing yield, r (Crack and Nawalkha [2001]). D does, however, fall slowly with rising yield for a standard bond with coupons, and this does contribute marginally to curvature.

Note that curvature (i.e., changing slope) of the plot does *not* always imply that the Macaulay duration of a bond is changing (Crack and Nawalkha [2001])! This is a common misconception (it is easy to misconstrue this in Fabozzi and Fabozzi [1995, pp97–98]). For example, consider a zero-coupon bond with ten years to maturity. The plot of bond value versus bond yield is downward sloping with curvature. Whatever the yield, however, the bond's duration is ten years because it is a ten-year zero

A mathematical explanation of the convexity: you know that the curve slopes downward, it goes to a vertical asymptote at yield -1 and a horizontal asymptote at yield infinity. You know that the curve must be smooth because the pricing relationship is simple and well-behaved. The only way to get a well-behaved smooth curve in this situation is to have it be convex.

Answer 3.7: This question is an interesting intersection of theory and empirical reality. The question is not necessarily well-posed, but you should do your best to answer it. I give what I think is the best answer possible.

If the empirical security market line (SML) is wholly above the theoretical one, this means that stocks are underpriced relative to the CAPM. I propose two possible causes: First, maybe there is only one risk factor (the Market), but market participants require higher compensation per unit of beta-risk than suggested by the CAPM; second, maybe there is more than one risk factor, and market participants require compensation for factors not mentioned by the CAPM. Conversely, if the empirical SML is wholly below the theoretical one, then stocks are overpriced relative to the CAPM. In this case, market participants do not require as much compensation per unit of beta-risk as theory suggests.

I think the best answer is to say that the CAPM does not account for all priced risk factors. It is likely, however, that beta is priced. It follows that stocks require a premium over and above that suggested by CAPM, and you could think of this as an empirical SML plotting above the theoretical one. For more on factor models and estimation, see Kritzman (1993a).

There have been several papers pronouncing the CAPM either dead or alive (Wallace [1980]; Fama and French [1992]; Black [1993]; Fama and French [1996]). For a friendly introduction to the CAPM, see Mullens (1992).

You should note that there are some theoretical problems with both the question and my answer. It is quite difficult (if not impossible) to get either of the empirical SML's mentioned. This is not because the CAPM is "correct," or because there is only one risk factor. Rather, it is because there is a *very* tight mathematical relationship between betas and returns (Sharpe [1964]; Roll [1977a]). You would certainly need that the market proxy is not mean-variance efficient to get the plots suggested. It is probably not sufficient to simply assume that there are risk factors not accounted for by the CAPM. Go with the answer above, but realize that there is more here than meets the eye.

Answer 3.8: This question is very similar to Question 3.3 (and is just as common). You should be able to do it in your head almost instantly. If you cannot, then go back and try Question 3.3 again before reading on.

The rate over the first year and the rate over the second year must average out to give the rate over the full two years. That is, the average of 7.15% and the unknown forward rate must give 7.60%. The unknown rate must be around 8.05% (remember, it is as far above 7.60% as 7.15% is below 7.60%).

In fact, if you work it out exactly, the forward rate is

$$\left[\frac{(1.0760)^2}{(1.0715)^1}\right] - 1 = 8.052\%.$$

You are making a "first-order" approximation when doing the simple averaging, and the answer is quite accurate.

Answer 3.9: This is introductory finance theory; it uses no-arbitrage and not much more. Assume for the sake of simplicity that interest rates are constant at r per unit time, today is time t, and the forward contract matures at time T. The forward price, $F(t, T)$, is related to the spot price, $S(t)$, as follows:

$$F(t, T) = S(t)e^{r(T-t)} \geq S(t).$$

The discount bond sells at a forward *premium* because of no-arbitrage.

The coupon bond is a different story. If you assume a continuous coupon of ρ per unit time, then the forward price, $F(t, T)$, is related to the spot price, $S(t)$, as follows:

$$F(t, T) = S(t)e^{(r-\rho)(T-t)} \leq S(t)$$

(the inequality follows because I assumed that $r \leq \rho$). The coupon bond sells at a forward *discount* because of no-arbitrage.

For a practitioner's view on futures, forwards, and hedging, see Kritzman (1993c).

Answer 3.10: This is a classic question, and a very good test of your dexterity with elementary finance theory. If you have not yet figured it out, and you are peeking at the answers for a hint, I strongly recommend that you go back to the question and try again; read no further. If you are still reading, here is a hint: think about your investment horizon, and an immunization strategy. Now go back and try again.

Your investment horizon is very short. You want to profit from the change in the relationship between short- and long-term rates. However, you want to protect yourself from the level of the yield curve. That is, you want your position to be insensitive to parallel shifts in the yield curve, but positively sensitive to a steepening. This suggests that you should go short long-term debt, go long short-term debt, and match both the duration and price of the positions (i.e., use very low coupon short-term debt and very high coupon long-term debt).[5]

You may think of this as a "zero-duration" portfolio (to match your horizon). However, in just the same way that a zero net investment stock portfolio has no well-defined beta but can still be market-neutral, a zero net investment bond portfolio has no well-defined duration but can still be insensitive to parallel shifts in the yield curve.

Traders tell me that this strategy originated with the Salomon Bond arbitrage ("bond-arb") group. However, it is now so well known that profits may be slim.

For more on "Yield Curve Strategies," see the excellent papers by Jones (1991) and Litterman and Scheinkman (1991). Jones describes the statistical relationship between changes in level, slope, and curvature of the yield curve.

Answer 3.11: For a standard bond, the Macaulay duration (Macaulay [1938]) is just the weighted-average term-to-maturity of the bond:

$$D \equiv \frac{\sum_{t=1}^{T} \frac{t \times C_t}{(1+r)^t}}{\sum_{s=1}^{T} \frac{C_s}{(1+r)^s}} = \sum_{t=1}^{T} \omega_t \times t, \quad \text{where } \omega_t \equiv \frac{\frac{C_t}{(1+r)^t}}{\sum_{s=1}^{T} \frac{C_s}{(1+r)^s}},$$

and C_t are the cash flows (both coupon and principal). The weights ω_t are applied to the timing of the bond's cash flows. Each weight is equal to the present value of the particular cash flow as a proportion of the total value of the bond. It follows that the duration of a zero-coupon bond equals its term-to-maturity—because the weight of the final cash flow is unity (i.e., +1). Duration is measured in units of time, as is the term-to-maturity.[6]

[5]If you cannot match durations of the positions, you can match on the product of duration × price. However, this will no longer be a zero net investment strategy.

[6]Duration is usually measured in years, but this is not essential. If the dummy variable t in the formula counts half-years (so $T = 20$ for a 10-year bond), then a 10-year zero will have duration 20 (half-years). The only reason I mention this is that when valuing semiannual bonds, you do sometimes count in half-years, and this can lead to confusion in the duration calculation. Obviously 20 half-years is the same as 10 years.

Duration is a measure of how sensitive a bond's price is to changes in interest rates. Duration is related to, but differs from, the slope of the plot of bond price versus yield-to-maturity.

I find the following construction to be an instructive way of understanding how duration works.[7] Suppose that you have a liability due in the future and that you buy a bond now with the intention of using the bond (and its accumulated coupons) to meet the liability (the maturity of the bond is assumed to be greater than or equal to the maturity of the liability). Suppose that the present value of the bond is identical to the present value of the liability. Suppose that you open a bank account that earns the market interest rate (the yield-to-maturity of the bond). You deposit all cash in-flows from the bond in the bank account and let them compound through time (with no taxes or transaction costs). When your liability falls due, you sell your bond and close your bank account. Call the proceeds of the bond sale together with your final bank balance the "Terminal Value."

Can you meet your liability with the Terminal Value? Well, there are two risks involved. A fall in interest rates immediately after you purchase the bond pushes up the price at which you are able to sell your bond. However, a fall in interest rates also decreases your final bank balance because you earn less interest on the coupons. The opposite obtains with a rise in interest rates. That is, higher interest rates decrease the price at which you can sell the bond, but your closing bank balance is higher because you earn more interest on the coupons. These two risks are known as *price risk* and *coupon reinvestment rate risk*, respectively.

Price risk and coupon reinvestment rate risk have opposite influences on the Terminal Value. The Terminal Value differs depending upon which influence is strongest. It can be proved that if your liability falls due before the weighted-average term-to-maturity of your bond, the price risk has the stronger influence on Terminal Value. If your liability falls due after the weighted-average term-to-maturity of your bond, the coupon reinvestment rate risk has the stronger influence on

[7]This construction using an artificial bank account is my own idea. I have not seen it anywhere in the literature. It is a natural construction, so I would not be surprised to find that someone else has already presented it. If you see a reference that uses this construction, please let me know.

Terminal Value. If your liability falls due precisely at the weighted-average term-to-maturity of the bond, the Terminal Value is relatively insensitive to an immediate change in interest rates.

By definition, the weighted-average term-to-maturity of the bond is just its Macaulay duration. This means that if you know when your liability falls due, you should provide for it by purchasing bonds with a duration equal to your investment horizon. You are "immunized" against a change in interest rates if the duration of your bond portfolio equals your investment horizon.[8]

I must emphasize that the bank account/Terminal Value construction is an artificial one. The fact that you must rebalance your position as time passes (in order to remain immunized) means that you cannot stick with the same bond until the liability is due. Indeed, the only bond that you can hold until the liability falls due (while remaining immunized) is a zero-coupon bond with maturity equal to the maturity of the liability; and in this case, the absence of coupons removes the need for the bank account in the construction.

Thus, the bank account/Terminal Value construction tells you how sensitive the Terminal Value is to an *immediate* parallel shift in the term-structure; for it is only in the *immediate* future that your chosen bond portfolio is immunized. This means that: if you can, open a bank account that pays the yield-to-maturity on your bond, purchase a coupon bond with duration and present value the same as those of the liability, and deposit all coupons in the bank account until the liability falls due—then, if there is one and only one parallel shock to the flat term-structure of interest rates between now and your liability falling due and if that single shift in interest rates occurs immediately, the Terminal Value will meet your liability. If anything else happens, you may be in trouble.

Other things being equal, duration increases with increasing term-to-

[8]Please note that an initial immunized position protects you from exactly one parallel shock to a *flat* term-structure. You are no longer immunized after a shock has hit. You get your planned future value only if no more shocks hit. You must re-balance after each shock to stay immunized. In fact, to stay immunized, you must rebalance even if no shocks hit. This is because changes in bond duration are not generally in lock-step with the passage of time. Thus, your horizon and your bond's duration decrease at different speeds, and you become non-immunized.

maturity.[9] Other things being equal, duration decreases with increasing coupons (larger cash flows early on decrease the proportional importance of the repayment of principal at maturity).

Compared to duration, convexity is a higher-order measure of sensitivity of bond price to interest rates. Convexity measures the rate at which the sensitivity of bond price to interest rates changes with changing interest rates.[10] Convexity is related to, but differs from, the rate of change of slope of the plot of bond price versus yield-to-maturity. See the summary in Table C.1. For a practitioner's view of Macaulay duration and convexity, see Kritzman (1992b).[11]

How do the definitions of duration and convexity arise? Suppose the price of a bond, P, is expanded in terms of yield-to-maturity, r, using a second-order Taylor series (that is, one that stops at the quadratic term):[12]

$$P(r+\Delta r) - P(r) \approx \frac{\partial P(r)}{\partial r} \times \Delta r + \frac{\frac{\partial^2 P(r)}{\partial r^2}}{2!} \times (\Delta r)^2.$$

Letting $\Delta P \equiv P(r+\Delta r) - P(r)$, use $P(r) = \sum_{t=1}^{T} \frac{C_t}{(1+r)^t}$ to find that[13]

$$\Delta P \approx \frac{-\Delta r}{1+r} \sum_{t=1}^{T} \frac{t \times C_t}{(1+r)^t} + \frac{(\Delta r)^2}{2!(1+r)^2} \sum_{t=1}^{T} \frac{t \times (t+1) \times C_t}{(1+r)^t}.$$

[9] Deeply discounted coupon bonds (bonds paying coupons far below current market rates) can be an exception (Fisher and Weil [1971, Table 4, p418]).

[10] Strictly speaking, this is not true. See Crack and Nawalkha (2001) for details.

[11] The standard Macaulay duration is a relatively simple concept. People on The Street expect you to know that they use more complex tools. For example, the standard Macaulay duration can be generalized to allow for immunization against parallel shifts in yield curves that are *not* flat. This generalization was originally proposed by Macaulay (1938), but was made popular by Fisher and Weil (1971). An even more sophisticated measure of duration is presented by Cox, Ingersoll, and Ross (1979). Duration measures for bonds with embedded options are also important (Mehran and Homaifar [1993]).

[12] Note that this is similar to expressing the change in the price of a call option (given a change in the level of the underlying) in terms of the "delta" and the "gamma." The delta is the rate of change of call price with respect to underlying, and the gamma measures the "convexity" of call price with respect to underlying.

[13] I used the result $\frac{\partial P(r)}{\partial r} = \sum_{t=1}^{T} \frac{\partial}{\partial r} \frac{C_t}{(1+r)^t} = \frac{-1}{1+r} \sum_{t=1}^{T} \frac{t \times C_t}{(1+r)^t}$ and an analogous result for $\frac{\partial^2 P(r)}{\partial r^2}$.

Now divide both sides by P to get

$$\frac{\Delta P}{P} \approx \frac{-\Delta r}{1+r} D + \frac{(\Delta r)^2}{2!} \mathcal{C},$$

where

$$D \equiv \frac{1}{P} \sum_{t=1}^{T} \frac{t \times C_t}{(1+r)^t}$$

is the standard Macaulay duration, and

$$\mathcal{C} \equiv \frac{1}{(1+r)^2 P} \sum_{t=1}^{T} \frac{t \times (t+1) \times C_t}{(1+r)^t}$$

is a measure of curvature, or "convexity," in the plot of bond price versus yield-to-maturity. Other things (i.e., duration and price) being equal, $\mathcal{C}$ increases with increasing coupons. Even a zero-coupon bond has positive convexity (because $C_1 = C_2 = \cdots = C_{T-1} = 0$, but $C_T = FACE > 0$).

In addition to immunization, duration and convexity enable you to estimate the impact on bond price of a change in interest rates. A "first-order" estimate uses duration; a "second-order" estimate uses duration and convexity. Higher-order approximations are more accurate.

Take a 20-year bond paying an annual coupon of 7%. Assume a face value of \$1,000. Assume that the term-structure is flat at 10%. The price of the bond is \$744.59 under these assumptions.

If the entire term-structure rises by one percentage point (i.e., 0.01), what is the new price of the bond? This can be estimated using the equation we derived previously:[14]

$$\frac{\Delta P}{P} \approx \frac{-\Delta r}{1+r} D + \frac{(\Delta r)^2}{2!} \mathcal{C}.$$

The Macaulay duration of this bond is calculated as 10.0018 years, the convexity $\mathcal{C}$ can be calculated as 130.04676, $\Delta r = +0.01$, $r = 0.10$, and

[14]Note that the term $\frac{D}{1+r}$ that multiplies $-\Delta r$ is often called the "modified duration," frequently denoted D^*. It follows that the first-order approximation using modified duration is $\Delta P \approx -\Delta r D^* P$.

$P = \$744.59$, thus:

$$\begin{aligned}
\Delta P &\approx \frac{-\Delta r}{1+r} D \times P + \frac{(\Delta r)^2}{2!}\, \mathcal{C} \times P \\
&= \frac{-0.01}{1.10} \times 10 \times \$744.59 + \frac{(0.01)^2}{2} \times 130.04676 \times \$744.59 \\
&= -\$67.69 + \$4.84 \\
&= -\$62.85.
\end{aligned}$$

Thus, $P(r + \Delta r) \approx P + \Delta P = \$744.59 - \$62.85 = \$681.74$. *Direct evaluation* gives the answer as \$681.47 (the estimate is 27 cents too high and would have been out by roughly \$5 if not for the convexity term).

If the entire term-structure falls by one percentage point (i.e., 0.01), the change in bond price is estimated as follows:

$$\begin{aligned}
\Delta P &\approx \frac{-\Delta r}{1+r} D \times P + \frac{(\Delta r)^2}{2!}\, \mathcal{C} \times P \\
&= \frac{+0.01}{1.10} \times 10 \times \$744.59 + \frac{(0.01)^2}{2} \times 130.04676 \times \$744.59 \\
&= +\$67.69 + \$4.84 \\
&= +\$72.53.
\end{aligned}$$

Thus, $P(r + \Delta r) \approx P + \Delta P = \$744.59 + \$72.53 = \$817.12$. *Direct evaluation* gives the answer as \$817.43 (the estimate is 31 cents too low and would have been out by roughly \$5 if not for the convexity term).[15]

Note that the "27 cents too high" and the "31 cents too low" in the above examples can be reduced to pennies (at least) by using a third term in the expansion—a measure of rate of change of convexity with respect to yield. Mehran and Homaifar (1993) refer to this third term

[15] Why am I *estimating* the change in bond price when direct evaluation gives the exact answer? For purposes of demonstration, it is convenient to be able to show you exactly how the duration and convexity measures work and where the approximations break down. This simple example is a good way to do that. In a real world situation, you might know the current value of your bond portfolio and its duration and convexity. It may be easier (and much faster) to *estimate* how your portfolio changes in value with changes in interest rates—using current value, duration, and convexity—than it is to *directly evaluate* each bond individually.

as "velocity." Thus, they represent change in bond price as a function of duration, convexity, and velocity—see Mehran and Homaifar (1993) for more details.[16]

For a standard bond, Macaulay duration $D = \frac{1}{P}\sum_{t=1}^{T}\frac{t\times C_t}{(1+r)^t}$ may be written in a closed-form formula (i.e., no summation term). With coupons $C_t = C$, a constant, for $t = 1, 2, \ldots, T-1$, and $C_T = C + F$, where F is face value, the standard Macaulay duration of the bond may be written as follows:

$$D = \frac{1+r}{r} - \frac{\left\{(1+r) + T\ \left[\frac{C}{F} - r\right]\right\}}{\frac{C}{F}\left[(1+r)^T - 1\right] + r}, \quad \text{where} \quad r \neq 0.$$

The proof of this result uses the standard closed-form formula for an annuity and, although not difficult, may be a little tedious—a similar type of expression exists for convexity.

