Derivative Securities
G63.2791, Fall 2004
Mondays 7:10–9:00pm
109 WWH

Instructor: Robert V. Kohn. Office: 612 Warren Weaver Hall. Phone: 998-3217. Email: kohn@cims.nyu.edu. Office hours: Mondays 5-6pm, Wednesdays 5-6pm , after class or by appointment. Web: www.math.nyu.edu/faculty/kohn.

Teaching Assistant: Paris Pender. Office: WWH 810. Phone: 998-3204. Email: pender@cims.nyu.edu. Office hours: Thursdays 12:15-1:15 and Fridays 5-6.

Content: An introduction to arbitrage-based pricing of derivative securities. Topics include: arbitrage; risk-neutral valuation; the log-normal hypothesis; binomial trees; the Black-Scholes formula and applications; the Black-Scholes partial differential equation; American options; one-factor interest rate models; swaps, caps, floors, swaptions, and other interest-based derivatives; credit risk and credit derivatives.

Lecture notes: Lecture notes, homework assignments, etc. will be posted on my web page in pdf format – normally within a day of when they are distributed. I'll build a fresh set of notes, homeworks, etc as we go along, but at the top of the course web page you'll find a link to the page built when I last taught the class, in Fall 2000. This fall's version will be similar, except that we'll do one or two lectures on credit risk and credit derivatives near the end of the semester.

Prerequisites: Calculus, linear algebra, and discrete probability. Concerning probability: students should be familiar with concepts such as expected value, variance, independence, conditional probability, the distribution of a random variable, the Gaussian distribution, the law of large numbers, and the central limit theorem. These topics are addressed early in most undergraduate texts on probability, for example K-L Chung and F. Aitsahlia, *Elementary probability theory : with stochastic processes and an introduction to mathematical finance* Springer 2003, on reserve in the CIMS library.

Course requirements: There will be approx 7 homework sets, one every couple of weeks. Collaboration on homework is encouraged (homeworks are not exams) but registered students must write up and turn in their solutions individually. There will be an in-class final exam.The first class is Monday Sept 13; the last class is Monday Dec 13; the final exam is Mon Dec 20.

Books: We will not follow any single book linearly. However to master the material of this course you should expect to do plenty of reading. I recommend purchasing at least these two books:

- J.C. Hull, *Options, futures and other derivative securities*, 5th edition.
- M. Baxter and A. Rennie, *Financial calculus: an intoduction to derivative pricing*, Cambridge University Press, 1996.

The NYU bookstore has ordered about 30 copies of each; you may be able to save money by buying them used. Earlier editions of Hull will be sufficient for this class, but the 5th edition has some new sections on advanced or rapidly-developing topics like credit. These two books go far beyond the scope of this course; roughly, they cover both Derivative Securities and its spring sequel Continuous Time Finance.

Here are some additional books you may wish to buy or at least consult:

- R. Jarrow and S. Turnbull, *Derivative securities*, Southwestern, 2nd edition
- M. Avellaneda and P. Laurence, *Quantitative Modeling of Derivative Securities*, CRC Press, 1999.
- P. Wilmott, S. Howison, and J. Dewynne, *The mathematics of financial derivatives - a student introduction*, Cambridge University Press, 1995
- S. Neftci, *An introduction to the mathematics of financial derivatives*, Academic Press, 2nd edition.
- S. Shreve, *Stochastic calculus for finance I: The binomial asset pricing model*, Springer-Verlag, 2004

All these books are on reserve in the CIMS library. Some brief comments: Jarrow-Turnbull has roughly the same goals as Hull. I find it clearer on some topics, though Hull is the industry standard. Wilmott-Howison-Dewynne is especially good for people with background in PDE but unfortunately it de-emphasizes risk neutral valuation. Neftci provides a good introduction to the most basic aspects of stochastic differential equations and the Ito calculus (the first edition is sufficient for this purpose). Shreve's book, hot off the presses, is a lot like the first part of Baxter-Rennie, and a lot like the first half of this course. (His *Stochastic calculus for finance II: continuous-time models* was just published; it corresponds roughly to our classes Stochastic Calculus and Continuous Time Finance.)

An FAQ about probability: Math finance students often ask me for suggestions how to enhance their knowledge of probability, for example in connection with the class Stochastic Calculus. Professor Goodman is teaching Stochastic Calculus this fall, and he'll undoubtedly provide his own reading list. But here are some suggestions of my own:

(a) CALCULUS-BASED PROBABILITY. This material (at the level usually taught to upper-level math majors) is a prerequisite for Stochastic Calculus. There are many good texts. The one by K-L Chung and F. Aitsahlia (*Elementary probability theory with stochastic processes and an introduction to mathematical finance*, Springer-Verlag, 2003) has the advantage of including some material at the end that overlaps with Derivative Securities. Earlier editions (by Chung alone) cover the probability without the finance; they're just as useful.

(b) MORE ADVANCED PROBABILITY BOOKS. Past students have found it useful to read parts of the book by Z. Brzezniak and T. Zastawniak (*Basic stochastic processes : a course through exercises*, Springer-Verlag, 1999) and/or the one by S. Resnick ((*A probability path*, Springer-Verlag, 1999). The former includes a lot of material on

Markov chains; the latter includes an introduction to measure theory as it interfaces with probability. Neither book covers stochastic calculus or its applications to finance.

(c) STOCHASTIC CALCULUS. Students with relatively little background should certainly look at S. Neftci's book (*An introduction to the mathematics of financial derivatives*, Academic Press), with the warning that it only scratches the surface. You might also find T. Mikosch's book (*Elementary stochastic calculus with finance in view*) helpful, but be warned that it's more a list of facts than an explanation of them. Students with sufficient background find J.M. Steele's book a pleasure to read (*Stochastic calculus and financial applications*, Springer-Verlag, 2001). The newest addition to the list is Volume II of S. Shreve's book (*Stochastic calculus for finance II: continuous-time models*, Springer-Verlag, 2004). Its first half corresponds to our Stochastic Calculus course; its second half is similar to our Continuous Time Finance course.

All the probability books suggested above are on reserve in the CIMS library (except Mikosch, which is on order; it will go on reserve when it arrives).

Derivative Securities – Fall 2004 – Section 1
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences.

Forwards, puts, calls, and other contingent claims. This section discusses the most basic examples of contingent claims, and explains how considerations of arbitrage determine or restrict their prices. This material is in Chapters 2 and 3 of Jarrow and Turnbull, and Chapters 1, 3 and 8 of Hull (5th edition). We concentrate for simplicity on European options rather than American ones, on forwards rather than futures, and on deterministic rather than stochastic interest rates.

****

The most basic instruments:

Forward contract with maturity T and delivery price K.

$$
\begin{aligned}
\text{buy a forward} &\leftrightarrow \text{hold a long forward}\\
&\leftrightarrow \text{holder is obliged to buy the}\\
&\quad\ \text{underlying asset at price K on date T.}
\end{aligned}
$$

European call option with maturity T and strike price K.

$$
\begin{aligned}
\text{buy a call} &\leftrightarrow \text{hold a long call}\\
&\leftrightarrow \text{holder is entitled to buy the}\\
&\quad\ \text{underlying asset at price K on date T.}
\end{aligned}
$$

European put option with maturity T and strike price K.

$$
\begin{aligned}
\text{buy a put} &\leftrightarrow \text{hold a long put}\\
&\leftrightarrow \text{holder is entitled to sell the}\\
&\quad\ \text{underlying asset at price K on date T.}
\end{aligned}
$$

These are *contingent claims*, i.e. their value at maturity is not known in advance. Payoff formulas and diagrams (value at maturity, as a function of S_T=value of the underlying) are shown in the Figure.

Any long position has a corresponding (opposite) *short* position:

Buyer of a claim has a long position ↔ seller has a short position.

Payoff diagram of short position = negative of payoff diagram of long position.

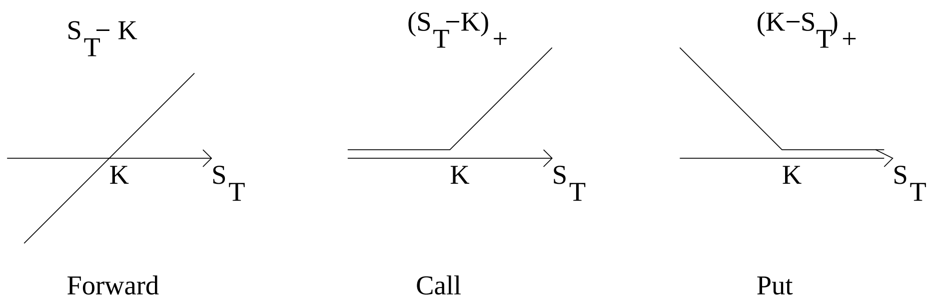


Figure 1: Payoffs of forward, call, and put options.

An *American* option differs from its European sibling by allowing early exercise. For example: the holder of an American call with strike K and maturity T has the right to purchase the underlying for price K at any time $0 \leq t \leq T$. A discussion of American options must deal with two more-or-less independent issues: the unknown future value of the underlying, and the optimal choice of the exercise time. By focusing initially on European options we'll develop an understanding of the first issue before addressing the second.

****

Why do people buy and sell contingent claims? Briefly, to *hedge* or to *speculate*. Examples of hedging:

- A US airline has a contract to buy a French airplane for a price fixed in FF, payable one year from now. By going long on a forward contract for FF (payable in dollars) it can eliminate its foreign currency risk.
- The holder of a forward contract has unlimited downside risk. Holding a call limits the downside risk (but buying a call with strike K costs more than buying the forward with delivery price K). Holding one long call and one short call costs less, but gives up some of the upside benefit:

$$(S_T - K_1)_+ - (S_T - K_2)_+ \qquad K_1 < K_2$$

  This is known as a "bull spread". (See the figure.)

Options are also frequently used as a means for speculation. Basic reason: the option is more sensitive to price changes than the underlying asset itself. Consider for example a European call with strike $K = 50$, at a time t so near maturity that the value of the option is essentially $(S_t - K)_+$. Let $S_t = 60$ now, and consider what happens when S_t increases by 10% to 66. The value of the option increases from about $60 - 50 = 10$

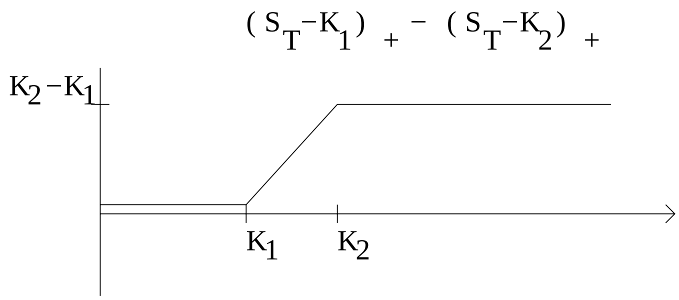


Figure 2: Payoff of a bull spread.

to about $66 - 50 = 16$, an increase of 60%. Similarly if S_t decreases by 10% to 54 the value of the option decreases from 10 to 4, a loss of 60%. This calculation isn't special to a call: almost the same calculation applies to stock bought with borrowed funds. Of course there's a difference: the call has more limited downside exposure.

We assumed the time t was very near maturity so we could use the payoff $(S_T - K)_+$ as a formula for the value of the option. But the idea of the preceding paragraph applies even to options that mature well in the future. We'll study in this course how the Black-Scholes analysis assigns a value $c = c[S_t; T-t, K]$ to the option, as a function of its strike K, its time-to-maturity $T - t$ and the current stock price S_t. The graph of c as a function of S_t is roughly a smoothed-out version of the payoff $(S_t - K)_+$.

Don't be confused: our assertion that "the option is more sensitive to price changes than the underlying asset itself" does *not* mean that $\partial c/\partial S$ is bigger than 1. This expression, which gives the sensitivity of the option to change in the underlying, is called Δ. At maturity the call has value $(S_T - K)_+$ so $\Delta = 1$ for $S_T > K$ and $\Delta = 0$ for $S_T < K$. Prior to maturity the Black-Scholes theory will tell us that Δ varies smoothly from nearly 0 for $S_t \ll K$ to nearly 1 for $S_t \gg K$.

****

Some pricing principles:

- If two portfolios have the same payoff then their present values must be the same.
- If portfolio 1's payoff is always at least as good as portfolio 2's, then present value of portfolio 1 $\geq$ present value of portfolio 2.

We'll see presently that these principles must hold, because if they didn't the market would support arbitrage.

First example: value of a forward contract. We assume for simplicity:

(a) underlying asset pays no dividend and has no carrying cost (e.g. a non-dividend-paying stock);

(b) time value of money is computed using compound interest rate r, i.e. a guaranteed income of D dollars time T in the future is worth $e^{-rT}D$ dollars now.

The latter hypothesis amounts to introducing one more investment option:

Bond worth D dollars at maturity T

$$\begin{aligned} \text{buy a bond} &\leftrightarrow \text{hold a long bond} \\ &\leftrightarrow \text{lend } e^{-rT}D \text{ dollars, to be repaid at time T with interest.} \end{aligned}$$

Consider these two portfolios:

Portfolio 1 – one long forward with maturity T and delivery price K, payoff $(S_T - K)$.

Portfolio 2 – long one unit of stock (present value S_0, value at maturity S_T) and short one bond (present value $-Ke^{-rT}$, value at maturity $-K$).

They have the same payoff, so they must have the same present value. Conclusion:

$$\text{Present value of forward} = S_0 - Ke^{-rT}.$$

In practice, forward contracts are normally written so that their present value is 0. This fixes the delivery price, known as the *forward price*:

$$\text{forward price} = S_0 e^{rT} \text{ where } S_0 \text{ is the spot price.}$$

We can see why the "pricing principles" enunciated above must hold. If the market price of a forward were different from the value just computed then there would be an arbitrage opportunity:

$$\begin{aligned} \text{forward is overpriced} &\rightarrow \text{sell portfolio 1, buy portfolio 2} \\ &\rightarrow \text{instant profit at no risk} \\ \text{forward is underpriced} &\rightarrow \text{buy portfolio 1, sell portfolio 2} \\ &\rightarrow \text{instant profit at no risk.} \end{aligned}$$

In either case, market forces (oversupply of sellers or buyers) will lead to price adjustment, restoring the price of a forward to (approximately) its no-arbitrage value.

Second example: put–call parity. Define

$$\begin{aligned} p[S_0, T, K] &= \text{price of European put when spot price is} \\ & \quad S_0\text{, strike price is } K\text{, maturity is } T \\ c[S_0, T, K] &= \text{price of European call when spot price is} \\ & \quad S_0\text{, strike price is } K\text{, maturity is } T. \end{aligned}$$

The Black-Scholes model gives formulas for p and c based on a certain model of how the underlying security behaves. But we can see now that p and c are related, without knowing anything about how the underlying security behaves (except that it pays no dividends and has no carrying cost). "Put-call parity" is the relation

$$c[S_0, T, K] - p[S_0, T, K] = S_0 - Ke^{-rT}.$$

To see this, compare

Portfolio 1 – one long call and one short put, both with maturity T and strike K; the payoff is $(S_T - K)_+ - (K - S_T)_+ = S_T - K$.

Portfolio 2 – a forward contract with delivery price K and maturity T. Its payoff is also $S_T - K$.

These portfolios have the same payoff, so they must have the same present value. This justifies the formula.

Third example: The prices of European puts and calls satisfy

$$c[S_0, T, K] \geq (S_0 - Ke^{-rT})_+ \quad \text{and} \quad p[S_0, T, K] \geq (Ke^{-rT} - S_0)_+.$$

To see the first relation, observe first that $c[S_0, T, K] \geq 0$ by optionality – holding a long call is never worse than holding nothing. Observe next that $c[S_0, T, K] \geq S_0 - Ke^{-rT}$, since holding a long call is never worse than holding the corresponding forward contract. Thus $c[S_0, T, K] \geq \max\{0, S_0 - Ke^{-rT}\}$, which is the desired conclusion. The argument for the second relation is similar.

****

Note some hypotheses underlying our discussion:

- no transaction costs; no bid-ask spread;
- no tax considerations;
- unlimited possibility of long and short positions; no restriction on borrowing.

These are of course merely approximations to the truth (like any mathematical model). More accurate for large institutions than for individuals.

Note also some features of our discussion: We are simply reaping consequences of the hypothesis of no arbitrage. Conclusions reached this way don't depend at all on what you think the market will do in the future. Arbitrage methods restrict the prices of (related) instruments. On the other hand they don't tell an individual investor how best to invest his money. That's the issue of portfolio optimization, which requires an entirely different type of analysis and is discussed in the course Capital Markets and Portfolio Theory.

****

A word about interest rates. In the real world interest rates change unpredictably. And the rate depends on maturity. In discussing forwards and European options this isn't particularly important: all that matters is the cost "now" of a bond worth one dollar at maturity T. Up to now we wrote this as e^{-rT}. When multiple borrowing times and maturities are being considered, however, it's clearer to use the notation

$$B(t,T) = \text{cost at time } t \text{ of a risk-free bond worth 1 dollar at time } T.$$

In a constant interest rate setting $B(t,T) = e^{-r(T-t)}$. If the interest rate is non-constant but deterministic – i.e. known in advance – then an arbitrage argument shows that $B(t_1,t_2)B(t_2,t_3) = B(t_1,t_3)$. If however interest rates are stochastic – i.e. if $B(t_2,t_3)$ is not known at time t_1 – then this relation must fail, since $B(t_1,t_2)$ and $B(t_1,t_3)$ are (by definition) known at time t_1.

Since our results on forwards, put-call parity, etc. used only one-period borrowing, they remain valid when the interest rate is nonconstant and even stochastic. For example, the value at time 0 of a forward contract with delivery price K is $S_0 - KB(0,T)$ where S_0 is the spot price.

****

Forwards versus futures. A future is a lot like a forward contract – its writer must sell the underlying asset to its holder at a specified maturity date. However there are some important differences:

- Futures are standardized and traded, whereas forwards are not. Thus a futures contract (with specified underlying asset and maturity) has a well-defined "future price" that is set by the marketplace. At maturity the future price is necessarily the same as the spot price.

- Futures are "marked to market," whereas in a forward contract no money changes hands till maturity. Thus the value of a future contract, like that of a forward contract, varies with changes in the market value of the underlying. However with a future the holder and writer settle up daily while with a forward the holder and writer don't settle up till maturity.

The essential difference between futures and forwards involves the timing of payments between holder and writer: daily (for futures) versus lump sum at maturity (for forwards). Therefore the difference between forwards and futures has a lot to do with the time value of money. If interest rates are constant – or even nonconstant but deterministic – then an arbitrage-based argument shows that the forward and future prices must be equal. (I like the presentation in Appendix 3A of Hull (5th edition). It is presented in the context of a constant interest rate, but the argument can easily be modified to handle a deterministically-changing interest rate.)

If interest rates are stochastic, the arbitrage-based relation between forwards and futures breaks down, and forward prices can be different from future prices. In practice they are different, but usually not much so.

****

A word about taxes. Tax considerations are not always negligible. Here are two examples, each closely related to put-call parity.

Constructive sales. An investor holds stock in XYZ Corp. His stock has appreciated a lot, and he thinks it's time to sell, but he wishes to postpone his gain till next year when he expects to have losses to offset them. Prior to 1997 he could have (1) kept his stock, (2) bought a put (one year maturity, strike K), (3) sold a call (one year maturity, strike K), and (4) borrowed Ke^{-rT}. The value of this portfolio at maturity is $S_T + (K - S_T)_+ - (S_T - K)_+ - K = 0$. Since his position at time T is valueless and risk-free, he would have effectively "sold" his stock. Since the present value of items (1)-(4) together is 0, the combined value of the long put, short call, and loan must be the present value of the stock. Thus the investor would have effectively sold the stock for its present market value, while postponing realization of the capital gain till the options matured.

The tax law was changed in 1997 to treat such a transaction as a "constructive sale," eliminating its attractiveness (the capital gain is no longer postponed). A related strategy is still available however: by combining puts and calls with different maturities, an investor can take a position that still has some risk (thus avoiding the constructive sale rule) while locking in most of the gain and avoiding any capital gains tax till the options mature.

Dick Cheney's Halliburton options. When he was nominated for Vice President in 2000, Dick Cheney held "executive stock options" from Halliburton corporation that matured well after he took office. Executive stock options are essentially call options, except (a) they can only be exercised, not sold; (b) if they expire in the money, the gain associated with their exercise is taxed as ordinary income, not as a capital gain. For simplicity let's concentrate on just part of Cheney's portfolio: an option to buy 100,000 shares at K=39.50/share in December 2002. Let's further take the fall 2000 stock price to be $S_0 = 53.00$/share (that's about right) and let's ignore the time value of money (a minor detail). We'll refer to fall 2000 as the "current time" since we're thinking about his situation at that time.

Cheney's problem was this: if continued to hold the options as Vice President he could be viewed as having a conflict of interest. But he could not sell the options prior to maturity. And simply disowning them meant disowning an asset that was likely to be very valuable once it matured.

A common-at-the-time (but flawed) suggestion was that Cheney enter into a forward contract to sell his shares at the time of maturity. Since we're ignoring the time-value of money, the forward price is the present market value, i.e. a contract to sell his stock at $S_0 = 53.00$/share in December 2002 had value 0 in fall 2000. This proposal however had two flaws:

(a) it did not fully eliminate his conflict of interest; and

(b) when taxes are taken into account it didn't even come close to eliminating his conflict of interest.

Concerning (a): the payoff at maturity of the call and forward together is

$$(S_T - K)_+ - (S_T - S_0) = \begin{cases} S_0 - S_T & \text{if } S_T < K \\ S_0 - K & \text{if } S_T > K \end{cases}$$

Thus Cheney's position would not be insensitive to S_T – at least not if $S_T < K$. The best possible result (for him) would be for Halliburton to go bankrupt ($S_T = 0$).

Concerning (b): the after-tax value of the combined call and forward position depends on S_T and also the details of Cheney's finances. The reason is that his profit on the option $(S_T - K)_+$ would be treated as ordinary income but his gain or loss on the forward contract would be treated as a capital gain or loss. Suppose the stock price went up, i.e. $S_T > S_0$. Then in 2002 Cheney would have ordinary income $S_T - K$ and a capital loss of $S_T - S_0$. US tax law treats the two very differently: individuals are permitted to offset at most $3000 of ordinary income by capital losses. Thus if Cheney had no capital gains on other investments to offset his capital losses on the forward contract then he would be taxed on essentially the full gain $S_T - K$ (though he could use the capital loss to offset capital gains in a future year).

Fixing (a) would be easy, if it were not for the constructive-sale rule: the problem is that a call is not equivalent to a foward. Rather, by put-call parity, it is equivalent to a forward plus a put. Of course Cheney could have gotten around the constructive-sale rule by taking a position whose value still had some dependence on S_T (e.g. by using a put with a strike price different from K) – but then he wouldn't have fixed his conflict of interest problem! Fixing (b) seems nearly impossible, since it depends on details of Cheney's financial position that have nothing to do with this transaction.

How did Cheney resolve this issue in fall 2000? I don't remember.

Derivative Securities – Section 2 – Fall 2004

Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences.

Binomial, trinomial, and more general one-period models. This section explores the implications of arbitrage for the pricing of contingent claims in a one-period setting. The first part, on the one-period binomial and trinomial markets, will be covered in class; the second part, on more general one-period markets, is provided for your information and enrichment only – it will not be covered in our lectures, homework, or exams.

The point of this discussion is to capture, in the most elementary possible setting, these important concepts: (a) market completeness (or lack thereof); (b) why the price of an option is the discounted risk-neutral expectation of its payoff; and (c) the link between risk-neutral pricing and the duality theory of linear programming.

The analysis of the binomial market is standard textbook material; I like Baxter-Rennie but you'll also find it in Jarrow-Turnbull and Hull. The analysis of trinomial and more general models and the link to linear programming duality is standard but not commonly considered textbook material. My treatment is more or less like the one at the beginning of Darrell Duffie's book *Dynamic Asset Pricing Theory* (which however is not easy reading). The one-period problem is discussed at length in John Cochrane's book *Asset Pricing*, with a viewpoint different from (complementary to) the one given here – rather than bounds, he emphasizes portfolio selection and the maximization of expected utility.

****

The binomial model. We consider an economy which has

- just two securities: a stock (paying no dividend, initial unit price per share s_1 dollars) and a bond (interest rate r, one bond pays one dollar at maturity).
- just one maturity date T
- just two possible states for the stock price at time T: s_2 and s_3, with $s_2 < s_3$

(see the figure).

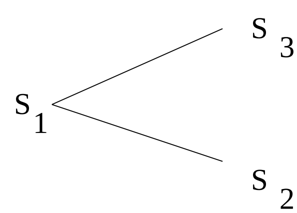


Figure 1: Prices in the one-period binomial market model.

We could suppose we know the probability p that the stock will be worth s_3 at time T. This would allow us to calculate the expected value of any contingent claim. However we will make no use of such knowledge. Pricing by arbitrage considerations makes no use of information about probabilities – it uses just the list of possible events.

The reasonable values of s_1, s_2, s_3 are not arbitrary: the economy should permit no arbitrage. This requires that

$$s_2 < s_1 e^{rT} < s_3.$$

It's easy to see that if this condition is violated then an arbitrage is possible. The converse is extremely plausible; a simple proof will be easy to give a little later.

In this simple setting a contingent claim can be specified by giving its payoff when $S_T = s_2$ and when $S_T = s_3$. For example, a long call with strike price K has payoff $f_2 = (s_2 - K)_+$ in the first case and $f_3 = (s_3 - K)_+$ in the second case. The most general contingent claim is specified by a vector $f = (f_2, f_3)$ giving its payoffs in the two cases.

Claim 1: In this model every contingent claim has a replicating portfolio. Thus arbitrage considerations determine the value of every contingent claim. (A market with this property is said to be "complete".)

In fact, consider the portfolio consisting of ϕ shares of stock and ψ bonds. Its initial value is

$$\phi s_1 + \psi e^{-rT}.$$

Its value at maturity replicates the contingent claim $f = (f_2, f_3)$ if

$$\begin{aligned} \phi s_2 + \psi &= f_2 \\ \phi s_3 + \psi &= f_3. \end{aligned}$$

This is a system of two linear equations for the two unknowns ϕ, ψ. The solution is

$$\phi = \frac{f_3 - f_2}{s_3 - s_2}, \quad \psi = \frac{s_3 f_2 - s_2 f_3}{s_3 - s_2}.$$

The initial value of the contingent claim f is necessarily the initial value of the replicating portfolio:

$$V(f) = \phi s_1 + \psi e^{-rT} = \frac{f_3 - f_2}{s_3 - s_2} s_1 + \frac{s_3 f_2 - s_2 f_3}{s_3 - s_2} e^{-rT}.$$

Claim 2: The value can conveniently be expressed as

$$V(f) = e^{-rT}[(1-q)f_2 + q f_3] \quad \text{where} \quad q = \frac{s_1 e^{rT} - s_2}{s_3 - s_2}.$$

Moreover, the condition that the market admit no arbitrage is $0 < q < 1$, which is equivalent to $s_2 < s_1 e^{rT} < s_3$.

The formula for $V(f)$ in terms of q is a matter of algebraic rearrangement. This simplification seems mysterious right now, but we'll see a natural reason for it later.

The condition that the market supports no arbitrage has two parts:

(i) a portfolio with nonnegative payoff must have a nonnegative value; and

(ii) a portfolio with nonnegative and sometimes positive payoff must have positive value.

In the binomial setting every payoff (f_2, f_3) is replicated by a portfolio, so we may replace "portfolio" by "contingent claim" in the preceding statement without changing its impact. Part (i) says $f_2, f_3 \geq 0 \Rightarrow (1-q)f_2 + qf_3 \geq 0$. This is true precisely if $0 \leq q \leq 1$. Part (ii) forces the sharper inequalities $q > 0$ and $q < 1$.