Finally, let me exorcise a myth. Most of the foregoing is predicated on parallel shifts in yield curves. Other things (i.e., price and duration) being equal, the higher the convexity of a bond, the better off you are if there is a parallel shift (up or down) in a yield curve: hence the myth that you should pay for convexity.[17] In reality, these shifts are anything but parallel (Jones [1991]; Litterman and Scheinkman [1991]). Other things (i.e., price and duration) being equal, if the yield curve steepens, additional convexity will probably hurt you. Whether additional convexity helps or hurts depends upon the bonds you consider, and the "twist" in the yield curve that occurs. Crack and Nawalkha (2000) derive simple expressions that allow bond portfolio managers to capture the combined effects of term-structure height, slope and curvature shifts on duration, convexity, and higher-order bond risk measures. See Kahn and Lochoff (1990) and Lacey and Nawalkha (1993) for empirical evidence.

[16]People on The Street tell me that duration measures accounting for embedded options are important. Mehran and Homaifar (1993) discuss duration and convexity for bonds with embedded options. Before looking at Mehran and Homaifar (1993), be sure that both your mathematics and finance are up to scratch. They have the ideas correct, but their notation is contrary to conventional symbolic mathematics.

[17]This win-win situation is not kosher. A model that allows only parallel shifts in the yield curve freely admits arbitrage opportunities: match on price and duration and go long high convexity and short low convexity (Lacey and Nawalkha [1993]).

The bond pays C_t for $t = 1, \ldots, T$, and has discretely compounded annual yield r.	
Bond Price	$P = \sum_{t=1}^{T} C_t(1+r)^{-t}$
Modified Duration	$\begin{aligned} D^* &\equiv \frac{-\frac{\partial P}{\partial r}}{P} \\ &= \frac{\sum_{t=1}^{T} t\, C_t(1+r)^{-(t+1)}}{P} = \frac{1}{(1+r)}\sum_{t=1}^{T} t\,\omega_t, \\ \text{where } \omega_t &\equiv \frac{C_t(1+r)^{-t}}{P} \;\; \& \;\; \sum_{t=1}^{T} \omega_t = 1. \end{aligned}$
Macaulay Duration	$\begin{aligned} D &= D^*(1+r) = \sum_{t=1}^{T} t\,\omega_t \\ \text{where } \omega_t &\equiv \frac{C_t(1+r)^{-t}}{P} \;\; \& \;\; \sum_{t=1}^{T} \omega_t = 1. \end{aligned}$
Bond Convexity	$\begin{aligned} \mathcal{C} &\equiv \frac{\frac{\partial^2 P}{\partial r^2}}{P} \\ &= \frac{\sum_{t=1}^{T} t\,(t+1)\, C_t(1+r)^{-(t+2)}}{P} \\ &= \frac{1}{(1+r)^2}\sum_{t=1}^{T} t\,(t+1)\,\omega_t, \\ \text{where } \omega_t &\equiv \frac{C_t(1+r)^{-t}}{P} \;\; \& \;\; \sum_{t=1}^{T} \omega_t = 1. \end{aligned}$
Slope of Price-Yield Curve	$\text{SLOPE} = \frac{\partial P}{\partial r} = -D^* P = -\frac{D}{(1+r)} P$
Curvature of Price-Yield Curve	$\text{CURVATURE} = \frac{\partial^2 P}{\partial r^2} = \mathcal{C}P$
Taylor Series	$\Delta P \approx \frac{\partial P}{\partial r}(\Delta r) + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}(\Delta r)^2 = -D^* P(\Delta r) + \frac{1}{2}\mathcal{C}P(\Delta r)^2$

Table C.1: Duration/Convexity Summary

In the table, D is Macaulay duration, D^* is modified duration, P is bond price, and $\mathcal{C}$ is convexity. Try proving that $\frac{\partial D^*}{\partial r} = (D^*)^2 - \mathcal{C}$. Many of these relationships simplify substantially when we use continuously compounded yields (e.g., D and D^* are identical using continuously compounded yields y, so $\frac{\partial D}{\partial y} = D^2 - \mathcal{C}$, which in turn equals zero if the bond is a pure discount bond (Crack and Nawalkha [2001])).

Answer 3.12: From empirical investigations, it is known that stock returns do not have constant variance through time and that periods of high (low) volatility tend to follow periods of high (low) volatility (Fama [1965]; Akgiray [1989]). The GARCH model attempts to capture this empirical fact.

Suppose you estimate a simple linear model like $r_{it} = \alpha_i + \beta_i m_t + u_{it}$ (return on stock i at time t is a constant plus a constant times return on the market plus a residual). If you do not take account of changes in the variance of u_{it} through time, you can draw faulty statistical inferences about α_i and β_i. Note that the standard ordinary least squares (OLS) regression does not account for changing variance. In remedying this problem, the GARCH estimation captures a portion of stock price behaviour that might otherwise be interpreted as non-Normality and might lead to faulty inferences.

The GARCH model is a generalization of the ARCH model first presented in Engle (1982).[18] The formal GARCH(1,1) model for the residuals of a market model of stock returns is as follows:[19]

$$\begin{aligned} r_{it} &= \alpha_i + \beta_i m_t + u_{it} \\ u_{it}|\mathcal{F}_{i\,t-1} &\sim \mathcal{N}(0, h_{it}) \\ h_{it} &= \gamma_{0i} + \gamma_{1i} u_{i\,t-1}^2 + \gamma_{2i} h_{i\,t-1}. \end{aligned}$$

The residuals, u_{it}, may be assumed to be independently distributed across stocks i. The market return, m_t, is assumed common to all stocks. $\mathcal{F}_{i\ t-1}$ is the information set relative to stock i available just prior to date t; $\mathcal{F}_{i\ t-1}$ contains $u_{i\,t-1}$, $h_{i\,t-1}$ and all past returns on stock i. Note that conditional Normality is not required for the GARCH model (Bollerslev [1987]).

The GARCH model estimation differs from a straightforward ordinary least squares (OLS) estimation; you do not have a nice closed-form expression for $\hat{\alpha}_i$ or $\hat{\beta}_i$. In the GARCH estimation, you typically run OLS to get an initial guess for α_i and β_i. Then you adjust guesses of the γ_{ji}'s, α_i and β_i until you obtain what seem to be the most

[18] The review paper by Bera and Higgins (1993) is the best overview of ARCH and GARCH models that I have seen. Following this, you might look at Bera et al. (1988) as an introduction to ARCH, and also as an introduction to Engle (1982). For an introduction to statistical models for financial market volatility, see Engle (1993) and his references. For a higher-level review of ARCH modelling in finance, see Bollerslev et al. (1992). For a concise overview of the broad econometric peculiarities of the ARCH(1) model, see Hendry (1986).

[19] If you remove the term $h_{i\,t-1}$ from the second moment of the GARCH(1,1) model, you get the ARCH(1) model.

likely parameter estimates. This is a "maximum likelihood estimation" technique.[20]

Answer 3.13: You reduce exposure (i.e., hedge) by shorting T-bond futures contracts. Each Chicago Board of Trade (CBOT) T-bond contract covers a face value of $100,000 of T-bonds. If the duration of your bond is the same as the duration of the cheapest-to-deliver (CTD) T-bond, then you short $\frac{\$50{,}000{,}000}{\$100{,}000} = 500$ contracts.[21] If the duration of your bond is different from the duration of the CTD T-bond, then you adjust for durations: go short $\frac{D_B}{D_F} \times \frac{\$50{,}000{,}000}{\$100{,}000} = \frac{D_B}{D_F} \times 500$ contracts, where D_B is the duration of your bonds, and D_F is the duration of the CTD T-bond.

A final note. You could hedge by shorting Eurodollar futures (the underlying is the interest rate on a three-month $1 million Eurodollar deposit). However, the short end of the yield curve does not move with the long end. It, therefore, makes sense to use a hedging instrument whose underlying interest has maturity as close as possible to the portfolio to be hedged.

Answer 3.14: "Brady bonds" are sovereign bonds issued by developing countries in exchange for previously rescheduled bank loans. They are either "Par" bonds or "Discount" bonds. The former were issued at the par value of the loans but carry a below-market interest rate; the latter were issued at a discount from the face value of the former loans but carry a (floating) market interest rate. About a quarter of the market value of Brady bonds is collateralized by US Treasury issues. The size of this collateralization means that Brady bonds are sensitive to changes in US interest rates. In fact, something like a quarter of the variation in price movements of Brady bonds is (statistically) explained by moves in US Treasuries (sometimes with a lag of one day).[22]

Let us assume that the yield on the Brady bond increases by 25 basis points (one quarter of the US Treasury yield change). If we assume

[20] See Berndt et al. (1974) for details on a good maximum likelihood estimation technique. See Bollerslev (1986) and Greene (1993) for more on the GARCH model.

[21] For more details on the CTD bond, see Hull (1997, pp92–93); for details on duration-based hedging, see Hull (1997, pp102–104).

[22] This summary benefited from an unpublished research report prepared for Merrill Lynch by a group of my former students at MIT.

that the duration of the Brady bond is about 15 years, that the bond is trading at around par of \$1,000, and that the Mexican yield curve is flat at around 8%, then the price response would be (denoting yield by y)

$$\begin{aligned}
\Delta P &\approx -DP\frac{\Delta y}{(1+y)} \\
&= -15 \times \$1,000 \times \frac{0.0025}{1.08} \\
&= -\$15,000 \times \frac{0.0025}{1.08} \\
&= -\frac{\$37.50}{1.08} \approx -\$35.
\end{aligned}$$

With these assumptions, my guess is that the Brady bond price goes down by about three or four percentage points.

Answer 3.15: This question is similar to Question 3.10. The zero-coupon corporate bond has the same duration as longer-term coupon-bearing treasuries. You should short the corporate bond and buy treasuries that have the same duration and value as the corporate. By matching on duration and value, you create a zero-net investment portfolio that reaps profits.

Answer 3.16: First of all, the 5/10 time span is not relevant. The same result holds for a 1/2-year time span. That is, if the one-year interest rate is 10%, and two-year interest rate is 15%, then the forward rate for the second year is close to, but strictly greater than, 20%. Second, the order of the rates is not important. That is, if the two-year rate is 15%, and the forward rate for the second year is 10%, then the one-year rate is close to, but strictly greater than, 20%. Third, the result holds for effective (i.e., simple) interest rates but does not hold for continuously compounded interest rates (for which the approximation is exact).

The argument relies upon the way in which the interest on your interest accumulates. If you are offered 10% for the first year and 20% for the second year, you will not do as well as if you are offered the average (15%) for two years. Although the interest on the principal is the same in both cases (and equal to 30%), the interest on the interest is not the same (15% of 15% equals 2.25% and exceeds 20% of 10%, which is only 2%). To avoid arbitrage, the "plug" rate has to exceed 20%.

Enough of the "plain English" approach. The result can be proved using math. Let R_1 and R_2 be two different interest rates, then

$$\begin{aligned}\left(\frac{R_1+R_2}{2}\right)^2 - R_1R_2 &= \frac{1}{4}(R_1^2+2R_1R_2+R_2^2-4R_1R_2)\\ &= \frac{1}{4}(R_1^2-2R_1R_2+R_2^2)\\ &= \frac{1}{4}(R_1-R_2)^2 \;>\; 0.\end{aligned}$$

It follows that $\left(\frac{R_1+R_2}{2}\right)^2 > R_1R_2$. This means that the interest on the interest is better at the average rate than at the product of rates—as stated above.

The result may also be written as $\left(\frac{R_1+R_2}{2}\right) > \sqrt{R_1R_2}$. This is a special case of a more general result that an arithmetic average exceeds a geometric average. This result is true beyond the case $n = 2$ and can be extended to encompass harmonic averages also. Let $\mathcal{A}$, $\mathcal{G}$, and $\mathcal{H}$ denote the arithmetic, geometric, and harmonic averages of the positive numbers $x_1, x_2, \ldots, x_n$ as follows:

$$\begin{aligned}\mathcal{A} &\equiv \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1+x_2+\cdots+x_n}{n},\\ \mathcal{G} &\equiv \sqrt[n]{\Pi_{i=1}^{n} x_i} = \sqrt[n]{x_1x_2\ldots x_n}, \quad \text{and}\\ \mathcal{H} &\equiv \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\ldots\frac{1}{x_n}}.\end{aligned}$$

Then the following result holds (Spiegel [1968]):[23]

$$\mathcal{A} \geq \mathcal{G} \geq \mathcal{H},$$

and the inequalities are equalities only in the special case where

$$x_1 = x_2 = \cdots = x_n.$$

[23]To help you remember the ranking $\mathcal{A} \geq \mathcal{G} \geq \mathcal{H}$, note that it is the same as the ranking of the letters A, G, and H in the Latin alphabet.

Answer 3.17: If the one-year rate is 12%, and the two-year rate is 18%, then the forward rate for the second year is 24% to a first-order approximation (it is exactly 24% if these are continuously compounded rates). Let us assume this is 12% per half-year in the second year. Then your discounted expected payoff to the game is approximately

$$\begin{aligned}\left(\frac{1}{2}\times -\$2\right)+\left(\frac{1}{2}\times\frac{\$7}{(1.12)(1.12)}\right) &\approx -\$1+\frac{\$3.50}{1.25}\\ &= -\$1+\frac{14}{5}\\ &= \$1.80.\end{aligned}$$

If you can play repeatedly, then you are risk-neutral, and you would pay anything up to about \$1.80 to play this game. If you can play only once, then you might argue that the amount is so small you are still risk-neutral. If you multiply everything by a factor of one million, then you'll need to add a risk premium to the discount rates, and you will not pay as much to play.[24]

> **Story:** *1. Announced she hadn't had lunch and proceeded to eat a hamburger and french fries in the interviewer's office. 2. Without saying a word, candidate stood up and walked out during the middle of the interview.*
>
> *Interview Horror Stories from Recruiters*
> *Reprinted by kind permission of* MBA Style Magazine
> ©*1996–2004 MBA Style Magazine,* www.mbastyle.com

Answer 3.18: No one wants to trade with the informed (i.e., insider) trader because you almost always lose to someone who is better informed than you are. The identity of the informed trader has not been announced. This means that *any* trade could be a losing trade. Traders will, therefore, be reluctant to trade. This leads directly to decreased trading volume.

Here is another way to look at it. Uncertainty over the identity of the informed trader means that traders widen their bid-ask spreads to com-

[24]The risk premium as a function of the size of the bet is discussed by Tversky and Kahneman (1981) and Kahneman and Tversky (1982). Tversky and Kahneman (1974) is an earlier article you might like to read before reading these two.

pensate (on average) for any potential losses. Wider bid-ask spreads is one component of a decrease in liquidity, and it is usually associated with a decreased volume of trade (Chordia and Subrahmanyam [1995]).

Appendix D

Statistics Answers

This appendix contains answers to the questions posed in Chapter 4.

Answer 4.1: This sort of question is common. Begin by calculating the expected payoff to the game. As usual, this is just the summation over the product of potential outcomes times their probability of occurrence:

$$\begin{aligned}\left(\frac{1}{6}\times \$1\right)+\left(\frac{1}{6}\times \$2\right) \quad &+ \quad \left(\frac{1}{6}\times \$3\right)+\left(\frac{1}{6}\times \$4\right)\\ &+ \quad \left(\frac{1}{6}\times \$5\right)+\left(\frac{1}{6}\times \$6\right)=\$3.50.\end{aligned}$$

If you are selling tickets to *repeated* plays of this game, you are effectively risk-neutral.[1] This means you should charge the expected payoff (\$3.50) plus a margin for profit. You choose how wide to make the margin—it depends on your overhead, monopoly power, greed, and so on. You cannot charge \$6.00 or above, since no one will play. If the

[1]This is an application of the "Weak Law of Large Numbers." The law says, essentially, that if you independently draw repeated observations from the same random distribution, then for very many drawings, the sample mean is very close to the population mean (DeGroot [1989, p229–231]). In other words, after many repeated plays of the game, the ticket seller can be sure that his average payout per game is very close to the expected payout per game. Because all that matters is the expected payout, not the variance of payouts, the ticket seller is effectively risk-neutral. Similarly, casinos are effectively risk-neutral. With repeated plays, and odds slightly in the favour of "the house," the casino expects to be the winner for sure in the long run.

game is to be played only once, then you are risk-averse. You should charge the expected value, plus your profit margin, plus a risk premium. The risk premium depends upon how risk-averse you are.