Notice the form of Claim 2. It says the present value of a contingent claim is obtained by taking its "expected final value" $(1-q)f_2 + qf_3$ then discounting (multiplying by e^{-rT}). However the "expected final value" has nothing to do with the probability of the stock going up or down. Instead it must be taken with respect to a special probability measure, assigning weight $1-q$ to state s_2 and q to state s_3, where q is determined by s_1, s_2, s_3 and r as above. This special probability measure is known as the "risk-neutral probability" associated with the market.

I like to view q as nothing more than a convenient way of representing $V(f)$. However the term "risk-neutral probability" can be understood as follows. Real-life investors prefer a guaranteed return at rate r to an uncertain one with expected rate r. We may nevertheless imagine the existence of "risk-neutral" investors, who are indifferent to risk. Such investors would consider these two alternatives to be equivalent. In a world where all investors were risk-neutral, investments would have to be valued so that their expected return agrees with the risk-free interest rate r. The formula for $V(f)$ has this form, if we assume in addition that the "expectations" of the risk-neutral investors are expressed by the probabilities $1-q$ and q:

$$\text{expected payoff} = (1-q)s_2 + qs_3 = e^{rT}V(f) = e^{rT} \cdot \text{initial investment.}$$

****

The trinomial model. This is the simplest example of an *incomplete* economy. It resembles the binomial model in having

- just two securities: a stock (paying no dividend, initial unit price per share s_1 dollars) and a bond (interest rate r, one bond pays one dollar at maturity).
- just one maturity date T.

However it differs by having three final states rather than two:

- The stock price at time T can take values s_2, s_3, or s_4, with $s_2 < s_3 < s_4$

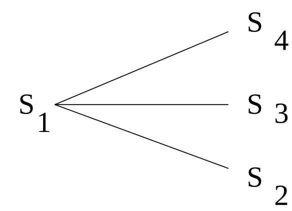


Figure 2: Prices in the one-period trinomial market model.

(see the figure). The reasonable values of $s_1, \ldots, s_4$ are not arbitrary: the economy should permit no arbitrage. This requires that

$$s_2 < s_1 e^{rT} \quad \text{and} \quad s_1 e^{rT} < s_4.$$

In other words, the stock must be able to do better than or worse than the risk-free return on an initial investment of s_1 dollars. It's easy to see that if this condition is violated then an arbitrage is possible.

In this case a contingent claim is specified by a 3-vector $f = (f_2, f_3, f_4)$; here f_j is the payoff at maturity if the stock price is s_j. Question: which contingent claims are replicatable? Answer: those for which the system

$$\begin{aligned} \phi s_2 + \psi &= f_2 \\ \phi s_3 + \psi &= f_3 \\ \phi s_4 + \psi &= f_4 \end{aligned}$$

has a solution. This specifies a two-dimensional space of f's. So the market is not complete, and "most" contingent claims are not replicatable.

If a contingent claim f is not replicatable then arbitrage does not specify its price $V(f)$. However arbitrage considerations still *restrict* its price:

$V(f) \leq$ the value of any portfolio whose payoff dominates f;

$V(f) \geq$ the value of any portfolio whose payoff is dominated by f.

In other words,

$$\left.\begin{aligned} \phi s_2 + \psi &\geq f_2 \\ \phi s_3 + \psi &\geq f_3 \\ \phi s_4 + \psi &\geq f_4 \end{aligned}\right. \Longrightarrow V(f) \leq \phi s_1 + e^{-rT}\psi$$

$$\left.\begin{aligned} \phi s_2 + \psi &\leq f_2 \\ \phi s_3 + \psi &\leq f_3 \\ \phi s_4 + \psi &\leq f_4 \end{aligned}\right. \Longrightarrow V(f) \geq \phi s_1 + e^{-rT}\psi.$$

We obtain the strongest possible consequences for $V(f)$ by solving a pair of linear programming problems:

$$\max_{\substack{\phi s_j + \psi \leq f_j \\ j=2,3,4}} \phi s_1 + e^{-rT}\psi \leq V(f) \leq \min_{\substack{\phi s_j + \psi \geq f_j \\ j=2,3,4}} \phi s_1 + e^{-rT}\psi.$$

These bounds capture *all* the information available from arbitrage concerning the price of the contingent claim f. (The actual price observed in the market must be determined by additional considerations besides arbitrage. The standard "equilibrium" approach to understanding prices uses utility maximization – to be discussed in the course Capital Markets and Portfolio Theory.)

Linear programming duality. In the binomial model $V(f)$ had a convenient expression in terms of a special "risk-neutral probability." To derive the analogous result here, we use the duality theory of linear programming. The following discussion should be accessible even to those who know nothing about linear programming – but it should make you want to learn something about this important topic. My favorite text is V. Chvatal, *Linear Programming*. A more recent text, more up-to-date on current developments such as interior point methods, is R. Vanderbei, *Linear programming: foundations and extensions*. Also of note is Peter Lax's book *Linear Algebra*, whose chapter on linear programming takes exactly the min-max viewpoint I use below.

Let us concentrate on the upper bound for $V(f)$. It is

$$\begin{aligned}
\min_{\phi s_j+\psi\geq f_j} \phi s_1 + e^{-rT}\psi &= \min_{\phi,\psi}\max_{\pi_j\geq 0}\ \phi s_1 + e^{-rT}\psi + \sum_{j=2}^{4}\pi_j(f_j-\phi s_j-\psi) \\
&= \max_{\pi_j\geq 0}\min_{\phi,\psi}\ \phi s_1 + e^{-rT}\psi + \sum \pi_j(f_j-\phi s_j-\psi) \\
&= \max_{\pi_j\geq 0}\min_{\phi,\psi}\ \phi(s_1-\sum\pi_j s_j) + \psi(e^{-rT}-\sum\pi_j) + \sum\pi_j f_j \\
&= \max_{\substack{\sum\pi_j s_j=s_1\\ \sum\pi_j=e^{-rT}\\ \pi_j\geq 0}} \sum \pi_j f_j.
\end{aligned}$$

The first line holds because

$$\max_{\pi_j\geq 0}\pi_j(f_j-\phi s_j-\psi) = \begin{cases} 0 & \text{if } \phi s_j+\psi\geq f_j \\ +\infty & \text{otherwise.}\end{cases}$$

The second line holds by the duality theorem of linear programming, which says in this setting that $\min\max = \max\min$. The third line is obtained by rearrangement, and the fourth line by an argument similar to the first.

The preceding argument is correct, but if you have no prior exposure to duality and/or convex analysis, the assertion that "min max = max min" may seem rather mysterious. To demystify it, let's explain by an entirely elementary argument why

$$\min_{\phi s_j+\psi\geq f_j}\phi s_1 + e^{-rT}\psi \geq \max_{\substack{\sum\pi_j s_j=s_1\\ \sum\pi_j=e^{-rT}\\ \pi_j\geq 0}} \sum\pi_j f_j.$$

(The opposite inequality is more subtle; the main point of linear programming duality theory is to prove it.) Indeed, consider any ϕ and ψ such that $\phi s_j + \psi \geq f_j$ for each $j = 2, 3, 4$; and consider any $\{\pi_j\}_{j=2}^4$ such that $\pi_j \geq 0$, $\sum_{j=2}^4 \pi_j s_j = s_1$, and $\sum_{j=2}^4 \pi_j = e^{-rT}$. Multiply each inequality $\phi s_j + \psi \geq f_j$ by π_j, then add and use the hypotheses on π_j to see that $\phi s_1 + e^{-rT}\psi \geq \sum \pi_j f_j$. Minimizing the left hand side (over all admissible ϕ, ψ) and maximizing the right hand side (over all admissible π_j) gives the desired inequality.

Making the minor change of variables $\hat{\pi}_j = e^{rT}\pi_j$, our duality argument has shown that

$$V(f) \leq \max\{e^{-rT}[\hat{\pi}_2 f_2 + \hat{\pi}_3 f_3 + \hat{\pi}_4 f_4] \ : \ \hat{\pi}_2 s_2 + \hat{\pi}_3 s_3 + \hat{\pi}_4 s_4 = e^{rT} s_1 \quad \hat{\pi}_2 + \hat{\pi}_3 + \hat{\pi}_4 = 1, \quad \hat{\pi}_j \geq 0\}.$$

The lower bound is handled similarly. The only difference is that we are maximizing in ϕ, ψ and minimizing in π_j. An argument parallel to the one given above shows

$$V(f) \geq \min\{e^{-rT}[\hat{\pi}_2 f_2 + \hat{\pi}_3 f_3 + \hat{\pi}_4 f_4] \ : \ \hat{\pi}_2 s_2 + \hat{\pi}_3 s_3 + \hat{\pi}_4 s_4 = e^{rT} s_1 \quad \hat{\pi}_2 + \hat{\pi}_3 + \hat{\pi}_4 = 1, \quad \hat{\pi}_j \geq 0\}.$$

Thus the upper and lower bounds on $V(f)$ are obtained by maximizing and minimizing the "discounted expected return" $e^{-rT}[\hat{\pi}_2 f_2 + \hat{\pi}_3 f_3 + \hat{\pi}_4 f_4]$ over an appropriate class of "risk-neutral probabilities" $(\hat{\pi}_2, \hat{\pi}_3, \hat{\pi}_4)$. The incompleteness of the market is reflected in the fact that there is more than one risk-neutral probability: in the present trinomial setting the 3-vector $(\hat{\pi}_2, \hat{\pi}_3, \hat{\pi}_4)$ is constrained by two inequalities, so the class of risk-neutral probabilities is one-dimensional (a line segment).

We noted before the condition $s_2 < s_1 e^{rT} < s_4$, which is required for the economy to be "reasonable" – i.e. not to admit an arbitrage. This is precisely the condition that there be at least one "risk-neutral probability", i.e. a vector $(\hat{\pi}_2, \hat{\pi}_3, \hat{\pi}_4)$ such that $\sum \hat{\pi}_j = 1$, $\sum \hat{\pi}_j s_j = e^{rT} s_1$, and $\hat{\pi}_j > 0$ for each j.

****

The general one-period market model. The binomial and trinomial models are special cases of a more general theory, which we now present. Main purposes of this discussion:

- deeper understanding of risk-neutral probabilities; and
- more careful treatment of the "principle of no arbitrage."

In considering one-period models with few assets and many states, we are close to the question of portfolio analysis: which of the many possible portfolios should an investor hold? We'll resist addressing this, however, concentrating instead on the narrower goal of understanding risk-neutral pricing and the role of arbitrage.

The general one-period market has

- N securities, $i = 1, \ldots, N$

- M final states, $\alpha = 1, \ldots, M$
- fixed initial values: one unit of security i is worth p_i dollars
- state-dependent final values: if the final state is α then one unit of security i is worth $D_{i\,\alpha}$.

An investor can hold any *portfolio*: θ_i units of security i. It has initial value $\langle p, \theta \rangle = \sum_i p_i \theta_i$. If the final state is α then its final value is $\langle \theta, D_{\cdot\,\alpha} \rangle = \sum_i \theta_i D_{i\,\alpha}$.

Examples:

Binomial model: $\quad p = (e^{-rT}, s_1), \quad D = \begin{pmatrix} 1 & 1 \\ s_2 & s_3 \end{pmatrix}$

Trinomial model: $\quad p = (e^{-rT}, s_1), \quad D = \begin{pmatrix} 1 & 1 & 1 \\ s_2 & s_3 & s_4 \end{pmatrix}$

In general, if security 1 is a riskless bond then

$$p = (e^{-rT}, p_2, \ldots, p_N), \quad D = \begin{pmatrix} 1 & \cdots & 1 \\ D_{21} & \cdots & D_{2M} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ D_{N1} & \cdots & D_{NM} \end{pmatrix}$$

Here's a careful statement of the **Principle of no arbitrage**:

(a) $\sum_i \theta_i D_{i\,\alpha} \geq 0$ for all $\alpha \implies \sum_i \theta_i p_i \geq 0$

(b) if we have both $\sum_i \theta_i D_{i\,\alpha} \geq 0$ for all α and $\sum_i \theta_i p_i = 0$ then we must have $\sum_i \theta_i D_{i\,\alpha} = 0$ for every α.

These capture with precision the informal statements that (a) a portfolio with nonnegative payoff has nonnegative value; and (b) a portfolio with nonnegative and sometimes positive payoff has strictly positive value.

The key result relating risk-neutral probabilities to lack of arbitrage is this:

Arbitrage Theorem: The economy satisfies (a) iff there exist $\pi_\alpha \geq 0$ such that

$$\sum_\alpha D_{i\,\alpha} \pi_\alpha = p_i, \quad i = 1, \ldots, N.$$

It satisfies both (a) and (b) if in addition the π_α can be chosen to be all strictly positive.

The theorem is trivial in one direction: assuming the existence of π_α we can easily prove the absence of arbitrage. In fact, for any portfolio θ_i we have

$$\begin{aligned} \sum_i \theta_i D_{i\,\alpha} \geq 0 \text{ for all } \alpha \quad &\implies \quad \sum_{i,\alpha} \theta_i D_{i\,\alpha} \pi_\alpha \geq 0 \\ &\implies \quad \sum_i \theta_i p_i \geq 0 \end{aligned}$$

since $\pi_\alpha \geq 0$. If $\pi_\alpha > 0$ for each α then the conclusion can hold with $=$ only if each hypothesis holds with $=$ rather than $\geq$. Thus existence of $\pi_\alpha \geq 0$ implies part (a) of the no-arbitrage principle; and if each π_α is strictly positive then we also get part (b) of the principle.

The other half of the theorem (no arbitrage implies existence of π_α) is decidedly nontrivial. We shall sketch the proof presently; but first let us make contact with what we did earlier in the trinomial setting. Suppose security 1 is a risk-less bond. Then $p_1 = e^{-rT}$ and the first row of $D_{i\,\alpha}$ is filled with 1's. The statement of the theorem becomes: the market permits no arbitrage iff there exist positive π_α such that

$$\begin{aligned} \pi_1 + \cdots + \pi_M &= e^{-rT} \\ \sum_\alpha \pi_\alpha D_{i\,\alpha} &= p_i, \quad i = 2, \ldots, N. \end{aligned}$$

Writing $\hat{\pi}_\alpha = e^{rT}\pi_\alpha$ we see that the this is equivalent to the existence of positive $\hat{\pi}_\alpha$ such that

$$\begin{aligned} \hat{\pi}_1 + \cdots + \hat{\pi}_M &= 1 \\ \sum_\alpha \hat{\pi}_\alpha D_{i\,\alpha} &= e^{rT} p_i, \quad i = 2, \ldots, N. \end{aligned}$$

These $\hat{\pi}_\alpha$ are the *risk-neutral probabilities*.

For the trinomial market, we showed how arbitrage considerations restrict the initial value of any contingent claim f. The same max/min argument works in general, for any market in which Security 1 is a riskless bond. The conclusion is

$$\min_{\substack{\text{risk-neutral} \\ \text{probs } \hat{\pi}}} e^{-rT} \sum_\alpha \hat{\pi}_\alpha f_\alpha \leq V(f) \leq \max_{\substack{\text{risk-neutral} \\ \text{probs } \hat{\pi}}} e^{-rT} \sum_\alpha \hat{\pi}_\alpha f_\alpha.$$

We immediately see that

$$\begin{aligned} \text{market completeness} \quad &\Leftrightarrow \quad \text{arbitrage determines the value of every contingent claim} \\ &\Leftrightarrow \quad \text{there is a unique risk-neutral probability.} \end{aligned}$$

****

Sketch of a proof of the Arbitrage Theorem. The rest of this section sketches a proof of (the nontrivial part of) the Arbitrage Theorem, based on the following

Fundamental Lemma of Linear Programming: If a collection of linear inequalities implies another linear inequality then it does so "trivially," i.e. the conclusion is a (nonnegative) linear combination of the hypotheses.

The name "fundamental lemma of linear programming" is my own; the proper name of this result is Farkas' Lemma. See e.g. V. Chvatal, *Linear Programming*, pg. 248, for this and related results.

Our first task is to show that part (a) of the no-arbitrage principle implies the existence of $\pi_\alpha \geq 0$. Now, part (a) says that the collection of linear inequalities $\sum_i \theta_i D_{i\,\alpha} \geq 0$ for $\alpha = 1, \ldots, M$ implies another linear inequality $\sum_i \theta_i p_i \geq 0$. By the Fundamental Lemma of Linear Programming this occurs only if there is a "trivial" proof, i.e. if there exists $\pi_\alpha \geq 0$ such that $\sum_i \theta_i p_i = \sum_{i,\alpha} \theta_i D_{i\,\alpha}\pi_\alpha$ for all θ_i. But that means $\sum D_{i\,\alpha}\pi_\alpha = p_i$.

Our second task is to show that if the economy satisfies both parts (a) and (b) of the no-arbitrage principle then we can take $\pi_\alpha > 0$ for all α. If the π_α already identified are all positive then we're done. If not, then renumbering states if necessary we may suppose $\pi_1, \ldots, \pi_{M'} > 0$ and $\pi_{M'+1} = \ldots = \pi_M = 0$.

Let's concentrate for a moment on index $M'+1$. If $D_{\cdot\,M'+1} = (D_{1\,M'+1}, \ldots, D_{N\,M'+1})$ is a linear combination of $D_{\cdot\,1}, \ldots, D_{\cdot\,M'}$ then we can easily modify π_α to make $\pi_{M'+1} > 0$. In fact, suppose $D_{\cdot\,M'+1} = b_1 D_{\cdot\,1} + \ldots + b_{M'} D_{\cdot\,M'}$. Then

$$\begin{aligned} p_i &= \sum_{\alpha=1}^{M'} D_{i\,\alpha}\pi_\alpha \\ &= \epsilon D_{i\,M'+1} + \sum_{\alpha=1}^{M'} D_{i\,\alpha}(\pi_\alpha - \epsilon b_\alpha), \end{aligned}$$

so replacing $\pi = (\pi_1, \ldots, \pi_{M'}, 0, \ldots, 0)$ with $(\pi_1 - \epsilon b_1, \ldots, \pi_{M'} - \epsilon b_{M'}, \epsilon, 0, \ldots, 0)$ does the trick when ϵ is sufficiently small.

Essentially the same argument shows that if *any* positive combination of $D_{\cdot\,M'+1}, \ldots, D_{\cdot\,M}$ lies in the span of $D_{\cdot\,1}, \ldots, D_{\cdot\,M'}$ then we can modify π_α to make additional components positive.

Applying the preceding argument finitely many times, we either arrive at a new π with strictly positive components, or we find ourselves in a situation (with a new value of M') where no positive combination of $D_{\cdot\,M'+1}, \ldots, D_{\cdot\,M}$ lies in the span of $D_{\cdot\,1}, \ldots, D_{\cdot\,M'}$. We claim the second alternative cannot happen when the economy has property (b).

This is another application of the Fundamental Lemma of Linear Programming. Our "second alternative" is that

$$\sum_{\alpha=M'+1}^{M} a_\alpha D_{\cdot\,\alpha} = \sum_{\alpha=1}^{M'} b_\alpha D_{\cdot\,\alpha}, \quad a_\alpha \geq 0 \implies a_\alpha = 0, \alpha = M'+1, \ldots, M.$$

The "trivial consequences" of the hypotheses are obtained by taking linear combinations. This amounts to taking the inner product with a vector $\theta \in R^N$. Thus the trivial consequences of the hypotheses are

$$\sum_{\alpha=M'+1}^{M} a_\alpha \langle D_{\cdot\,\alpha}, \theta\rangle = \sum_{\alpha=1}^{M'} b_\alpha \langle D_{\cdot\,\alpha}, \theta\rangle.$$

For this (coupled with $a_\alpha \geq 0$) to give a trivial proof that $a_\alpha = 0$ we must have

$$\begin{aligned}
\langle D_{\cdot\,\alpha}, \theta\rangle = \sum_i \theta_i D_{i\,\alpha} &= 0 \quad \alpha = 1, \ldots, M' \\
\langle D_{\cdot\,\alpha}, \theta\rangle = \sum_i \theta_i D_{i\,\alpha} &> 0 \quad \alpha = M'+1, \ldots, M.
\end{aligned}$$

But then θ represents a portfolio with no downside, some upside, and value 0 since $\sum_i \theta_i p_i = \sum_i \sum_{\alpha=1}^{M'} \theta_i D_{i\,\alpha} \pi_\alpha = 0$. This contradicts our assumption that the economy admits no arbitrage.

Derivative Securities – Section 3 – Fall 2004
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences.

Multiperiod Binomial Trees. We turn to the valuation of derivative securities in a time-dependent setting. We focus for now on multi-period binomial models, i.e. binomial trees. This setting is simple enough to let us do everything explicitly, yet rich enough to approximate many realistic problems.

The material covered in this section is very standard (and very important). The treatment here is essentially that of Baxter and Rennie (Chapter 2). The same material is also in Jarrow and Turnbull (Chapter 5) and Hull (Chapter 10 in the 5th edition). Jarrow and Turnbull has the advantage of including many examples. In the next section of notes we'll discuss how the parameters should be chosen to mimic the conventional (Black-Scholes) hypothesis of lognormal stock prices, and we'll pass to the continuous-time limit.

Binomial trees are widely used in practice, in part because they are easy to implement numerically. (Also because the scheme can easily be adjusted to price American options.) For a nice discussion of alternative numerical implementations, see the article "Nine ways to implement the binomial method for option valuation in Matlab," by D.J. Higham, SIAM Review 44, no. 4, 661-677.

****

The multi-period binomial model generalizes the single-period binomial model we considered in Section 2. It has

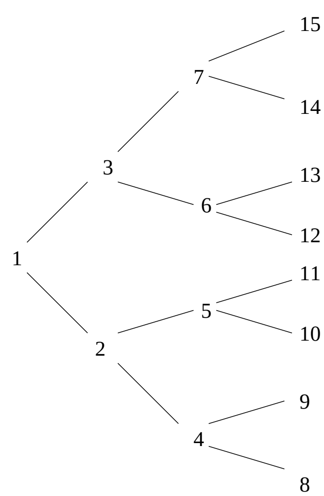


Figure 1: States of a non-recombining binomial tree.

- just two securities: a risky asset (a "stock," paying no dividend) and a riskless asset ("bond");
- a series of times $0, \delta t, 2\delta t, \ldots, N\delta t = T$ at which trades can occur;

- interest rate r_i during time interval i for the bond;
- a binomial tree of possible states for the stock prices.

The last statement means that for each stock price at time $j\delta t$, there are two possible values it can take at time $(j+1)\delta t$ (see Figure 1).

The interest rate environment is described by specifying the interest rates r_i. We restrict our attention for now to the case of a constant interest rate: $r_i = r$ for all i.

The stock price dynamics is described by assigning a price s_j to each state in the tree. Strictly speaking we should also assign (subjective) probabilities p_j to the branches (the two branches emerging from a given node should have probabilities summing to 1): see Figure 2.

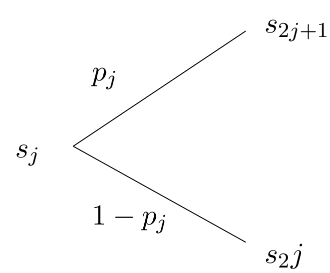


Figure 2: One branch of a binomial tree.

Actually, we will make no use of the subjective probabilities p_j; our arguments are based on arbitrage, so they depend only on the list of possible states not on their probabilities. However our pricing formula will make use of *risk-neutral probabilities* q_j. These "look like" subjective probabilities, except that they are determined by the stock prices and the interest rate.

Our stock prices must be "reasonable" in the sense that the market support no arbitrage. Motivated by the one-period model, we (correctly) expect this condition to take the form:

- starting from any node, the stock price may do better than or worse than the risk-free rate during the next period.

In other words, $s_{2j} < e^{r\delta t} s_j < s_{2j+1}$ for each j.

The tree in Figure 1 is the most general possible. At the nth time step it has 2^n possible states. That's a lot of states, especially when n is large. It's often convenient to let selected states have the same prices in such a way that the list of distinct prices forms a *recombinant tree*. Figure 3 gives an example of a 4-stage recombinant tree, with stock prices marked for each state: (A recombinant tree has just $n+1$ possible states at time step n.)

A special class of recombinant trees is obtained by assuming the stock price goes up or down by fixed multipliers u or d at each stage: see Figure 4. This last class may seem

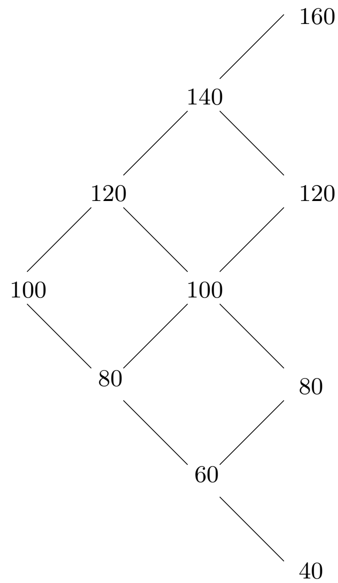


Figure 3: A simple recombinant binomial tree.

terribly special relative to the general binomial tree. But we shall see it is general enough for many practical purposes – just as a random walk (consisting of many steps, each of fixed magnitude but different in direction) can approximate Brownian motion. And it has the advantage of being easy to specify – one has only to give the values of u and d.

It may seem odd that we consider a market with just one stock, when real markets have many stocks. But our goal is to price contingent claims based on considerations of arbitrage. If we succeed using just these two instruments (the stock and the riskless bond) then our conclusions necessarily apply to any larger market containing both instruments.

****

Our goal is to determine the value (at time 0) of a contingent claim. We will consider American options later; for the moment we consider only European ones, i.e. early redemption is prohibited. The most basic examples are European calls and puts (payoffs: $(S_T - K)_+$ and $(K - S_T)_+$ respectively). However our method is much more general. What really matters is that *the payoff of the claim depends entirely on the state of stock process at time T.*

Let's review what we found in the one-period binomial model. Our multiperiod model consists of many one-period models, so it is convenient to introduce a flexible labeling scheme. Writing "now" for what used to be the initial state, and "up, down" for what used to be the two final states, our risk-neutral valuation formula was

$$f_{\text{now}} = e^{-r\delta t}[qf_{\text{up}} + (1-q)f_{\text{down}}]$$

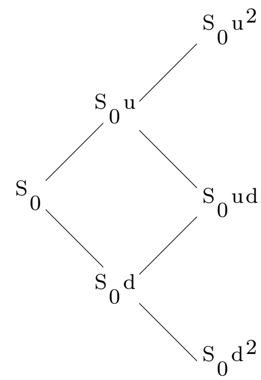


Figure 4: A multiplicative recombinant binomial tree.

where

$$q = \frac{e^{r\delta t} s_{\text{now}} - s_{\text{down}}}{s_{\text{up}} - s_{\text{down}}}.$$

Here we're writing f_{now} for what we used to call $V(f)$, the (present) value of the contingent claim worth f_{up} or f_{down} at the next time step if the stock price goes up or down respectively. This formula was obtained by replicating the payoff with a combination of stock and bond; the replicating portfolio used

$$\phi = \frac{f_{\text{up}} - f_{\text{down}}}{s_{\text{up}} - s_{\text{down}}}$$

units of stock, and a bond worth $f_{\text{now}} - \phi s_{\text{now}}$.

The valuation of a contingent claim in the multiperiod setting is an easy consequence of this formula. We need only "work backward through the tree," applying the formula again and again.

Consider, for example, the four-period recombinant tree shown in Figure 3. (This example, taken straight from Baxter and Rennie, has the nice feature of very simple arithmetic.) Suppose the interest rate is $r = 0$, for simplicity. Then $q = 1/2$ at each node (we chose the prices to keep this calculation simple). Let's find the value of a European call with strike price 100 and maturity $T = 3\delta t$. Working backward through the tree:

- The values at maturity are $(S_T - 100)_+ = 60$, 20, 0, 0 respectively.
- The values one time step earlier are 40, 10, and 0 respectively, each value being obtained by an application of the one-period formula.
- The values one time step earlier are 25 and 5.
- The value at the initial time is 15.

Easy. But is it right? Yes, because these values can be replicated. However the replication strategies are more complicated than in the one-period case: the replicating portfolio must be adjusted at each trading time, taking into account the new stock price.

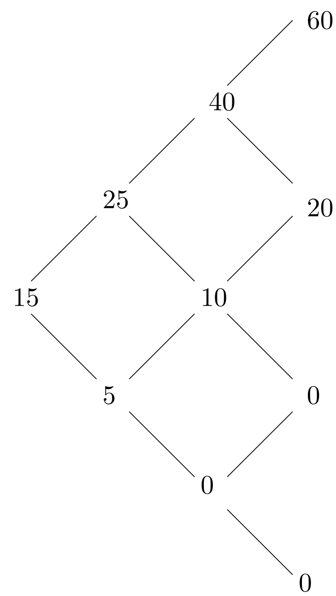


Figure 5: Value of the option as a function of stock price state.

Let's show this in the example. Using our one-period rule, the replicating portfolio starts with $\phi = (25 - 5)/(120 - 80) = .5$ units of stock, worth $.5 \times 100 = 50$ dollars, and a bond worth $15 - 50 = -35$. (Viewed differently: the investment bank that sells the option collects 15 dollars; it should borrow another 35 dollars, and use these 15+35=50 dollars to buy 1/2 unit of stock.) The claim is that by trading intelligently at each time-step we can adjust this portfolio so it replicates the payoff of the option no matter what the stock price does. Here is an example of a possible history, and how we would handle it:

Stock goes up to 120 The new ϕ is $(40 - 10)/(140 - 100) = .75$, so we need another .25 units of stock. We must buy this at the present price, 120 dollars per unit, and we do it by borrowing 30 dollars. Thus our debt becomes 65 dollars.

Stock goes up again to 140 The new ϕ is $(60 - 20)/(160 - 120) = 1$, so we buy another .25 unit at 140 dollars per unit. This costs another 35 dollars, bringing our debt to 100 dollars.

Stock goes down to 120 At maturity we hold one share of stock and a debt of 100. So our portfolio is worth 120-100 = 20, replicating the option. (Put differently: if the investment bank that sold the option followed our instructions, it could deliver the unit of stock, collect the 100, pay off its loan, and have neither a loss nor a gain.)

That wasn't a miracle. We'll explain it more formally pretty soon. But here's a second example – a different possible history – to convince you:

Stock goes down to 80 The revised ϕ is $(10 - 0)/(100 - 60) = .25$. So we should sell 1/4 unit stock, receiving 80/4=20. Our debt is reduced to 15.

Stock goes up to 100 The new ϕ is $(20-0)/(120-80) = .5$. So we must buy 1/4 unit stock, spending 100/4=25. Our debt goes up to 40.

Stock goes down again to 80 We hold 40 dollars worth of stock and we owe 40. Our position is worth 40-40=0, replicating the option, which is worthless since it's out of the money. (Different viewpoint: the investment bank that sold the option can liquidate its position, selling the stock at market and using the proceeds to pay off the loan. This results in neither a loss nor a gain.)

Notice that the portfolio changes from one time to the next but the changes are *self-financing* – i.e. the total value of the portfolio before and after each trade are the same. (The investment bank neither receives or spends money except at the initial time, when it sells the option.)

Our example shows the importance of tracking ϕ, the number of units of stock to be held as you leave a given node. It characterizes the replicating portfolio (the "hedge"). Its value is known as the *Delta* of the claim. Thus:

$$\Delta_{\text{now}} = \text{our } \phi = \frac{f_{\text{up}} - f_{\text{down}}}{s_{\text{up}} - s_{\text{down}}}.$$

To understand the meaning Δ, observe that as you leave node j,

$$\text{value of claim at a node j} = \Delta_j s_j + b_j$$

by definition of the replicating portfolio (here b_j is the value of the bond holding at that moment). If the value of the stock changes by an amount ds while the bond holding stays fixed, then the value of the replicating portfolio changes by Δds. Thus Δ is a sort of derivative:

$$\Delta = \text{rate of change of replicating portfolio value, with respect to change of stock price.}$$

****

Our valuation algorithm is easy to implement. But in the one-period setting we had more than an algorithm: we also had a *formula* for the value of the option, as the discounted expected value using a risk-neutral probability. A similar formula exists in the multiperiod setting. To see this, it is most convenient to work with a general binomial tree. Consider, for example, a tree with two time steps. The risk-neutral probabilities $q_j, 1-q_j$ are determined by the embedded one-period models. (Remember, the risk-neutral probabilities are characteristic of the market; they don't depend on the contingent claim under consideration.) In this case:

$$q_1 = \frac{e^{r\delta t}s_1 - s_2}{s_3 - s_2}, \quad q_2 = \frac{e^{r\delta t}s_2 - s_4}{s_5 - s_4}, \quad q_3 = \frac{e^{r\delta t}s_3 - s_6}{s_7 - s_6}.$$

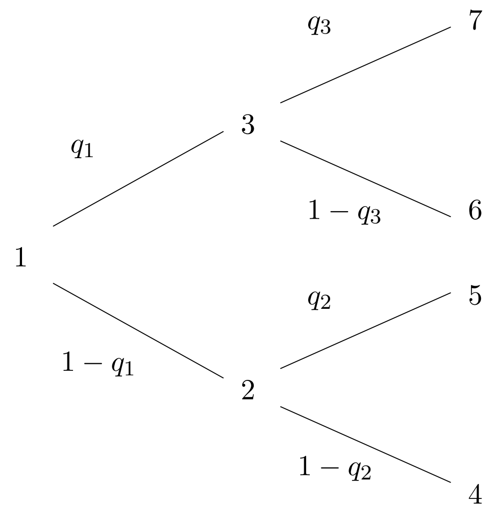


Figure 6: General binomial tree with two time steps.

As we work backward through the tree, we get a formula for the value of the contingent claim at each node, as a discounted weighted average of its values at maturity. In fact, writing $f(j)$ for the value of the contingent claim f at node j,

$$f(3) = e^{-r\delta t}[q_3 f(7) + (1-q_3)f(6)]$$

and

$$f(2) = e^{-r\delta t}[q_2 f(5) + (1-q_2)f(4)]$$

so

$$\begin{aligned} f(1) &= e^{-r\delta t}[q_1 f(3) + (1-q_1)f(2)] \\ &= e^{-2r\delta t}[q_1 q_3 f(7) + q_1(1-q_3)f(6) + (1-q_1)q_2 f(5) + (1-q_1)(1-q_2)f(4)]. \end{aligned}$$

It should be clear now what happens, for a binomial tree with any number of time periods:

$$\text{initial value of the claim} = e^{-rN\delta t} \sum_{\text{final states}} [\text{probability of the associated path}] \times [\text{payoff of state}],$$

where the probability of any path is the product of the probabilities of the individual risk-neutral probabilities along it. (Thus: the different risk-neutral probabilities must be treated as if they described independent random variables.)

A similar rule applies to recombinant trees, since they are just special binomial trees in disguise. We must simply be careful to count the paths with proper multiplicities. For example, consider a two-period model with a recombinant tree and $s_{\text{up}} = us_{\text{now}}$, $s_{\text{down}} = ds_{\text{now}}$. In this case the formula becomes

$$f(1) = e^{-2r\delta t}[(1-q)^2 f(4) + 2q(1-q)f(5) + q^2 f(6)]$$

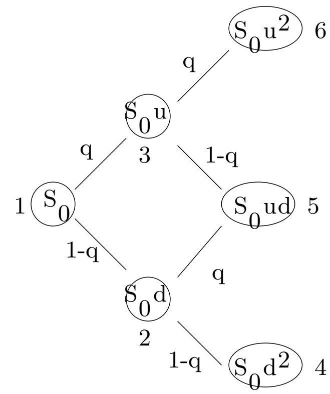


Figure 7: Price states in a multiplicative recombinant tree.

with $q = (e^{r\delta t} - d)/(u - d)$, since there are two distinct paths leading to node 5.

The preceding calculation extends easily to recombinant trees with any number of time steps. The result is one of the most famous and important results of the theory: an explicit formula for the value of a European option. This is in a sense the binomial tree version of the Black-Scholes formula. (To really use it, of course, we'll need to know how to specify the parameters u and d; we'll come to that soon.) Consider an N-step recombinant stock price model with with $s_{\text{up}} = us_{\text{now}}$, $s_{\text{down}} = ds_{\text{now}}$, and s_0=initial spot price. Then the present value of an option with payoff $f(S_T)$ is

$$e^{-rN\delta t} \sum_{j=0}^{N} \left[\binom{N}{j} q^j (1-q)^{N-j} f(s_0 u^j d^{N-j}) \right].$$

with $q = (e^{r\delta t} - d)/(u - d)$. This holds because there are $\binom{N}{j}$ different ways of accumulating j ups and $N - j$ downs in N time-steps (just as there are $\binom{N}{j}$ different ways of getting heads exactly j times out of N coin flips.) Making this specific to European puts and calls: a call with strike price K has present value

$$e^{-rN\delta t} \sum_{j=0}^{N} \left[\binom{N}{j} q^j (1-q)^{N-j} (s_0 u^j d^{N-j} - K)_+ \right];$$

a put with strike price K has present value

$$e^{-rN\delta t} \sum_{j=0}^{N} \left[\binom{N}{j} q^j (1-q)^{N-j} (K - s_0 u^j d^{N-j})_+ \right].$$

Derivative Securities – Section 4 – Fall 2004
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences.

Lognormal price dynamics and passage to the continuum limit. After a brief recap of our pricing formula, this section introduces the lognormal model of stock price dynamics, and explains how it can be approximated using binomial trees. Then we use these binomial trees to price contingent claims. The Black-Scholes analysis is obtained in the limit $\delta t \to 0$. As usual, Baxter–Rennie captures the central ideas concisely yet completely (Section 2.4) while Jarrow–Turnbull has a more liesurely treatment (Chapter 4, supplemented by Section 5.5). Hull has a lot of information about the lognormal model scattered through Chapters 10 and 11.

****

Recap of the multiperiod option pricing formula. Recall what we achieved at the end of the last section: if the risk-free rate is constant and the risky asset price evolution is described by a multiplicative binomial tree with $s_{\text{up}} = us_{\text{now}}$ and $s_{\text{down}} = ds_{\text{now}}$ then the value at time 0 of a contingent claim with maturity $T = N\delta t$ and payoff $f(s_T)$ is

$$V(f) = e^{-rT} \cdot E_{\text{RN}}[f(s_T)]$$

where $E_{\text{RN}}[f(s_T)]$ is the expected final payoff, computed with respect to the risk-neutral probability:

$$E_{\text{RN}}[f(s_T)] = \sum_{j=0}^{N} \binom{N}{j} q^j (1-q)^{N-j} f(s_0 u^j d^{N-j}),$$

with $q = (e^{r\delta t} - d)/(u - d)$. Let's check this assertion for consistency and gain some intuition by making a few observations:

What if the contingent claim pays the stock price itself? This is the case $f(s_T) = s_T$. It is replicated by the portfolio consisting of one unit of stock (no bond, no trading). So the present value should be s_0, the price of the stock now. Let's verify that this is the same result we get by "working backward through the tree." It's enough to show that if $f(s) = s$ for every possible price s at a given time then the same relation holds at the time just before. To see this, let "now" refer to any possible stock price at the time just before. We are assuming $f(s_{\text{up}}) = s_{\text{up}}$ and $f(s_{\text{down}}) = s_{\text{down}}$ and we want to show $f(s_{\text{now}}) = s_{\text{now}}$. By definition,

$$f(s_{\text{now}}) = e^{-r\delta t}[qs_{\text{up}} + (1-q)s_{\text{down}}]$$

with $q = \frac{e^{r\delta t}s_{\text{now}} - s_{\text{down}}}{s_{\text{up}} - s_{\text{down}}}$. Simple algebra confirms the expected result $f(s_{\text{now}}) = s_{\text{now}}$.

There is of course an equivalent calculation involving risk-neutral expectation. The formula for q in a multiplicative tree gives

$$qu + (1-q)d = e^{r\delta t}$$

and taking the Nth power gives

$$\sum_{j=0}^{N} \binom{N}{j} q^j (1-q)^{N-j} u^j d^{N-j} = e^{rN\delta t} = e^{rT}.$$

Multiplying both sides by s_0 gives

$$e^{-rT} E_{\text{RN}}[s_T] = s_0$$

as desired.

What if the contingent claim is a forward contract with strike price K? Under our standing constant-interest-rate hypothesis we know the present value should be $s_0 - e^{-rT}K$ if the maturity is $T = N\delta t$. Let's verify that any binomial tree gives the same result. The payoff is $f(s_T) = s_T - K$. Our formula

$$e^{-rT} E_{\text{RN}}[f(s_T)]$$

is linear in the payoff. Also $E_{\text{RN}}[K] = K$, i.e. the total risk-neutral probability is 1; this can be seen from the fact that $(q + [1-q])^N = 1$. Thus our formula for the value of a forward is

$$e^{-rT} E_{\text{RN}}[s_T - K] = e^{-rT} E_{\text{RN}}[s_T] - e^{-rT} E_{\text{RN}}[K] = s_0 - e^{-rT} K$$

as expected.

What if the contingent claim is a European call with strike price $K \gg s_0$? We expect such a call to be worthless, or nearly so. This is captured by the model, since only a few exceptional paths (involving an exceptional number of "ups") will result in a postive payoff.

What if the contingent claim is a European call with strike price $K \ll s_0$? We expect such a call to be worth about the same as a forward with strike price K. This too is captured by the model, since only a few exceptional paths (involving an exceptional number of "downs") will result in a payoff different from that of the forward.

Analogous observations hold for European puts.

****

Lognormal stock price dynamics. Our simple model of a risk-free asset has a constant interest rate. A bond worth ψ_0 dollars at time 0 is worth $\psi(t) = \psi_0 e^{rt}$ dollars at time t. The quantity that's constant is not the growth rate $\frac{d\psi}{dt}$ but rather the interest rate $r = \frac{1}{\psi}\frac{d\psi}{dt} = \frac{d \log \psi}{dt}$.

Our stock is risky, i.e. its evolution is unknown and appears to be random. We can still describe its dynamics in terms of an equivalent interest rate for each time period. Breaking time up into intervals of length δt, the equivalent interest rate for $j\delta t < t < (j+1)\delta t$ is r_j if $s((j+1)\delta t) = e^{r_j \delta t} s(j\delta t)$, i.e.

$$r_j = \frac{\log s((j+1)\delta t) - \log s(j\delta t)}{\delta t}.$$

Standard terminology: r_j is the **return** of the stock over the relative time interval. Note that to calculate the stock price change over a longer interval you just add the exponents:

$$s(k\delta t) = e^{(r_j \delta t + r_{j+1}\delta t + \ldots + r_{k-1}\delta t)} s(j\delta t), \quad \text{for } j < k.$$

Beware of the following linguistic fine point: some people would say that r_j is the "rate of return" over the jth period, and the actual "return" over that period is $e^{r_j \delta t}$. But "rate of return" is a mouthful, so I prefer to use the word "return" for r_j itself.

Since the stock price is random so is each r_j. The lognormal model of stock price dynamics specifies their statistics:

- The random variables $r_j \delta t$ are independent, identically distributed, Gaussian random variables with mean $\mu\delta t$ and variance $\sigma^2 \delta t$, for some constants μ and σ.

The constant μ is called the **expected return** (though actually, the expected return over a time interval of length δt is $\mu\delta t$). The constant σ is called the "volatility of return," or more briefly just **volatility**. These constants are assumed to be the same regardless of the length of the interval δt. Thus we really mean the following slightly stronger statement:

- For any time interval (t_1, t_2), $\log s(t_2) - \log s(t_1)$ is a Gaussian random variable with mean $\mu(t_2 - t_1)$ and variance $\sigma^2(t_2 - t_1)$.

- The Gaussian random variables associated with disjoint time intervals are independent.

In particular (for those who know what this means) $\log s(t)$ executes a Brownian motion with drift. Strictly speaking σ has units of $1/\sqrt{\text{time}}$, however it is common to call σ the "volatility per year".

Why should we believe this hypothesis about stock prices? Perhaps it would be more credible to suppose that the daily (or hourly or minute-by-minute) return is determined by a random event (arrival of news, perhaps) which we can model by flipping a coin. The lognormal model is the limit of such dynamics, as the time-frequency of the coin-flips tends to zero. We'll discuss this in detail presently.

Different type of justification: compare the hypothesis to data. It does reasonably well, though not perfectly. See Hull for a discussion. Another recent survey: J. Case, "The modeling and analysis of financial time series", *Amer. Math. Monthly* 105 (May 1998) 401–411.

The lognormal hypothesis will lead us to a formula for the present value of a derivative security – but it's important to remember that the formula is no better than the stock price model it's based on. The formula doesn't agree perfectly with what one finds in the marketplace; the main reason is probably that the lognormal model isn't a perfect model of real stock prices. Much work has been done on improving it – for example by letting the volatility itself be random rather than constant in time.

****

Lognormal dynamics and the limit of multiperiod binomial trees. We claim that lognormal dynamics can be approximated by dividing time into many intervals, and flipping a coin to determine the return for each interval.

The coin can be fair or biased; to keep things as simple as possible let's concentrate on the fair case first. To simulate a lognormal process with expected return μ and volatility σ the return should be

$$\begin{array}{ll} \mu\delta t + \sigma\sqrt{\delta t} & \text{if heads (probability 1/2)} \\ \mu\delta t - \sigma\sqrt{\delta t} & \text{if tails (probability 1/2).} \end{array}$$

In other words, given δt we wish to consider the recombinant binomial tree with with

$$s_{\text{up}} = s_{\text{now}} e^{\mu\delta t + \sigma\sqrt{\delta t}}, \quad s_{\text{down}} = s_{\text{now}} e^{\mu\delta t - \sigma\sqrt{\delta t}}$$

and with each branch assigned (subjective) probability 1/2.

Consider any time t. What is the probability distribution of stock prices at time t? Let's assume for simplicity that t is a multiple of δt, specifically $t = n\delta t$. If in arriving at this time you got heads j times and tails $n - j$ times, then the stock price is

$$s(0) \exp\left[n\mu\delta t + j\sigma\sqrt{\delta t} - (n-j)\sigma\sqrt{\delta t}\right] = s(0) \exp\left[\mu t + (2j - n)\sigma\sqrt{\delta t}\right].$$

We should be able to understand the probability distribution (asymptotically as $\delta t \to 0$), since we surely understand the results of flipping a coin many times. Briefly: if you make a histogram of the proportion of heads, it will resemble (as $n \to \infty$) a Gaussian distribution centered at 1/2. We'll get the variance straight in a minute. (What we're really using here is the central limit theorem.)

To proceed more quantitatively it's helpful to use the notation of probability. Recognizing that j is a random variable, let's change notation to make it look like one by calling it X_n:

$$X_n = \text{number of times you get heads in } n \text{ flips of a fair coin.}$$

Since X_n is the sum of n independent random variables (one for each coin-flip), each taking values 0 and 1 with probability 1/2, one easily sees that

$$\text{Expected value of } X_n = n/2, \quad \text{Variance of } X_n = n/4.$$

Thus our histogram, which was the distribution function of $\frac{1}{n}X_n$, tended to a Gaussian with mean $\frac{1}{2}$ and variance $\frac{1}{4n}$. It's easy to see from this that

$$\frac{2X_n - n}{\sqrt{n}}$$

tends to a Gaussian with mean value 0 and variance 1. Since $\sqrt{\delta t} = \sqrt{t}/\sqrt{n}$ our formula for the final stock price can be expressed as

$$s(t) = s(0)\exp\left[\mu t + \sigma\sqrt{t}\frac{2X_n - n}{\sqrt{n}}\right].$$

Thus asymptotically, as $\delta t \to 0$ and $n \to \infty$ with $t = n\delta t$ held fixed,

$$s(t) = s(0)\exp\left[\mu t + \sigma\sqrt{t}Z\right]$$

where Z is a random variable with mean 0 and variance 1. In particular $\log s(t) - \log s(0)$ is a Gaussian random variable with mean μt and variance $\sigma^2 t$, as expected.

Our assertion of lognormal dynamics said a little more: that $\log s(t_2) - \log s(t_1)$ was Gaussian with mean $\mu(t_2 - t_1)$ and variance $\sigma^2(t_2 - t_1)$ for all $t_1 < t_2$. The justification is the same as what we did above – it wasn't really important that we started at 0.

Notice that our calculation used only the mean and variance of X_n, since it was based on the Central Limit Theorem. Our particular way of choosing the tree – with $s_{\text{up}} = s_{\text{now}}e^{\mu\delta t + \sigma\sqrt{\delta t}}$, $s_{\text{down}} = s_{\text{now}}e^{\mu\delta t - \sigma\sqrt{\delta t}}$, and with each choice having probability 1/2, was not the only one possible. A more general approach would take $s_{\text{up}} = s_{\text{now}}u$ with probability p, $s_{\text{down}} = s_{\text{now}}d$ with probability $1 - p$, and choose the three parameters u, d, p to satisfy two constraints associated with the mean and variance. Evidently one degree of freedom remains. Thus once p is fixed the other parameters are determined.

****

Implication for pricing options. We attached subjective probabilities (always equal to 1/2) to our binomial tree because we wanted to recognize lognormal dynamics as the limit of a coin-flipping process. Now let us consider one of those binomial trees – for some specific δt near 0 – and use it to price options.

The structure of the tree remains relevant (particularly the factors u and d determining $s_{\text{up}} = us_{\text{now}}$ and $s_{\text{down}} = ds_{\text{now}}$.) The subjective probabilities (1/2 for every branch) are irrelevant because our pricing is based on arbitrage. But we know a formula for the price of the option with payoff $f(s(T))$ at time maturity T:

$$V(f) = e^{-rT} \cdot E_{\text{RN}}[f(s_T)]$$

where E_{RN} denotes the expected value relative to the risk-neutral probability. And using the risk-neutral probability instead of the subjective probability just means our coin is no longer fair. Instead it is biased, with probability of heads (stock goes up) q and probability of tails (stock goes down) $1-q$, where

$$q = \frac{e^{r\delta t} - d}{u - d} = \frac{e^{r\delta t} - e^{\mu\delta t - \sigma\sqrt{\delta t}}}{e^{\mu\delta t + \sigma\sqrt{\delta t}} - e^{\mu\delta t - \sigma\sqrt{\delta t}}}.$$

One verifies (using the Taylor expansion of e^x near $x = 0$) that this is close to 1/2 when δt is small, and in fact

$$q = \frac{1}{2}\left(1 - \sqrt{\delta t}\frac{\mu - r + \frac{1}{2}\sigma^2}{\sigma}\right) + \text{terms of order } \delta t.$$

Our task is now clear. All we have to do is find the distribution of final values $s(T)$ when one uses the q-biased coin, then take the expected value of $f(s(T))$ with respect to this distribution. We can use a lot of what we did above: writing X_n for the number of heads as before, we still have

$$s(t) = s(0)\exp\left[\mu t + \sigma\sqrt{t}\frac{2X_n - n}{\sqrt{n}}\right].$$

But now X_n is the sum of n independent random variables with mean q and variance $q(1-q)$, so X_n has mean nq and variance $nq(1-q)$. So

$$\begin{aligned}\text{mean of } \frac{2X_n - n}{\sqrt{n}} &= (2q-1)\sqrt{n} \\ &\approx -\sqrt{t}\left(\frac{\mu - r + \frac{1}{2}\sigma^2}{\sigma}\right)\end{aligned}$$

and

$$\text{variance of } \frac{2X_n - n}{\sqrt{n}} \approx 1.$$

The central limit theorem tells us the limiting distribution is Gaussian, and the preceding calculation tells us its mean and variance. In summary: as $\delta t \to 0$, when using the biased coin associated with the risk-neutral probabilities,

$$s(t) = s(0) \exp\left[\mu t + \sigma\sqrt{t} Z'\right]$$

where Z' is a Gaussian random variable with mean $\sqrt{t}\left(\frac{r-\mu-\frac{1}{2}\sigma^2}{\sigma}\right)$ and variance 1. Equivalently, writing $Z' = Z + \sqrt{t}\left(\frac{r-\mu-\frac{1}{2}\sigma^2}{\sigma}\right)$,

$$s(t) = s(0) \exp\left[(r - \tfrac{1}{2}\sigma^2)t + \sigma\sqrt{t} Z\right]$$

where Z is Gaussian with mean 0 and variance 1. Notice that the statistical distribution of $s(t)$ depends on σ and r but not on μ (we'll return to this point soon).

The value of the option is the e^{-rT} times the expected value of the payoff relative to this probability distribution. Using the distribution function of the Gaussian to evaluate the expected value, we get:

$$V(f) = e^{-rT} E\left[f(s_0 e^X)\right]$$

where X is a Gaussian random variable with mean $(r - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$, or equivalently

$$V(f) = e^{-rT} \int_{-\infty}^{\infty} f(s_0 e^x) \frac{1}{\sigma\sqrt{2\pi T}} \exp\left[\frac{-(x - [r - \sigma^2/2]T)^2}{2\sigma^2 T}\right] dx.$$

This (when specialized to puts and calls) is the famous Black-Scholes relation.

We'll talk later about evaluating the integral. For now let's be satisfied with working backward through the binomial tree obtained with a specific (small) value of δt. Reviewing what we found above: given a lognormal stock process with return μ and volatility σ, and given a choice of δt, the tree should be constructed so that $s_{\text{up}} = u s_{\text{now}}$, $s_{\text{down}} = d s_{\text{now}}$ with

$$u = e^{\mu\delta t + \sigma\sqrt{\delta t}}, \quad d = e^{\mu\delta t - \sigma\sqrt{\delta t}}.$$

These determine the risk-neutral probability q by the formula given above. Working backward through the tree is equivalent to finding the discounted expected value of $f(s(T))$ relative to the risk-neutral probability.

Let us return to the observation, made above, that the statistics of $s(t)$ relative to the risk-neutral probability depend on σ (volatility of the stock) and r (risk-free return) but not on μ. It follows that for pricing derivative securities the value of μ isn't really needed. More precisely: in the limit $\delta t \to 0$ the lognormal stock models with different μ's but the same σ all assign the same values to options. So we may choose μ any way we please – there's no reason to require that it match the actual expected return of the stock under consideration. The two most common choices are

1. choose μ to be the expected return of the stock nevertheless; or
2. choose μ so that $\mu - r + \frac{1}{2}\sigma^2 = 0$, i.e. $\mu = r - \frac{1}{2}\sigma^2$.

The latter choice has the advantage that it puts q even closer to $1/2$. This is the selection favored by Jarrow–Turnbull and many other authors.

It may seem strange that the value of an option doesn't depend on μ. Heuristic argument why this should be so: we are using arbitrage considerations, so it doesn't matter whether the stock tends to go up or down, which is (mainly) what μ tells us.

Here's a more limited but less heuristic argument why the option pricing formula should not depend on μ. We start from the observation that in the special case $f(s) = s$, i.e. if the payoff is just the value-at-maturity of the stock, then the value of the option at time 0 must be s_0. We discussed this at length at the beginning of this section. Of course it should be valid also in the continuous-time limit. (The payoff $f(s_T) = s_T$ is replicated by a very simple trading strategy – namely hold one unit of stock and never trade – whether time is continuous or discrete.) Now consider the analysis we just completed, passing to the continuum limit via binomial trees. It tells us that when $f(s) = s$, the value of the option is

$$e^{-rT} E\left[s_0 e^X\right]$$

where X is Gaussian with mean $rT - \frac{1}{2}\sigma^2 T$ and variance $\sigma^2 T$. The two calculations are consistent only if for such X

$$e^{-rT} E\left[e^X\right] = 1.$$

Are the two calculations consistent? The answer is *yes*. Moreover, *if* you accept the existence of a pricing formula $V(f) = e^{-rT} E\left[f(s_0 e^X)\right]$, with X a Gaussian random variable with variance $\sigma^2 T$, then this consistency test *forces* the mean of X to be $rT - \frac{1}{2}\sigma^2 T$.