Answer 4.2: Another die-rolling question; they are very popular. You want to get as many dollars as possible. You let me roll once and look at which number comes up. You must compare this number to the possible payoffs on the remaining two rolls. If it seems likely that you can do better by not stopping the game, then you proceed, otherwise you stop me.[2]

You must work backwards to deduce the best strategy. This is analogous to pricing an American-style option using a tree method. So, suppose that you have seen the second roll and are trying to decide whether to ask for a third. You must compare the outcome of the second roll to the distribution of possible outcomes on the third roll:

Maximum Payoff	\$1	\$2	\$3	\$4	\$5	\$6
Probability	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Table D.1: Distribution of Payoff to Third Roll of a Die

The expected value of the distribution in Table D.1 is \$3.50; the variance is \$2.92; the standard deviation is \$1.71. If you see a 4 or higher on the second roll, you might stop the game because you probably will not do better. If you get a 3 or lower, you might continue because you expect to do better.

Now, stepping backwards again, suppose that you have just seen the first roll. You must decide whether to ask for a second roll (which may lead to a third). You must compare the outcome of the first roll to the distribution of possible outcomes if you proceed to a second (and possibly third) roll.

If you ask for the second roll, there is one-half a chance that it yields a 1, 2, or 3, and one-half a chance that it yields a 4, 5, or 6. Using the argument above, in the first case (1, 2, or 3 on roll two) you proceed

[2] I thank Bingjian Ni for suggesting the solution technique. I am responsible for any errors.

to a third roll; in the second case (4, 5, or 6 on roll 2) you do not proceed. There is thus one-half a chance that you proceed to a third roll (expected value \$3.50 from Table D.1), and one-half a chance that you stop the game at roll two (expected value $\frac{\$4+\$5+\$6}{3}$). It follows that the expected value of asking for a second roll is as follows:

$$\left(\frac{1}{2} \times \$3.50\right) + \left[\frac{1}{2} \times \left(\frac{\$4 + \$5 + \$6}{3}\right)\right] = \$4.25.$$

Thus, you would ask for a second roll only if you get a 1, 2, 3, or 4 on roll one. If you have a 5 or 6 on roll one, you should stop the game.

In simple terms, then, the strategy is to stop the game at roll number one if a 5 or 6 appears (probability $\frac{1}{3}$), otherwise continue (probability $\frac{2}{3}$). If you continue, stop the game at roll number two if a 4, 5, or 6 appears, otherwise continue.

Please note that my argument involving expected payoffs assumes that you are risk-neutral; your stopping rule might use lower acceptable payoffs if you are risk-averse, or higher payoffs if you are risk-loving.[3]

The overall expected value of the game may now be calculated.

$$\text{Value} = \left(\frac{2}{3} \times \$4.25\right) + \left[\frac{1}{3} \times \left(\frac{\$5 + \$6}{2}\right)\right] = \frac{\$14}{3} \approx \$4.67.$$

If you are charging entry to repeated plays of this game, you are effectively risk-neutral.[4] You charge the expected value (\$4.67) plus a commission. You add a risk premium to the ticket price if there is only one or a few plays of the game; the more plays, the lower the risk premium. You would never charge more than six dollars because the player can never earn more than six dollars.

In the amended game (where I roll the die three times and pay you the maximum number of the three rolls), you need the distribution of the maximum payoff to three rolls of a die; this distribution is given in Table D.2.[5]

[3] In addition, you should question my treatment of discreteness. For example, although you cannot roll a "$3\frac{1}{2}$" or a "$4\frac{1}{4}$," I use these as cutoff points when deciding whether to proceed or not.

[4] This is the "Weak Law of Large Numbers" again. See Footnote 1 (on page 207 above)

Maximum Payoff	\$1	\$2	\$3	\$4	\$5	\$6
Probability	$\frac{1}{216}$	$\frac{7}{216}$	$\frac{19}{216}$	$\frac{37}{216}$	$\frac{61}{216}$	$\frac{91}{216}$

Table D.2: Distribution of Maximum Payoff in Three Rolls of a Die

The mean of the distribution of the maximum payoff from three rolls of the die is $\$\frac{1071}{216} = \$4.96$; the variance is $\$\frac{61047}{46656} = \$1.31$; and the standard deviation is \$1.14 (all calculated using information in Table D.2). You should, therefore, charge a ticket price of \$4.96 plus some profit margin for repeated plays. Again, you cannot charge more than six dollars because no one will play the game.

The second game is more expensive than the first game ($\$4.96 > \$4.67$) because it strictly dominates it. That is, the payoff to the second game is never less than, and often exceeds, the payoff to the first game. This is because the second game *guarantees* the maximum of three rolls without risk, but the first game does not.

Answer 4.3: This has been a very popular question. Assume that neither of you peek into your envelopes. Assume that you have $\$X$ in your envelope, where $\$X$ has a fifty-fifty chance of being either $\$m$ or $\$2m$. This means that your opponent's envelope has a fifty-fifty chance of containing $\$2X$ or $\$\frac{1}{2}X$. The expected value of switching is

$$\left(\frac{1}{2} \times \$2X\right) + \left(\frac{1}{2} \times \$\frac{1}{2}X\right) = \$1.25X.$$

The expected *benefit* of switching is, therefore, $\$0.25X$. On this basis, it looks as though you should switch envelopes. Of course, if your opponent does not peek, and she has $\$Y$ in her envelope, exactly the same argument shows that she has an expected benefit to switching of $\$0.25Y$. So, it looks as though she should switch also. This is the first part of the "Exchange Paradox": it seems that you *both* benefit from switching.

and DeGroot (1989, p229–231).

[5]Can you use elementary statistics to prove that this probability distribution is described by $\text{Prob}(Max = m) = \frac{m^3-(m-1)^3}{216} = \frac{3m(m-1)+1}{216}$, where m is the maximum of three rolls of the die? If you cannot, you need to work on your statistics.

Now, suppose that neither of you peek and that you do switch envelopes once. If you still do not peek, then a repeat of exactly the same argument suggests an expected benefit of 0.25 of the contents of your envelope if you switch again. The same applies to your opponent. This is the second part of the "Exchange Paradox": it seems that you could happily switch forever (like a dog chasing its own tail). The foregoing is the naive answer.

The problem is twofold: First, you are assuming that value is expected payoff (this is so only if you are genuinely risk-neutral);[6] second, your "prior" beliefs are that you have a fifty-fifty chance of having either $\$m$ or $\$2m$. The first problem is a function of your individual risk preferences and is difficult to address. The second problem can be tackled using two approaches: the first approach is to reconsider the nature of your prior; the second approach is to "update" your prior probability assessment (this is "Bayesian" statistics as opposed to "classical" statistics).

The first approach is to reconsider the nature of your priors. Our previous (paradoxical) calculation yielded $\$1.25X$ as the expected payoff to switching. However, this assumes that for any given X, it is equally likely that your opponent has $\$2X$ or $\$\frac{1}{2}X$. If you do not peek, then you are assuming a "diffuse level prior" because you assume this equality of likelihood for *any* X. Your prior is, therefore, not a valid probability density function (pdf) because the probabilities—across X---do not sum to **1**. However, for any *particular* m, it is equally likely that you received one of $\$m$ or $\$2m$. Thus, for any particular m, your priors are a pdf and any paradoxes should disappear. The expected value of switching should be zero. This is easily demonstrated. Let $P(\$m)$ denote the probability that *you* got $\$m$ (the lower amount); let $E(V)$

[6]An aside is in order. In corporate finance, the present value of a projected random payout is the discounted expected cash flow. The discounting is done at a rate that incorporates risk (e.g., using the CAPM), and the expectation is a mathematical one using real world probabilities (Brealey and Myers [1991]). An alternative to the real world expected cash flow coupled with the risk-adjusted discount rate is a risk-neutral world expected cash flow coupled with a riskless discount rate. The former is popular in corporate finance; the latter is popular in option pricing (see Arnold and Crack [2003]). With no discounting (e.g., the envelope question), value is expected payoff only if you are risk-neutral.

denote the expected value to switching; then $E(V)$ is given as follows:

$$\begin{aligned} E(V) &= [E(V|\$m) \times P(\$m)] + [E(V|\$2m) \times P(\$2m)] \\ &= \left(+\$m \times \frac{1}{2}\right) + \left(-\$m \times \frac{1}{2}\right) \\ &= \$0. \end{aligned}$$

The expected value is zero, and you are thus indifferent—resolving the paradox.[7] Note that $E(V|\$m) = +\$m$ because, conditional on your having been given the envelope containing only $\$m$, you gain $\$m$ by switching.

The second approach is to update your prior. To update your prior, you need information. The most obvious source of information is to peek into your envelope. So, assume that both you and your opponent peek into your envelopes. Now it gets subjective. If you see an amount that *seems* very high, then you update your prior probabilities: the probability that you have the high-value envelope increases, and the probability that you have the low-value envelope decreases. You no longer see value in switching envelopes.[8] If you see an amount that *seems* very low, then you see value in switching. The problem now is that you must subjectively assess the amount in the envelope as being either "low" or "high." The "Bayesian Resolution of the Exchange Paradox" is covered in detail in Christensen and Utts (1992).

If you have both peeked, and you do switch, then you will not switch again. This is because one of you gained, and that person will not want to lose by switching back. A similar question (but with an upper bound on the quantities possible) appears in Dixit and Nalebuff (1991, Chapter 13). The Dixit and Nalebuff book on strategic thinking is well worth a look.

Answer 4.4: The interviewee said that this question sets you up to think that the answer is difficult, when in fact it is straightforward. I do not think that is entirely fair, but as one interviewee said, "you never

[7]I thank Andres Almazan for suggesting this type of solution technique. I am responsible for any errors.

[8]However, you might argue that if you see an amount that seems so high that even one-half of it is more money than you can comprehend, you might switch envelopes just for the hell of it; it is worth the gamble.

know when they are going to bring out the guy in the chicken suit." That is, you never know what is going to happen next, or exactly what you should expect. The interviewers try to put you in a stressful or confusing situation just to see how you perform.

The easiest answer is that you bet \$100 on one team. If they win, you win \$100; if they lose, you lose \$100. Everything else is a "red herring."

Answer 4.5: This is elementary statistics, and one of the easiest questions in this book. The rules of the game have effectively removed the 1 from the sample space (i.e., the collection of possible outcomes). It follows that there are five possible outcomes (2 to 6), and each is equally likely. The expected outcome is simply

$$\frac{\sum_{i=2}^{i=6} i}{5} = \$4.$$

To do the sum in your head, remember that the dots on the opposing faces of the die add to seven. The sum must be three times seven, less one to give 20. Now divide by five to get the expected payoff of \$4.

Answer 4.6: The naive answer is that the probability is just $\frac{2}{52} \approx 4\%$. This is incorrect. There are four chances that the first card dealt to you (out of a deck of 52) is a King. Conditional on the first card being a King, there are three chances that the second card dealt to you (out of the remaining deck of 51) is a King. Conditional probability says that

$$\begin{aligned} &P(\text{Both are Kings}) = \\ &\qquad P(\text{Second is a King} \mid \text{First is a King}) \times P(\text{First is a King}) \end{aligned}$$

where " $|$ " is read as "conditional upon," or "given." This is a special case of the more general conditional probability result:

$$P(A \cap B) = P(A \mid B) \times P(B).$$

Thus, $P(\text{Both are Kings}) = \frac{3}{51} \times \frac{4}{52} = \frac{1}{17} \times \frac{1}{13} = \frac{1}{221} \approx 0.5\%$. Therefore, you have roughly one chance in 200 of getting exactly two Kings dealt to you.

I wish you to avoid a common form of confusion. Please note that although you multiply probabilities to get the answer, and such multiplication is often done when dealing with independent events, the events here (King on first card, and King on second card given King on first card) are *dependent*, not independent. That is, you calculate the probability that the second card is a King given that, or *dependent upon*, the first card being a King.

I wish to emphasize that the above procedure is different from that for figuring out the probability that, for example, you get two heads in two tosses of a fair coin (this probability is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$). The outcomes of the coin tosses are genuinely *independent*, and this is why you can multiply their probabilities directly. That is, the probability of a head on the second coin toss is not influenced by the event that you get a head on the first coin toss. However, the probability that you get a King on the second card dealt *is* influenced by the event that you get a King on the first card dealt. That is why the conditional probability theory is used. Be sure you understand the distinction and how and where to apply each method. If it is not clear, go to your favourite statistics book for a review (e.g., see Feller [1968, Chapter V]).

Answer 4.7: The "Let's Make a Deal" or "Monty Hall" problem is very frequently asked. In addition to being very frequent, it is quite difficult.

Warning: If you hear a question that sounds like Question 4.7, and you assume that my answer is the answer to the question you hear, then you are naive. Although this warning applies to all questions, it applies to this one in particular. This question is so popular that several different versions exist. You might be asked a similar sounding (but slightly differently worded) question that has a different answer.

Assume that you choose Door 3. The host opens Door 2 and offers you the chance to switch to Door 1. Should you do it? If you have decided that it does not matter whether you switch doors or not (indifference), or that you should definitely not switch (aversion), then you should go back and think again before reading any further. Stop here and try again.

At first glance, you might think that the random placement of the prize and the impartiality of the game show host means that you are indifferent between switching or not. In fact, the best decision is to switch. If you are to play this game repeatedly, two-thirds of the time you profit by switching, and one-third of the time you lose by switching. The chance of profiting exceeds the chance of losing, and you should switch. The details follow.

Let B_k denote the event that the prize is behind Door number k ("B" for behind). Let H_j denote the event that you see the host open Door number j ("H" for host).

The unconditional probabilities of the location of prizes (probabilities calculated without conditioning on which door the host opens) are simply $P(B_1) = P(B_2) = P(B_3) = \frac{1}{3}$. What you need to know is the conditional probability $P(B_1|H_2)$. That is, the probability that the prize is behind Door 1 given that you see (or "conditional on") the host open Door 2. We use a straightforward application of conditional expectations and Bayes' Theorem (see Feller [1968, Chapter V]), as follows:

$$P(B_1|H_2) = \frac{P(B_1 \cap H_2)}{P(H_2)} = \frac{P(H_2 \cap B_1)}{P(H_2)} = \frac{P(H_2|B_1) \times P(B_1)}{P(H_2)}$$

You know that $P(B_1) = \frac{1}{3}$, but what about $P(H_2|B_1)$ and $P(H_2)$? You know that the host is going to show you an empty door other than the one you choose (assume through all of this that it is Door 3 that you choose). The host's door must be revealed empty and cannot be the same door that you choose. Therefore, it must be that if you choose Door 3, then $P(H_2|B_1) = 1$.

Now, $P(H_2)$ is given by

$$\begin{aligned} P(H_2) &= [P(H_2|B_1) \times P(B_1)] + [P(H_2|B_2) \times P(B_2)] \\ &+ [P(H_2|B_3) \times P(B_3)], \end{aligned}$$

so some extra terms need to be calculated to get $P(H_2)$.

Well, the host's door must be shown to be empty, so it must be that $P(H_2|B_2) = 0$. The host is impartial, so it must be that $P(H_2|B_3) = \frac{1}{2}$ [and $P(H_1|B_3) = \frac{1}{2}$]. Thus, $P(H_2)$ is given as follows:

$$\begin{aligned} P(H_2) &= [P(H_2|B_1) \times P(B_1)] + [P(H_2|B_2) \times P(B_2)] \\ &+ [P(H_2|B_3) \times P(B_3)] \\ &= \left(1 \times \frac{1}{3}\right) + \left(0 \times \frac{1}{3}\right) + \left(\frac{1}{2} \times \frac{1}{3}\right) = \frac{1}{2}. \end{aligned}$$

It follows that the probability of finding the prize if you switch doors is two-thirds:

$$P(B_1|H_2) = \frac{P(H_2|B_1) \times P(B_1)}{P(H_2)} = \frac{1 \times \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}.$$

The summary in Table D.3 may clarify matters further. You choose Door 3. The host must choose an empty door to open. If the prize is behind Door 1, he *must* open Door 2 [$P(H_2|B_1) = 1$]. However, if the prize is behind Door 3, he can *choose* between Doors 1 and 2 [$P(H_2|B_3) = \frac{1}{2}$]. If you see Door 2, it is either because the prize is behind Door 1, and the host had no choice, or it is because the prize is behind Door 3, and the host randomly chose between Doors 1 and 2. It, therefore, follows that if you choose Door 3, and Door 2 is revealed empty by the host, the prize is twice as likely to be behind Door 1 as it is to be behind Door 3. Continuing along this line of thought, we may

Assume You Choose Door 3			
Prize Location B_j	Host Opens H_i	Unconditional Probability $P(H_i \cap B_j)$	Conditional Probability $P(H_i\|B_j)$
1	2	$\frac{1}{3}$	1
2	1	$\frac{1}{3}$	1
3	1	$\frac{1}{6}$	$\frac{1}{2}$
	2	$\frac{1}{6}$	$\frac{1}{2}$

Table D.3: The Monty Hall Problem

take a frequentist approach. Suppose you play the game repeatedly and always choose Door 3. If you look at all the times the host reveals Door 2 empty, you will find that two-thirds of the time the prize lies behind Door 1, and one-third of the time it is behind Door 3. Seeing Door 2 empty is thus a stronger signal that Door 1 has the prize than it is that Door 3 has it. This argument is more general, of course. Whichever door you choose, seeing the host reveal an empty door is a signal that you should switch.