It remains to justify our assertion of consistency. This follows easily from the following fact:

Lemma: If X is a Gaussian random variable with mean m and standard deviation s then

$$E\left[e^X\right] = e^{m+\frac{1}{2}s^2}$$

Proof: We start from the formula

$$E\left[e^X\right] = \frac{1}{s\sqrt{2\pi}} \int_{-\infty}^{\infty} e^x e^{-\frac{(x-m)^2}{2s^2}}\, dx.$$

Complete the square:

$$x - \frac{(x-m)^2}{2s^2} = m + \tfrac{1}{2}s^2 - \frac{(x-[m+s^2])^2}{2s^2}.$$

Therefore the expected value of e^X is

$$e^{m+\frac{1}{2}s^2} \frac{1}{s\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[\frac{-[x-(m+s^2)]^2}{2s^2}\right] dx.$$

Making the change of variable $u = (x - [m+s^2])/s$ this becomes

$$e^{m+\frac{1}{2}s^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2}\, du = e^{m+\frac{1}{2}s^2}$$

as desired.

Derivative Securities – Section 5 – Fall 2004
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences.

The Black-Scholes formula and its applications. This Section deduces the Black-Scholes formula for a European call or put, as a consequence of risk-neutral valuation in the continuous time limit. Then we discuss the Delta, Gamma, Vega, and Rho of a portfolio, and their significance for hedging. Our treatment is closest to Jarrow and Turnbull, however Hull's treatment of this material is also excellent. Hedging is a very important topic, and these notes don't do justice to it; see e.g. chapter 14 of Hull's 5th edition for further, more practical discussion. We assume throughout these notes that the underlying asset pays no dividend and has no carrying cost. (We'll remove these assumptions later in the semester. Briefly: if the underlying asset pays a continuous dividend, e.g. if it is the value of a broad stock index, or a foreign exchange rate, the continuous-time "risk neutral process" is like what we found in the no-dividend case but with r replaced by $r - q$ where q is the dividend rate.)

****

The Black-Scholes formula for a European call or put. The upshot of Section 4 is this: the value at time t of a European option with payoff $f(s_T)$ is

$$V(f) = e^{-r(T-t)} E_{\mathrm{RN}}[f(s_T)].$$

Here $E_{\mathrm{RN}}[f(s_T)]$ is the expected value of the price at maturity with respect to a special probability distribution – the risk-neutral one. This distribution is determined by the property that

$$s_T = s_t \exp\left[(r - \tfrac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{T-t}Z\right]$$

where s_t is the spot price at time t and Z is Gaussian with mean 0 and variance 1. Equivalently: $\log[s_T/s_t]$ is Gaussian with mean $(r - \frac{1}{2}\sigma^2)(T-t)$ and variance $\sigma^2(T-t)$.

This formula can be evaluated for any payoff f by numerical integration. But for special payoffs – including the put and the call – we can get explicit expressions in terms of the "cumulative distribution function"

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du.$$

($N(x)$ is the probability that a Gaussian random variable with mean 0 and variance 1 has value $\leq x$.) The explicit formulas have advantages over numerical integration: besides being easy to evaluate, they permit us to see quite directly how the value and the hedge portfolio depend on strike price, spot price, risk-free rate, and volatility.

It's sufficient, of course, to consider $t = 0$. Let

$$\begin{aligned} c[s_0, T; K] &= \text{value at time 0 of a European call with strike } K \\ &\quad \text{and maturity } T\text{, if the spot price is } s_0; \\ p[s_0, T; K] &= \text{value at time 0 of a European put with strike } K \\ &\quad \text{and maturity } T\text{, if the spot price is } s_0. \end{aligned}$$

The explicit formulas are:

$$\begin{aligned} c[s_0, T; K] &= s_0 N(d_1) - K e^{-rT} N(d_2) \\ p[s_0, T; K] &= K e^{-rT} N(-d_2) - s_0 N(-d_1) \end{aligned}$$

in which

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T}} \left[\log(s_0/K) + (r + \tfrac{1}{2}\sigma^2)T\right] \\ d_2 &= \frac{1}{\sigma\sqrt{T}} \left[\log(s_0/K) + (r - \tfrac{1}{2}\sigma^2)T\right] = d_1 - \sigma\sqrt{T}. \end{aligned}$$

To derive these formulas we use the following result. (The calculation at the end of Section 4 was a special case.)

Lemma: Suppose X is Gaussian with mean μ and variance σ^2. Then for any real numbers a and k,

$$E\left[e^{aX} \text{ restricted to } X \geq k\right] = e^{a\mu + \frac{1}{2}a^2\sigma^2} N(d)$$

with $d = (-k + \mu + a\sigma^2)/\sigma$.
Proof: The left hand side is defined by

$$E\left[e^{aX} \text{ restricted to } X \geq k\right] = \frac{1}{\sigma\sqrt{2\pi}} \int_k^\infty e^{ax} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] dx.$$

Complete the square:

$$ax - \frac{(x-\mu)^2}{2\sigma^2} = a\mu + \tfrac{1}{2}a^2\sigma^2 - \frac{[x - (\mu + a\sigma^2)]^2}{2\sigma^2}.$$

Thus

$$E\left[e^{aX} \mid X \geq k\right] = e^{a\mu + \frac{1}{2}a^2\sigma^2} \cdot \frac{1}{\sigma\sqrt{2\pi}} \int_k^\infty \exp\left[\frac{-[x - (\mu + a\sigma^2)]^2}{2\sigma^2}\right] dx.$$

If we set $u = [x - (\mu + a\sigma^2)]/\sigma$ and $\kappa = [k - (\mu + a\sigma^2)]/\sigma$ this becomes

$$\begin{aligned} e^{a\mu + \frac{1}{2}a^2\sigma^2} \cdot \frac{1}{\sqrt{2\pi}} \int_\kappa^\infty e^{-u^2/2}\, du &= e^{a\mu + \frac{1}{2}a^2\sigma^2} [1 - N(\kappa)] \\ &= e^{a\mu + \frac{1}{2}a^2\sigma^2} N(d) \end{aligned}$$

where $d = -\kappa = (-k + \mu + a\sigma^2)/\sigma$.

We apply this to the European call. Our task is to evaluate

$$e^{-rT} \int_{-\infty}^\infty (s_0 e^x - K)_+ \frac{1}{\sigma\sqrt{2\pi T}} \exp\left[\frac{-(x - [r - \sigma^2/2]T)^2}{2\sigma^2 T}\right] dx.$$

The integrand is nonzero when $s_0 e^x > K$, i.e. when $x > \log(K/s_0)$. Applying the Lemma with $a = 1$ and $k = \log(K/s_0)$ we get

$$e^{-rT} \int_k^\infty s_0 e^x \frac{1}{\sigma\sqrt{2\pi T}} \exp\left[\frac{-(x - [r - \sigma^2/2]T)^2}{2\sigma^2 T}\right] dx = s_0 N(d_1);$$

applying the Lemma again with $a = 0$ we get

$$e^{-rT} \int_k^\infty K \frac{1}{\sigma\sqrt{2\pi T}} \exp\left[\frac{-(x - [r - \sigma^2/2]T)^2}{2\sigma^2 T}\right] dx = Ke^{-rT} N(d_2);$$

combining these results gives the formula for $c[s_0, T; K]$.

The formula for the value of a European put can be obtained similarly. Or – easier – we can derive it from the formula for a call, using put-call parity:

$$\begin{aligned} p[s_0, T; K] &= c[s_0, T; K] + Ke^{-rT} - s_0 \\ &= Ke^{-rT}[1 - N(d_2)] - s_0[1 - N(d_1)] \\ &= Ke^{-rT} N(-d_2) - s_0 N(-d_1). \end{aligned}$$

For options with maturity T and strike price K, the value at any time t is naturally $c[s_t, T - t; K]$ for a call, $p[s_t, T - t; K]$ for a put.

****

Hedging. We know how to hedge in the discrete-time, multiperiod binomial tree setting: the payoff is replicated by a portfolio consisting of $\Delta = \Delta(0, s_0)$ units of stock and a (long or short) bond, chosen to have the same value as the derivative claim. At time δt the stock price changes to $s_{\delta t}$ and the value of the hedge portfolio changes by $\Delta(s_{\delta t} - s_0)$. The new value of the hedge portfolio is also the new value of the option, so

$$\Delta(0, s_0) = \frac{\text{change in value of option from time 0 to } \delta t}{\text{change in value of stock from time 0 to } \delta t}.$$

The replication strategy requires a self-financing trade at every time step, adjusting the amount of stock in the portfolio to match the new value of Δ.

In the real world prices are not confined to a binomial tree, and there are no well-defined time steps. We cannot trade continuously. So while we can pass to the continuous time limit for the value of the option, we must still trade at discrete times in our attempts to replicate it. Suppose, for simplicity, we trade at equally spaced times with interval δt. What to use for the initial hedge ratio Δ? Not being clairvoyant we don't know the value of the stock at time δt, so we can't use the formula given above. Instead we should use its continuous-time limit:

$$\Delta(0, s_0) = \frac{\partial(\text{value of option})}{\partial(\text{value of stock})}.$$

There's a subtle point here: if the stock price changes continuously in time, but we only rebalance at discretely chosen times $j\delta t$, then we cannot expect to replicate the option perfectly using self-financing trades. Put differently: if we maintain the principle that the value of the hedge portfolio is equal to that of the option at each time $j\delta t$, then our trades will no longer be self-financing. We will address this point soon, after developing the continuous-time Black-Scholes theory. We'll show then that (if transaction costs are ignored) the expected cost of replication tends to 0 as $\delta t \to 0$. (In practice transaction costs are *not* negligible; deciding when, really, to rebalance, taking into account transaction costs, is an important and interesting problem – but one beyond the scope of this course.)

For the European put and call we can easily get formulas for Δ by differentiating our expressions for c and p: at time T from maturity the hedge ratio should be

$$\Delta = \frac{\partial}{\partial s_0} c[s_0, T; K] = N(d_1)$$

for the call, and

$$\Delta = \frac{\partial}{\partial s_0} p[s_0, T; K] = -N(-d_1)$$

for the put. The "hard way" to see this is an application of chain rule: for example, in the case of the call,

$$\frac{\partial}{\partial s_0} c = N(d_1) + s_0 N'(d_1) \frac{\partial d_1}{\partial s} - K e^{-rT} N'(d_2) \frac{\partial d_2}{\partial s}.$$

But $d_2 = d_1 - \sigma\sqrt{T}$, so $\partial d_1 / \partial s = \partial d_2 / \partial s$; also $N'(x) = \frac{1}{\sqrt{2\pi}} \exp[-x^2/2]$. It follows with some calculation that

$$s_0 N'(d_1) \frac{\partial d_1}{\partial s} - K e^{-rT} N'(d_2) \frac{\partial d_2}{\partial s} = 0,$$

so finally $\partial c / \partial s_0 = N(d_1)$ as asserted. There is however an easier way: differentiate the original formula expressing the value as a discounted risk-neutral expectation. Passing the derivative under the integral, for a call with strike K:

$$\begin{aligned}
\Delta &= \frac{\partial}{\partial s_0} e^{-rT} \int_{-\infty}^{\infty} (s_0 e^x - K)_+ \frac{1}{\sigma\sqrt{2\pi T}} \exp\left[\frac{-(x - [r - \sigma^2/2]T)^2}{2\sigma^2 T}\right] dx \\
&= e^{-rT} \int_{-\infty}^{\infty} \frac{\partial (s_0 e^x - K)_+}{\partial s_0} \frac{1}{\sigma\sqrt{2\pi T}} \exp\left[\frac{-(x - [r - \sigma^2/2]T)^2}{2\sigma^2 T}\right] dx \\
&= e^{-rT} \int_{\log(K/s_0)}^{\infty} e^x \frac{1}{\sigma\sqrt{2\pi T}} \exp\left[\frac{-(x - [r - \sigma^2/2]T)^2}{2\sigma^2 T}\right] dx \\
&= N(d_1).
\end{aligned}$$

****

Since we have explicit formulas for the value of a put or call, we can differentiate them to learn the dependence on the underlying parameters. Some of these derivatives have names:

Definition	**Call**	**Put**
Delta $= \frac{\partial}{\partial s_0}$	$N(d_1) > 0$	$-N(-d_1) < 0$
Gamma $= \frac{\partial^2}{\partial s_0^2}$	$\frac{1}{s_0\sigma\sqrt{2\pi T}}\exp(-d_1^2/2) > 0$	$\frac{1}{s_0\sigma\sqrt{2\pi T}}\exp(-d_1^2/2) > 0$
Theta $= \frac{\partial}{\partial t}$	$-\frac{s_0\sigma}{2\sqrt{2\pi T}}\exp(-d_1^2/2) - rKe^{-rT}N(d_2) < 0$	$-\frac{s_0\sigma}{2\sqrt{2\pi T}}\exp(-d_1^2/2) + rKe^{-rT}N(-d_2) \gtrless 0$
Vega $= \frac{\partial}{\partial \sigma}$	$\frac{s_0\sqrt{T}}{\sqrt{2\pi}}\exp(-d_1^2/2) > 0$	$\frac{s_0\sqrt{T}}{\sqrt{2\pi}}\exp(-d_1^2/2) > 0$
Rho $= \frac{\partial}{\partial r}$	$TKe^{-rT}N(d_2) > 0$	$-TKe^{-rT}N(-d_2) < 0.$

These formulas apply at time $t = 0$; the formulas applicable at any time t are similar, with T replaced by $T - t$. These are obviously useful for understanding how the value of the option changes with time, volatility, etc. But more: they are useful for designing improved hedges. For example, suppose a bank sells two types of options on the same underlying asset, with different strike prices and maturities. As usual the bank wants to limit its exposure to changes in the stock price; but suppose in addition it wants to limit its exposure to changes (or errors in specification of) volatility. Let $i = 1, 2$ refer to the two types of options, and let n_1, n_2 be the quantities held of each. (These are negative if the bank sold the options.) The bank naturally also invests in the underlying stock and in risk-free bonds; let n_s and n_b be the quantities held of each. Then the value of the bank's initial portfolio is

$$V_{\text{total}} = n_1 V_1 + n_2 V_2 + n_s s_0 + n_b.$$

We already know how the stock and bond holdings should be chosen if the bank plans to replicate (dynamically) the options: they should satisfy

$$V_{\text{total}} = 0$$

and

$$n_1\Delta_1 + n_2\Delta_2 + n_s = 0.$$

Notice that the latter relation says $\partial V_{\text{total}}/\partial s_0 = 0$: the value of the bank's holdings are insensitive (to first order) to changes in the stock price.

If we were dealing in just one option there would be no further freedom: we would have two homogeneous equations in three variables n_1, n_s, n_b, restricting their values to a line – so that n_1 determines n_s and n_b. That's the situation we're familiar with. But if we're dealing in two (independent) options then we have the freedom to impose one additional linear equation. For example we can ask that the portfolio be insensitive (to first order) to changes in σ by imposing the additional condition

$$n_1 \text{Vega}_1 + n_2 \text{Vega}_2 = 0.$$

Thus: by selling the two types of assets in the proper proportions the bank can reduce its exposure to change or misspecification of volatility.

If the bank sells three types of options then we have room for yet another condition – e.g. we could impose first-order insensitivity to changes in the risk-free rate r. And so on. It is not actually necessary that the bank use the underlying stock as one of its assets. Each option is *equivalent* to a portfolio consisting of stock and risk-free bond; so a portfolio consisting entirely of options and a bond position will function as a hedge portfolio so long as its total Δ is equal to 0.

Replication requires dynamic rebalancing. The bank must change its holdings at each time increment to set the new Δ to 0. In the familiar, one-option setting this was done by adjusting the stock and bond holdings, keeping the option holding fixed. In the present, two-option setting, maintaining the additional condition $\text{Vega}_{\text{total}} = 0$ will require the ratio between n_1 and n_2 to be dynamically updated as well, i.e. the bank will have to sell or buy additional options as time proceeds.

Derivative Securities – Section 6 – Fall 2004
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences.

Stochastic differential equations and the Black-Scholes PDE. We derived the Black-Scholes formula by using no-arbitrage-based (risk-neutral) valuation in a discrete-time, binomial tree setting, then passing to a continuum limit. We started that way because binomial trees are very explicit and transparent. However the power of the discrete framework as a conceptual tool is rather limited. Therefore we now begin developing the more powerful continuous-time framework, via the Ito calculus and the Black-Scholes differential equation. This material is discussed in many places. Baxter & Rennie emphasize risk-neutral expectation, avoiding almost completely the discussion of PDE's. The "student guide" by Wilmott, Howison, & Dewynne takes almost the opposite approach: it emphasizes PDE's, avoiding almost completely the discussion of risk-neutral expectation. Neftci's book provides a good introduction to Brownian motion, the Ito calculus, stochastic differential equations, and their relation to option pricing, at a level that should be accessible to students in this class. (Students taking Stochastic Calculus will learn this material and much more over the course of the semester.) A brief survey of stochastic calculus (similar in spirit to what's here, but with more detail and more examples) can be found at the top of my Spring 2003 PDE for Finance course notes (on my web page).

****

Why work in continuous time?. Our discrete-time approach has the advantage of being very clear and explicit. However there is a different approach, based on Taylor expansion, the Ito calculus, and the "Black-Scholes differential equation." It has its own advantages:

- Passing to the continuous time limit is clearly legitimate for describing the stock price process. But is it legitimate for describing the value of the option, as determined by arbitrage? This is less clear, since a continuous-time hedging strategy is unattainable in practice. In what sense can we "approximately replicate" the option by trading at discrete times? The Black-Scholes differential equation will help us answer these questions.

- The differential equation approach gives fresh insight and computational flexibility. Imagine trying to understand the implications of compound interest without using the differential equation $df/dt = rf$! (Especially: imagine how stuck you'd be if r depended on f.)

- Differential-equation-based methods lead to efficient computational schemes (and even explicit solution formulas in some cases) not only for European options, but also for more complicated instruments such as barrier options.

Brownian motion. Recall our discussion of the lognormal hypothesis for stock price dynamics. It says that $\log[s(t_2)/s(t_1)]$ is a Gaussian random variable with mean $\mu(t_2 - t_1)$ and variance $\sigma^2(t_2 - t_1)$, and disjoint intervals give rise to independent random variables.

A time-dependent random variable is called a *stochastic process*. The lognormal hypothesis is related to Brownian motion $x(t)$, also known as the Wiener process, which satisfies:

(a) $x(t_2) - x(t_1)$ is a Gaussian random variable with mean value 0 and variance $t_2 - t_1$;

(b) distinct intervals give rise to independent random variables;

(c) $x(0) = 0$.

It can be proved that these properties determine a unique stochastic process, i.e. they uniquely determine the probability distribution of any expression involving $x(t_1)$, $x(t_2)$, $\ldots$, $x(t_N)$. Also: for almost any realization, the function $t \mapsto x(t)$ is continuous but not differentiable. The process $x(t)$ can be viewed as a limit of suitably scaled random walks (we showed this in Section 4). Another important fact: writing $x(t_2) - x(t_1) = \Delta x$ and $t_2 - t_1 = \Delta t$,

$$E\left[|\Delta x|^j\right] = C_j |\Delta t|^{j/2}, \quad j = 1, 2, 3, \ldots.$$

Our lognormal hypothesis can be reformulated as the statement that

$$s(t) = s(0) \exp[\mu t + \sigma x(t)].$$

Stochastic differential equations and Ito's lemma. Let's first review ordinary differential equations. Consider the ODE $dy/dt = f(y,t)$ with initial condition $y(0) = y_0$. It is a convenient mnemonic to write the equation in the form

$$dy = f(y,t)dt.$$

This reminds us that the solution is well approximated by its (explicit) finite difference approximation $y([j+1]\delta t) - y(j\delta t) = f(y(j\delta t), j\delta t)\delta t$, which we sometimes write more schematically as

$$\Delta y = f(y,t)\Delta t.$$

An extremely useful aspect of ODE's is the ability to use chain rule. From the ODE for $y(t)$ we can easily deduce a new ODE satisfied by any function of $y(t)$. For example, $z(t) = e^{y(t)}$ satisfies $dz/dt = e^y dy/dt = zf(\log z, t)$. In general $z = A(y(t))$ satisfies $dz/dt = A'(y)dy/dt$. The mnemonic for this is

$$dA(y) = \frac{dA}{dy}dy = \frac{dA}{dy}f(y,t)dt.$$

It reminds us of the proof, which boils down to the fact that (by Taylor expansion)

$$\Delta A = A'(y)\Delta y + \text{error of order } |\Delta y|^2.$$

In the limit as the timestep tends to 0 we can ignore the error term, because $|\Delta y|^2 \leq C|\Delta t|^2$ and the sum of $1/\Delta t$ such terms is small, of order $|\Delta t|$.

OK, now stochastic differential equations. We consider only the simplest class of stochastic differential equations, namely

$$dy = g(y,t)dx + f(y,t)dt, \quad y(0) = y_0,$$

where $x(t)$ is Brownian motion. The solution is a stochastic process, the limit of the processes obtained by the (explicit) finite difference scheme

$$y([j+1]\delta t) - y(j\delta t) = g(y(j\delta t), t)\,(x([j+1]\delta t) - x(j\delta t)) + f(y(j\delta t), j\delta t)\delta t,$$

which we usually write more schematically as

$$\Delta y = g(y,t)\Delta x + f(y,t)\Delta t.$$

Put differently (this is how the rigorous theory begins): we can understand the stochastic differential equation by rewriting it in integral form:

$$y(t') = y(t) + \int_t^{t'} f(y(\tau),\tau)\,d\tau + \int_t^{t'} g(y(\tau),\tau)dx(\tau)$$

where the first integral is a standard Riemann integral, and the second one is a *stochastic integral*:

$$\int_t^{t'} g(y(\tau),\tau)\,dx(\tau) = \lim_{\Delta\tau\to 0} \sum g(y(\tau_i),\tau_i)[x(\tau_{i+1}) - x(\tau_i)]$$

where $t = \tau_0 < \ldots < \tau_N = t'$.

It's easy to see that when μ and σ are constant, $y(t) = \mu t + \sigma x(t)$ solves

$$dy = \sigma dx + \mu dt.$$

The analogue of the chain rule calculation done above for ODE's is known as Ito's lemma. It says that if $dy = gdx + fdt$ then $z = A(y)$ satisfies the stochastic differential equation

$$dz = A'(y)dy + \tfrac{1}{2}A''(y)g^2dt = A'(y)gdx + \left[A'(y)f + \tfrac{1}{2}A''(y)g^2\right]dt.$$

Here is a heuristic justification: carrying the Taylor expansion of $A(y)$ to second order gives

$$\begin{aligned}\Delta A &= A'(y)\Delta y + \tfrac{1}{2}A''(y)(\Delta y)^2 + \text{error of order } |\Delta y|^3\\ &= A'(y)(g\Delta x + f\Delta t) + \tfrac{1}{2}A''(y)g^2(\Delta x)^2 + \text{errors of order } |\Delta y|^3 + |\Delta x||\Delta t| + |\Delta t|^2.\end{aligned}$$

One can show that the error terms are negligible in the limit $\Delta t \to 0$. For example, the sum of $1/\Delta t$ terms of order $|\Delta x||\Delta t|$ has expected value of order $\sqrt{\Delta t}$. Thus

$$\Delta A \approx A'(y)(g\Delta x + f\Delta t) + \tfrac{1}{2}A''(y)g^2(\Delta x)^2.$$

Now comes the subtle part of Ito's Lemma: the assertion that we we can replace $(\Delta x)^2$ in the preceding expression by Δt. This is sometimes mistakenly justified by saying "$(\Delta x)^2$ behaves deterministically as $\Delta t \to 0$" – which is certainly not true; in fact $(\Delta x)^2 = u^2\Delta t$ where u is a Gaussian random variable with mean value 0 and variance 1.

So why can we substitute Δt for $(\Delta x)^2$? This can be thought of as an extension of the law of large numbers. When we solve a difference equation (to approximate a differential equation) we must add the terms corresponding to different time intervals. So we're really interested in *sums* of the form

$$\sum_{j=1}^{N} A''(y(t_j))g^2(t_j)(\Delta x)_j^2$$

with $(\Delta x)_j = x(t_{j+1}) - x(t_j)$ and $N = T/\Delta t$. If A'' and g^2 were constant then, since the $(\Delta x)_j$ are independent, $\sum_{j=1}^{N}(\Delta x)_j^2 = \sum_{j=1}^{N} u_j^2 \Delta t$ would have mean value $N\Delta t = T$ and variance of order $N(\Delta t)^2 = T\Delta t$. Thus the sum would have standard deviation $\sqrt{T\Delta t}$, i.e. it is asymptotically deterministic. The rigorous argument is different, of course, since in truth $A''(y)g^2$ is not constant; but the essential idea is similar.

The version of Ito's Lemma stated and justified above is not the most general one – though it has all the main ideas. Similar logic applies, for example, if A is a function of both y and t. Then $z = A(y,t)$ satisfies

$$dz = \frac{\partial A}{\partial y}dy + \frac{\partial A}{\partial t}dt + \tfrac{1}{2}\frac{\partial^2 A}{\partial y^2}g^2 dt = \frac{\partial A}{\partial y}g dx + \left[\frac{\partial A}{\partial y}f + \frac{\partial A}{\partial t} + \tfrac{1}{2}\frac{\partial^2 A}{\partial y^2}g^2\right]dt.$$

Let's apply Ito's lemma to find the stochastic differential equation for the stock price process $s(t)$. The lognormal hypothesis says $s = e^y$ where $dy = \sigma dx + \mu dt$. Therefore $ds = e^y(\sigma dx + \mu dt) + \frac{1}{2}e^y\sigma^2 dt$, i.e.

$$\frac{1}{s}ds = \sigma dx + (\mu + \tfrac{1}{2}\sigma^2)dt.$$

(Warning: our conventions are those of Jarrow–Turnbull; however many books, including Wilmott-Howison-Dewynne and Hull, use a different notational convention. They assume that the price process solves the stochastic differential equation $\frac{1}{s}ds = \sigma dx + \mu dt$. That's OK – but then the quantity we've been calling the mean return is $\mu - \frac{1}{2}\sigma^2$ rather than μ.)

The Black-Scholes partial differential equation. Consider a European option with payoff $f(s_T)$ at maturity T. We have a formula for its value at time t, from Section 4:

$$\text{value at time t} = e^{-r(T-t)}\frac{1}{\sigma\sqrt{2\pi(T-t)}}\int_{-\infty}^{\infty} f(s_t e^x)\exp\left[\frac{-(x-[r-\frac{1}{2}\sigma^2](T-t)^2}{2\sigma^2(T-t)}\right]dx.$$

Notice that the value is a function of the present time t and the present stock price s_t, i.e. it can be expressed in the form:

$$\text{value at time t} = V(s_t, t).$$

for a suitable function $V(s,t)$ defined for $s > 0$ and $t < T$. It's obvious from the interpretation of V that

$$V(s,T) = f(s).$$

The Black-Scholes differential equation says that

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 s^2\frac{\partial^2 V}{\partial s^2} + rs\frac{\partial V}{\partial s} - rV = 0.$$

It offers an alternative procedure for evaluating the value of the option, by solving the PDE "backwards in time" numerically, using $t = T$ as the initial time.

Recall that in the setting of binomial trees we had two ways of evaluating the value of an option: one by expressing it as a weighted sum over all paths; the other by working backward through the tree. Evaluating the integral formula is the continuous-time analogue of summing over all paths. Solving the Black-Scholes PDE is the continuous-time analogue of working backward through the tree. Recall also that working backward through the tree was a little more flexible – for example it didn't require that the interest rate be constant. Similarly, the Black-Scholes equation can easily be solved numerically even when the interest rate and volatility are (deterministic) functions of time.

Where does the equation come from? We'll give two (related, but different) justifications, both based on Ito's formula. Examining these derivations you'll be able to see how the Black-Scholes PDE generalizes to more complicated market models (for example when the volatility and drift depend on stock price). However for simplicity we'll present the arguments in the usual constant-volatility, constant-drift setting

$$ds = \sigma s dx + (\mu + \tfrac{1}{2}\sigma^2) s dt$$

and we'll continue to assume that the interest rate is constant.

Derivation by considering a hedging strategy. Remember that when hedging in the discrete-time setting, we rebalance the portfolio so that it contains ϕ units of stock and the rest a risk-free bond, then we let the stock price jump to the new value. (I write ϕ not Δ to avoid confusion, because we have been using Δ for increments.) The analogous procedure in the continuous-time setting is to rebalance at successive at time intervals of length δt, then pass to the limit $\delta t \to 0$. Suppose that after rebalancing at time $j\delta t$ the portfolio contains $\phi = \phi(s(j\delta t), j\delta t)$ units of stock. Consider the value of the option less the value of the stock during the next time interval:

$$\Pi = V - \phi s.$$

Its increment $d\Pi = \Pi([j+1]\delta t) - \Pi(j\delta t)$ is approximately

$$\begin{aligned} dV - \phi ds &= \frac{\partial V}{\partial s} ds + \frac{\partial V}{\partial t} dt + \tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma^2 s^2 dt - \phi ds \\ &= \left(\frac{\partial V}{\partial s} - \phi\right)\sigma s dx + \left(\frac{\partial V}{\partial s} - \phi\right)(\mu + \tfrac{1}{2}\sigma^2) s dt + \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma^2 s^2\right) dt. \end{aligned}$$

Note that we do not differentiate ϕ because it is being held fixed during this time interval. We know enough to expect that the right choice of ϕ is $\phi(s,t) = \partial V/\partial s$. But if we didn't already know, we'd discover it now: this is the choice that eliminates the dx term on the right hand side of the the last equation. Fixing ϕ this way, we see that

$$dV - \phi ds = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma^2 s^2\right) dt \quad \text{is deterministic.}$$

Now, the principle of no arbitrage says that a portfolio whose return is deterministic must grow at the risk-free rate. In other words, for this choice of ϕ we must have

$$dV - \phi ds = r(V - \phi s)dt.$$

Combining these equations gives

$$\left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma^2 s^2\right) = r(V - \phi s)$$

with $\phi = \partial V / \partial s$. This is precisely the Black-Scholes PDE.

Derivation from the risk-neutral pricing formula. We learned in Section 3 that the value at time t of an option with payoff f is

$$e^{-r(T-t)} E_{\mathrm{RN}}[f(s(T))]$$

where the right hand side is the discounted expected final value using the risk-neutral process. We also learned in Section 4 that in the continuous time limit the risk-neutral process is

$$s(T) = s(t)\exp[(r - \tfrac{1}{2}\sigma^2)(T-t) + \sigma(x(T) - x(t))]$$

where x is a Brownian motion process. Now we know a different way to say the same thing: the risk-neutral process solves the stochastic differential equation

$$ds = rsdt + \sigma sdx$$

times between t and T, with initial data $s(t)$ that's known at the time t when we wish to value the option. We shall show that if V solves the Black-Scholes PDE with final-value f at time T, and if the stock price is $s(t)$ at time t, then $V(t, s(t))$ is indeed equal to the discounted risk-neutral expectation.

We begin again by using Ito's formula, this time applying it to the function $U(t,s) = \exp[r(T-t)]V(t,s)$, evaluated at $s = s(t)$. Evidently

$$\begin{aligned} dU &= e^{r(T-t)}\left[-rVdt + V_t dt + V_s ds + \tfrac{1}{2}V_{ss}(ds)^2\right] \\ &= e^{r(T-t)}\left[(-rV + V_t + rsV_s + \tfrac{1}{2}\sigma^2 s^2 V_{ss})dt + \sigma s V_s dx\right]. \end{aligned}$$

So far we could have used *any* smooth function $V(t,s)$. But if we use the solution of the Black-Scholes PDE then the right hand side simplifies a lot, and we get

$$dU = e^{r(T-t)}\sigma s V_s dx.$$

Strictly speaking a stochastic differential equation is shorthand for an integral equation; this is shorthand for

$$U(t') - U(t) = \int_t^{t'} e^{r(T-\tau)}\sigma s(\tau) V_s(\tau, s(\tau))\, dx(\tau)$$

for any $t' > t$. The crucial point is that *the right hand side has expected value* 0. In fact, any integral of the form $\int_a^b g dx$ has expected value 0 when x is Brownian motion, because it is the limit of expressions of the form $\sum g(\tau_i)[x(\tau_{i+1}) - x(\tau_i)]$ and each term in the sum has mean value 0. Applying this fact with $t' = T$ gives

$$E[U(T)] = E[U(t)].$$

But $U(t)$ is known with certainty at time t, so

$$E[U(t)] = U(t) = e^{r(T-t)} V(t, s(t)).$$

And V is known at time T, namely $V(T, s) = f(s)$ for all s, so

$$E[U(T)] = E[f(s(T)]$$

is the (undiscounted) expected final value of the option. Thus we have show that

$$e^{r(T-t)} V(t, s(t)) = E_{\mathrm{RN}}[f(s(T)]$$

as expected. (We wrote E rather than E_{RN} above, not to clutter our notation, but *all* the expectations taken above were in the risk-neutral setting, where the price process solves $ds = rs\, dt + \sigma s\, dx$.)

Derivative Securities – Section 7 – Fall 2004
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences.

Topics in this section: (a) further discussion of SDE's, including some examples and applications; (b) reduction of Black-Scholes PDE to the linear heat equation; and (c) discussion of what happens when you hedge discretely rather than continuously in time.

When I taught this class in Fall 2000 I discussed barrier options at this point. This time around I prefer to postpone that discussion. But you now have enough background to read about barrier options if you like; see Section 7 of my Fall 2000 notes, or the discussion in the "student guide" by Dewynne, Howison, and Wilmott.

****

Further discussion of stochastic differential equations. Several students requested more information on examples of SDEs' and how they can be used. Therefore the discussion that follows goes somewhat beyond the bare minimum we'll be using in this class. (Everything here is, however, relevant to financial applications.) For simplicity, we restricted the discussion to problems with a "single source of randomness," i.e. scalar SDE's of the form

$$dy = f(y(t), t)\, dt + g(y(t), t)\, dw \tag{1}$$

where w is a scalar-valued Brownian motion. The main things we will use about stochastic integrals and SDE's are the following:

(1) *Ito's lemma.* We discussed in the Section 6 notes the fact that if A is a smooth function of two variables and y solves (1) then $z = A(t, y(t))$ solves the SDE

$$dz = A_t\, dt + A_y\, dy + \tfrac{1}{2} A_{yy}\, dy\, dy = (A_t + A_y f + \tfrac{1}{2} A_{yy} g^2)\, dt + A_y g\, dw.$$

We'll also sometimes use this generalization: if y_1 and y_2 solve SDE's using the *same* Brownian motion w, say

$$dy_1 = f_1\, dt + g_1\, dw \quad \text{and} \quad dy_2 = f_2\, dt + g_2\, dw,$$

and $A(t, y_1, y_2)$ is a smooth function of three variables, then $z = A(t, y_1(t), y_2(t))$ solves the SDE

$$dz = A_t\, dt + A_1\, dy_1 + A_2\, dy_2 + \tfrac{1}{2} A_{11}\, dy_1\, dy_1 + A_{12}\, dy_1\, dy_2 + \tfrac{1}{2} A_{22}\, dy_2\, dy_2$$

with the understanding that

$$A_{ij} = \partial^2 A / \partial y_i \partial y_j \quad \text{and} \quad dy_i\, dy_j = g_i g_j\, dt.$$

The heuristic Taylor-expansion-based explanation is exactly parallel to the one sketched in Section 6.

(2) *A stochastic integral* $\int_a^b F\,dw$ *has mean value zero.* We used (and explained) this assertion at the end of Section 6, but perhaps we didn't emphasize it enough. The explanation is easy. The integrand $F = F(t, y(t))$ can be any function of t and $y(t)$. (The key point: its value at time t should depend only on information available at time t.) The stochastic integral is the limit of the Riemann sums

$$\sum F(t_j, y(t_j))[w(t_{j+1}) - w(t_j)]$$

and each term of this sum has mean value zero, since the increment $w(t_{j+1}) - w(t_j)$ has mean value 0 and is independent of $F(t_j, y(t_j))$.

(3) *Calculating the variance of a* dw *integral.* We just showed that $\int_a^b F\,dw$ has mean value 0. What about its variance? The answer is simple:

$$E\left[\left(\int_a^b F(s, y(s))\,dw\right)^2\right] = \int_a^b E[F^2(s, y(s))]\,ds. \tag{2}$$

The justification is easy. Just approximate the stochastic integral as a sum. The square of the stochastic integral is approximately

$$\left(\sum_{i,j} F(s_i, y(s_i))[w(s_{i+1}) - w(s_i)]\right)\left(\sum_{i,j} F(s_j, y(s_j))[w(s_{j+1}) - w(s_j)]\right)$$
$$= \sum_{i,j} F(s_i, y(s_i))F(s_j, y(s_j))[w(s_{i+1}) - w(s_i)][w(s_{j+1}) - w(s_j)] \quad .$$

For $i \neq j$ the expected value of the i, jth term is 0 (for example, if $i < j$ then $[w(s_{j+1}) - w(s_j)]$ has mean value 0 and is independent of $F(s_i, y(s_i))$, $F(s_j, y(s_j))$, and $[w(s_{i+1}) - w(s_i)]$). For $i = j$ the expected value of the i, jth term is $E[F^2(s_i, y(s_i))][s_{i+1} - s_i]$. So the expected value of the squared stochastic integral is approximately

$$\sum_i E[F^2(y(s_i), s_i)][s_{i+1} - s_i],$$

and passing to the limit $\Delta s \to 0$ gives the formula (2).

The following examples have been extracted from the "Stochastic Calculus Primer" posted at the top of my Spring 2003 PDE for Finance notes.

Log-normal dynamics with time-dependent drift and volatility. Suppose

$$dy = \mu(t)y dt + \sigma(t)y dw \tag{3}$$

where $\mu(t)$ and $\sigma(t)$ are (deterministic) functions of time. What stochastic differential equation describes $\log y$? Ito's lemma gives

$$\begin{aligned} d(\log y) &= y^{-1}dy - \tfrac{1}{2}y^{-2}dydy \\ &= \mu(t)dt + \sigma(t)dw - \tfrac{1}{2}\sigma^2(t)dt. \end{aligned}$$

Remembering that $y(t) = e^{\log y(t)}$, we see that

$$y(t_1) = y(t_0)e^{\int_{t_0}^{t_1}(\mu-\sigma^2/2)ds + \int_{t_0}^{t_1}\sigma dw}.$$

When μ and σ are constant in time we recover the formula (which we already knew):

$$y(t_1) = y(t_0)e^{(\mu-\sigma^2/2)(t_1-t_0)+\sigma(w(t_1)-w(t_0))}.$$

Stochastic stability. Consider once more the solution of (3). It's natural to expect that if μ is negative and σ is not too large then y should tend (in some average sense) to 0. This can be seen directly from the solution formula just derived. But an alternative, instructive approach is to consider the second moment $\rho(t) = E[y^2(t)]$. From Ito's formula,

$$d(y^2) = 2ydy + dydy = 2y(\mu ydt + \sigma ydw) + \sigma^2 y^2 dt.$$

Taking the expectation, we find that

$$E[y^2(t_1)] - E[y^2(t_0)] = \int_{t_0}^{t_1}(2\mu + \sigma)E[y^2]ds$$

or in other words

$$d\rho/dt = (2\mu + \sigma)\rho.$$

Thus $\rho = E[y^2]$ can be calculated by solving this deterministic ODE. If the solution tends to 0 as $t \to \infty$ then we conclude that y tends to zero in the mean-square sense. When μ and σ are constant this happens exactly when $2\mu+\sigma < 0$. When they are functions of time, the condition $2\mu(t) + \sigma(t) \leq -c$ is sufficient (with $c > 0$) since it gives $d\rho/dt \leq -c\rho$.

An example related to Girsanov's theorem. Suppose $\gamma(t)$ depends only on information up to time t. (For example, it could have the form $\gamma(t) = F(t, y(t))$ where y solves an SDE of the form (1).) Then

$$E\left[e^{\int_a^b \gamma(s)dw - \frac{1}{2}\int_a^b \gamma^2(s)ds}\right] = 1.$$

In fact, this is the expected value of $e^{z(b)}$, where

$$dz = -\tfrac{1}{2}\gamma^2(t)dt + \gamma(t)dw, \quad z(a) = 0.$$

Ito's lemma gives

$$d(e^z) = e^z dz + \tfrac{1}{2}e^z dzdz = e^z\gamma dw.$$

So

$$e^{z(b)} - e^{z(a)} = \int_a^b e^z \gamma dw.$$

The right hand side has expected value zero, so

$$E[e^{z(b)}] = E[e^{z(a)}] = 1.$$

Notice the close relation with the previous example "lognormal dynamics": all we've really done is identify the conditions under which $\mu = 0$ in (3).

[Comment for those taking Stochastic Calculus: this example is related to Girsanov's theorem, which gives the relation between the measure on path space associated with drift γ and the measure on path space associated with no drift. The expression

$$e^{\int_a^b \gamma(s)dw - \frac{1}{2}\int_a^b \gamma^2(s)ds}$$

is the Radon-Nikodym derivative relating these measures. The fact that it has expected value 1 reflects the fact that both measures are probability measures.]

The Ornstein-Uhlenbeck process. You should have learned in calculus that the deterministic differential equation $dy/dt + Ay = f$ can be solved explicitly when A is constant. Just multiply by e^{At} to see that $d(e^{At}y)/dt = e^{At}f$ then integrate both sides in time. So it's natural to expect that linear stochastic differential equations can also be solved explicitly. We focus on one important example: the "Ornstein-Uhlenbeck process," which solves

$$dy = -cydt + \sigma dw, \quad y(0) = x$$

with c and σ constant. (This is *not* a special case of (3), because the dw term is not proportional to y.) Ito's lemma gives

$$d(e^{ct}y) = ce^{ct}ydt + e^{ct}dy = e^{ct}\sigma dw$$

so

$$e^{ct}y(t) - x = \sigma \int_0^t e^{cs}dw,$$

or in other words

$$y(t) = e^{-ct}x + \sigma \int_0^t e^{c(s-t)}dw(s).$$

Now observe that $y(t)$ is a Gaussian random variable – because when we approximate the stochastic integral as a sum, the sum is a linear combination of Gaussian random variables. (We use here that a sum of Gaussian random variables is Gaussian; also that a limit of Gaussian random variables is Gaussian.) So $y(t)$ is entirely described by its mean and variance. They are easy to calculate: the mean is

$$E[y(t)] = e^{-ct}x$$

since the "dw" integral has expected value 0. To calculate the variance we use the formula (2). It gives

$$\begin{aligned} E\left[(y(t) - E[y(t)])^2\right] &= \sigma^2 E\left[\left(\int_0^t e^{c(s-t)}dw(s)\right)^2\right] \\ &= \sigma^2 \int_0^t e^{2c(s-t)}ds \\ &= \sigma^2 \frac{1 - e^{-2ct}}{2c}. \end{aligned}$$

We close this example with a brief discussion of the relevance of the Ornstein-Uhlenbeck process. One of the simplest interest-rate models in common use is that of Vasicek, which supposes that the (short-term) interest rate $r(t)$ satisfies

$$dr = a(b - r)dt + \sigma dw$$

with a, b, and σ constant. Interpretation: r tends to revert to some long-term average value b, but noise keeps perturbing it away from this value. Clearly $y = r - b$ is an Ornstein-Uhlenbeck process, since $dy = -aydt + \sigma dw$. Notice that $r(t)$ has a positive probability of being negative (since it is a Gaussian random variable); this is a reminder that the Vasicek model is not very realistic. Even so, its exact solution formulas provide helpful intuition.

****

Reduction of the Black-Scholes PDE to the linear heat equation. The linear heat equation $u_t = u_{xx}$ is the most basic example of a parabolic PDE; its properties and solutions are discussed in every textbook on PDE's. The Black-Scholes equation is really just this standard equation written in special variables. This fact is very well-known; my discussion follows the book by Dewynne, Howison, and Wilmott.

Recall that the Black-Scholes PDE is

$$V_t + \tfrac{1}{2}\sigma^2 s^2 V_{ss} + rsV_s - rV = 0;$$

we assume in the following that r and σ are constant. Consider the preliminary change of variables from (s,t) to (x,τ) defined by

$$s = e^x, \quad \tau = \tfrac{1}{2}\sigma^2(T - t),$$

and let $v(x,\tau) = V(s,t)$. An elementary calculation shows that the Black-Scholes equation becomes

$$v_\tau - v_{xx} + (1-k)v_x + kv = 0$$

with $k = r/(\frac{1}{2}\sigma^2)$. We've done the main part of the job: reduction to a constant-coefficient equation. For the rest, consider $u(x,t)$ defined by

$$v = e^{\alpha x + \beta \tau} u(x,\tau)$$

where α and β are constants. The equation for v becomes an equation for u, namely

$$(\beta u + u_\tau) - (\alpha^2 u + 2\alpha u_x + u_{xx}) + (1-k)(\alpha u + u_x) + ku = 0.$$

To get an equation without u or u_x we should set

$$\beta - \alpha^2 + (1-k)\alpha + k = 0, \quad -2\alpha + (1-k) = 0.$$

These equations are solved by

$$\alpha = \frac{1-k}{2}, \quad \beta = -\frac{(k+1)^2}{4}.$$

Thus,

$$u = e^{\frac{1}{2}(k-1)x + \frac{1}{4}(k+1)^2\tau} v(x, \tau)$$

solves the linear heat equation $u_\tau = u_{xx}$.

What good is this? Well, it can be used to give another proof of the integral formula for the value of an option (using the fundamental solution of the linear heat equation). It can also be used to understand the sense in which the value of an option at time $t < T$ is obtained by "smoothing" the payoff. Indeed, the solution of the linear heat equation at time t is obtained by "Gaussian smoothing" of the initial data.

****

Discrete-time hedging. My discussion of this topic follows the beginning of a paper by H. E. Leland, *Option pricing and replication with transaction costs*, J. Finance 40 (1985) 1283-1301 (available online through JSTOR). A thoughtful, quite readable discussion of this topic is the paper by E. Omberg, *On the theory of perfect hedging*, Advances in Futures and Options Research 5 (1991) 1-29 (not available online; I'll put a copy on reserve in the CIMS library in the green box with my name).

Suppose an investment bank sells an option and tries to replicate it dynamically, but the bank trades only at evenly spaced time intervals $j\delta t$. (Now δt is positive, not infinitesimal). The bank follows the standard trading strategy of rebalancing to hold $\phi = \partial V/\partial s$ units of stock each time it trades, where V is the value assigned by the Black-Scholes theory. As we shall see in a moment, this strategy is no longer self-financing – but it is *nearly so*, in a suitable stochastic sense, in the limit $\delta t \to 0$.

People often ask, when examining the derivation of the Black-Scholes PDE by examination of the hedging strategy, "Why do we apply Ito's lemma to $V(s(t), t)$ but not to Δ, even though the choice of Δ also depends on $s(t)$?" The answer, of course, is that the hedge portfolio is held fixed from t to $t + \delta t$. The following discussion – in which δt is small but not infinitesimal – should help clarify this point.

OK, let's return to that investment bank. The question is: how much additional money will the bank have to spend over the life of the option as a result of its discrete-time (rather than continuous-time) hedging? We shall answer this by considering each discrete time interval, then adding up the results.

The bank holds a short position on the option and a long position in the replicating portfolio. The value of its position just after rebalancing at any time $t = j\delta t$ is (by hypothesis)

$$0 = -V(s(t), t) + \phi s(t) + [V(s(t), t) - \phi s(t)] = \text{short option} + \text{stock position} + \text{bond position}$$

with $\phi = \frac{\partial V}{\partial s}(s(t), t)$. The value of its position just before the next rebalancing is

$$-V(s(t+\delta t), t+\delta t) + \phi s(t+\delta t) + [V(s(t), t) - \phi s(t)]e^{r\delta t}.$$

The cost (or benefit) of rebalancing at time $t + \delta t$ is minus the value of the preceding expression. Put differently: it is the difference between the two preceding expressions. So it equals

$$\delta V - \phi \delta s - [V - \phi s](e^{r\delta t} - 1).$$

If we estimate δV by Taylor expansion keeping just the terms one normally keeps in Ito's lemma, we get (remembering that $\phi = \partial V/\partial s$)

$$\frac{\partial V}{\partial s}\delta s + \tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}(\delta s)^2 + \frac{\partial V}{\partial t}\delta t - \frac{\partial V}{\partial s}\delta s - rV\delta t + rs\frac{\partial V}{\partial s}\delta t.$$

Notice that the first and fourth terms cancel. Also notice that the substitution $(\delta s)^2 = \sigma^2 s^2 \delta t$ leads to an expression that vanishes, according to the Black-Scholes equation. Thus, the failure to be self-financing is attributable to two sources: (a) errors in the approximation $(\delta s)^2 \approx \sigma^2 s^2 \delta t$, and (b) higher order terms in the Taylor expansion. Our task is to estimate the associated costs.

Collecting the information obtained so far: if the investment bank re-establishes the "replicating portfolio" demanded by the Black-Scholes analysis at each multiple of δt then it incurs cost

$$\tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}(\delta s)^2 + \frac{\partial V}{\partial t}\delta t - rV\delta t + rs\frac{\partial V}{\partial s}\delta t$$

at each time step, plus an error of magnitude $|\delta t|^{3/2}$ due to higher order terms in the Taylor expansion. Using the Black-Scholes PDE, this cost has the alternative expression

$$\tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}[(\delta s)^2 - \sigma^2 s^2 \delta t] \quad \text{plus an error of order } |\delta t|^{3/2}.$$

It can be shown that when $ds = (\mu + \frac{1}{2}\sigma^2)s\,dt + \sigma s\,dw$,

$$\delta s = \sigma s u \sqrt{\delta t} + (\mu + \tfrac{1}{2}\sigma^2)s\delta t \quad \text{plus an error of order } |\delta t|^{3/2}$$

where u is Gaussian with mean 0 and variance 1 (this is closely related to our our discussion of Ito's lemma). Therefore

$$(\delta s)^2 = \sigma^2 s^2 u^2 \delta t \quad \text{plus an error of order } |\delta t|^{3/2}.$$

Thus neglecting the error terms, the cost of refinancing at any given timestep is

$$\tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma^2 s^2 (u^2 - 1)\delta t$$

where u is Gaussian with mean value 0 and variance 1. This expression is obviously random; its expected value is 0 and its standard deviation is of order δt. Moreover the contributions associated with different time intervals are independent. Notice that the distribution of refinancing costs is *not* Gaussian, since it is proportional to $u^2 - 1$ not u.

Pulling this together: since the expected value of $u^2 - 1$ is zero, the *expected cost* of refinancing at any given timestep is at most of order $|\delta t|^{3/2}$, due entirely to the "error terms." However the *actual cost* (or benefit) of refinancing is larger, a random variable of order δt. But the picture changes when we consider many time intervals. Over $n = T/\delta t$ intervals, the terms $\frac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma^2 s^2(u^2 - 1)\delta t$ accumulate to a sum

$$\sum_{j=1}^{n} \tfrac{1}{2}\sigma^2 s^2(t_j)\frac{\partial^2 V}{\partial s^2}(s(t_j), t_j)(u_j^2 - 1)\delta t$$

with mean 0 and standard deviation of order $\sqrt{n\delta t^2} = \sqrt{T\delta t}$; the sum is still random, but it's small, statistically speaking, if δt is close to zero, by a sort of law-of-large-numbers. (Notice the resemblance of this argument to our explanation of Ito's lemma. That's no accident: we are in essence deriving Ito's formula all over again.) We've been ignoring the error terms – but they cause no trouble, because they too accumulate to a term of order $\sqrt{\delta t}$, because $n(\delta t)^{3/2} = T\sqrt{\delta t}$.

Final conclusion: the errors of refinancing tend to self-cancel, by a sort of law-of-large-numbers, since their mean value is 0. The net effect, when δt is small, is random but small — in the sense that its mean and standard deviation are of order $\sqrt{\delta t}$.

We have argued that the cost of refinancing tends to zero as $\delta t \to 0$. A recent article by A. Lo, D. Bertsimas, and L. Kogan goes further, examining the statistical distribution of refinancing costs when δt is small. (The relation between their work and the preceding discussion is like the relation between the central limit theorem and the law of large numbers.) The reference is: J. Financial Economics 55 (2000) 173-204 (available online through sciencedirect.com).

Derivative Securities – Section 8 – Fall 2004
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences.

American options; continuous dividend yield. These notes show how the theory developed so far applies, with minor modifications, to (a) American options, and (b) options on stock indices or foreign currency. We continue to assume that the risk-free interest rate is constant – a reasonable approximation for options with relatively short maturities, on assets other than interest-based instruments.

****

American options. Up to now we have discussed only European options, which (by definition) can only be exercised at a specified maturity date T. American options are different in that they permit early exercise: the holder of an American option can exercise it at any time up to the maturity T. Of the options actually traded in the market, the majority are American rather than European.

Clearly an American option is at least as valuable as the analogous European option, since the holder has the option to keep it to maturity.

Fact: An American call written on a stock that earns no dividend has the same value as a European call; early exercise is never optimal. To see why, suppose the strike price is K and consider the value of the American option "now," at some time $t < T$. Exercising the option now achieves a value at time t of $s_t - K$. Holding the option to maturity achieves a value at time t equal to that of a European call, $c[s_t, K, T - t]$. Without using the Black-Scholes formula (thus without assuming lognormal stock dynamics) we know the value of a European call is at least that of a forward with the same strike and maturity. Thus holding the option to maturity achieves a value at time t of at least $s_t - e^{-r(T-t)}K$. If $r > 0$ this is larger than $s_t - K$. So early exercise is suboptimal, as asserted.

The preceding is in some sense a fluke. When the underlying asset pays a dividend early exercise of a call can be optimal. But the simplest example where early exercise occurs is that of a put on a non-dividend-paying stock:

Fact: An American put written on a stock that earns no dividend can have a value greater than that of the associated European put; early exercise can be optimal. To see why, consider once again the value of the American option "now," at some time $t < T$. Exercising the option now achieves value $K - s_t$. Holding it to maturity achieves a value at time t equal to that of a European put, $p[s_t, K, T - t]$. Assuming lognormal stock price dynamics, p is given by the Black-Scholes formula, and its graph as a function of spot price s_t is shown in the figure.

The important point is that $p[s_t, K, T - t]$ is strictly less than $K - s_t$ when $s_t \ll K$. This is immediate from the Black-Scholes formula, since $p = Ke^{-r(T-t)}N(-d_2) - s_t N(-d_1) \approx Ke^{-r(T-t)} - s_t$ when $s_t \ll K$, since $d_1 \to -\infty$ and $d_2 \to -\infty$ as $s_t/K \to 0$. Briefly: if $s_t \ll K$ then the put is deep in the money, and the (risk-neutral) probability of it being out

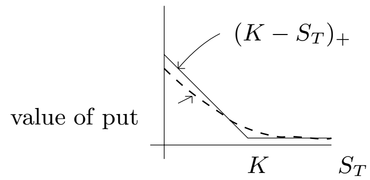


Figure 1: The value of a European put lies below the payoff when $s \ll K$.

of the money at time T is vanishingly small; therefore the value of the put is almost the same as the value of a short forward. In such a situation we are better off exercising the option at time t than holding it to maturity. (This does not show that exercise at time t is optimal, but it does show holding the option to maturity is not optimal.)

For European options we have three different (but related) valuation techniques: (a) working backward through the binomial tree; (b) evaluating the discounted expected payoff (using the risk-neutral version of the price process); and (c) solving the Black-Scholes PDE. Each of the three viewpoints can be extended to American options. We assume for simplicity that the underlying asset pays no dividends.