Answer 4.8: This question is solved most efficiently by trying a few possible combinations, not by some time-consuming feat of constrained linear optimization. You should begin with extreme distributions, or with symmetrical distributions. It is in the extremes or in symmetry that solutions to such problems usually lie.

The probability of selecting a white marble is maximized (at almost $\frac{3}{4}$) by placing one white marble in one jar and the remaining 99 marbles in the other. The probability of selecting a white marble is minimized (at $\frac{1}{4}$) by placing all 100 marbles in one jar (assuming you do not get a second chance if the jar you choose is empty). If zero marbles in one jar is not an acceptable answer to you, then you minimize the probability of a white marble (at just over $\frac{1}{4}$) by maximizing the probability of a black one. That is, put one black marble in one jar and the remaining 99 marbles in the other.

Answer 4.9: This is a tough "game theory" problem.[9] Although I give a full solution, you would not be expected to do so in an interview. However, you should be able to identify the issues that contribute to the solution, and you should be able to understand the solution once presented.[10] A summary of the key ideas appears on page 222. Go directly to the summary and then come back for the finer points of the argument—if you can stand them (personally, I hate this stuff).

The goal is to maximize the probability of survival. Assume that each player chooses a "pure strategy" at the start of the game: shoot at one target only until the first kill; after the first kill, the two remaining players shoot it out. The initial target can be one of the opponents, or the sky.

Why would anyone shoot at the sky? Well, you are the worst shot: Mr. 10. The other players might see you as very little threat and choose to shoot it out amongst themselves before trying to get you. However, if you shoot at and hit Mr. 60, then Mr. 30 gets the next shot. He is

[9] Game theory was invented by the physicist John von Neumann. Von Neumann is credited with inventing (or co-inventing) the digital computer (Bernstein [1996]). Von Neumann also worked with J. Robert Oppenheimer, Enrico Fermi, Edward Teller, Niels Bohr, and Richard P. Feynman on the Manhattan Project (building the "gadget" at Los Alamos). Would you believe he had a dog named "Inverse"?

[10] I thank Olivier Ledoit for this solution technique. Any errors are mine.

going to shoot at you; he is a much better shot than you; and he gets to shoot at you before you get to shoot at him. However, if you leave them alone and shoot at the sky, you might have a better chance of survival. This is because you get the first shot after either Mr. 30 or Mr. 60 is hit (as long as it was not you who pulled the trigger). Being first in the final shoot-out may be very valuable when your opponent is a much better shot than you.

Does Mr. 30 choose Mr. 60 or Mr. 10 (that's you) as a target (or does he choose the sky)? He needs to know how likely is his survival if he hits Mr. 60 and then has to shoot it out with Mr. 10 (you). He compares that to how likely is survival if he hits Mr. 10 (you) and has to shoot it out with Mr. 60. In either case, the remaining opponent gets to shoot at Mr. 30 first.

Consider a general case. Suppose two players, Mr. P and Mr. Q, are in a duel. Assume that Mr. P has probability p of hitting any target, and Mr. Q has probability q of hitting any target. Let $\mathcal{P}(P|P,Q)$ denote the probability that Mr. P survives a shoot-out between Mr. P and Mr. Q, assuming Mr. P gets to shoot first. Similarly, let $\mathcal{P}(P|Q,P)$ denote the probability that Mr. P survives a shoot-out between Mr. P and Mr. Q, assuming Mr. Q gets to shoot first. Then elementary probability theory yields

$$
\begin{aligned}
\mathcal{P}(P|P,Q) &= p + (1-p)(1-q)p + (1-p)^2(1-q)^2p + \ldots \\
&= p\sum_{k=0}^{\infty}[(1-p)(1-q)]^k \\
&= \frac{p}{p+q-pq}.
\end{aligned}
$$

Thus, the probability that Mr. Q survives a shoot-out with Mr. P, where Mr. P goes first, is as follows:

$$
\mathcal{P}(Q|P,Q) = 1 - \mathcal{P}(P|P,Q) = \frac{q-pq}{p+q-pq}.
$$

For example, if Mr. 60 and Mr. 30 end up in a shoot-out, and Mr. 60 goes first, the probability of Mr. 60's survival is given by $\mathcal{P}(60|60,30) = 0.83333$. However, if Mr. 30 goes first, the probability is $\mathcal{P}(60|30,60) =$

0.58333 (a 25 percentage points worse chance of survival for Mr. 60 if Mr. 30 gets to shoot first).

We begin by ignoring the actions of the other players and focusing just on who each player is going to shoot at first.

Mr. 30's Decision: If Mr. 30 shoots at Mr. 10 (that's you!), his overall probability of survival is $0.30 \times \mathcal{P}(30|60, 30) + 0.70 \times \mathcal{P}(30|MISS10)$. $\mathcal{P}(30|60, 30)$ is the probability of Mr. 30's survival if he hits Mr. 10 and ends up in a shoot-out with Mr. 60 (and Mr. 60 gets to shoot first). The other term is the probability of Mr. 30's survival if he misses Mr. 10: $\mathcal{P}(30|MISS10)$.

If Mr. 30 shoots at Mr. 60, his overall probability of survival is $0.30 \times \mathcal{P}(30|10, 30) + 0.70 \times \mathcal{P}(30|MISS60)$. $\mathcal{P}(30|10, 30)$ is the probability of Mr. 30's survival if he hits Mr. 60 and ends up in a shoot-out with Mr. 10 (that's you, and you get to shoot first).

You can work out that $\mathcal{P}(30|60, 30) = 0.16667$ and $\mathcal{P}(30|10, 30) = 0.72973$. It must be that $\mathcal{P}(30|MISS10) \leq \mathcal{P}(30|MISS60)$ because whatever else is going on, missing the target just pushes you to the next round of shooting, and shooting at Mr. 60 is an otherwise healthier strategy for Mr. 30 than shooting at Mr. 10. It follows that Mr. 30 has a much higher chance of survival if he shoots at Mr. 60 than if he shoots at Mr. 10.

Mr. 60's Decision: If Mr. 60 shoots at Mr. 10, his overall probability of survival is $0.60 \times \mathcal{P}(60|30, 60) + 0.40 \times \mathcal{P}(60|MISS10)$. $\mathcal{P}(60|30, 60)$ is the probability of Mr. 60's survival if he hits Mr. 10 and ends up in a shoot-out with Mr. 30 (and Mr. 30 gets first shot). If Mr. 60 shoots at Mr. 30, his overall probability of survival is $0.60 \times \mathcal{P}(60|10, 60) + 0.40 \times \mathcal{P}(60|MISS30)$. You can work out that $\mathcal{P}(60|30, 60) = 0.58333$ and $\mathcal{P}(60|10, 60) = 0.84375$. Because $\mathcal{P}(60|MISS10) \leq \mathcal{P}(60|MISS30)$, it follows that for Mr. 60, shooting at Mr. 30 dominates shooting at Mr. 10.

Assume for the moment that neither Mr. 30 nor Mr. 60 choose to shoot at the sky; I shall return to this issue.

Mr. 10's Decision: As the poor shot (Mr. 10), how do you choose what to do? So far it looks as if Mr. 30 and Mr. 60 are going to shoot it out amongst themselves. If you interfere and hit one of them, you end up in a shoot-out with the remaining one—and he gets to shoot first, and he is a better shot.

You must choose between doing nothing, until either Mr. 30 or Mr. 60 is killed, or shooting at one of them. Well, suppose your strategy is to shoot at Mr. 30. Your overall probability of survival is $0.10 \times \mathcal{P}(10|60,10) + 0.90 \times \mathcal{P}(10|MISS30)$. If your strategy is to shoot at Mr. 60, your overall probability of survival is $0.10 \times \mathcal{P}(10|30,10) + 0.90 \times \mathcal{P}(10|MISS60)$. You can work out that $\mathcal{P}(10|60,10) = 0.0625$ and $\mathcal{P}(10|30,10) = 0.18919$. Thus, shooting at Mr. 60 dominates shooting at Mr. 30 ($\mathcal{P}(10|MISS30) = \mathcal{P}(10|MISS60)$ because missing just pushes you to the next round and no other player is shooting at you yet).

Now suppose that you (Mr. 10) shoot at the sky. Well, with Mr. 30 and Mr. 60 shooting it out, and Mr. 30 shooting first, there is a probability of $\mathcal{P}(30|30,60) = 0.41667$ that you end up with Mr. 30 as an opponent (if this happens you have a probability of $\mathcal{P}(10|10,30) = 0.27027$ of surviving). Similarly, there is a probability of $\mathcal{P}(60|30,60) = 0.58333$ of you ending up with Mr. 60 as an opponent (if this happens you have a probability of $\mathcal{P}(10|10,60) = 0.15625$ of surviving). Your overall probability of survival is thus

$$(0.41667 \times 0.27027) + (0.58333 \times 0.15625) = 0.20378.$$

You know that shooting at Mr. 60 dominates shooting at Mr. 30. So, you now compare shooting at Mr. 60 to shooting at the sky. What is the probability of survival if you shoot at Mr. 60? It is $\mathcal{P}(10|MISS60) = 0.20378$ because if you miss, the probability is 0.41667 that you end up in a shoot-out with Mr. 30 (chance of survival 0.27027), and the probability is 0.58333 that you end up in a shoot-out with Mr. 60 (chance of survival 0.15625), as above. Thus, the probability of survival if you shoot at Mr. 60 is $(0.10 \times 0.18919) + (0.90 \times 0.20378) = 0.20230$. So, shooting at the sky (overall probability of survival 0.20378) only just dominates shooting at Mr. 60 (overall probability of survival 0.20230).

Thus, as Mr. 10, the poor shot, you should shoot at the sky until Mr. 30 or Mr. 60 is knocked out of the competition. Then you have a shoot-out with the other survivor (probability of survival of 20.378%).

Of course, if everyone shot at the sky, everyone would have a probability of 100% of survival. This is a "corner solution" that is unlikely. Mr. 30 shoots at the sky only if he thinks Mr. 60 is going to do so—if Mr. 60 chooses a strategy of shooting at Mr. 30, Mr. 30 should definitely shoot back. There would need to be a pre-arranged cooperative pact if shooting in the air were to be optimal for everyone.[11]

If the order of shooting is reversed so that it is Mr. 10, Mr. 60, Mr. 30, Mr. 10, and so on, then the strategy changes. It still turns out to be optimal for Mr. 30 and Mr. 60 to begin by shooting at each other. However, Mr. 60 gets to shoot first. This means that if you shoot into the sky, you now have a chance of survival of only

$$\begin{aligned}
&\mathcal{P}(30|60,30) \times \mathcal{P}(10|10,30) + \mathcal{P}(60|60,30) \times \mathcal{P}(10|10,60) \\
&\quad = (0.16667 \times 0.27027) + (0.83333 \times 0.15625) \\
&\quad = 0.17525.
\end{aligned}$$

In this revised case, you are marginally better off shooting at Mr. 60, with a 10% chance of hitting him (chance of survival 0.18919 in the subsequent shoot-out with Mr. 30) and with a 90% chance of missing

[11]This unlikely "corner solution" is similar to a "prisoners' dilemma." Two people are detained in prison, suspected of a crime. If both prisoners keep their mouths shut, they each get sentences of two years. However, the police offer them a deal individually as follows:

Sentences (A,B)	B: Mouth Shut	B: Implicate A
A: Mouth Shut	(2,2)	(10,0)
A: Implicate B	(0,10)	(5,5)

If exactly one prisoner implicates the other, the implicated one gets 10 years, while the implicator goes free. If each implicates the other, each gets five years. If Suspect A says nothing, B gets two years if he says nothing or zero if he implicates A; similarly, if Suspect B says nothing, A gets two years if he says nothing or zero if he implicates B. The dilemma is whether to keep your mouth shut or implicate your accomplice. Without a pre-arranged cooperative pact, the best thing to do is implicate. The paradox is that the suspects end up worse off by doing the "best" thing (implicating) than if they had kept their mouths shut. The prisoners' dilemma can be presented in several different ways. The solution is a "Nash equilibrium," named after famous mathematician and 1994 Nobel prize winner, John Forbes Nash (see Nasar [1998] for Nash's riveting story). For more on the prisoners' dilemma and on game theory in general, see the introductory chapters of Fudenberg and Tirole (1991).

him (chance of survival 0.17525): for an overall probability of survival of

$$(0.10 \times 0.18919) + (0.90 \times 0.17525) = 0.17664.$$

Summary of the key ideas: Mr. 30 and Mr. 60 are going to shoot at each other because they do not see you as an immediate threat; you do not die first because Mr. 60 and Mr. 30 are shooting it out; you do not want to be put into a shoot-out where your opponent is a very good shot and gets to shoot first; if Mr. 30 gets to shoot before Mr. 60, it is less likely that you end up facing Mr. 60 than if Mr. 60 gets to shoot first, so you shoot in the air; if the direction of play is reversed, and Mr. 60 gets to shoot before Mr. 30, then you should help out Mr. 30 (and yourself) by shooting at Mr. 60 also, otherwise, leave it to Mr. 30; the cost of stepping in and shooting at Mr. 60 is that if you hit Mr. 60, you lose your chance to shoot first in the final shootout with Mr. 30; the benefit of stepping in and shooting at Mr. 60 is that you increase the likelihood of your facing Mr. 30 rather than Mr. 60 in the final shoot-out; there is a delicate balance between leaving it to Mr. 30 and stepping in to help him out, and it changes with the direction of play. Finally, there is a slim chance that everyone shoots at the sky, but this requires some sort of cooperation.

Answer 4.10: If you take the three-point shot, you have a 40% chance of winning. If you take the two-point shot, you have a 70% chance of a tie, and conditional on a tie you have a 50% chance of winning in overtime. Informally, the probability of winning if you take the two-point shot is thus 70% multiplied by 50%, which is 35%. This is lower than for the 40% for the three-point shot, so you should take the three-pointer.

More formally, let "W" denote winning, let "2" denote taking the two-point shot, let "T" denote sinking the two-pointer and getting a tie, and let "T^C" denote missing the two-pointer and not getting the tie (the "C" is for complement—the remainder of the sample space). Then

$$\begin{aligned} P(W|2) &= P(W|T)P(T|2) + P(W|T^C)P(T^C|2) \\ &= (0.50 \times 0.70) + (0 \times 0.50) \\ &= 0.35. \end{aligned}$$

Answer 4.11: This is one of the easier problems. If the cost is \$1.50 per spin, and you may play as often as you want, then yes, you should play.

The expected payoff is \$1.80 per spin ($\sum_{i=1}^{i=5} \text{Payoff}_i \times \frac{1}{5} = \$1.80$). If you can play as often as you want, you are risk-neutral (in the long run, your average payoff will equal the expected payoff), and you expect to make \$0.30 per spin on average.

If you get only one spin, then whether you play or not depends upon whether the expected \$0.30 gain is sufficient to compensate you for the risk of losing \$0.50 (the \$1.50 cost less the \$1.00 worst possible payoff). With amounts this small, you would probably take the bet. It is like spending \$1.50 on a lottery ticket—it is too small to care about. If the numbers were larger, say everything multiplied by one billion, and if your job is lost if you lose, then you are significantly more risk-averse, and your boss would not want you to take the bet.

Answer 4.12: Assuming no special information on your part, each sports match presents a fifty-fifty chance of winning. Assuming each match is independent of each other, then winning is analogous to tossing a fair coin four times in a row and trying to get four heads. This probability is only $\left(\frac{1}{2}\right)^4 = \frac{1}{16}$. The odds of winning are thus much worse than the odds offered by the bookie, and you should not play unless you are a risk-seeker. If the odds were raised to 25-to-1, this would be an attractive bet.

Answer 4.13: The standard deviation is just the square root of the expected squared deviation from the mean. Assuming equally likely probabilities, the mean of (1, 2, 3, 4, 5) is 3. The squared deviations are 4, 1, 0, 1, 4. The expected squared deviation is 2. The standard deviation is thus $\sqrt{2} \approx 1.4142$. I expect you to know $\sqrt{2}$ to four decimal places.

Answer 4.14: All you need is simple statistics. What happens if you ask the interviewer to shoot without spinning again? The first time the trigger was pulled, no bullet was found. It follows that *that* empty chamber will not be the next chamber. Also, if the first chamber was empty, then it certainly did not hold the first of the two contiguous bullets, bullet #1, so you will not meet the second of the two bullets, bullet #2. Thus, there are only four chambers that you might meet: three empty and one containing bullet #1. You have one chance in four of not having to talk about your resume.

If you do ask the interviewer to spin the barrel again, then you have the same chance you had when you sat down initially. That is, there is one chance in three that you do not have to talk about your resume. It follows that you are better off not spinning.

In summary, because the first chamber did not contain a bullet, then it was not bullet #1, so you know you will not see bullet #2. You face only one possible bullet from the remaining four chambers. However, spinning the barrel again puts both bullets into play, and that is not a choice you want to make.