Valuation using a binomial tree. This is perhaps the simplest approach, conceptually and numerically. We can use the same recombinant binomial tree as for a European option. (Remember: pricing is done using the risk-neutral process. If the underlying asset is lognormal with volatility σ then a convenient choice of the parameters defining the tree is $u = \exp[(r - \frac{1}{2}\sigma^2)\delta t + \sigma\sqrt{\delta t}]$, $d = \exp[(r - \frac{1}{2}\sigma^2)\delta t - \sigma\sqrt{\delta t}]$, where r is the risk-free rate). But since early exercise is permitted, we must ask at each node: is the option worth more "alive" or "dead"? If the option is worth more dead, then it should be exercised (by its holder) whenever the price arrives at that node. For example, consider the pricing of an American option with payoff $f(s)$ using a two-period recombinant tree:

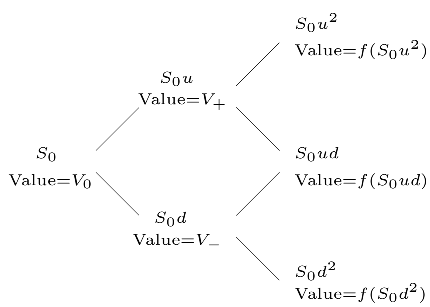


Figure 2: Valuation of an American option using a binomial tree.

When the stock price is s_0u the option is worth

$$f(s_0u) \text{ dead, and } e^{-r\delta t}[qf(s_0u^2) + (1-q)f(s_0ud)] \text{ alive.}$$

Allowing for both possibilities the value of the option at $s_0 u$ is

$$V_+ = \max\{f(s_0 u), e^{-r\delta t}[qf(s_0 u^2) + (1-q)f(s_0 ud)]\}.$$

Similarly, when the stock price is $s_0 d$ the value is

$$V_- = \max\{f(s_0 d), e^{-r\delta t}[qf(s_0 ud) + (1-q)f(s_0 d^2)]\}.$$

The value at the initial time is obtained by repeating the process:

$$V_0 = \max\{f(s_0), e^{-r\delta t}[qV_+ + (1-q)V_-]\}.$$

Our example has only two time periods, but a binomial tree of any size is handled similarly.

Valuation using the discounted expected payoff. For a European option, we saw that the value assigned by the binomial tree was expressible in the form $e^{-rT}E_{\mathrm{RN}}[f(s(T))]$. A similar calculation applies to the American option – however $f(s(T))$ must be replaced by the value realized *at exercise*: the value of the option is $E_{\mathrm{RN}}\left[e^{-r\tau}f(s(\tau))\right]$ where τ is the exercise time. Once we've worked backward through the tree we know how to determine τ – for each realization of the risk-neutral process, it's the first time that realization reaches a node of the tree associated with early exercise (or T, if that realization does not reach an "early-exercise" node).

Actually, this viewpoint can also be used, at least conceptually, to *determine* the early-exercise criterion, without working backward through the tree. In fact,

$$\text{Value} = \max_{\text{exercise rules}} E_{\mathrm{RN}}\left[e^{-r\tau}f(s(\tau))\right].$$

In other words the exercise rule selected by backsolving the binomial tree is the one that maximizes the discounted expected payoff. An honest proof of this fact is not trivial – mainly because it requires formalization of what one means by an "exercise rule." But heuristically: any exercise rule determines a hedging strategy, i.e. a synthetic option that is available in the marketplace. So the max over exercise rules gives a lower bound for the value of the option. Our strategy of working backward through the tree gives an upper bound. The two bounds agree since the value obtained by working backward through the tree is associated with a special exercise rule.

Valuation using a PDE. (This material is not in Jarrow-Turnbull or Hull; you can find a brief summary in Wilmott or Wilmott-Howison-Dewynne.) For a European option the continuous-time limit of working backward through the tree amounts to solving the Black-Scholes PDE for $t < T$, with final data $f(s)$ at $t = T$. There is an analogous statement for an American option, however the PDE is replaced by a *free boundary problem*:

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma^2 s^2 + rs\frac{\partial V}{\partial s} - rV \le 0,$$

$$V(s,t) \ge f(s),$$

and

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma^2 s^2 + rs\frac{\partial V}{\partial s} - rV = 0 \quad \text{or} \quad V(s,t) = f(s).$$

The logic behind the first inequality is this: in our derivation of the Black-Scholes PDE, the crucial juncture was when we saw that the choice $\phi = \partial V/\partial s$ made $d(V - \phi s)$ deterministic:

$$d(V - \phi s) = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma^2 s^2\right) dt.$$

We concluded, by the principle of no arbitrage, that this must equal $r(V - \phi s)dt$. But that arbitrage argument assumed that you continued to hold the option. In the present context, where early exercise is permitted, the absence of arbitrage gives a weaker conclusion: the deterministic portfolio $(V - \phi s)$ can grow *no faster* than the risk-free rate. Thus

$$\left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma^2 s^2\right) \leq r\left(V - \frac{\partial V}{\partial s}s\right);$$

this is our first inequality. The logic behind the second inequality is obvious: the value is no smaller than can be realized by immediate exercise. The third relation simply says that one of the first two relations always holds – because for any given (s,t) the optimal strategy involves either holding the option a little longer (in which case the Black-Scholes equation applies) or exercising it immediately.

We call this a free-boundary problem because the value is still governed by the Black-Scholes PDE in *some* region of the (s,t) plane – the region where immediate exercise isn't optimal – however this region isn't given as data but must be found as part of the problem. Schematically:

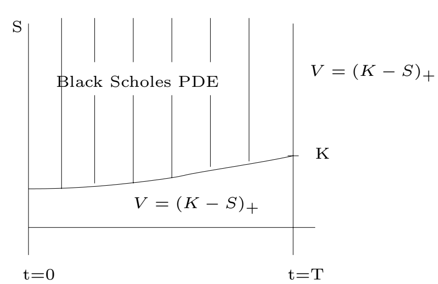


Figure 3: Schematic of the free boundary problem whose solution values an American put.

One can show that V and $\Delta = \partial V/\partial s$ are both continuous across the free boundary. Of course, on the "exercise" side of the boundary $V = f(s)$ and $\partial V/\partial s = f'(s)$ are known, giving two boundary conditions. If the domain of the PDE were known then just one boundary condition would be permitted; but the domain isn't known, and the extra boundary condition serves to fix the free boundary.

What if the underlying asset pays dividends at discrete times? The distribution of dividends (of predetermined sizes at predetermined times) is easily handled by minor modification of the techniques explained above. For an American call, exercise can only be optimal just before the distribution of a dividend. So we can value the option by working backward in time, using Black-Scholes or a (European) binomial tree to pass from one dividend date to the next, but taking the maximum of the value (a) if exercised, and (b) if not exercised, at each dividend date. For an American put we must still check for possible exercise at each time if using a tree – or we must still solve a free boundary problem between exercise dates if using the PDE. A point to watch out for: at the moment when a dividend is declared, the value of the underlying asset drops discontinuously by an amount equal to the dividend.

****

Black-Scholes analysis with constant dividend yield. (This topic is treated in Jarrow-Turnbull chapters 11 and 12, and in Hull chapter 12.) We return to the case of a European option, and consider what happens when the underlying asset is a foreign currency or a stock index. These two cases are essentially identical (you learned this in Homework 1, problem 1): a foreign currency earns interest, while a stock index pays dividends (which we may choose to reinvest). Both cases are described, to a reasonable approximation, by supposing that the risky asset *pays dividends at a fixed rate* D. In other words, if you hold ϕ units of the risky asset now, and you do no trading, then you'll hold ϕe^{Dt} units after time t.

The bottom line is simple: options on such an asset can be priced using the Black-Scholes framework, however the risk-neutral process is different: it has drift $r - D - \frac{1}{2}\sigma^2$ rather than $r - \frac{1}{2}\sigma^2$. Less ambiguous, perhaps: the risk-neutral process solves the stochastic differential equation $ds = (r - D)sdt + \sigma sdw$ rather than $ds = rsdt + \sigma sdw$. We shall explain this assertion in two different ways: (a) using binomial trees, and (b) using the Black-Scholes PDE.

Using binomial trees. Suppose the subjective price process is lognormal, say $\log s(t) = \log s(0) + \mu t + \sigma w(t)$. (For foreign currency $s(t)$ is the exchange rate, in dollars per unit foreign currency; for a stock index $s(t)$ is the price *without* reinvestment of dividends). The value of the option at maturity T is a function of $s(T)$. Our strategy is the same as used earlier in the semester: choose a binomial tree that mimics the subjective price process; then find the appropriate risk-neutral probabilities; then value the option by working backward in the tree.

The constant dividend yield D changes only the middle part of this program – the formula for the risk-neutral probabilities. To see how, we need only consider a single time period.

Consider the portfolio consisting at time 0 of ϕ units of risky asset and ψ dollars worth of risk-free asset. Its initial value is $\phi s_0 + \psi$, and its value at the next time period is either $\phi s_0 u e^{D\delta t} + \psi e^{r\delta t}$ or $\phi s_0 d e^{D\delta t} + \psi e^{r\delta t}$, depending whether the risky asset went up or down. The situation is identical with that of a non-dividend-paying binomial market with

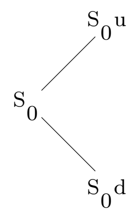


Figure 4: Standard one-period binomial tree.

$s_{\text{up}} = s_0 u e^{D\delta t}$ and $s_{\text{down}} = s_0 d e^{D\delta t}$. So by our previous analysis of (non-dividend-paying) binomial trees, the risk-free probability associated with the "up" state is

$$q = \frac{e^{r\delta t} - de^{D\delta t}}{ue^{D\delta t} - de^{D\delta t}} = \frac{e^{(r-D)\delta t} - d}{u - d},$$

and the value of the option at time 0 is

$$f_0 = e^{-r\delta t}[qf_{\text{up}} + (1-q)f_{\text{down}}]$$

where f_{up} and f_{down} are its values in the "up" and "down" states respectively. The extension to a multiperiod tree is obvious: we still work backward in the tree, and the value of the option is still the discounted expected payoff; the only difference is that we must use $q = (e^{(r-D)\delta t} - d)/(u-d)$ when we work backward in the tree. In the continuous-time limit: if the underlying asset has continuous dividend yield D then the value of a European option with payoff $f(s(T))$ is

$$e^{-rT}E[f(s_0e^X)] \quad \text{where } X \text{ is Gaussian with mean } (r - D - \tfrac{1}{2}\sigma^2)T \text{ and variance } \sigma^2 T.$$

The preceding argument shows, in essence, that the risk-neutral process solves the stochastic differential equation $ds = (r-D)sdt + \sigma sdw$. Here's an easy way to remember this. When there is no dividend, the risk-neutral process is $ds = rsdt + \sigma sdw$. It has has expected return r, in the sense that $(d/dt)E_{\text{RN}}[ds] = rs(t)$. When the asset pays continuous dividend yield D, the risk-neutral process has expected return $r - D$ *without counting the dividend yield*, so it has expected return r *if we include the dividend yield.*

An alternative analysis using the stochastic differential equation. We can obtain the same result by reconsidering the continuous-time hedging argument that led to the Black-Scholes PDE. Our lognormal hypothesis is equivalent to the stochastic differential equation

$$ds = \sigma sdx + (\mu + \tfrac{1}{2}\sigma^2)sdt, \quad s(0) = s_0.$$

Let $V(s(t), t)$ be the value of the option at stock price s and time t. Arguing as in Section 7, we consider a portfolio consisting of a short position in the option and a long position in the hedge portfolio (which consists of $\phi = (\partial V/\partial s)(s(t), t)$ units of stock, and $V - \phi s$ dollars risk-free). Its value at time t is

$$-V + \phi s + (V - \phi s) = 0$$

where $V = V(s(t), t)$ and $s = s(t)$; its value at time $t + \delta t$ is

$$-(V + \delta V) + \phi(s + \delta s + Ds\delta t) + (V - \phi s)(1 + r\delta t)$$

using the approximations $e^{D\delta t} \approx 1 + D\delta t$ and $e^{r\delta t} \approx 1 + r\delta t$. The only new term is the one associated with the dividends. To get to the PDE we must (a) use Ito's formula to estimate δV, then (b) set the value at time $t + \delta t$ to 0. Writing ds rather than δs, as we usually do for the Ito calculus, this gives

$$(V_t + V_s ds + \tfrac{1}{2} V_{ss} \sigma^2 s^2 dt) - (\phi ds + \phi D s dt) - (V - \phi s) r dt = 0$$

which becomes, after the subsitution $\phi = V_s$,

$$V_t + (r - D) s V_s + \tfrac{1}{2} \sigma^2 V_{ss} - rV = 0.$$

This is the Black-Scholes PDE for options on an asset with continuous dividend yield D. It is of course consistent with the one we obtained using binomial trees: the solution of this PDE (with $V(s, T) = f(s)$ at the final time T) satisfies

$$V(s(t), t) = e^{-r(T-t)} E_{\text{RN}}[f(s(T)]$$

if we understand the risk-neutral process to be the one solving $ds = (r - D) s dt + \sigma s dw$. The proof is essentially the same as the argument we gave at the end of the Section 6 notes.

A shortcut to deriving explicit solutions. It is not necessary to derive new solution formulas. Instead we can use our existing solution formulas (derived for a non-dividend-paying asset) together with the following simple rule: *To value an option on an asset with continuous dividend yield D and maturity T, reduce the spot price by* $e^{-D(T-t)}$ *then apply the no-dividend-yield formula at this reduced spot price.* Thus, for example, the value at time 0 of a call with strike K is

$$s_0 e^{-DT} N(d_1) - K e^{-rT} N(d_2)$$

where

$$d_1 = \frac{\log(s_0 e^{-DT}/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} = \frac{\log(s_0/K) + (r - D + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

and

$$d_2 = \frac{\log(s_0 e^{-DT}/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} = \frac{\log(s_0/K) + (r - D - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}.$$

Similarly the value of a put is $K e^{-rT} N(-d_2) - s_0 e^{-DT} N(-d_1)$.

The justification of this rule is easy. We may consider just $t = 0$. Starting from the valuation formula as the discounted risk-neutral expectation, we observe that if X Gaussian with mean $(r - D - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$ then

$$s_0 e^X = s_0 e^{(r - D - \frac{1}{2}\sigma^2)T + \sigma w(T)} = (s_0 e^{-DT}) e^Y$$

where Y has mean $(r - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$. Thus the value of the option is

$$e^{-rT}E[f(s_0 e^X)] = e^{-rT}E[(f(s_0 e^{-DT}e^Y)]$$

and the right hand side is the "ordinary" (non-dividend-paying) Black-Scholes formula evaluated at the reduced spot price.

Of course we can also give a justification based on the Black-Scholes PDE. Let V solve the PDE derived above, and consider the change of variables

$$\bar{s} = se^{-D(T-t)}, \quad \bar{V}(\bar{s}, t) = V(s, t).$$

One verifies by an elementary calculation that $\bar{V}$ solves the no-dividend Black-Scholes equation. Moreover its final-time data is still the option payoff, since $\bar{s} = s$ at $t = T$. Since $\bar{s}(0) = e^{-DT}s(0)$, we conclude that

$$\text{Option value} = V(s(0), 0) = \bar{V}(e^{-DT}s(0), 0),$$

and the right hand side is again the "ordinary" (non-dividend-paying) Black-Scholes formula evaluated at the reduced spot price.

Black's formula. It's confusing to have so many different formulas. Fisher Black observed that the situation is simpler if we focus on the *forward price* rather than the spot price. Recall that if our underlying asset has spot price s_0 at time 0, then its *forward price* for delivery at time T is

$$F_0 = s_0 e^{(r-D)T}.$$

The value of an option with payoff $f(s(T)$ is evidently

$$e^{-rT}E[f(F_0 e^Z)] \quad \text{where } Z \text{ has mean } -\tfrac{1}{2}\sigma^2 T \text{ and variance } \sigma^2 T.$$

The advantage of this formula is that neither r nor D enters explicitly, except in the discount factor e^{-rT}. Instead, they are taken into account by use of the forward price F_0. Viewed this way, our formulas for the value of a call and a put become:

$$\text{call} = e^{-rT}[F_0 N(d_1) - KN(d_2)], \quad \text{put} = e^{-rT}[KN(-d_2) - F_0 N(-d_1)]$$

with

$$d_1 = \frac{\log(F_0/K) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \quad d_2 = \frac{\log(F_0/K) - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}$$

A final note. The financial intuition behind these results can be a bit confusing. It's natural to ask: why can't we just assume the dividends are reinvested, effectively changing the stochastic differential equation from $ds = (\mu + \frac{1}{2}\sigma^2)sdt + \sigma s dw$ to $ds = (\mu + \frac{1}{2}\sigma^2 + D)sdt + \sigma s dw$, then use the Black-Scholes framework (which is anyway insensitive to the drift)? The answer is this: we must be careful to recognize that the dividends are paid to those who hold the risky asset – but not to those who simply hold options on it.

Derivative Securities – Section 9 – Fall 2004
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences.

Futures, and options on futures. Martingales and their role in option pricing. A brief introduction to stochastic interest rates. But first some comments about the final exam:

- The exam will be Monday December 20, in the normal class hour (7:10-9pm).
- The exam will be closed-book, however you may bring two pages of notes (8.5×11, both sides, any font).
- The exam questions will focus on fundamental ideas and examples covered in the lectures and homeworks.

****

Futures. We discussed *forward contracts* in detail in Section 1. A forward contract with delivery price K has payoff $s(T) - K$. We saw in Section 1 that value of this payoff is determined by arbitrage. No model of asset dynamics is needed, the market need not be complete, and no multiperiod trading is involved. In a constant interest rate environment the value of the forward at time 0 is $s_0 - Ke^{-rT}$. The hypothesis of constant interest rate is not really needed: the cash-and-carry argument involves no trading, so all that matters is the value at time 0 of a dollar received at time T. Denoting this by $B(0,T)$, the value of the forward contract is $s_0 - KB(0,T)$.

We also discussed the *forward price* in Section 1. It is the special delivery price for which the present value of the forward contract is 0. By elementary arithmetic, in a constant interest rate environment the forward price at time 0 is $F_0 = s_0 e^{rT}$. Similarly the forward price at time t is $F_t = s_t e^{r(T-t)}$. In a variable interest rate environment the forward price is $F_t = s_t / B(t,T)$.

A *futures contract* is similar to a forward – but not quite the same. Briefly, the investor who holds a long futures contract:

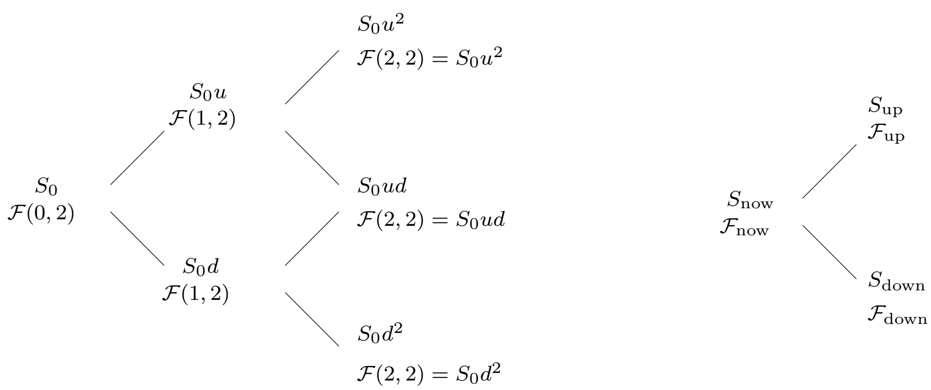
(a) pays nothing to acquire the contract – here it resembles a forward contract, with the forward price as delivery price;

(b) pays or receives funds daily as the value of the underlying asset varies – here it is quite different from a forward, where no cash flow occurs till maturity;

(c) buys the underlying asset at its market value s_T when the contract matures.

Thus the essential difference between a forward (with delivery price = forward price) and a future is that the settlement of the forward occurs entirely at maturity, while the settlement payoff of the future takes place daily. There are some other differences: futures are standardized, and they are bought and sold by financial institutions – thus they are liquid, traded instruments, which forwards are not. They are also often more liquid than the underlying asset itself. Thus while we often think of replicating an option by a (time-dependent) portfolio of the underlying asset and a cash account, in practice it is often better to use a (time-dependent) portfolio of futures on the underlying asset and a cash account.

Let us explain further how futures contracts work. Practical details aside (see Hull or Jarrow-Turnbull for those), this means explaining the *futures price* and the role it plays in *settlement*. We shall discuss this in the context of a binomial-tree market, following Section 5.7 of Jarrow-Turnbull. It's sufficient to consider the two-period tree in the figure: once we understand it, the multiperiod extension will be obvious. We write $\mathcal{F}(i,j)$ for the futures price at time i of a futures contract which matures at time j. Our goal is to understand a specific futures contract – maturing in this example at time 2 – so the maturity $j = 2$ will be fixed throughout our discussion; we shall determine its futures price $\mathcal{F}(i,2)$ by working backward through the tree, starting at $i = 2$ and ending at $i = 0$.



As usual, the essential calculation involves a single-period binomial model branch. Suppose the stock price is s_{now} and it can go up to s_{up} or down to s_{down}. Suppose further that the futures price is already known to be $\mathcal{F}_{\text{up}}$ and $\mathcal{F}_{\text{down}}$ in the up and down states. To determine the futures price now, $\mathcal{F}_{\text{now}}$, let's look for a portfolio consisting of ϕ units of stock and ψ dollars risk-free that replicates the futures contract. The value of the futures contract now is 0, since one pays nothing to acquire a futures contract. Its value in the up state is $\mathcal{F}_{\text{up}} - \mathcal{F}_{\text{now}}$, and its value in the down state is $\mathcal{F}_{\text{down}} - \mathcal{F}_{\text{now}}$, since the settlement procedure involves a cash payment of $\mathcal{F}_{\text{new}} - \mathcal{F}_{\text{old}}$ at each time period. So the replicating portfolio must satisfy

$$\begin{aligned}
\phi s_{\text{now}} + \psi &= 0 \\
\phi s_{\text{up}} + \psi e^{r\delta t} &= \mathcal{F}_{\text{up}} - \mathcal{F}_{\text{now}} \\
\phi s_{\text{down}} + \psi e^{r\delta t} &= \mathcal{F}_{\text{down}} - \mathcal{F}_{\text{now}}.
\end{aligned}$$

We know, from our treatment of binomial markets, that the last two equations alone give

$$\begin{aligned}
\phi s_{\text{now}} + \psi &= e^{-r\delta t}[q(\mathcal{F}_{\text{up}} - \mathcal{F}_{\text{now}}) + (1-q)(\mathcal{F}_{\text{down}} - \mathcal{F}_{\text{now}})] \\
&= e^{-r\delta t}[q\mathcal{F}_{\text{up}} + (1-q)\mathcal{F}_{\text{down}} - \mathcal{F}_{\text{now}}]
\end{aligned}$$

where

$$q = \frac{e^{r\delta t} s_{\text{now}} - s_{\text{down}}}{s_{\text{up}} - s_{\text{down}}}$$

is the risk-neutral probability of the up state. Now the first equation (the condition that the futures contract have value 0 now) determines $\mathcal{F}_{\text{now}}$:

$$0 = \phi s_{\text{now}} + \psi = e^{-r\delta t}[q\mathcal{F}_{\text{up}} + (1-q)\mathcal{F}_{\text{down}} - \mathcal{F}_{\text{now}})]$$

whence

$$\mathcal{F}_{\text{now}} = q\mathcal{F}_{\text{up}} + (1-q)\mathcal{F}_{\text{down}}.$$

This formula can of course be iterated over multiple time periods to give

$$\text{futures price at time } t = E_{\text{RN}}\left[\text{futures price at time } t'\right]$$

for any $t' > t$. At the time when the futures contract matures, its price is (by definition) the spot price of the underlying asset.

For a standard muliplicative binomial tree in a constant interest rate environment the risk-neutral probability is $q = (e^{r\delta t} - d)/(u - d)$, the same at every branch. Thus for the two-period tree shown in the preceding figure

$$\mathcal{F}(1,2) = \begin{cases} qs_0u^2 + (1-q)s_0ud & \text{if the price is } s_0u \\ qs_0ud + (1-q)s_0d^2 & \text{if the price is } s_0d \end{cases}$$

and

$$\mathcal{F}(0,2) = q^2 s_0 u^2 + 2q(1-q)s_0ud + (1-q)^2 s_0 d^2.$$

If the risk-free rate is constant then the futures price is equal to the forward price. This is true for any market (see Hull or Jarrow/Turnbull for a proof). In the special case of a binomial market it follows easily from the results just derived, together with the crucial feature of the risk-neutral probabilities that

$$s_0 = e^{-rT} E_{\text{RN}}\left[s(T)\right]$$

(this was clear from Section 3, and explicit at the beginning of Section 4). Using these facts: the futures price at time 0 for contracts maturing at time T is

$$\mathcal{F}_0 = E_{\text{RN}}\left[s(T)\right] = e^{rT} s_0$$

which is precisely the forward price.

Why have we fussed so much over futures prices, if in the end they are simply the same as forward prices? The answer is two-fold: (a) to clarify the essential nature of a futures contract (in particular its periodic settlement); (b) to prepare for modeling of interest-based instruments, and non-constant *stochastic* risk-free rates.

Notice the similarity – and the difference – between the results

$$\text{futures price at time } 0 = E_{\text{RN}}\left[\text{futures price at } t\right]$$

for a futures contract maturing at time $T > t$, and

$$\text{value at time } 0 = e^{-rt} E_{\text{RN}} [\text{value at time } t]$$

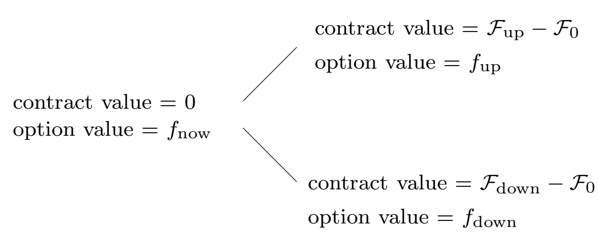
for a tradeable asset such as the underlying stock or any option. People sometimes get confused, and ask why there is no factor of e^{-rt} in the formula for the futures price. The answer is that these two formulas are intrinsically different. The futures price is not the value of a tradeable asset.

****

Options on futures. Now consider a European option with maturity T, whose payoff is a function of a futures price $\mathcal{F}(T, T')$ for some $T' > T$. A typical example – a call with strike K – would give its holder the right to receive, at time T, one futures contract with maturity T', plus $\mathcal{F}(T, T') - K$ in cash. This instrument has payoff $(\mathcal{F}(T, T') - K)_+$, since the futures contract itself has value 0.

Options on futures are attractive because they involve only delivery of futures contracts and cash upon exercise – nobody has to purchase or sell the underlying asset. This is convenient, since the futures contracts are usually more liquid than the underlying asset. Options are traded on many kinds of futures, but options on interest rate futures have particularly active markets. We'll nevertheless continue to assume that the risk-free rate is constant for the moment, returning to the case of options on interest rate futures in a future lecture.

If the payoff of an option is determined by a futures price, then it's natural to value the option using the tree of futures prices rather than the tree of stock prices. As usual the value of the option is determined by working backward in the tree, so the heart of the matter is the handling of a single-period binomial market. Suppose the futures price now is $\mathcal{F}_{\text{now}}$ and at the next period the futures prices are $\mathcal{F}_{\text{up}}$ and $\mathcal{F}_{\text{down}}$. Assume further that the option's value at the next period is already known to be f_{up} in the up state, and f_{down} in the down state.



Rather than replicate the option using the underlying asset, it's convenient to replicate it using futures contracts. The value of the futures contract upon entry is 0, and its value at the next time period is $\mathcal{F}_{\text{up}} - \mathcal{F}_{\text{now}}$ or $\mathcal{F}_{\text{down}} - \mathcal{F}_{\text{now}}$ due the settlement procedure, so the relevant price tree is as shown in the figure.

Our results on binomial trees determine the option price as

$$f_{\text{now}} = e^{-r\delta t} [p f_{\text{up}} + (1 - p) f_{\text{down}}]$$

where p is the relevant risk-neutral probability, determined by

$$0 = e^{-r\delta t}\left[p(\mathcal{F}_{\text{up}} - \mathcal{F}_{\text{now}}) + (1-p)(\mathcal{F}_{\text{down}} - \mathcal{F}_{\text{now}})\right].$$

The last relation amounts to $\mathcal{F}_{\text{now}} = p\mathcal{F}_{\text{up}} + (1-p)\mathcal{F}_{\text{down}}$, so we recogize that $p = q$ is the *same* risk-neutral probability we used to determine the futures prices. However let us forget this fact for a moment, and consider pricing the option using *only* the tree of forward prices. Then the convenient definition of p is

$$p = \frac{\mathcal{F}_{\text{now}} - \mathcal{F}_{\text{down}}}{\mathcal{F}_{\text{up}} - \mathcal{F}_{\text{down}}},$$

and the option price is determined by

$$f_{\text{now}} = e^{-r\delta t}\left[pf_{\text{up}} + (1-p)f_{\text{down}}\right].$$

Working backward in the tree, we obtain (if the interest rate is constant) that the option value is its discounted expected payoff:

$$\text{option value at time } 0 = e^{-rT}E_{\text{RN}}\left[\text{payoff at time } T\right].$$

Let's compare this result to the one obtained long ago for pricing ordinary options on lognormal assets. There the option price was the discounted risk-neutral expected value, using risk-neutral measure $\pi = \frac{e^{r\delta t}s_{\text{now}} - s_{\text{down}}}{s_{\text{up}} - s_{\text{down}}}$. Here the option price is the discounted risk-neutral value, using risk-neutral measure $p = \frac{\mathcal{F}_{\text{now}} - \mathcal{F}_{\text{down}}}{\mathcal{F}_{\text{up}} - \mathcal{F}_{\text{down}}}$. We see that the two situations are parallel, except that we must set $r = 0$ in the definition of the risk-neutral probability. Passing as usual to the continuous-time limit, we conclude that *an option on a futures contract can be valued using Black's formula*: if the futures price $\mathcal{F}_t$ is lognormal with volatility σ then an option with maturity T and payoff $f(\mathcal{F}_T)$ has value

$$e^{-rT}E[f(\mathcal{F}_0 e^Z)] \quad \text{where } Z \text{ has mean } -\tfrac{1}{2}\sigma^2 T \text{ and variance } \sigma^2 T.$$

In particular, for a call (payoff $(\mathcal{F}_T - K)_+$) or a put (payoff $(K - \mathcal{F}_T)_+$) we get the value

$$\text{call} = e^{-rT}[F_0N(d_1) - KN(d_2)], \quad \text{put} = e^{-rT}[KN(-d_2) - F_0N(-d_1)]$$

with

$$d_1 = \frac{\log(F_0/K) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \quad d_2 = \frac{\log(F_0/K) - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}.$$

Comparing these results with those of Section 8, we see that pricing an option on a futures price is analogous to pricing an option on an asset with continuous dividend yield equal to the risk-free rate.

We avoided making reference to the original stock price tree in the preceding argument. In particular we avoided using the fact that when interest rates are constant, forward and future prices are the same: $\mathcal{F}_t = F_t = s_t e^{r(T-t)}$. But we could alternatively have based our analysis on this fact. Indeed, it easily implies the crucial relation

$$\frac{\mathcal{F}_{\text{now}} - \mathcal{F}_{\text{down}}}{\mathcal{F}_{\text{up}} - \mathcal{F}_{\text{down}}} = \frac{e^{r\delta t}s_{\text{now}} - s_{\text{down}}}{s_{\text{up}} - s_{\text{down}}}$$

from which all else follows. (In our prior notation this relation says that $p = q$.)

There is of course an alternative route to Black's formula using stochastic PDE's. Let us stop distinguishing between the forward price F and the futures price $\mathcal{F}$, since this discussion is restricted to the constant-interest-rate environment where they are the same. The critical assertion is that if $V(F(t), t)$ is the value of the option as a function of futures price, then V solves $V_t + \frac{1}{2}\sigma^2 F^2 V_{FF} - rV = 0$ with final value $f(F)$, where f is the payoff. (This is equivalent to the assertion that $V(F_0, 0) = e^{-rT} E\left[f(F(T))\right]$ where $dF = \sigma F dw$, and its solution is given by Black's formula.) Arguing as in the last section, consider a portfolio consisting of a short position in the option and a long position in the hedge portfolio (which consists of $\phi = (\partial V/\partial F)(F(t), t)$ futures and $V(F(t), t)$ dollars risk-free). Its value at time t is

$$-V + \phi \cdot 0 + V = 0$$

since a futures contract costs nothing at the time of acquisition. Its value at time $t + \delta t$ is

$$-(V + \delta V) + \phi \delta F + V(1 + r\delta t);$$

the middle term $\phi \delta F$ represents the cost of settlement at time $t + \delta t$. To get the PDE we must (a) use Ito's formula to evaluate δV, then (b) set the value at time $t + \delta t$ to 0. We suppose the futures price solves $dF = \mu F dt + \sigma F dw$ for some μ. Then steps (a) and (b) lead to

$$(V_t dt + V_F dF + \tfrac{1}{2} V_{FF} \sigma^2 F^2 dt) - V_F dF - rV dt.$$

The dF terms cancel, and what remains is the desired relation

$$V_t + \tfrac{1}{2}\sigma^2 F^2 V_{FF} - rV = 0.$$

****

Martingales. The basic prescription for working backward in a binomial tree is this: if f is the value of a tradeable security (such as an option) then

$$f_{\text{now}} = e^{-r\delta t}[q f_{\text{up}} + (1 - q) f_{\text{down}}] = e^{-r\delta t} E_{\text{RN}}[f_{\text{next}}]$$

and if $\mathcal{F}$ is the futures price of a tradeable security then

$$\mathcal{F}_{\text{now}} = [q \mathcal{F}_{\text{up}} + (1 - q) \mathcal{F}_{\text{down}}] = E_{\text{RN}}[\mathcal{F}_{\text{next}}],$$

where q is the risk-neutral probability, defined by

$$s_{\text{now}} = e^{-r\delta t}[q s_{\text{up}} + (1 - q) s_{\text{down}}] = e^{-r\delta t} E_{\text{RN}}[s_{\text{next}}].$$

When the risk-free rate is constant the factors of $e^{-r\delta t}$ don't bother us – we just bring them out front. When the risk-free rate is stochastic, however, we must handle them differently. To this end it is convenient to introduce a *money market account* which earns interest at the risk-free rate. Let $A(t)$ be its balance, with $A(0) = 1$. In the constant interest rate setting obviously $A(t) = e^{rt}$; in the variable interest rate setting we still have $A(t + \delta t) = e^{r\delta t} A(t)$, however r might vary from time to time, and even (if interest rates are stochastic) from one binomial subtree to another. With this this convention, the prescription for determining the price of a tradeable security becomes

$$f_{\text{now}}/A_{\text{now}} = E_{\text{RN}}[f_{\text{next}}/A_{\text{next}}]$$

since $A_{\text{now}}/A_{\text{next}} = e^{-r\delta t}$ where r is the risk-free rate. (This relation is valid even if the risk-free rate varies from one subtree to the next). Working backward in the tree, this relation generalizes to one relating the option value at any pair of times $0 \le t < t' \le T$:

$$f(t)/A(t) = E_{\text{RN}}[f(t')/A(t')].$$

Here, as usual, the risk-neutral expectation weights each state at time t' by the probability of reaching it via a coin-flipping process starting from time t — with independent, biased coins at each node of the tree, corresponding to the risk-neutral probabilities of the associated subtrees.

The preceding results say, in essence, that certain processes are *martingales*. Concentrating on binomial trees, a "process" is just a function g whose values are defined at every node. A process is said to be a *martingale* relative to the risk-neutral probabilities if it satisfies

$$g(t) = E_{\text{RN}}[g(t')]$$

for all $t < t'$. The risk-neutral probabilities are determined by the fact that

- $s(t)/A(t)$ is a martingale relative to the risk-neutral probabilities

where $s(t)$ is the stock price process. Option prices are determined by the fact that

- $f(t)/A(t)$ is a martingale relative to the risk-neutral probabilities

if f is the value of a tradeable asset. Futures prices are determined by the fact that

- $\mathcal{F}(t)$ is a martingale relative to the risk-neutral probabilities.

One advantage of this framework is that it makes easy contact with the continuous-time theory. The central connection is this: in continuous time, the solution of a stochastic differential equation $dy = f dt + g dw$ is a martingale if $f = 0$. Indeed, the expected value of a dw-stochastic integral is 0, so for any $t < t'$ we have $E[y(t)] = E\left[\int_t^{t'} f(\tau)\, d\tau\right] = \int_t^{t'} E[f(\tau)]\, d\tau$; for the right hand side to vanish (for all $t < t'$) we must have $E[f] = 0$. If f is deterministic then this condition says simply that $f = 0$.

We can use this insight to explain and/or confirm some results previously obtained by other means. We return here to the constant-interest-rate environment, so $A(t) = e^{rt}$.

Question: why does the risk-neutral stock price process satisfy $ds = rsdt + \sigma sdw$? Answer: because the risk-neutral stock price has the property that $s(t)/A(t) = s(t)e^{-rt}$ is a martingale. Explanation: if we assume that the risk-neutral price process has the form $ds = fdt + gdw$ for some f, we easily find that

$$d(se^{-rt}) = e^{-rt}ds - re^{-rt}sdt = (f - rs)dt + e^{-rt}gdw.$$

So se^{-rt} is a martingale exactly if $f = rs$. (You may wonder why the risk-neutral stock price process has the same *volatility* as the subjective stock price process. This is because changing the drift has the effect of re-weighting the probabilites of paths, without actually changing the set of "possible" paths; changing the volatility on the other hand has the effect of considering an entirely different set of "possible paths." This is the essential content of Girsanov's theorem, which is discussed and applied in the course Continuous-Time Finance.)

Question: why does the option price satisfy the Black-Scholes PDE? Answer: because the option price normalized by $A(t)$ must be a martingale. Explanation: suppose the option price has the form $V(s(t),t)$ for some function $V(s,t)$. Then

$$\begin{aligned} d\left(V(s(t),t)e^{-rt}\right) &= e^{-rt}dV - re^{-rt}Vdt \\ &= e^{-rt}(V_t dt + V_s ds + \tfrac{1}{2}V_{ss}\sigma^2 s^2 dt) - re^{-rt}Vdt \\ &= e^{-rt}(V_t + rsV_s + \tfrac{1}{2}\sigma^2 s^2 V_{ss} - rV)dt + e^{-rt}\sigma s V_s dw. \end{aligned}$$

For this to be a martingale the coefficient of dt must vanish. That is exactly the Black-Scholes PDE.

Question: why does the solution of the Black-Scholes PDE give the discounted expected payoff of the option? Answer: because the option price normalized by $A(t)$ is a martingale. Explanation: suppose V solves the Black-Scholes PDE, with final value $V(s,T) = f(s)$. We have shown that $e^{-rt}V(s(t),t)$ is a martingale. Therefore

$$V(s(0),0) = E_{\mathrm{RN}}\left[e^{-rt}V(s(t),t)\right]$$

for any $t > 0$. Bringing e^{-rt} out of the expectation and setting $t = T$ gives

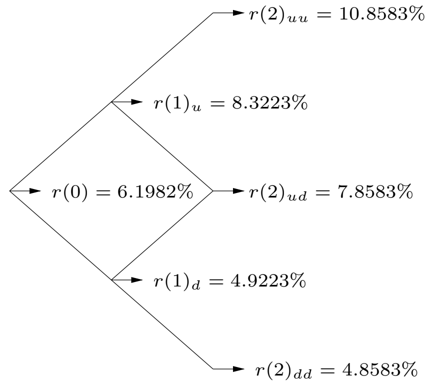
$$V(s(0),0) = e^{-rT}E_{\mathrm{RN}}\left[V(s(T),T)\right] = e^{-rT}E_{\mathrm{RN}}\left[f(s(T))\right]$$

as asserted.

****

A brief introduction to stochastic interest rates. When considering interest-based instruments, the essential source of randomness is the interest rate itself. Let us briefly explain how the binomial-tree framework can be used for modeling stochastic interest rates, following Jarrow-Turnbull Section 15.2. The basic idea is shown in the figure: each node of the tree is assigned a risk-free rate, different from node to node; it is the one-period risk-free rate for the binomial subtree just to the right of that node.

What probabilities should we assign to the branches? It might seem natural to start by figuring out what the subjective probabilities are. But why bother? All we really need



for option pricing are the risk-neutral probabilities. Moreover we know (from Homework 3) that there is some freedom in the choice of the risk-neutral probability q, and that for lognormal dynamics it is always possible to set $q = 1/2$. So the usual practice is to

- restrict attention to the risk-neutral interest rate process.
- assume the risk-neutral probability is $q = 1/2$ at each branch, and
- choose the interest rates at the various nodes so that the long-term interest rates associated with the tree match those observed in the marketplace.

The last bullet – the *calibration* of the tree to market information – is of course crucial. We'll return to it in a few weeks. For now let's just be sure we understand what it means. In other words let's be sure we understand how such a tree determines long-term interest rates. As an example let's determine $B(0,3)$, the value at time 0 of a dollar received at time 3, for the tree shown in the figure. (Put differently: $B(0,3)$ is the price at time 0 of a zero-coupon bond which matures at time 3.) We take the convention that $\delta t = 1$ for simplicity.

Consider first time period 2. The value at time 2 of a dollar received at time 3 is $B(2,3)$; it has a different value at each time-2 node. These values are computed from the fact that

$$B(2,3) = e^{-r\delta t}[\tfrac{1}{2}B(3,3)_{\text{up}} + \tfrac{1}{2}B(3,3)_{\text{down}}] = e^{-r\delta t}$$

since $B(3,3) = 1$ in every state, by definition. Thus

$$B(2,3) = \begin{cases} e^{-r(2)_{uu}} = .897104 & \text{at node } uu \\ e^{-r(2)_{ud}} = .924425 & \text{at node } ud \\ e^{-r(2)_{dd}} = .952578 & \text{at node } dd. \end{cases}$$

Now we have the information needed to compute $B(1,3)$, the value at time 1 of a dollar received at time 3. Applying the rule

$$B(1,3) = e^{-r\delta t}[\tfrac{1}{2}B(2,3)_{\text{up}} + \tfrac{1}{2}B(2,3)_{\text{down}}]$$

at each node gives

$$B(1,3) = \begin{cases} e^{-r(1)_u}(\frac{1}{2} \cdot .897104 + \frac{1}{2} \cdot .924425) = .838036 & \text{at node } u \\ e^{-r(1)_d}(\frac{1}{2} \cdot .924425 + \frac{1}{2} \cdot .952578) = .893424 & \text{at node } d. \end{cases}$$

Finally we compute $B(0,3)$ by applying the same rule:

$$\begin{aligned} B(0,3) &= e^{-r\delta t}[\tfrac{1}{2}B(1,3)_{\text{up}} + \tfrac{1}{2}B(1,3)_{\text{down}}] \\ &= e^{-r(0)}[\tfrac{1}{2} \cdot .838036 + \tfrac{1}{2} \cdot .893424] = .8137. \end{aligned}$$

Derivative Securities – Section 10 – Fall 2004
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences.

Interest-based instruments: bonds, forward rate agreements, and swaps. This section provides a fast introduction to the basic language of interest-based instruments, then introduces some specific, practically-important examples, including forward rate agreements and swaps. This material can be found in both Hull (chapters 5 and 6) and Jarrow-Turnbull (chapters 13 and 14); personally I find Hull easier to read.

****

Bond prices and term structure. The time-value of money is expressed by the *discount factor*

$$B(t, T) = \text{value at time } t \text{ of a dollar received at time } T.$$

This is, by its very definition, the price at time t of a zero-coupon bond which pays one dollar at time T. If interest rates are stochastic then $B(t, T)$ will not be known until time t. Prior to time t it is random – just as in our discussion of equity prices, $s(t)$ was random. Note however that $B(t, T)$ is a function of *two* variables, the initiation time t and the maturity time T. Its dependence on T reflects the *term structure* of interest rates. We usually take the convention that the present time is $t = 0$; thus what is observable now is $B(0, T)$ for all $T > 0$.

There are several equivalent ways to represent the time-value of money. The *yield* $Y(t, T)$ is defined by

$$B(t, T) = e^{-Y(t,T)(T-t)};$$

it is the unique constant interest rate that would have the same effect as $B(t, T)$ under continuous compounding. The *term rate* $R(t, T)$ is defined by

$$B(t, T) = \frac{1}{1 + R(t, T)(T - t)};$$

it is the unique interest rate that would have the same effect as $B(t, T)$ with no compounding. The *instantaneous forward rate* $f(t, T)$ is defined by

$$B(t, T) = e^{-\int_t^T f(t,\tau)\, d\tau};$$

it is unique deterministic time-varying interest rate that describes all the discount factors with initial time t and various maturities.[1] We can easily solve for $Y(t, T)$, $R(t, T)$, or $f(t, T)$ in terms of $B(t, T)$. Therefore each contains the same information as $B(t, T)$ as t and T vary. (Let us also mention one more: the discount rate $I(t, T)$, defined by $B(t, T) = 1-I(t, T)(T-t)$. It has little conceptual importance; however interest rates for US Treasury bills are usually presented by tabulating these discount rates.)

Most long-term bonds have *coupon payments* as well as a *final payment*. The value of the bond at time 0 is the sum of the present values of all future payments. For a *fixed-rate* bond

[1] Do not confuse this with the forward term rate, introduced below and called $f_0(t, T)$.

the coupon payments (amount c_j at time t_j) are fixed in advance, as is the final payment (amount F at time T). The value of the bond at time t is thus

$$\text{cash price} = \sum c_j B(t, t_j) + FB(t, T).$$

This is known as the *cash price*; it is a consequence of the principle of no arbitrage. Notice that the cash price is a discontinuous function of time: it rises gradually between coupon payments, then falls abruptly at each coupon date t_j because the holder of the bond collects the coupon payment. The cash price is not the value you'll see quoted in the newspaper. What you find there is the difference between the cash price and the interest accrued since the last coupon date:

$$\text{quoted price} = \sum c_j B(t, t_j) + FB(t, T) - \text{accrued interest}.$$

Notice that the quoted price is a continuous function of time, since the accrued interest is discontinuous (it resets to zero at each coupon payment) and the two discontinuities cancel.

A *floating-rate* bond is one whose interest rate (coupon rate) is reset at each coupon date. By definition, after each coupon payment its value returns to its *face value*. A typical example is a one-year floating-rate note with semiannual payments and face value one dollar, pegged to the LIBOR (London Interbank Offer) rate. Suppose at date 0 the LIBOR term interest rate for six-month-maturity is 5.25 percent per annum, but at the six-month reset the LIBOR six-month-maturity rate has changed to 5.6 percent per annum. Then the coupon payment due at six months is $.0525/2 = .02625$, and the coupon payment due at one year is $.056/2 = .028$; in addition the face value (one dollar) is repaid since the bond matures. Note the convention: interest is paid at the end of each period, using the interest rate set at the beginning of the period.

The value of the fixed-rate bond was the discounted value of its future income stream. The same is true of the floating-rate bond, provided that we *discount using the LIBOR rate*. In other words for this purpose $B(t, T)$ should be the value at time t of a LIBOR contract worth one dollar at time T. In fact, the value of the floating-rate bond at six months (just after the first coupon payment) is the value at that time of the payments to be made at one year. If t is six months and T is 1 year then this is

$$B(t, T)(.028 + 1) = \frac{1}{1 + .028}(.028 + 1) = 1.$$

The bond could be sold for this value – so holding it at six months is exactly the same as having one dollar of income at six months. The value of the bond at time 0 is similarly

$$B(0, t)(\text{first coupon} + \text{value at six months}) = \frac{1}{1 + .02625}(.02625 + 1) = 1.$$

Our calculation is clearly not special to the example; it resides in the fact that $B(t, T) = 1/(1 + R(t, T)(T - t))$.

****

Forward rates and forward rate agreements. When interest rates are deterministic $B(0,t)B(t,T) = B(0,T)$ (this was on Homework 1). When they are random this is clearly not the case, since $B(0,t)$ and $B(0,T)$ are known at time 0 while $B(t,T)$ is not. However the ratio

$$F_0(t,T) = B(0,T)/B(0,t)$$

still has an important interpretation: it is the discount factor (for time-t borrowing, with maturity T) that can be locked in now, at no cost, by a combination of market positions. In fact, consider the following portfolio:

(a) long a zero-coupon bond worth one dollar at time T (present value $B(0,T)$), and

(b) short a zero-coupon bond worth $B(0,T)/B(0,t)$ at time t (present value $-B(0,T)$).

Its present value is 0, and its holder pays $B(0,T)/B(0,t)$ at time t and receives one dollar at time T. Thus the holder of this portfolio has "locked in" $F_0(t,T)$ as his discount rate for borrowing from time t to time T.

This discussion makes reference to just three times: 0, t and T. So it is natural and conventional to work with term rates rather discount rates. Defining $f_0(t,T)$ by

$$F_0(t,T) = \frac{1}{1 + f_0(t,T)(T-t)}$$

we have shown that $f_0(t,T)$ is the *forward term rate* for borrowing from time t to time T.[2] In other words, an agreement now to borrow or lend later (at time t, with maturity T) has present value zero, if it stipulates that the term rate is $f_0(t,T)$.

What about a contract to borrow or lend at a rate R_K other than $f_0(t,T)$? This is known as a *forward rate agreement*. We can value it by an easy modification of the argument used above. Suppose the principal (the amount to be borrowed at time t) is L. Then the contract provides a payment at time T of

$$(1 + R_K \Delta T)L = (1 + f_0 \Delta T)L + (R_K - f_0)\Delta T L$$

where $f_0 = f_0(t,T)$ and $\Delta T = T - t$. So the contract is equivalent to a forward rate agreement at rate $f_0(t,T)$ on principal L *plus* an additional payment of $(R_K - f_0)\Delta T \cdot L$ at time T. The forward agreement at rate f_0 has present value 0, so the contract's present value (to the lender) is

$$B(0,T)(R_K - f_0)\Delta T \cdot L.$$

The following observations are useful in connection with swaps (which we'll discuss shortly):

(1) *A forward rate agreement is equivalent to an agreement that the lending party may pay interest at the market rate $R(t,T)$ but receive interest at the contract rate R_K.* Indeed, the lender pays L at time t and receives $(1 + R_K \Delta T)L$ at time T. We may suppose that the payment at time t is borrowed at the market rate. Then the lender

[2] Do not confuse this with the instantaneous forward rate discussed earlier.

is (a) borrowing L at the market rate R at time t, repaying $(1 + R\Delta T)L$ at time T, and (b) lending L to the counterparty at time t, receiving repayment $(1 + R_K \Delta T)L$ at time T. Briefly: the lender is *exchanging* the market rate R for the contract rate R_K.

(2) *A forward rate agreement can be priced by assuming that the market rate $R(t,T)$ will be the forward rate $f_0(t,T)$.* Indeed, the pair of loans just considered have net cash flow 0 at time t, and the lender receives $(R_K - R)\Delta T \cdot L$ at time T. The value of R is not known at time 0. But substitution of f_0 in place of R gives the correct value of the contract at time 0.

****

Swaps. A *swap* is an exchange of one income or payment stream for another. The most basic example is a (plain vanilla) interest rate swap, which exchanges the cash flow of a floating-rate debt for that of a fixed-rate debt with the same principal. We shall restrict our attention to this case.

A swap is, in a sense, the floating-rate bond analogue of a forward contract. It permits the holder of a floating-rate bond to eliminate his interest-rate risk. This risk arises because the future interest payments on a floating-rate bond are unknown. It can be eliminated by entering into a swap contract, exchanging the income stream of the floating-rate bond for that of a fixed-rate bond. What fixed rate to use? Any rate is possible – but in general the associated swap contract will have some (positive or negative) value. However at any given time there is a fixed rate that sets the present value of the swap to 0. This is rate that would normally be used. Jarrow and Turnbull call it the *par swap rate*.

It is clear from the definition that a swap is equivalent to a portfolio of two bonds, one short and the other long, one a fixed-rate bond and the other a floating-rate bond. Real bonds would have coupon payments then would return the principal at maturity. In a swap the coupon payments don't match, so there is a cash flow at each coupon date; however the principals do match, so there is no net cash flow at maturity. But the principal of the associated bonds isn't irrelevant – we need it to calculate the interest payments. It is called the *notional principal* of the swap.

A swap can also be viewed as a collection of forward rate agreements. Indeed, we showed above that the value of a floating-rate bond is equal to its principal just after each reset. So being short the floating-rate bond and long the fixed-rate bond is equivalent to paying the market interest rate and receiving the fixed interest rate. This amounts to a collection of forward rate agreements – one for each coupon payment – all with the same principal (the notional principal of the swap) and the same interest rate (the fixed rate of the swap).

Valuing a swap is easy: it suffices to value each associated bond then take the difference. (An alternative, equivalent procedure is to value each associated forward rate agreement and add them up.) The following example is a slightly modified version of the one in Jarrow-Turnbull Section 14.1. Suppose an institution receives fixed payments at 7.15% per annum and floating payments determined by LIBOR. We assume there are two payments per year, the maturity is two years, and the notional principal is N. To value the fixed side of the swap we must find the present value of the future coupon payments. It is natural to use the LIBOR discount rate for $B(0,T)$ (Hull makes this choice) though it would be possible to use the treasure-bill discount rates instead (Jarrow-Turnbull makes this choice). Let us assume

$$B(0,t_1) = .9679, \quad B(0,t_2) = .9362, \quad B(0,t_3) = .9052, \quad B(0,t_4) = .8749$$

where $t_1 = 182$ days, $t_2 = 365$ days, $t_3 = 548$ days, and $t_4 = 730$ days are the precise payment dates. The value of the fixed side of the swap is then

$$\begin{aligned} V_{\text{fix}} &= N\{.9679 \times .0715 \times (182/365) + .9362 \times .0715 \times (183/365) \\ &\quad +.9052 \times .0715 \times (183/365) + .8749 \times .0715 \times (182/365)\} = (0.1317)N \end{aligned}$$

Notice that we counted only the coupon payments, with no final payment of principal.

Now let's value the floating side of the swap. Of course we cannot know its cash flows at each time – this would require knowledge of $B(t_i, t_{i+1})$ for each i, which cannot be known at time 0. However to value the swap all we really need to know is $B(0,t_4)$. Indeed, the value of the floating bond at time 0 is just its notional principal N. But we did not count the return of principal V_{fix}, so we must not count it here either. Thus the value of the floating side of the swap is

$$V_{\text{float}} = N - B(0,t_4)N = (0.1251)N.$$

The value of the swap is the difference, namely

$$V_{\text{swap}} = V_{\text{fix}} - V_{\text{float}} = (0.0066)N.$$

This is, of course, the value of the swap to the party receiving the fixed rate and paying the floating rate. The value to the other party is $-(0.0066)N$.

OK, that was easy. But the answer didn't come out zero. What fixed rate could have been used to make the answer come out zero – in other words, what is the par swap rate? That's easy: we must replace .0715 in the above by a variable x, set the value of the swap to 0, and solve for x. This gives

$$x \cdot \{.9679 \times (182/365) + .9362 \times (183/365) + .9052 \times (183/365) + .8749 \times (182/365)\} = 0.1251,$$

which simplifies to $1.8421x = 0.1251$ whence $x = .0679$. Thus the par swap rate is 6.79% per annum.

We have discussed only the simplest kind of swap – a "plain vanilla interest rate swap". But the general principle should be clear. Another widely used instrument is the "plain vanilla foreign currency swap," which exchanges a fixed-rate income stream in a foreign currency for a fixed-rate income stream in dollars. Such an instrument can be used to eliminate foreign currency risk. See Jarrow-Turnbull section 14.2 for a discussion of its valuation.