Answer 4.15: Before we look at the formal math, let's use some informal intuition. There is one chance in a thousand (unconditionally) that you plucked the two-headed coin (which would certainly explain 10 heads in a row). There is also about one chance in a thousand that a fair coin would give 10 heads in a row (because $\left(\frac{1}{2}\right)^{10} = \frac{1}{1024} \approx \frac{1}{1000}$). Looking at the event (10 heads), I'd have to say that the coin is roughly equally likely to be two-headed or fair.

Now turn to the formal math – a direct application of Bayes' Theorem. Let "TH" denote the event that your coin is the two-headed one. Let "$10H$" denote the event that you toss one of the pennies and get 10 heads. Let X^c denote the complement of an event X. Then

$$\begin{aligned} P(TH|10H) &= \frac{P(TH \cap 10H)}{P(10H)} \\ &= \frac{P(10H|TH)P(TH)}{P(10H|TH)P(TH) + P(10H|TH^c)P(TH^c)} \\ &= \frac{1 \times \frac{1}{1000}}{\left[1 \times \frac{1}{1000}\right] + \left[\left(\frac{1}{2}\right)^{10} \times \frac{999}{1000}\right]} \\ &\approx \frac{1}{2}, \end{aligned}$$

where I used the facts that $2^{10} = 1024 \approx 1000$, and $\frac{999}{1000} \approx 1$. So, given the 10 heads, you have about a half a chance that you have the two-headed coin—as per our intuition.

Answer 4.16: If you turn over water and earth, you win. If you turn over fire, you lose. Wind is effectively absent from the sample space—it does not affect your chances of winning or losing. Ignoring the wind card completely, you turn over two cards, and you win only if the third card is fire. This happens with probability one-third.

Answer 4.17: Assuming that the players have fifty-fifty probabilities of playing Red or Blue,[12] each player has the same expected payoff: \$1. Player B has a variance of payoffs given by

$$\left[(0-1)^2 \times \frac{1}{2}\right] + \left[(2-1)^2 \times \frac{1}{2}\right] = 1,$$

whereas player A has a variance of payoffs given by:

$$\left[(1-1)^2 \times \frac{1}{4}\right] + \left[(3-1)^2 \times \frac{1}{4}\right] + \left[(0-1)^2 \times \frac{1}{2}\right] = 1.5.$$

Thus, if you are risk averse, player B's position is favoured (it offers the same expected return, but less risk).

Answer 4.18: We seek $F_{P|H}(p) = P(P \leq p|H) = P(A|H)$, where "$A$" denotes the event that $P \leq p$, and "H" denotes the event that you get a head. Let $f(u) \equiv 1$, $0 \leq u \leq 1$ denote the unconditional pdf of P. We apply Bayes' Theorem directly for $p \in [0,1]$ to get[13]

$$\begin{aligned} F_{P|H}(p) &= P(A|H) \\ &= \frac{P(A \cap H)}{P(H)} \\ &= \frac{\int_0^p uf(u)du}{\int_0^1 uf(u)du} \\ &= \frac{\left(\frac{p^2}{2}\right)}{\left(\frac{1}{2}\right)} = p^2. \end{aligned}$$

As $p \to 1$, $F_{P|H}(p) \to 1$, and as $p \to 0$, $F_{P|H}(p) \to 0$ (just checking). This cdf produces the pdf $f_{P|H}(p) = 2p$ that is left-skewed and has a mean of 2/3—slightly above 1/2 as you might have expected.

[12] A mixed strategy (Nash) equilibrium exists where B plays Red with probability $\frac{1}{4}$ and A plays Red with probability $\frac{1}{2}$. In this case, the expected payoff to playing Red equals the expected payoff to playing Blue for each player. A's expected payoff is $\frac{3}{4}$, whereas B's is 1. Thus, B is favoured. I thank Alex Butler for this argument. Any errors are mine.

[13] I thank Alex Vigodner for this answer.

Let "$750H/1000$" denote the event that you flip the coin 1,000 times and get 750 heads. In this case, the (conditional) distribution function is going to look much like the step function:

$$F_{(P|750H/1000)}(p) \approx \begin{cases} 0, & 0 \leq p < 0.75, \\ 1, & 0.75 \leq p \leq 1. \end{cases}$$

This conclusion relies upon a Weak Law of Large Numbers argument (see Footnote 1 [on page 207 above], and DeGroot [1989, p229–231]). The naive answer is to work it out mathematically, using binomial distributions and such like, but it quickly gets very messy, and the result should be essentially the same.

Answer 4.19: Both games have the same expected payoff: \$3.5 million. However, the second game has much less volatility than the first. The Weak Law of Large Numbers says that your actual payoff will be much closer to the expected payoff in Game Two. As a risk-averse individual, you choose Game Two.

Appendix E

Non-Quantitative Answers (Selected)

This appendix contains answers to those questions appearing in Chapter 5 that I deem worthy of response.

Answer 5.2.10: A "tombstone" is of course an advertisement that lists (like the names on a tombstone) the underwriters associated with a public issue of a security. The particular placement of the underwriters' names on the tombstone carries with it implications for the perceived status of the underwriters on the deal.

A student came to see me recently. He told me that he was flying to Chicago the next day for a job interview with an investment bank. I did not recognize the name of the bank. He asked me what sort of non-quantitative questions he might face, so I pulled out my book and tried several on him. When I got to the tombstone question, I stopped and asked him if he knew the definition of a tombstone. I pulled out that day's *Wall Street Journal* (WSJ) to see if there was a tombstone in the third section. The page at which I opened the WSJ contained a tombstone from the bank he was going to interview with the next day! I clipped it out and gave it to him, and he talked about it in his interview. It is worth keeping your eye on the tombstones in the third section of the WSJ in the weeks leading up to your interviews.

Answer 5.3.6: This is basic macroeconomics, and you should be fully familiar with it. The two forms of macroeconomic policy are monetary policy and fiscal policy. Monetary policy tries to achieve the broad objectives of economic policy through control of the monetary system and by operating on the supply of money, the level and structure of interest rates, and other conditions affecting the supply of credit (Pearce [1984, p291]). With monetary policy, the Federal Reserve Bank ("the Fed") sell bonds and reduces the money supply—an "open market operation." This increases interest rates (the cost of money) and makes capital expenditures more costly. This in turn slows down growth in the economy and should fight the inflationary threat.

In addition to open market operations, the Fed implements monetary policy by managing the discount rate (the rate the Fed charges banks for loans), adjusting the Fed funds rate (the rate banks charge each other for loans of federal funds), managing reserve requirements for banks (the proportion of a bank's assets required to be held in Treasury securities), and operations in the government repo (i.e., repurchase) market.[1]

Fiscal policy refers to the use of taxation and government expenditure to regulate the aggregate level of economic activity (Pearce [1984, p160]). Increasing taxes and decreasing government spending should slow down growth in the economy and fight inflationary fears. Go to any standard macroeconomic text if you want more details on fiscal or monetary policy.

Answer 5.3.11: The "Dow Jones Dogs strategy" involves buying the "Dow Jones Dogs" at the start of the year. These are the 10 Dow Jones Industrial Average (DJIA) stocks with the highest dividend yield. They are dogs because you get a relatively high dividend yield by having a low price relative to dividends. You are supposed to rebalance the portfolio every year. Historically this has been a very profitable strategy. In January 1999, the CBOE introduced options on the Dow Jones Dogs (ticker symbol "MUT").

Answer 5.4.8: The answer given by the interviewer was that if you are Avis or Hertz, cars are inventory. The same would seem to apply to GM, Chrysler, or Ford (or any of their distributors).

Answer 5.4.16: No. FCF does not include interest payments or repay-

[1]A "repo" is a repurchase agreement. It is an agreement to repurchase a security in the future. You give up the security now in exchange for cash, agreeing to repurchase the security at a later date for a larger amount of cash. A repo is thus a collateralized loan. A reverse repo is the other side of the deal—you purchase securities now with an agreement to sell them later. Repos range in maturity from overnight ("O/N") to as long as five years; shorter-term repos are the most popular.

ment of principal because FCF is the cash flows generated by net assets and available to the owners of the company (both debt and equity holders). The tax benefit of interest payments is recognized in a lower after tax cost of debt in the WACC. Finally, financing costs are not cash outflows. They do not reduce cash available to owners. To the contrary, they *are* cash payments to owners and, therefore, have no net effect on cash flows available to owners (i.e., FCF).

Answer 5.5.2: The "how many somethings are there somewhere" questions are common. There is no precise algebraic solution routine. You make several rough assumptions and hope the errors cancel. For example, the US population is about 275,000,000 (late 1998). The population of Bloomington, Indiana, is about 100,000. There are about six McDonald's in Bloomington. I calculate $\frac{275,000,000}{100,000} \times 6 = 2,750 \times 6 = 16,000$. However, Bloomington is a college town, and students eat more junk food than the general populace, so I adjust my answer downwards to the range 10,000 to 14,000 McDonald's outlets in the US. The *Wall Street Journal* (June 3, 1998, pA6) quotes 12,400 McDonald's outlets in the US. My ballpark figure is in the ballpark.

In general, you grab something you know, scale it up, and adjust for any biases. Let us try it again a different way. I cannot believe anyone over 30 or under five would eat in a McDonald's. If lifespans are uniformly distributed between zero and 75 years, then only one-third of the population (100,000,000) is eligible for eating at McDonald's. Half of these are health nuts. That leaves 50,000,000 customers. Suppose they eat four meals per week. That works out to about 30,000,000 meals served per day. If one outlet sells a burger every 30 seconds, that is 120 an hour and about 3,000 per day. 30,000,000 meals served per day at 3,000 per outlet implies about 10,000 outlets. This is still in the ballpark.

Whether it is ping-pong balls in a 747, barbers in Chicago, or elevators in the US, find something you know and scale it up. Be sure to know the population of the Earth, the US, the city you live in, and the city you interview in.

Answer 5.5.4: The answer given by the interviewer was that you should threaten to kill yourself by hitting your head against the wall. The administrative nightmare that would follow would ruin the guard's upcoming weekend. He would have to give you a cigarette.

Answer 5.5.7: You have to figure that the coin is not fair. The probability of another head is essentially one. See Huff's book, "How to Lie with Statistics," for related arguments (Huff [1982, Chapter 3]).

Answer 5.5.16: Aeroplanes can fly because of the curved shape of the wing. The wing, taken in cross section, has more surface area on the top surface than the bottom surface. The air must flow more quickly over the top of the wing than the bottom. This partial vacuum creates lift. A helicopter's blades are the same shape in cross section.

You may have noticed that at takeoff and landing of a jet, the pilot pushes forward a "leading edge" and lowers a "trailing edge" on the wing to give it even more curvature and thus even more lift. You can hear this in the cabin (watch the nervous flyers who think the noise is the plane malfunctioning), and you see this if you sit by a window near the wing.

You can test the principle in the safety of your own home by folding a sheet of paper down the middle and setting it up like a tent. Blow through the inside of the "tent" and the walls are "lifted" inward. Similarly, if you walk quickly through a doorway, the rush of wind will "lift" the door closer to closure (if it is sufficiently well oiled and only half open).

Answer 5.5.19: There are several possible responses that make sense. An obvious reason is safety: a round cover cannot fall down a round hole. Whereas if both hole and cover are either square or rectangular or oval, the cover can easily fall down the hole if lifted vertically and turned diagonally and dropped. Incidentally, I noticed while travelling through New Zealand that some of their manholes have rectangular covers. However, in this case, the covers are hinged and attached to a frame that is immovable—thus preventing the cover from falling.

Another reason for being round is that the (very heavy) covers may be rolled easily. Similarly, a (very heavy) round cover need not be manipulated before being returned to its hole—it may be replaced in any orientation. Finally, and with some sarcasm, manhole covers are round because the holes that they cover are round.[2]

[2] I might add that the holes themselves may be round because it is easier to drill a round hole in the street than a square one. Have you ever tried drilling a *square* hole in anything?

References for Further Research

1. Akgiray, Vedat, 1989, "Conditional Heteroscedasticity in the Time Series of Stock Returns: Evidence and Forecasts," *The Journal of Business*, Vol 62 No 1, (January), pp55–80.

2. Allen, Jeffrey G., 2000, *The Complete Q&A Job Interview Book*, Third Edition, John Wiley and Sons: New York, NY.

3. Arnold, Tom and Timothy Falcon Crack, 2003, "Option Pricing in the Real World: A Generalized Binomial Model With Applications to Real Options," Working Paper, University of Richmond, (April 15), 56pp.

4. Anton, Howard, 1988, *Calculus with Analytical Geometry*, Third Edition, John Wiley and Sons: New York, NY.

5. Bachelier, Louis, 1900, "Théorie de la Spéculation," *Annales de l'Ecole Normale Supéieure*, Series 3, XVII, pp21–86, Gauthier-Villars: Paris. Note that an English translation by A. James Boness appears in Cootner (1964).

6. Ball, Clifford A. and Walter N. Torous, 1985, "On Jumps in Common Stock Prices and Their Impact on Call Option Pricing," *The Journal of Finance*, Vol 40 No 1, (March), pp155–173.

7. Barone-Adesi, G. and Robert Whaley, 1987, "Efficient Analytic Approximation of American Option Values," *The Journal of Finance*, Vol 42 No 2, (June), pp301–320.

8. Bates, David S., 1995, "Testing Option Pricing Models," Working Paper, The Wharton School, University of Pennsylvania.

9. Baxter, Martin and Andrew Rennie, 1998, *Financial Calculus*, Cambridge University Press: Cambridge, England.

10. Bera, Anil, Edward Bubnys, and Hun Park, 1988, "Conditional Heteroscedasticity in the Market Model and Efficient Estimates of Betas," *The Financial Review*, Vol 23 No 2, (May), pp201–214.

11. Bera, Anil and Matthew L. Higgins, 1993, "ARCH models: Properties, Estimation, and Testing," *Journal of Economic Surveys*, Vol 7 No 4, pp305–366.

12. Berndt, Ernst K., Bronwyn Hall, Robert Hall, and Jerry A. Hausman, 1974, "Estimation and Inference in Nonlinear Structural Models," *Annals of Economic and Social Measurement*, Vol 3 No 4, (October), pp653–665.

13. Bernoulli, Daniel, 1738, "Speciman Theoriae Novae de Mensura Sortis," *Papers of the Imperial Academy of Sciences in Petersburg*, Vol V, pp175–192. Note that an English Translation appears in Bernoulli (1954).

14. Bernoulli, Daniel, 1954, "Exposition of a New Theory on the Measurement of Risk," *Econometrica*, Vol 22 No 1, (January), pp23–36 (Translated from the Latin by Louise Sommer).

15. Bernstein, Peter L., 1992, *Capital Ideas: The Improbable Origins of Modern Wall Street*, The Free Press: New York, NY.

16. Bernstein, Peter L., 1996, *Against the Gods: The Remarkable Story of Risk*, John Wiley and Sons: New York, NY.

17. Biger, Nahum and John Hull, 1983, "The Valuation of Currency Options," *Financial Management*, Vol 12 No 1, (Spring), pp24–28.

18. Black, Fischer and Myron Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, Vol 81 No 3, (May/June), pp637–659.

19. Black, Fischer, 1975, "Fact and Fantasy in the Use of Options," *The Financial Analysts Journal*, Vol 31 No 4, (July/August), pp36–41, 61–72.

20. Black, Fischer, 1976, "Studies of Stock Price Volatility Changes," *Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section of the American Statistical Association*, pp177–181.

21. Black, Fischer, 1989, "How We Came Up With the Option Formula," *The Journal of Portfolio Management*, Vol 15 No 2, (Winter), pp4–8.

22. Black, Fischer, 1993, "Beta and Return," *The Journal of Portfolio Management*, Vol 20 No 1, (Fall), pp8–18.

23. Black, Fischer and Myron S. Scholes, 1972, "The Valuation of Option Contracts and a Test of Market Efficiency," *The Journal of Finance*, Vol 27 No 2, (May), pp399–417.

24. Black, Fischer and Myron Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, Vol 81 No 3, (May/June), pp637–659.

25. Bollerslev, Tim, 1986, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, Vol 31 No 3, (April), pp307–327.

26. Bollerslev, Tim, 1987, "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics*, Vol 69 No 3, (August), pp542–547.

27. Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner, 1992, "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, Vol 52 No 1/2, (April/ May), pp5–59.

28. Boyle, Phelim P., 1977, "Options: A Monte Carlo Approach," *The Journal of Financial Economics*, Vol 4 No 3, (May), pp323–338.

29. Brealey, Richard A. and Stewart C. Myers, 1991, *Principles of Corporate Finance*, Fourth Edition, McGraw-Hill: New York, NY.

30. Brenner, Menachem and Marti G. Subrahmanyam, 1988, "A Simple Formula to Compute the Implied Standard Deviation," *The Financial Analysts Journal*, Vol 44 No 5, (September/October), pp80–83.

31. Brock, William A., David A. Hsieh, and Blake LeBaron, 1991, *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*, MIT Press: Cambridge, MA.

32. Brown, Robert, 1828, "A Brief Account of Microscopical Observations Made in the Months of June, July, and August, 1827, on the Particles Contained in the Pollen of Plants; and on the General Existence of Active Molecules in Organic and Inorganic Bodies," *The London and Edinburgh Philosophical Magazine and Annals of Philosophy*, Vol 4 No 21, pp161–173.

33. Chance, Don M., 1994, "Translating the Greek: The Real Meaning of Call Option Derivatives," *The Financial Analysts Journal*, Vol 50 No 4, (July/August), pp43–49.

34. Chesney, Marc and Louis Scott, 1989, "Pricing European Currency Options: A Comparison of the Modified Black-Scholes Mod-el and a Random Variance Model," *The Journal of Financial and Quantitative Analysis*, Vol 24 No 3, (September), pp267–284.

35. Chordia, Tarun and Avanidhar Subrahmanyam, 1995, "Market Making, the Tick Size, and Payment-for-Order-Flow: Theory and Evidence," *Journal of Business*, Vol 68 No 4, (October), pp543–575.

36. Christensen, Peter Ove, and Bjarne Sørensen. 1994. "Duration, Convexity, and Time Value." *The Journal of Portfolio Management*, Vol 20 No 2, (Winter), pp51–60.

37. Christensen, Ronald and Jessica Utts, 1992, "Bayesian Resolution of the 'Exchange Paradox'," *The American Statistician*, Vol 46 No 4, (November), pp274–276.

38. Conze, Antoine and Viswanathan, 1991, "Path Dependent Options: The Case of Lookback Options," *The Journal of Finance*, Vol 46 No 5, (December), pp1893–1907.

39. Cootner, Paul H., ed., 1964, *The Random Character of Stock Market Prices*, MIT Press: Cambridge, MA.

40. Cox, J. C., J. Ingersoll, and S. Ross, 1979, "Duration and the Measurement of Basis Risk," *The Journal of Business*, Vol 52 No 1, (January), pp51–61.

41. Cox, J. C. and S. Ross, 1976, "The Valuation of Options for Alternative Stochastic Processes," *The Journal of Financial Economics*, Vol 3 No 1/2, (January/March), pp145–166.

42. Cox, J. C., S. Ross, and M. Rubinstein, 1979, "Option Pricing: A Simplified Approach," *The Journal of Financial Economics*, Vol 7 No 3, (September), pp229–263.

43. Cox, J. C. and Mark Rubinstein, 1985, *Options Markets*, Prentice-Hall: Englewood Cliffs, NJ.

44. Crack, Timothy Falcon, 1999, *Derivatives Securities Pricing*, MBA Course Notes, Indiana University, Kelley School of Business, (Spri-ng #II), 108pp.

45. Crack, Timothy Falcon, 2004, *Basic Black-Scholes: Option Pricing and Trading*. See `www.BasicBlackScholes.com`, and the advertisement on the last page of this book, for details.

46. Crack, Timothy Falcon and Olivier Ledoit, 1996, "Robust Structure Without Predictability: The 'Compass Rose' Pattern of the Stock Market," *The Journal of Finance*, Vol 51 No 2, (June), pp751–762.

47. Crack, Timothy Falcon and Olivier Ledoit, 1998, "Asymptotic Distributions of Sample Statistics for a Gaussian AR(1) with Applications to Auto-Correlated Equity Returns," Working Paper, Indiana University, Kelley School of Business, Finance Department.

48. Crack, Timothy Falcon and Sanjay K. Nawalkha, 2000, "Interest Rate Sensitivities of Bond Risk Measures," *The Financial Analysts Journal*, Vol. 56 No. 1, (January/February), pp34–43.

49. Crack, Timothy Falcon and Sanjay K. Nawalkha, 2001, "Common Misunderstandings Concerning Duration and Convexity," *Journal of Applied Finance*, Vol. 1, (October), pp82–92.

50. DeGroot, Morris H., 1989, *Probability and Statistics*, Addison-Wesley: Reading, MA.

51. Derman, Emanuel and Iraj Kani, 1993, "The Ins and Outs of Barrier Options," *Goldman Sachs Quantitative Strategies Research Notes*, Goldman, Sachs, June.

52. Derman, Emanuel and Iraj Kani, 1994, "The Volatility Smile and its Implied Tree," *Goldman Sachs Quantitative Strategies Research Notes*, Goldman, Sachs, January.

53. Dixit, Avinash K. and Barry J. Nalebuff, 1991, *Thinking Strategically: The Competitive Edge in Business, Politics, and Everyday Life*, Norton: New York, NY.

54. Edwards, Franklin R. and Cindy W. Ma, 1992, *Futures and Options*, McGraw-Hill: New York, NY.

55. Einstein, A., 1905, "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen" (On the Molecular Kinetic Theory of the Heat-Generated Motion of Particles Suspended in Fluid), *Annalen der Physik*, Series 4, Vol 17, pp549–560.

56. Engle, Robert F., 1982, "Autoregressive Conditional Heterosceda-sticity with Estimates of the Variance of United Kingdom Infl-ation," *Econometrica*, Vol 50 No 4, (July), pp987–1007.

57. Engle, Robert F., 1993, "Statistical Models for Financial Volatility," *The Financial Analysts Journal*, Vol 49 No 1, (January/ February), pp72–78.

58. Evans, Merran, Nicholas Hastings, and Brian Peacock, 1993, *Statistical Distributions*, Second Edition, John Wiley and Sons: New York, NY.

59. Fabozzi, Frank J. and T. Dessa Fabozzi, 1995, *The Handbook of Fixed Income Securities*, Irwin: New York, NY.

60. Fama, Eugene, F., 1965, "The Behavior of Stock Market Prices," *The Journal of Business*, Vol 38 No 1, (January), pp34–105.

61. Fama, Eugene F. and Kenneth R. French, 1988, "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy*, Vol 96 No 2, (April), pp246–273.

62. Fama, Eugene F. and Kenneth R. French, 1992, "The Cross-Section of Expected Stock Returns," *The Journal of Finance*, Vol 47 No 2, (June), pp427–465.

63. Fama, Eugene F. and Kenneth R. French, 1996, "The CAPM is Wanted, Dead or Alive," *The Journal of Finance*, Vol 51 No 5, (December), pp1947–1958.

64. Farlow, Stanley J., 1993, *Partial Differential Equations for Scientists and Engineers*, Dover: New York, NY.

65. Feller, William, 1968, *An Introduction to Probability Theory and its Applications*, Volume I, Third Edition, John Wiley and Sons: New York, NY.

66. Fisher, L. and R. Weil, 1971, "Coping with the Risk of Interest Rate Fluctuations: Returns to Bondholders from Naive and Optimal Strategies," *The Journal of Business*, Vol 44 No 4, (October), pp408–431.

67. Fleming, Jeff, 1998, "The Quality of Market Volatility Forecasts Implied by S&P100 Index Option Prices," *The Journal of Empirical Finance*, Vol 5 No 4, (October), pp317–345.

68. Fraser, Michael K., 1993, "What It Takes to Excel in Exotics," *Global Finance*, Vol 7 No 3, (March), pp44–49.

69. Fudenberg, Drew and Jean Tirole, 1991, *Game Theory*, The MIT Press: Cambridge, MA.

70. Garman, Mark B. and Steven W. Kohlhagen, 1983, "Foreign Currency Option Values," *Journal of International Money and Finance*, Vol 2 No 3, (December), 231–237.

71. Geske, Robert, 1979, "A Note on an Analytic Valuation Formula for Unprotected American Call Options on Stocks with Known Dividends," *Journal of Financial Economics*, Vol 7 No 4, (December) pp375–380.

72. Girsanov, I.V., 1960, "On Transforming a Certain Class of Stocha-stic Processes by Absolutely Continuous Substitution Measures," *Theory of Probability and its Applications*, Vol 5, pp285–301.

73. Gleick, James, 1987, *Chaos: Making a New Science*, Penguin: New York, NY.

74. Gleick, James, 1993, *Genius: The Life and Science of Richard Feynman*, Vintage Books: New York, NY.

75. Goldman, B.M., H.B. Sosin, and M.A. Gatto, 1979, "Path Dependent Options: 'Buy at the Low, Sell at the High'," *The Journal of Finance*, Vol 34 No 5, (December), pp1111–1127.

76. Goldman, B.M., H.B. Sosin, and L.A. Shepp, 1979, "On Contingent Claims That Insure Ex-Post Optimal Stock Market Timing," *The Journal of Finance*, Vol 34 No 2, (May), pp401–414.

77. Grabbe, J. Orlin, 1983, "The Pricing of Put and Call Options on Foreign Exchange," *Journal of International Money and Finance*, Vol 2 No 3, (December), 239–253.

78. Greene, William H., 1993, *Econometric Analysis*, Second Edition, Macmillan: New York, NY.

79. Hammer, Jerry A., 1989, "On Biases Reported in Studies of the Black-Scholes Option Pricing Model," *Journal of Economics and Business*, Vol 41 No 2, (May), pp153–169.

80. Harrison, J. Michael, 1985, *Brownian Motion and Stochastic Flow Systems*, John Wiley and Sons: New York, NY.

81. Harrison, J.M. and S.R. Pliska, 1981, "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Their Applications*, Vol 11, pp215–260.

82. Haug, Espen Gaarder, 1997, *The Complete Guide to Option Pricing Formulas*, McGraw-Hill: New York, NY.

83. Haug, Espen Gaarder, 2001, "The Options Genius," *Wilmott Magazine*, (July), pp1–4.

84. Haung, Chi-fu, 1992, *Theory of Financial Markets*, Unpublished incomplete book manuscript, Department of Finance, Sloan Scho-ol of Management, MIT, Cambridge, MA 02142.

85. Hendry, David F., 1986, "An Excursion into Conditional Varianceland," *Econometric Reviews*, Vol 5 No 1, pp63–69.

86. Holland, A.S.B., 1973, *Introduction to the Theory of Entire Functions*, Academic Press: New York, NY.

87. Huff, Darrell, 1982, *How to Lie with Statistics*, Norton: New York, NY.

88. Hull, John C., 1997, *Options, Futures, and Other Derivatives*, Third Edition, Prentice-Hall: Englewood Cliffs, NJ.

89. Hull, John C., 1998, *Introduction to Futures and Options Markets*, Third Edition, Prentice-Hall: Englewood Cliffs, NJ.

90. Hull, John and Alan White, 1987, "The Pricing of Options on Assets with Stochastic Volatilities," *The Journal of Finance*, Vol 42 No 2, (June), pp281–300.

91. Hull, John and Alan White, 1993, "Efficient Procedures for Valuing European and American Path-Dependent Options," *The Journal of Derivatives*, Vol 1, (Fall), pp21–31.

92. Hunter, William C. and David W. Stowe, 1992, "Path-Dependent Options: Valuation and Applications," *Economic Review (Federal Reserve Bank of Atlanta)*, Vol 77 No 4, (July/August), pp30–43.

93. Jarrow, Robert and Andrew Rudd, 1983, "Approximate Option Valuation for Arbitrary Stochastic Processes," *Journal of Financial Economics*, Vol 10 No 3, (November), pp347–369.

94. Jarrow, Robert and Stuart Turnbull, 1996, *Derivative Securities*, South-Western College Publishing: Cincinnati, OH.

95. Jones, Frank J., 1991, "Yield Curve Strategies," *The Journal of Fixed Income*, Vol 1 No 2, (September), pp43–51.

96. Kahn, Ronald N. and Roland Lochoff, 1990, "Convexity and Exceptional Return," *The Journal of Portfolio Management*, Vol 16 No 2, (Winter), pp43–47

97. Kahneman, David and Amos Tversky, 1982, "The Psychology of Preferences," *Scientific American*, Vol 246, pp160–173.

98. Kotz, Samuel and Norman L. Johnson (editors-in-chief), and Campbell B. Read (associate editor), 1982, *Encyclopedia of Statistical Sciences*, Vol 6, John Wiley and Sons: New York, NY.

99. Krämer, Walter and Ralf Runde, 1997, "Chaos and the Compass Rose," *Economics Letters*, Vol 54 No 2, (February), pp113–118.

100. Krause, Robert (editor), 1998, *Global Equity and Derivative Market Risk*, Morgan Stanley Dean Witter Quantitative Strategies Group, Morgan Stanley and Co.: New York, NY.

101. Kritzman, Mark, 1992a, "What Practitioners Need to Know About Utility," *The Financial Analysts Journal*, Vol 48 No 3, (May/June), pp17–20.

102. Kritzman, Mark, 1992b, "What Practitioners Need to Know About Duration and Convexity," *The Financial Analysts Journal*, Vol 48 No 6, (November/December), pp17–20.

103. Kritzman, Mark, 1993a, "What Practitioners Need to Know About Factor Methods," *The Financial Analysts Journal*, Vol 49 No 1, (January/February), pp12–15.

104. Kritzman, Mark, 1993b, "What Practitioners Need to Know About the Term Structure of Interest Rates," *The Financial Analysts Journal*, Vol 49 No 4, (July/August), pp14–18.

105. Kritzman, Mark, 1993c, "What Practitioners Need to Know About Hedging," *The Financial Analysts Journal*, Vol 49 No 5, (September/October), pp22–26.

106. Kunitomo, Naoto and Masayuki Ikeda, 1992, "Pricing Options with Curved Boundaries," *Mathematical Finance*, Vol 2 No 4, (October), pp275–298.

107. Lacey, Nelson J. and Sanjay K. Nawalkha, 1993, "Convexity, Risk, and Returns," *The Journal of Fixed Income*, Vol 3 No 3, (December), pp72–79.

108. Latané, Henry A. and Richard J. Rendleman, Jr., 1976, "Standard Deviations of Stock Price Ratios Implied in Option Prices," *The Journal of Finance*, Vol 31 No 2, (May), pp369–381.

109. Lewis, William Dodge, Henry Seidel Canby, and Thomas Kite Brown (editors), 1942, *The Winston Dictionary*, The John C. Winston Company: Philadelphia. PA.

110. Lewis, Michael M., 1990, *Liar's Poker: Rising Through the Wreckage of Wall Street*, Penguin Books: New York, NY.

111. Litterman, Robert, and José Scheinkman, "Common Factors Affecting Bond Returns," *The Journal of Fixed Income*, Vol 1 No 1, (June), pp54–61.

112. Lo, Andrew W. and A. Craig MacKinlay, 1988, "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test," *The Review of Financial Studies*, Vol 1 No 1, (Spring), pp41–66.

113. Lo, Andrew W. and Jiang Wang, 1995, "Implementing Option Pricing Models When Asset Returns are Predictable," *The Journal of Finance*, Vol 50 No 1, (March), pp87–129.

114. Lorenz, Edward N., 1963, "Deterministic Nonperiodic Flow," *Jou-rnal of the Atmospheric Sciences*, Vol 20 No 2, (March), pp130–140.

115. Macaulay, Frederick Robertson, 1938, *Some Theoretical Problems Suggested by the Movements of Interest Rates and Stock Prices in the United States Since 1856*, National Bureau of Economic Research: New York, NY.

116. MacMillan, L.W., 1986, "Analytic Approximation for the American Put Option," *Advances in Futures and Options Research*, Vol 1 Part A, pp119–139.

117. Marler, Patty and Jan Bailey Mattia, 1995, *Job Interviews Made Easy*, VGM Career Horizons: Chicago, IL.

118. Mehran, Jamshid and Ghassem Homaifar, 1993, "Duration and Convexity for Bonds with Embedded Options: The Case of Convertibles," *The Journal of Business Finance and Accounting*, Vol 20 No 1, (January), pp107–113.

119. Melino, Angelo and Stuart Turnbull, 1990, "Pricing Foreign Currency Options with Stochastic Volatility," *The Journal of Econometrics*, Vol 45 No 1/2, (July/August), pp239–265.

120. Merton, Robert C., 1973, "Rational Theory of Option Pricing," *Bell Journal of Economics and Management Science*, Vol 4 No 1, (Spring), pp141–183. Note that this appears as Chapter 8 in Merton (1992).

121. Merton, Robert C., 1976, "Option Pricing When Underlying Stock Returns Are Discontinuous," *The Journal of Financial Economics*, Vol 3 No 1, (January/March), pp125–144.

122. Merton, Robert C., 1992, *Continuous-Time Finance*, Blackwell: Cambridge, MA.

123. Minton, Bernadette A., 1997, "An Empirical Examination of Valuation Models for Plain Vanilla U.S. Interest Rate Swaps," *The Journal of Financial Economics*, Vol 44 No 2, (May), pp251–277.

124. John Mongan and Noah Suo Suojanen, 2000, *Programming Interviews Exposed: Secrets to Landing Your Next Job*, John Wiley and Sons: New York, NY.

125. Mullens, David W., 1982, "Does the Capital Asset Pricing Model Work?," *Harvard Business Review*, Vol 60 No 1, (January/Februa-ry), pp105–114.

126. Murphy, Gareth, 1994, "When Options Price Theory Meets the Volatility Smile," *Euromoney*, No 299, (March), pp66–74.

127. Musiela, Marek and Marek Rutkowski, 1997, *Martingale Methods in Financial Modelling*, Springer-Verlag: Berlin.

128. Naik, Vasanttilak and Moon Lee, 1990, "General Equilibrium Pricing of Options on the Market Portfolio with Discontinuous Returns," *The Review of Financial Studies*, Vol 3 No 4, pp493–521.

129. Nasar, Sylvia, 1998, *A Beautiful Mind*, Simon and Schuster: New York, NY.

130. Natenberg, Sheldon, 1994, *Option Volatility and Pricing: Advanced Trading Strategies and Techniques*, Irwin: Chicago, IL.

131. Options Clearing Corporation, 1993, *Understanding Stock Options*, (September), The Options Clearing Corporation, 440 S. LaSalle St., Suite 2400, Chicago, IL 60605.

132. Parkinson, Michael, 1977, "Option Pricing: The American Put," *The Journal of Business*, Vol 50 No 1, (January), pp21–39.

133. Pearce, David, W., 1984, *The Dictionary of Modern Economics*, The MIT Press: Cambridge, MA.

134. Peterson, Richard L., Christopher K. Ma, and Robert J. Ritchey, 1992, "Dependence in Commodity Prices," *The Journal of Futures Markets*, Vol 12 No 4, (August), pp429–446.

135. Poterba, James and Lawrence Summers, 1988, "Mean Reversion in Stock Returns: Evidence and Implications," *Journal of Financial Economics*, Vol 22 No 1, (October), pp27–60.

136. Rendleman, Richard J., Jr. and Brit J. Bartter, 1979, "Two-State Option Pricing," *The Journal of Finance*, Vol 34 No 5, (December), pp1093–1110.

137. Richardson, Matthew, 1993, "Temporary Components in Stock Prices: A Skeptic's View," *Journal of Business and Economic Statistics*, Vol 11 No 2, (April), pp199–207.

138. Ritchken, Peter, L. Sankarasubramanian, and Anand M. Vijh, 1993, "The Valuation of Path Dependent Contracts on the Average," *Management Science*, Vol 39 No 10, (October), pp1202–213.

139. Roll, R., 1977a, "A Critique of the Asset Pricing Theory's Tests: Part I: On Past and Potential Testability of the Theory," *The Journal of Financial Economics*, Vol 4 No 2, (March), pp129–176.

140. Roll, R., 1977b, "An Analytical Formula for Unprotected American Call Options on Stocks with Known Dividends," *The Journal of Financial Economics*, Vol 5 No 2, (November), pp251–258.

141. Samuelson, Paul A., 1965, "Rational Theory of Warrant Pricing," *Industrial Management Review*, Vol 6 No 2, (Spring), pp13–31.

142. Samuelson, Paul A., 1973, "Mathematics of Speculative Price," *SIAM Review*, Vol 15 No 1, (January), pp1–42.

143. Scott, Louis O., 1987, "Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application," *The Journal of Financial and Quantitative Analysis*, Vol 22 No 4, (December), pp419–438.

144. Sharpe, W.F., 1964, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *The Journal of Finance*, Vol 19 No 3, (September), pp425–442.

145. Sharpe, William F., 1978, *Investments*, Prentice-Hall: Englewood Cliffs, N.J.

146. Smith, Clifford W., Jr., 1976, "Option Pricing: A Review," *The Journal of Financial Economics*, Vol 3 No 1/2, (January/March), pp3–51.

147. Snyder, Gerard L., 1969, "Alternative Forms of Options," *The Financial Analysts Journal*, Vol 25 No 5, (September/October), pp93–101.

148. Spiegel, Murray R., 1956, *College Algebra*, McGraw-Hill: New York, NY.

149. Spiegel, Murray R., 1968, *Mathematical Handbook*, McGraw-Hill: New York, NY.

150. Spiegel, Murray R., 1981, *Advanced Calculus*, McGraw-Hill: New York, NY.

151. Sprenkle, Case M., 1961, "Warrant Prices as Indicators of Expectations and Preferences," *Yale Economic Essays*, Vol 1 No 2, (Fall), pp178–231. Note that this paper appears in Cootner (1964).

152. Sullivan, Edward J. and Timothy M. Weithers, 1991, "Louis Bachelier: The Father of Modern Option Pricing Theory," *Journal of Economic Education*, Vol 22 No 2, (Spring), pp165–171.

153. Sullivan, Sara, 1993, "Risk reversals," *Euromoney Treasury Manager*, (December 3), p15.

154. Sundaresan, Suresh, 1997, *Fixed Income Markets and Their Deri- vatives*, South-Western College Publishing: Cincinnati, OH.

155. Thorp, Edward O., 1973, "Extensions of the Black-Scholes Option Model," *Proceedings of the 39th Session of the International Statistical Institute, Vienna*, appearing in *Bulletin of the International Statistical Institute*, Vol 45 Book 2, pp522–529.

156. Tian, Yisong, 1993, "A Modified Lattice Approach to Option Pricing," *The Journal of Futures Markets*, Vol 13 No 5, (August), pp563–577.

157. Trippi, Robert R., Edward A. Brill, and Richard B. Harriff, 1992, "Pricing Options on an Asset with Bernoulli Jump-Diffusion Returns," *The Financial Review*, Vol 27 No 1, (February), pp59–79.

158. Tversky, Amos and Daniel Kahneman, 1974, "Judgement under Uncertainty: Heuristics and Biases," *Science*, Vol 185, (September 27), pp1124–1131.

159. Tversky, Amos and Daniel Kahneman, 1981, "The Framing of Decisions and the Psychology of Choice," *Science*, Vol 211, (January 30), pp453–458.

160. Wallace, Anise, 1980, "Is Beta Dead?," *Institutional Investor*, Vol 14 No 7, (July), pp23–30.

161. Whaley, Robert, 1981, "On the Valuation of American Call Options on Stocks with Known Dividends," *The Journal of Financial Economics*, Vol 9 No 2, (June), pp207–211.

162. Wiggins, James B., 1987, "Option Values under Stochastic Volati-lity," *The Journal of Financial Economics*, Vol 19 No 2, (December), pp351–372.

163. Wilmott, Paul, 1998, *Derivatives: The Theory and Practice of Financial Engineering*, John Wiley and Sons: Chichester, England.

164. Wilmott, Paul, Jeff Dewynne, and Sam Howison, 1993, *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press: Oxford, England.

165. Wilmott, Paul, Sam Howison, and Jeff Dewynne, 1997, *The Mathematics of Financial Derivatives: A Student Introduction*, Cambridge University Press: Cambridge, England.

166. Xu, Xinzhong and Stephen J. Taylor, 1995, "Conditional Volatility and the Informational Efficiency of the PHLX Currency Options Market," *The Journal of Banking and Finance*, Vol 19 No 5, (August), pp803–821.

Table E.1: Alphabets and Numerical Equivalences

Greek[a]				NATO Phonetic		Roman (Latin)[a]	
α	A	Alpha	1	A	Alpha	A	50; 500
β	B	Beta	2	B	Bravo	B	300
γ	Γ	Gamma	3	C	Charlie	C	100
δ	Δ	Delta	4	D	Delta	D	500
ϵ	E	Epsilon	5	E	Echo	E	250
ζ	Z	Zeta	7	F	Foxtrot	F	40
η	H	Eta	8	G	Golf	G	400
θ	Θ	Theta	0	H	Hotel	H	200
ι	I	Iota	10	I	India	I	1
κ	K	Kappa	20	J	Juliett	J	–[b]
λ	Λ	Lambda	30	K	Kilo	K	250
μ	M	Mu	40	L	Lima	L	50
ν	N	Nu	50	M	Mike	M	1,000
ξ	Ξ	Xi	60	N	November	N	90
o	O	Omicron	70	O	Oscar	O	11
π	Π	Pi	80	P	Papa	P	400
ρ	R	Rho	100	Q	Quebec	Q	90; 500
σ	Σ	Sigma	200	R	Romeo	R	80
τ	T	Tau	300	S	Sierra	S	7;70
υ	Υ	Upsilon	400	T	Tango	T	160
ϕ	Φ[c]	Phi	500	U	Uniform	U	–[d]
χ	X[c]	Chi	700	V	Victor	V	5
ψ	Ψ[c]	Psi	700	W	Whiskey	W	–[e]
ω	Ω	Omega	800	X	X-Ray	X	10
				Y	Yankee	Y	150
				Z	Zulu	Z	2,000

[a] Some information from Lewis et al. (1942, p1161). The book is out of print and the publisher defunct.

[b] Originally the same as I.

[c] The Greek letters Φ, X, and Ψ were not needed in the medieval Latin alphabet. However, the Romans used them as numerical symbols, writing D (or M), X, and L, respectively.

[d] Originally the same as V.

[e] Not used in medieval Latin.

Index

$\mathcal{A}$
a.e. (almost everywhere), 154
acute angle, 97
affine function, 121
Against the Gods (book), 232
Agostini, Giulio, vi
Akgiray, Vedat, 199, 231
Allen, Jeffrey G., 3, 47, 231
Almazan, Andres, vi, 212
American call, 28, 132, 165
  approx. pricing formula, 166
  c.f. European call, 132, 165
  early exercise, 132, 133, 165
  exact pricing formulae, 166
  perpetual, 168
    exact pricing formula, 168
American option
  early exercise, 105
  exercise boundary, 105
  valuation analogy, 105, 208
American put
  approx. pricing formulae, 166
  early exercise, 165
  high contact condition, 166
  no exact pricing formula, 165, 166
  perpetual, 34
    exact pricing formula, 168
analytic function, 15, 74
  as a power series, 15
angular velocity, 91, 92
annulus, 105
Anton, Howard, 97, 100, 231
ARCH model, 200
  formula, 200
Archimedes' Principle, 83
arithmetic average, 203
arithmetic Brownian motion, 30, 33, 139, 142–144, 187
Arnold, Tom, vi, 60, 115, 179, 211, 231
Asian option, 36, 184
asset-or-nothing option, 27, 33, 126, 127, 129, 164
asymptote, 138, 189
asymptotic, 138
autocorrelation, 131
  and mean reversion, 170
  and option pricing, 131

$\mathcal{B}$
Bachelier's formula
  generalization (without $S = X$), 141
  generalization (without $S = X$, or $r = 0$), 144
  original ($S = X$, $r = 0$), 143
Bachelier, Louis, 142, 145, 231
  birth place and date, 142
  c.f. Einstein, Albert, 142
Ball, Clifford A., 138, 231

Barclays Global Investors, vi
Barings, 186
Barone-Adesi, G., 166, 231
barrier option, 27, 33, 118, 126, 128, 163
    parity relationship, 126, 163
Bartter, Brit J., 136, 243
basis change, 96
basis points, 149
Bates, David S., 170, 231
Bates, Mary Chris, vi
Baxter, Martin, 180, 232
Bayes' Theorem, 215, 224, 225
BDS Test, 188
Bennett, Grahame, 102
Bera, Anil, 200, 232
Berndt, Ernst K., 201, 232
Bernoulli, Daniel, 37, 232
Bernstein, Peter L., 217, 232
bet (digital option), 126
bid-ask spread, 37, 186, 204
Biger, Nahum, 232
binary (i.e., digital) option, 27, 33, 126, 127, 129
binomial coefficient, 160
binomial option pricing, 136, 176
binomial tree, 124
Black, Fischer, 29, 122, 123, 131, 132, 135–137, 139, 142, 162, 165, 166, 178, 190, 232, 233
Black-Scholes formula
    approximation, 143–145
    for call, 122
    for put, 126
    implied volatility, 148
    in your head, 31, 144
    summary of parts, 125
    with continuous dividend, 122, 176
Black-Scholes PDE, 118, 120, 121, 166, 175, 182, 184
Bohr, Niels, 217
Bollerslev, Tim, 200, 201, 233
bond
    promised yield, 187
    yield-to-maturity, 187
bond velocity, 197, 198
bootstrapping
    spot curve, 172
    swap curve, 173
    zero-coupon yield curve, 173
Boston Stock Exchange, 135
boundary conditions, 121, 166–168, 182
Boyle, Phelim P., 124, 233
Brady bonds, 201
Brealey, Richard A., 211, 233
Brenner, Menachem, 143, 233
Brill, Edward A., 245
Brock, William A., 188, 234
Brown, Robert, 234
Brown, Thomas Kite, 241
Brownian motion, 176, 181
BSE, 135
Bubnys, Edward, 232
Buff, Klara, vi
Butler, Alex, vi, 225

C

callable bonds, 151
call option: *see American call, Black-Scholes formula, European call, perpetual ...*
Canby, Henry Seidel, 241
Capital Ideas (book), 232
CAPM, 38, 54, 137, 179, 189, 190

instantaneous, 137
Cartesian coordinates, 96
cash-or-nothing option, 27, 33, 126, 127, 129, 164
casinos
effectively risk-neutral, 207
CBOE, 135, 228
CBOT, 135, 201
Chance, Don M., 127, 146, 234
Chang, Jinpeng, vi, 61
chaos theory, 38, 187, 188
BDS test, 188
Chaput, Scott, vi
characteristic equation, 89
cheapest-to-deliver (CTD) T-bond, 201
Chesney, Marc, 162, 234
Chicago Board of Trade, 135, 201
Chicago Board Options Exchange, 135
Chicago Stock Exchange, 135
Chordia, Tarun, 205, 234
Chou, Ray Y., 233
Christensen, Peter Ove, 234
Christensen, Ronald, 212, 234
collateralized mortgage obligations (CMO), 151
Comparison test, 101
complex
function, 15, 75
plane, 15, 75
compression to par, 151
conditional probability, 213–215
condom question, 24
connected sets, 105
path, 105
simply, 105
continuously compounded returns, 28, 31, 37, 143
assumed process, 139
convexity
callable bonds, 151
mortgage-backed securities, 32, 151
of bonds, 38, 39, 189, 195–198
myth, 198
of options, 116, 195
formula, 117
summary table, 199
Conze, Antoine, 124, 234
Cootner, Paul H., 142, 231, 234
cost of carry, 172
coupon reinvestment rate risk, 193
Cox and Ross technique, 181
Cox, John C., 133, 136, 178, 195, 235
Crack, Timothy Falcon, 1, 25, 115, 123, 125, 126, 132, 136, 158, 165, 171, 176, 179, 181, 187–189, 195, 198, 211, 231, 235
Crash of 1987, 153
CSE, 135
currency translated option, 177
Curry, Sean, vi

$\mathcal{D}$

de Sorbon, Robert, 142
de-trended prices, 180
deep discount bond, 195
default risk, 39, 149, 186
DeGroot, Morris H., 97, 172, 207, 210, 226, 236
delta of an option, 26–28, 117, 175
defined, 124
for a knock-out, 27, 128, 130

illustrated, 119
intuition, 132
numerically, 124
delta-hedge, 26, 29, 116, 118
Derman, Emanuel, 27, 160, 162, 236
determinant of a matrix, 96, 97
Dewynne, Jeff, 246
differential equation, 19, 89, 90, 118, 166, 167
diffuse prior, 211
digital option, 27, 33, 126, 127, 129, 164
discount bond, 39
dividend capture, 133
Dixit, Avinash K., 212, 236
double-barrier knock-out option, 33, 163
Dow Jones Dogs, 54, 228
down barrier, 33
down-and-out, 27, 128, 130
Dunedin Stock Exchange, 135
duration, 39, 191–198, 201, 202
as a weighted average, 192
closed-form formula, 198
common misconception, 189
continuous case, 196
effect of term-structure shifts on, 198
formula, 192
modified, 196
summary table, 199
units of measurement, 192
duration and convexity
summary table, 199
dynamic replication, 178
technical requirement, 178

$\mathcal{E}$

early exercise, 132
of call, 133, 165
of put, 165
Edwards, Franklin R., 236
efficient markets hypothesis, 180
Einstein, Albert, 60, 142, 236
c.f. Bachelier, Louis, 142
Engle, Robert F., 200, 236
entire function, 15
equivalent martingale measure, 180
equivalent measure, 180
errata, 10
eurodollar forward, 31, 149
eurodollar futures, 31, 149, 173, 201
European call
formula, 122
perpetual, 168
European put
formula, 126
perpetual, 168
Evans, Merran, 236
event risk, 147
Exchange Paradox
the question, 41
the solution, 210–212
exercise boundary, 105
exotic option, 27, 123, 124
pricing summary, 170
extension risk, 150

$\mathcal{F}$

Fabozzi, Frank J., 189, 236
Fabozzi, T. Dessa, 189, 236
factorial, 24, 112
Fama, Eugene, F., 170, 190, 199, 236, 237
Farhadi, Allesio, vi

Farlow, Stanley J., 120, 237
Fed discount rate, 228
Fed funds rate, 228
Feller, William, 185, 186, 214, 215, 237
Fermi, Enrico, 217
Feynman, Richard P., 68, 217
Figures
  call price versus stock and futures, 140
  call price, delta, & gamma, 119
  call vega, 129
  lighthouse setup, 92
  number of cubes on chessboard (A), 94
  number of cubes on chessboard (Q), 20
  power calls, 161
  Road Race Analogy, 81
  S-E-N Problem, 111
  time value, 134
  two triangles, 83
fiscal policy, 227, 228
Fisher, Lawrence, 195, 237
Fleming, Jeff, 174, 237
forex options
  Garman-Kohlhagen formula, 169
forex swap, 173
forward contract, 31, 35, 39, 149, 172, 191
  and convenience yield, 172
  and cost of carry, 172
  and dividends, 172
  and storage costs, 172
forward rate, 37, 38, 40, 187, 190, 202, 204
fractals, 188
Fraser, Michael K., 124, 237
free cash flows, 55, 228
  and debt, 55, 228
French, Kenneth R., 170, 190, 237
Fudenberg, Drew, 221, 237
Fundamental Theorem of Algebra, 84
futures contract, 191
  eurodollar, 31, 149, 173, 201

$\mathcal{G}$
game theory, 217, 221
gamma of an option, 25, 116–118, 175, 195
  formula, 117
gamma-hedge, 118
GARCH model, 39, 199–201
  formula, 200
Garman, Mark B., 237
Garman-Kohlhagen formula
  for forex, 169
Gatto, M.A., 124, 238
Gauss' test, 101
Gaussian
  process, 33
GBM, 32, 183
Genius (book), 68, 238
geodesic, 62
geometric average, 203
geometric Brownian motion, 30, 32, 36, 139, 142, 143, 183
Geske, Robert, 166, 237
Girsanov, I.V., 238
Gleick, James, 68, 188, 238
Goldman, B.M., 124, 238
Goodman, Victor W., vi
Grabbe, J. Orlin, 238
great circle, 62
Greek alphabet, 248

Greene, William H., 201, 238

$\mathcal{H}$

Hammer, Jerry A., 132, 160, 238
harmonic average, 203
Harriff, Richard B., 245
Harrison and Kreps technique, 181
Harrison, J. Michael, 142, 238
Harvard, 166
Hastings, Nicholas, 236
Haug, Espen Gaarder, 163, 184, 238
Haung, Chi-fu, 239
Hendry, David F., 200, 239
Heron's Formula, 18
Higgins, Matthew, 200, 232
high contact condition, 166
historical volatility, 35, 172
    estimator, 172
Hoel, Tim, vi, 104
Holland, A.S.B., 15, 74, 75, 239
Homaifar, Ghassem, 195, 197, 198, 242
How to Lie with Statistics (book), 230, 239
Howison, Sam, 246
HP12C, 261
HP17B, 261
HP19B, 261
Hsieh, David A., 234
Huff, Darrell, 230, 239
Hull, John C., 124, 137, 150, 152, 153, 162, 163, 166, 201, 232, 239
Hunter, William C., 124, 239
hypotenuse, 63, 97

$\mathcal{I}$

Ikeda, Masayuki, 163, 240
illegal questions, 39, 51
immunization, 191, 194–196
implied standard deviation, 35, 174
implied vol, 148
implied volatility, 31, 33, 35, 147, 148, 159, 160, 174
    definition, 148
inclusion-exclusion formula, 85
Indiana University, v, vi
indicator function, 125, 165
induction, 87, 88
infinite gamma
    knock-out option, 118
    standard option, 118
Ingersoll, Jonathan E., 195, 235
integral test, 100
intended audience, v
internal rate of return, 187
interview books, v, 47
intrinsic value, 28
Itô's Lemma, 155, 181, 182
    two dimensional, 183
Itô, Kiyoshi, 155

$\mathcal{J}$

Jacobian, 96, 97
Jarrow, Robert, 33, 159, 178, 239
Jones, Frank J., 192, 198, 239
jump diffusion process, 29, 136, 137
jump process, 29, 135–138, 178
    cf. stochastic volatility, 162

$\mathcal{K}$

Kahn, Ronald N., 198, 240
Kahneman, David, 204, 240, 245
Kani, Iraj, 27, 160, 162, 236
Klymchuk, Taras, vi, 85, 153, 155
knock-out option, 27, 33, 126, 128

- double-barrier, 33, 163
- gamma, 118
- parity relationship, 126, 163

Kohlhagen, Steven W., 237
Kotz, Samuel, 240
Krämer, Walter, 188, 240
Krause, Robert, 162, 240
Kritzman, Mark, 185, 187, 190, 191, 195, 240
Kroner, Kenneth F., 233
Kunitomo, Naoto, 163, 240
kurtosis, 162

$\mathcal{L}$

Lacey, Nelson J., 198, 241
Latané, Henry A., 148, 241
Latin alphabet, 203, 248
lattice pricing, 176
LEAPS option, 146
LeBaron, Blake, 234
Ledoit, Olivier, vi, 171, 186, 188, 217, 235
Lee, Moon, 137, 242
Lehman Bros., 32
lemma, 73
leptokurtosis, 162
Let's Make a Deal (Monty Hall)
- the question, 43
- the solution, 214

Lewis, Michael M., 32, 241
Lewis, William Dodge, 241, 248
Liar's Poker (book), 32, 241
LIBOR, 31, 34, 53, 149, 172
Lin, Victor H., vi, 104
Listed Options Quotations, 133
literal numbers, 84
Litterman, Robert, 192, 198, 241
Lo, Andrew W., 131, 170, 171, 241
Lochoff, Roland, 198, 240
Longstaff, Francis, 147
Lorenz, Edward N., 187, 241
lottery tickets, 146
lowest common multiple, 79
- finding it, 79

Lown, Cecily, vi
Lown, Marianne, vi

$\mathcal{M}$

Ma, Christopher K., 243
Ma, Cindy W., 236
Macaulay, Frederick Robertson, 192, 195, 241
MacKinlay, A. Craig, 131, 170, 171, 241
MacMillan, L.W., 166, 241
macroeconomic policy, 54, 227
Mandelbrot, Benoit, 188
Manhattan Project, 217
Marcus, Alan J., vi, 120
Marler, Patty, 47, 241
martingale, 178, 180
- method
  - Cox and Ross, 181
  - Harrison and Kreps, 181

Maslen, David, vi, 103, 108
MathWorks Inc, vi
MATLAB, vi, 112, 128
Mattia, Jan Bailey, 47, 241
maximum likelihood estimation, 201
mean reversion, 34, 170
- and autocorrelation, 170
- and hypothesis testing, 171

measure, 180
- equivalent, 180
- equivalent martingale, 180

Mehran, Jamshid, 195, 197, 198, 242
Melino, Angelo, 162, 242
Merton, Robert C., 29, 122, 128, 131, 135–138, 155, 166, 178, 181, 242
Metallgesellschaft, 186
Minton, Bernadette A., 173, 242
MIT, v, vi, 15, 28, 44, 46, 47, 68, 153, 155, 187, 201
mode, 174
Modigliani, Franco, 153
monetary policy, 227, 228
Mongan, John, 3, 242
Monte-Carlo simulation, 32, 123
Monty Hall problem, 43, 214
mortgage-backed securities, 32, 150–152
    CMO, 151
    compression to par, 151
    convexity, 32
    extension risk, 150
    pass-through, 151
    prepayment risk, 150
Moshkevich, Vince, vi, 102
Mullens, David W., 190, 242
Murphy, Gareth, 160, 242
Musiela, Marek, 163, 180, 242
MUT, 228
Myers, Stewart C., 211, 233

$\mathcal{N}$

nth root test, 99
Naik, Vasanttilak, 137, 242
Nalebuff, Barry J., 212, 236
Nasar, Sylvia, 221, 243
Nash equilibrium, 221, 225
Nash, John Forbes, 221
Natenberg, Sheldon, 243
NATO phonetic alphabet, 248
Nawalkha, Sanjay K., 189, 195, 198, 235, 241
New York Stock Exchange, 135
Ni, Bingjian, vi, 61, 70, 208
Nikkei, 4, 53
no-arbitrage technique, 29, 120, 135–137, 178
    versus equilibrium argument, 137
Nobel Prize
    1994, 221
non-central moment, 154
numeraire, 179
NYSE, 135

$\mathcal{O}$

ODE, 19, 89, 90, 166, 167
    characteristic equation, 89
    linear homogeneous, 89
    nonhomogeneous, 90
open market operation, 228
Oppenheimer, J. Robert, 217
option pricing
    summary table (exotic), 170
    summary table (plain vanilla), 169
options
    stochastic volatility, 162
    term-structure of volatility, 162
    volatility smile, 160
Options Clearing Corporation, 146, 243
overnight repo, 228

$\mathcal{P}$

Paris Bourse, 135, 142
Park, Hun, 232

Parkinson, Michael, 166, 243
parlay card, 44
partial sums, 72, 98
Pascal's Triangle, 106
pass-through, 151
path dependence, 163, 170
path-dependent option, 33, 163
payoff diagram, 26, 126, 142
PDE, 120, 124, 166, 176, 181
  Black-Scholes, 118, 120, 121, 166, 175, 182, 184
  favourite book on, 120
  pricing approach, 182
Peacock, Brian, 236
Pearce, David, W., 228, 243
perfect squares, 67
perpetual
  American call, 168
  American put, 34, 168
  European call, 168
  European put, 168
  European up-and-out, 25
Peterson, Richard L., 131, 171, 243
PIBOR, 53
Pliska, S.R., 238
Poisson event, 29
Poisson process, 29, 136
  pure, 136
polar coordinates, 74, 96
polynomials, 84, 106
  and Pascal's Triangle, 106
  degree of, 84
  roots of, 84, 89
  zeroes of, 84
Porro, Eva, vi, 68
positively skewed, 121
Poterba, James, 170, 243
power option, 32, 156–159, 161
  approx. pricing formula, 159
  call delta, 157
  exact pricing formulae, 156
    with continuous dividend, 159
  payoff diagrams, 161
powered option, 33, 159
  case ($\alpha = 2$), 160
  general formula, 160
PPE (physical plant and equipment), 55
prepayment risk, 150
price risk, 193
Price, Katie, vi
prime number, 23, 24, 109, 114
prisoners' dilemma, 221
product call, 36, 176, 182
promised yield, 187
Prymas, Wolfgang, vi
put-call parity, 26, 27, 117, 122, 145
  stated, 126
put option: *see American put, Black-Scholes formula, European put, perpetual ...*
Pythagoras' Theorem, 63, 83

$\mathcal{Q}$

quadratic formula, 89
quanto option, 177

$\mathcal{R}$

Raabe's test, 99
radially symmetric, 104
radians, 64, 91, 92
Radon-Nikodym derivative, 180
Rakotomalala, Marc, vi
rate of return on a bond, 187
ratio test, 99

Rendleman, Richard J., 136, 148, 241, 243
Rennie, Andrew, 180, 232
replicating portfolio, 29, 123, 124, 178
repo, 228
repurchase agreement, 228
reverse repo, 228
Richardson, Matthew, 171, 243
Riemann calculus, 153
right skewed, 121
risk-neutral
  pricing, 137
    bullet point review, 178–182
  probabilities, 180
  world
    country specific views, 177
risk-neutrality, 207, 209, 223
Ritchey, Robert J., 243
Ritchken, Peter, 124, 243
Roll, Richard, 166, 190, 243, 244
Roman
  alphabet, 248
  numerals, 248
Ross, Stephen A., 136, 195, 235
Roth, Jason, vi, 86
Rubinstein, Mark, vi, 18, 133, 136, 178, 235
Rudd, Andrew, 178, 239
Runde, Ralf, 188, 240
Rutkowski, Marek, 163, 180, 242

S

S-E-N problem, 110
safe-cracking, 14, 67
Salomon
  bond-arb group, 192
Samuelson, Paul A., 142, 143, 155, 244
Sankarasubramanian, L., 124, 243
Sato, Naoki, vi, 23, 73
Scheinkman, José, 192, 198, 241
Scholes, Myron S., 122, 123, 131, 132, 135, 137, 139, 142, 162, 165, 178, 232, 233
Scott, Louis, 162, 234, 244
SEC, 40
Security Market Line, 38, 189
sequences, 98
  convergence, 98
serial correlation, 131
series, 98–101
  absolute convergence, 101
  Comparison test, 101
  conditional convergence, 101
  convergence, 98
  convergence tests, 99–101
  Gauss' test, 101
  integral test, 100
  n^{th} root test, 99
  Raabe's test, 99
  ratio test, 99
Sharpe, William F., 136, 190, 244
Shen, Yi, vi, 61, 72
Shepp, L.A., 124, 238
similarity solution, 176, 184
simple jump process, 136, 178
singleton set, 62
skewness, 162
Smelyansky, Valeri, vi
smile curve, 160
Smith, Clifford W., 29, 142, 244
Snyder, Gerard L., 128, 244
SOH-CAH-TOA, 97
Sommer, Louise, 232

Sorbonne, 142
Sørensen, Bjarne, 234
Sosin, H.B., 124, 238
Spiegel, Murray R., 84, 101, 102, 203, 244
spot curve, 172
Sprenkle, Case M., 244
St. Petersburg, Game, 37, 185
St. Petersburg, Paradox, 185
standard Brownian motion, 32, 139, 154
Statistical Distributions (book), 236
stochastic calculus, 32, 153, 155, 178
stochastic volatility
 cf jumps, 162
stochastic volatility options models, 162
stock buy back, 55
stock-numeraire world, 180
 probabilities
  in $N(d_1)$, 180
storage costs, 172
Stowe, David W., 124, 239
straddle, 31, 148
Subrahmanyam, Avanidhar, 205, 234
Subrahmanyam, Marti G., 143, 233
Sullivan, Edward J., 143, 245
Sullivan, Sara, 160, 245
Summers, Lawrence, 170, 243
Sundaresan, Suresh, 151, 152, 245
Suojanen, Noah Suo, 3, 242
swaps
 and default risk, 172, 173
 buyer, 172
 curve, 172, 173
 FOREX, 173
 rates, 172
 rates vs default-risky debt, 174
 rates vs treasuries, 174
 who is long the swap, 172

$\mathcal{T}$

T-bond futures, 201
Tamir, Dahn, vi, 77, 81
Taylor series, 75, 195
Taylor, Stephen J., 174, 246
Ted spread, 149
Teller, Edward, 217
Tenorio, Juan, vi, 61
term-structure, 187, 194, 196, 197
 height, slope, curvature shifts, 198
 implied forward rate, 37, 38, 187, 190
term-structure of volatility, 162
theta of an option, 175
 and time decay, 117
Thinking Strategically (book), 212, 236
Thorp, Edward O., 132, 245
Tian, Yisong, 245
tick size, 135
time decay (theta), 117
time value, 132, 133, 139
 defined, 29
 illustrated, 134
Tirole, Jean, 221, 237
tombstone in WSJ, 52, 227
Torous, Walter N., 138, 231
tranche, 151
trigonometric functions, 97
Trippi, Robert R., 138, 245
Turnbull, Stuart, 33, 159, 162, 239, 242
Turner, Paul, vi, 106

Tversky, Amos, 204, 240, 245
Twelve Balls Problem
  the question, 12
  the solution, 61

$\mathcal{U}$
under water, 147
Unique Factorization Theorem, 84
University of Paris, 142
up barrier, 33
up-and-out, 27, 130
utility, 185
Utts, Jessica, 212, 234

$\mathcal{V}$
variance ratio test, 131
vega of an option, 129
velocity of a bond, 197
Vigodner, Alex, vi, 225
Vijh, Anand, 124, 243
Viswanathan, S., 124, 234
Vivian, Nick, vi
volatility
  skew, 174
  smile, 162, 174
  term-structure, 162
von Neumann, John, 217
Voropaev, Mikhail, vi, 144

$\mathcal{W}$
Wallace, Anise, 190, 245
Wang, Jiang, 131, 170, 171, 241
Watson, Thomas C., vi, 73, 83
Weak Law of Large Numbers, 185, 207, 209, 226
  fails, 185
web pages
  `www.BasicBlackScholes.com`, 1, 25, 235, 261
  `www.InvestmentBankingJobInterviews.com`, 4, 10
  `www.OneChicago.com`, 30
Weil, Roman, 195, 237
Weithers, Timothy M., 143, 245
Whaley, Robert, 166, 231, 245
White, Alan, 162, 163, 239
Wiener process, 29, 176, 177
Wiener, Norbert, 155
Wiggins, James B., 162, 245
Wilmott, Paul, 121, 166, 175, 182, 246
World Series problem, 42, 212

$\mathcal{X}$
Xu, Xinzhong, 174, 246

$\mathcal{Y}$
Yale University, 188
yield curve, 39, 191, 192, 195, 198, 201
  convexity myth, 198
  height, slope, curvature shifts, 198
  strategies, 192
  twists, 198
  zero-coupon, 172
yield on a bond, 187
yield-to-maturity, 187

$\mathcal{Z}$
Z-score
  in $N(d_1)$, 180
  in $N(d_2)$, 180
zero-coupon bonds, 39, 151, 189, 192
zero-coupon yield curve, 172
zero-duration portfolio, 192

BY THE SAME AUTHOR

Basic Black-Scholes: Option Pricing and Trading

Timothy Falcon Crack

BSc (HONS 1^{st} Class), PGDipCom, MCom, PhD (MIT), IMC

This new book gives extremely clear explanations of Black-Scholes option pricing theory, and discusses direct applications of the theory to option trading. The presentation does not go far beyond basic Black-Scholes for three reasons: First, a novice need not go far beyond Black-Scholes to make money in the options markets; Second, all high-level option pricing theory is simply an extension of Black-Scholes; and Third, there already exist many books that look far beyond Black-Scholes without first laying the firm foundation given here. The trading advice does not go far beyond elementary call and put positions because more complex trades are simply combinations of these. The appendix includes Black-Scholes option pricing code for the HP17B, HP19B, and HP12C. An accompanying spreadsheet allows the user to forecast transactions costs for option positions using simple models.

PUBREFEHOTS:08.01.2004.702,665.OU adapted from ISBN 0-9700552-3-4