****

Forwards versus futures. There is a well-developed market for futures contracts on treasury bonds. At first this may seem surprising, since there are so many different types of bonds and a futures contract must refer to a well-defined underlying. In practice this difficulty is avoided by rules that permit a variety of similar bonds to be delivered when the contract matures, with cash adjustments depending on the specific bond delivered. This feature makes the futures market complicated and interesting.

Here however we wish to focus on a different issue, namely the relationship between forward and future prices. (This discussion follows section 12.3 of the book by Avellaneda and Laurence. Note however that what they call $F(t,T)$ is what we are calling $f_0(t,T)$.) Our purpose is partly to emphasize that the two are different, and partly to get a handle on how the dynamics of interest rates determines forward rates.

We have in mind the binomial-tree, risk-neutral-expectation setting explained at the end of Section 9. However we shall use the notation of a continuous-time model (mainly: integrals rather than sums) since this is less cumbersome. Our starting point is the fact that

$$B(t,T)/A(t) = E_{\text{RN}}\left[1/A(T)\right]$$

where $A(t)$ is the value of the money-market fund at time t. In the continuous time setting $A(T) = A(t)\exp\int_t^T r(s)\,ds$ so the preceding formula becomes

$$B(t,T) = E_{\text{RN}}\left[e^{-\int_t^T r(s)\,ds}\right].$$

A typical interest rate future involves 3-month Eurodollar contracts: at the contract's maturity the holder must make a 3-month loan to the counterparty, at interest rate equal to the 3-month-term LIBOR rate. We have called this rate $R(t,T)$, where $T{=}t+3$ months, and t is the maturity date of the futures contract. We know from Section 9 that the associated futures price $\tilde{f}_0(t,T)$ – which determines the daily settlements during the course of the contract – is a martingale, in other words

$$\tilde{f}_0(t,T) = E_{\text{RN}}\left[R(t,T)\right].$$

Let us seek a similar representation for the forward term rate $f_0(t,T)$, defined as above by

$$\frac{1}{1+f_0(t,T)\Delta T} = F_0(t,T) = \frac{B(0,T)}{B(0,t)}$$

with $\Delta T = T - t$. Solving for $f_0(t,T)$ gives

$$f_0(t,T) = \frac{1}{\Delta T\cdot B(0,T)}(B(0,t) - B(0,T)).$$

Rewriting the expression in parentheses as a risk-neutral expectation gives

$$\begin{aligned}
f_0(t,T) &= \frac{1}{\Delta T\cdot B(0,T)}E_{\text{RN}}\left[e^{-\int_0^t r(s)\,ds} - e^{-\int_0^T r(s)\,ds}\right]\\
&= \frac{1}{B(0,T)}E_{\text{RN}}\left[e^{-\int_0^t r(s)\,ds}\cdot\frac{1-e^{-\int_t^T r(s)\,ds}}{\Delta T}\right]\\
&= \frac{1}{B(0,T)}E_{\text{RN}}\left[e^{-\int_0^t r(s)\,ds}\cdot\frac{1-B(t,T)}{\Delta T}\right],
\end{aligned}$$

making use in the last step of the fact that risk-neutral expectations are determined working backward in time. Now, the relation $B(t,T) = 1/[1 + R(t,T)\Delta T]$ can be rewritten as

$$\frac{1 - B(t,T)}{\Delta T} = R(t,T)B(t,T),$$

so we have shown that

$$\begin{aligned} f_0(t,T) &= \frac{1}{B(0,T)} E_{\mathrm{RN}} \left[e^{-\int_0^t r(s)\,ds} R(t,T)B(t,T) \right] \\ &= \frac{1}{B(0,T)} E_{\mathrm{RN}} \left[e^{-\int_0^t r(s)\,ds} R(t,T) e^{-\int_t^T r(s)\,ds} \right], \end{aligned}$$

using once more the fact that risk-neutral expectations are determined working backward in time. Combining the two exponential terms, we conclude finally that

$$f_0(t,T) = \frac{E_{\mathrm{RN}}\left[R(t,T) e^{-\int_0^T r(s)\,ds} \right]}{E_{\mathrm{RN}}\left[e^{-\int_0^T r(s)\,ds} \right]}.$$

Thus the forward rate $f_0(t,T)$ is *not* the risk-neutral expectation of the term rate $R(t,T)$. Rather it is the expectation of $R(t,T)$ with respect to a different probability measure, the one obtained by weighting each path by $\exp\left(-\int_0^T r(s)\,ds\right)$.

It is clear from this calculation that forward rates and futures prices are different. We can also see something about the relation between the two. In fact, writing $R = R(t,T)$ and $D = \exp\left(-\int_0^T r(s)\,ds\right)$ we have

$$\text{forward rate} - \text{futures price} = \frac{E[RD] - E[R]E[D]}{E[D]}.$$

where E represents risk-neutral expectation. If R and D were independent the right hand side would be zero and forward rates would equal futures prices. In general however we should expect R and D to be negatively correlated, since R is a term interest rate and D is a discount factor. Recognizing that $E[RD] - E[R]E[D]$ is the covariance of R and D, we conclude that this expression should normally be negative, implying that

$$\text{forward rate} < \text{futures price}.$$

This is in fact what is observed (the difference is relatively small). A scheme for adjusting the futures price to obtain the forward rate is sometimes called a "convexity adjustment". It should be clear from our analysis that different models of stochastic interest rate dynamics lead to different convexity adjustment rules.

****

Caps, floors, and swaptions. Easiest first: a swaption is just an option on a swap. When it matures, its holder has the right to enter into a specified swap contract. He'll do so of course only if this swap contract has positive value. Since a swap is equivalent to a pair of bonds, a swaption can be viewed as an option on a pair of bonds. Similarly, since a swap is equivalent to a collection of forward rate agreements, a swaption can be viewed as an option on a collection of forward rate agreements.

Now let's discuss caps. The borrower in a floating-rate loan does not know his future expenses, since they depend on the floating interest rate. He could eliminate this uncertainty entirely by entering into a swap agreement. But suppose all he wants is insurance against the worst-case scenario of a high interest rate. The cap was invented for him: it pays the difference between the market interest rate and a specified cap rate at each coupon date, if this difference is positive. By purchasing a cap, the borrower insures in effect that he'll never have to pay an interest rate aboved the cap rate. The cap can be viewed as a collection of caplets, one associated with each coupon payment. Each caplet amounts to an option on a bond. It is roughly speaking a call option on the market rate at the coupon time.

A floor is like a cap, but it insures a sufficiently high interest rate rather than a sufficiently low one. It can be viewed as a collection of floorlets, one associated with each coupon payment. Each floorlet is again an option on a bond – roughly speaking a put option on the market rate at the coupon time.

There is a version of put-call parity in this setting: cap-floor=swap, if the fixed rate specified by all three instruments is the same.

Thus caps and floors are collections of options on bonds; swaptions are options on collections of bonds. We'll discuss them in more detail in the next section, and we'll explain how they can be priced using a variant of Black's formula.

Derivative Securities – Section 11
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences.

Options on interest-based instruments: pricing of bond options, caps, floors, and swaptions. The most widely-used approach to pricing options on caps, floors, swaptions, and similar instruments is Black's model. We discuss how this model works, why it works, and when it is appropriate. The main alternative to Black's model is the use of a suitable interest rate tree (or a continuous-time model of the risk-free interest rate dynamics). We discuss briefly the calibration of a binomial tree – essentially, a special case of the Black-Derman-Toy model.

****

Black's model. My discussion of Black's model and its applications follows mainly chapter 20 of Hull, augmented by some examples from Clewlow and Strickland.

The essence of Black's model is this: consider an option with maturity T, whose payoff $\phi(V_T)$ is determined by the value V_T of some interest-related instrument (a discount rate, a term rate, etc). For example, in the case of a call $\phi(V_T) = (V_T - K)_+$. Black's model stipulates that

(a) the value of the option today is its discounted expected payoff.

No surprise there – it's the same principle we've been using all this time for valuing options on stocks. If the payoff occurs at time T then the discount factor is $B(0,T)$ so statement (a) means

$$\text{option value} = B(0,T)E_*[\phi(V_T)].$$

We write E_* rather than E_{RN} because in the stochastic interest rate setting this is *not* the risk-neutral expectation; we'll explain why E_* is different from the risk-neutral expectation later on. For the moment however, we concentrate on making Black's model computable. For this purpose we simply specify that (under the distribution associated with E_*)

(b) the value of the underlying instrument at maturity, V_T, is lognormal; in other words, V_T has the form e^X where X is Gaussian.

(c) the mean $E_*[V_T]$ is the forward price of V (for contracts written at time 0, with delivery date T).

We have not specified the variance of $X = \log V_T$; it must be given as data. It is customary to specify the "volatility of the forward price" σ, with the convention that

$$\log V_T \text{ has standard deviation } \sigma\sqrt{T}.$$

Notice that the Gaussian random variable $X = \log V_T$ is fully specified by knowledge of its standard deviation $\sigma\sqrt{T}$ and the mean of its exponential $E_*[e^X]$, since if X has mean m then $E_*[e^X] = \exp(m + \frac{1}{2}\sigma^2 T)$.

Most of the practical examples involve calls or puts. For a call, with payoff $(V_T - K)_+$, hypothesis (b) gives

$$E_*[(V_T - K)_+] = E_*[V]N(d_1) - KN(d_2)$$

where

$$d_1 = \frac{\log(E_*[V_T]/K) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \quad d_2 = \frac{\log(E_*[V_T]/K) - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}.$$

This is a direct consequence of the lemma we used long ago (in Section 5) to evaluate the Black-Scholes formula. Using hypotheses (a) and (c) we get

$$\text{value of a call} = B(0,T)[F_0 N(d_1) - KN(d_2)]$$

where F_0 is the forward price of V today, for delivery at time T, and

$$d_1 = \frac{\log(F_0/K) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \quad d_2 = \frac{\log(F_0/K) - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}.$$

These formulas are nearly identical to the ones we obtained in Sections 8 and 9 for pricing options on foreign currency rates, options on stocks with continuous dividend yield, and options on futures. The only apparent difference is the discount factor: in the constant interest rate setting of Sections 8 and 9 it was e^{-rT}; in the present stochastic interest rate setting it is $B(0,T)$.

It is by no means obvious that Black's formula is correct in a stochastic interest rate setting. We'll give the honest justification a little later. But here is a rough, heuristic justification. Since the value of the underlying security is stochastic, we may think of it as having its own lognormal dynamics. If we treat the risk-free rate as being constant then Black's formula can certainly be used. Since the payoff takes place at time T, the only reasonable constant interest rate to use is the one for which $e^{-rT} = B(0,T)$, and this leads to the version of Black's formula given above.

Black's model applied to options on bonds. Here is an example, taken from Clewlow and Strickland (section 6.6.1). Let us price a one-year European call option on a 5-year discount bond. Assume:

- The current term structure is flat at 5 percent per annum; in other words $B(0,t) = e^{-.05t}$ when t is measured in years.
- The strike of the option is 0.8; in other words the payoff is $(B(1,5) - 0.8)_+$ at time $T = 1$.
- The forward bond price volatility σ is 10 percent.

Then the forward bond price is $F_0 = B(0,5)/B(0,1) = .8187$ so

$$d_1 = \frac{\log(.8187/.8000) + \frac{1}{2}(0.1)^2(1)}{(0.1)\sqrt{1}} = 0.2814, \quad d_2 = d_1 - \sigma\sqrt{T} = 0.2814 - 0.1\sqrt{1} = .1814$$

and the discount factor for income received at the maturity of the option is $B(0,1) = .9512$. So the value of the call option now, at time 0, is

$$.9512[.8187N(.2814) - .8N(.1814)] = .0404.$$

Black's formula can also be used to value options on coupon-paying bonds; no new principles are involved, but the calculation of the forward price of the bond must take into account the coupons and their payment dates; see Hull's Example 20.1.

One should avoid using the same σ for options with different maturities. And one should never use the same σ for underlyings with different maturities. Here's why: suppose the option has maturity T and the underlying bond has maturity $T' > T$. Then the value V_t of the underlying is known at both $t = 0$ (all market data is known at time 0) and at $t = T'$ (all bonds tend to their par values as t approaches maturity). So the variance of V_t vanishes at both $t = 0$ and $t = T'$. A common model (if simplified) model says the variance of V_t is $\sigma_0^2 t(T' - t)$ with σ_0 constant, for all $0 < t < T'$. In this case the variance of V_T is $\sigma_0^2 T(T' - T)$, in other words $\sigma = \sigma_0\sqrt{T' - T}$. Thus σ depends on the time-to-maturity $T' - T$. In practice σ – or more precisely $\sigma\sqrt{T}$ – is usually inferred from market data.

Black's model applied to caps. A cap provides, at each coupon date of a bond, the difference between the payment associated with a floating rate and that associated with a specified cap rate, if this difference is positive. The ith caplet is associated with the time interval (t_i, t_{i+1}); if $R_i = R(t_i, t_{i+1})$ is the term rate for this interval, R_K is the cap rate, and L is the principal, then the ith caplet pays

$$L \cdot (t_{i+1} - t_i) \cdot (R_i - R_K)_+$$

at time t_{i+1}. Its value according to Black's formula is therefore

$$B(0, t_{i+1})L\Delta_i t[f_i N(d_1) - R_K N(d_2)].$$

Here $\Delta_i t = t_{i+1} - t_i$; $f_i = f_0(t_i, t_{i+1})$ is the forward term rate for time interval under consideration, defined by

$$\frac{1}{1 + f_i\Delta_i t} = \frac{B(0, t_{i+1})}{B(0, t_i)};$$

and

$$d_1 = \frac{\log(f_i/R_K) + \frac{1}{2}\sigma_i^2 t_i}{\sigma_i\sqrt{t_i}}, \quad d_2 = \frac{\log(f_i/R_K) - \frac{1}{2}\sigma_i^2 t_i}{\sigma_i\sqrt{t_i}} = d_1 - \sigma_i\sqrt{t_i}.$$

The volatilities σ_i must be specified for each i; in practice they are inferred from market data. The value of a cap is obtained by adding the values of its caplets.

A floor is to a cap as a put is to a call: using the same notation as above, the ith floorlet pays

$$L\Delta_i t(R_K - R_i)_+$$

at time t_{i+1}. Its value according to Black's formula is therefore

$$B(0, t_{i+1})L\Delta_i t[R_K N(-d_2) - f_i N(-d_1)]$$

where d_1 and d_2 are as above. The value of a floor is obtained by adding the values of its floorlets.

Here's an example, taken from Section 20.3 of Hull. Consider a contract that caps the interest on a 3-month, \$10,000 loan one year from now; we suppose the interest is capped at 8% per annum (compounded quarterly). This is a simple caplet, with $t_1 = 1$ year and $t_2 = 1.25$ years. To value it, we need:

- The forward term rate for a 3-month loan starting one year from now; suppose this is 7% per annum (compounded quarterly).
- The discount factor associated to income 15 months from now; suppose this is .9220.
- The volatility of the 3-month forward rate underlying the caplet; suppose this is 0.20.

With this data, we obtain

$$d_1 = \frac{\log(.07/.08) + \frac{1}{2}(0.2)^2(1)}{0.2\sqrt{1}} = -0.5677, \quad d_2 = d_1 - 0.2\sqrt{1} = -0.7677$$

so the value of the caplet is, according to Black's formula,

$$(.9220)(10,000)(1/4)[.07N(-.5677) - .08N(-.7677)] = 5.19 \text{ dollars.}$$

Problem 3 of HW6 is very much like the preceding example, except that the necessary data is partly hidden in financial jargon. Here's some help interpreting that problem. The relevant term rate is LIBOR 3-month rate, 9 months from now. The statement that "the 9-month Eurodollar futures price is 92" implies (if we ignore the difference between futures and forwards) that the present 3-month forward term rate for borrowing 9 months from now is 8% per annum. The statement that "the interest rate volatility implied by a 9-month Eurodollar option is 15 percent per annum" gives $\sigma = .15\%$.

Black's model applied to swaptions. A swaption is an option to enter into a swap at some future date T (the maturity of the option) with a specified fixed rate R_K. To be able to value it, we must first work a bit to represent its payoff.

Let R_{swap} be the par swap rate at time T, when the option matures. If $t_1 < \ldots < t_N$ are the coupon dates of the swap and $t_0 = T$ then R_{swap} is characterized (see Problem 2b of HW6) by

$$\sum_{i=1}^{N} B(T, t_i) R_{\text{swap}}(t_i - t_{i-1}) L = (1 - B(T, t_N)) L$$

where L is the notional principal. Moreover the left hand side is the value at time T of the fixed payments at rate R_{swap} while the right hand side is the value of the variable payments. Suppose the swaption gives its holder the right to pay the fixed rate R_K and receive the floating rate. Then it will be in the money if $R_{\text{swap}} > R_K$, and in that case its value to the holder at time T is

$$\begin{aligned} V_{\text{float}} - V_{\text{fixed}} &= (1 - B(T, t_N))L - \sum_{i=1}^{N} B(T, t_i) R_K (t_i - t_{i-1}) L \\ &= \sum_{i=1}^{N} B(T, t_i) R_{\text{swap}} (t_i - t_{i-1}) L - \sum_{i=1}^{N} B(T, t_i) R_K (t_i - t_{i-1}) L \\ &= (R_{\text{swap}} - R_K) \sum_{i=1}^{N} B(T, t_i)(t_i - t_{i-1}) L. \end{aligned}$$

The ith term is the payoff of an option on R_{swap} with maturity T and cash flow

$$L(t_i - t_{i-1})(R_{\text{swap}} - R_K)_+$$

received at time t_i. Black's formula gives the time-0 value of this option as

$$B(0, t_i) L(t_i - t_{i-1})[F_{\text{swap}} N(d_1) - R_K N(d_2)]$$

where F_{swap} is the forward swap rate and

$$d_1 = \frac{\log(F_{\text{swap}}/R_K) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \quad d_2 = \frac{\log(F_{\text{swap}}/R_K) - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}.$$

The forward swap rate is obtained by taking the definition of the par swap rate, given above, and replacing $B(T, t_i)$ by the forward rate $F_0(T, t_i) = B(0, t_i)/B(0, T)$ for each i. To get the value of the swap itself we sum over all i:

$$\text{value of swap} = LA[F_{\text{swap}} N(d_1) - R_K N(d_2)] \quad \text{where } A = \textstyle\sum_{i=1}^{N} B(0, t_i)(t_i - t_{i-1}).$$

Here's an example, taken from Clewlow and Strickland section 6.6.1. Suppose the yield curve is flat at 5 percent per annum (continuously compounded). Let us price an option that matures in 2 years and gives its holder the right to enter a one-year swap with semiannual payments, receiving floating rate and paying fixed term rate 5 percent per annum. We suppose the volatility of the forward swap rate is 20% per annum.

The first step is to find the forward swap rate F_{swap}. It satisfies

$$\sum_{i=1}^{2} \frac{B(0, t_i)}{B(0, T)} F_{\text{swap}}(1/2) = \left(1 - \frac{B(0, t_2)}{B(0, T)}\right)$$

with $T = 2$, $t_1 = 2.5$, and $t_2 = 3.0$. Since the yield curve is flat at 5% compounded continuously, we have

$$\frac{B(0, 2.5)}{B(0, 2)} = e^{-(.05)(.5)} = .9753, \quad \frac{B(0, 3)}{B(0, 2)} = e^{-(.05)(1)} = .9512$$

and simple arithmetic gives $F_{\text{swap}} = .0506$, in other words 5.06%. Now

$$d_1 = \frac{\log(.0506/.0500) + \frac{1}{2}(0.2)^2(2)}{0.2\sqrt{2}} = 0.1587, \quad d_2 = d_1 - 0.2\sqrt{2} = -.0971,$$

and

$$\sum_{i=1}^{2} B(0, t_i)(t_i - t_{i-1}) = \tfrac{1}{2}(e^{-(.05)(2.5)} + e^{-(.05)(3)}) = .8716,$$

so the value of the swaption is

$$.8716L[.0506N(.1587) - .05N(-.0971)] = .0052L$$

where L is the notional principal of the underlying swap.

****

When and why is Black's model correct? Black's model is widely-used and appropriate for pricing European-style options on bonds, and analogous instruments such as caps, floors, and swaptions. It has two key advantages: (a) simplicity, and (b) directness. By simplicity I mean not that Black's model is easy to understand, but rather that it requires just one parameter (the volatility) to be inferred from market data. By directness I mean that we model the underlying instrument directly – the basic hypothesis of Black's model is the lognormal character of the underlying.

The main alternative to Black's model is the use of an interest-rate tree (or a continuous-time analogue thereof). Such a tree models the risk-neutral interest-rate process, which can then be used to value bonds of all types and maturities, and options of all types and maturities on these bonds. Interest-rate trees are not "simple" in the sense used above: to get started we must calibrate the entire tree to market data (e.g. the yield curve). And they are not "direct" in the sense used above: we are modeling the risk-neutral interest rate process, not the underlying instrument itself; thus there are two potential sources of modeling error: one in modeling the value of the underlying instrument, the other in modeling how the option's value depends on that of the underlying instrument.

The simplicity and directness of Black's model are also responsible for its disadvantages. Black's model must be used separately for each class of instruments – we cannot use it, for example, to hedge a cap using bonds of various maturities. For consistent pricing and hedging of multiple instruments one must use a more fundamental model such as an interest rate tree. Another restriction of Black's model: it can only be used for European-style options, whose maturity date is fixed in advance. Many bond options permit early exercise – sometimes American-style (permitting exercise at any time) but more commonly Bermudan (permitting exercise at a list of specified dates, typically coupon dates). Black's model does not allow for early exercise. Trees are much more convenient for this purpose, since early exercise is easily accounted for as we work backward in the tree.

Now we turn to the question of *why* Black's model is correct. The explanation involves "change of numeraire". (The following is a binomial-tree version of Hull's section 19.5.) The word numeraire refers to a choice of units.

Up to now our numeraire has been cash (dollars). Its growth as a function of time is described by the money-market account introduced in Section 9. The money-market account has balance is $A(0) = 1$ initially, and its balance evolves in time by $A_{\rm next} = e^{r\delta t} A_{\rm now}$. We are accustomed to finding the value f of a tradeable instrument (such as an option) by working backward in the tree using the risk-neutral probabilities. At each step this amounts to

$$f_{\rm now} = e^{-r\delta t}[q f_{\rm up} + (1-q) f_{\rm down}]$$

where q and $1-q$ are the risk-neutral probabilities of the up and down states. As we noted in Section 9, this can be expressed as

$$f_{\rm now}/A_{\rm now} = E_{\rm RN}[f_{\rm next}/A_{\rm next}],$$

and it can be iterated in time to give

$$f(t)/A(t) = E_{\rm RN}[f(t')/A(t')] \quad \text{for } t < t'.$$

This is captured by the statement that "$f(t)/A(t)$ is a martingale relative to the risk-neutral probabilities."

But sometimes the money-market account is not the convenient comparison. In fact we may use *any* tradeable security as the numeraire – though when we do so we must also change the probabilities. Indeed, for any tradeable security g there is a choice of probabilities on the tree such that

$$\frac{f_{\rm now}}{g_{\rm now}} = \left[q_* \frac{f_{\rm up}}{g_{\rm up}} + (1-q_*) \frac{f_{\rm down}}{g_{\rm down}} \right].$$

This is an easy consequence of the two relations

$$f_{\rm now} = e^{-r\delta t}[q f_{\rm up} + (1-q) f_{\rm down}] \quad \text{and} \quad g_{\rm now} = e^{-r\delta t}[q g_{\rm up} + (1-q) g_{\rm down}],$$

which hold (using the risk-neutral q) since both f and g are tradeable. A little algebra shows that these relations imply the preceding formula with

$$q_* = \frac{q g_{\rm up}}{q g_{\rm up} + (1-q) g_{\rm down}}.$$

(The value of q_* now varies from one binomial subtree to another, even if q was uniform throughout the tree.) Writing E_* for the expectation operator with weight q_*, we have defined q_* so that

$$f_{\rm now}/g_{\rm now} = E_*[f_{\rm next}/g_{\rm next}].$$

Iterating this relation gives (as in the risk-neutral case)

$$f(t)/g(t) = E_*[f(t')/g(t')] \quad \text{for } t < t';$$

in other words "$f(t)/g(t)$ is a martingale relative to the probability associated with E_*." In particular

$$f(0)/g(0) = E_*[f(T)/g(T)]$$

where T is the maturity of an option we may wish to price.

Let us apply this result to explain Black's formula. For simplicity we focus on options whose maturity T is also the time the payment is received. (This is true for options on bonds, not for caplets or swaptions – but the modification needed for caplets and swaptions is straightforward.) The convenient choice of g is then

$$g(t) = B(t,T).$$

Since $g(T) = 1$ this choice gives

$$f(0) = g(0)E_*[f(T)] = B(0,T)E_*[f(T)].$$

We shall apply this twice: once with f equal to the value of the underlying instrument, what we called V_t on page one of these notes; and a second time with f equal to the value of the option. The first application gives

$$E_*[V_T] = V_0/B(0,T)$$

and we recognize the right hand side as the *forward price* of the instrument. For this reason the probability distribution associated with this E_* is called *forward risk-neutral.* The second application gives

$$\text{option value} = B(0,T)E_*[\phi(V_T)]$$

where $\phi(V_T)$ is the payoff of the option – for example $\phi(V_T) = (V_T - K)_+$ if the option is a call.

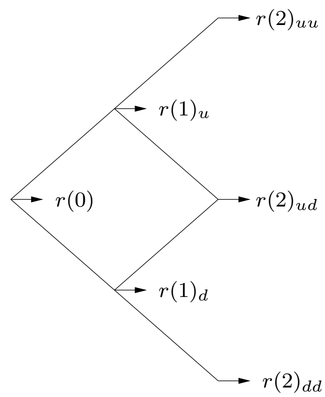
This explains Black's formula, except for one crucial feature: the hypothesis that V_T is lognormal with respect to the distribution associated with E_* (the forward-risk-neutral distribution). This is of course only asserted in the continuous-time limit, and only if the risk-neutral interest rate process is itself lognormal. The assertion is most easily explained using continuous-time (stochastic differential equation) methods, and we will not attempt to address it here.

****

Interest rate trees. We have already discussed the limitations of Black's model. Consistent pricing and hedging of diverse interest-based instruments requires a different approach. So does the pricing of American or Bermudan options, which permit early exercise.

A typical alternative is the use of a binomial tree. We explained briefly how this works in Section 9, where we discussed how to pass from the tree to the various discount factors $B(0,t)$. Valuing options on the tree is also easy (just work backward). So is hedging (each binomial submarket is complete, so a risky instrument can be hedged using any pair of zero-coupon bonds). These topics are discussed very clearly in Chapter 15 of Jarrow & Turnbull and I recommend reading them there.

To make this a practical alternative, however, we must say something about how to calibrate the tree. Jarrow and Turnbull aren't very clear on this; for an excellent treatment see the



presentation of the Black-Derman-Toy model in Chapter 8 of Clewlow and Strickland. To give the general flavor, I'll discuss just the simplest version – calibration of the tree to the yield curve, with constant volatility – for a two-period tree of the type discussed in Section 9 (see the figure).

The basic ansatz is this:

- at time 0: $r(0) = a_0$;
- at time 1: $r(1)_u = a_1 e^{\sigma\sqrt{\delta t}}$ and $r(1)_d = a_1 e^{-\sigma\sqrt{\delta t}}$;
- at time 2: $r(2)_{uu} = a_2 e^{2\sigma\sqrt{\delta t}}$, $r(2)_{ud} = a_2$, and $r(2)_{dd} = a_2 e^{-2\sigma\sqrt{\delta t}}$.

More generally: at time j, the possible values of $r(j)$ are $a_j u^k d^{j-k}$ with $u = e^{\sigma\sqrt{\delta t}}$, $d = e^{-\sigma\sqrt{\delta t}}$, and k ranging from 0 to j. The parameter σ is the volatility of the spot rate; we assume it is known (e.g. from market data) and constant. The parameters a_0, a_1, a_2, etc. represent a time-dependent drift in the spot rate (more precisely $\mu_i = (1/\delta t)\log(a_i)$ is the drift in the spot rate, since $a_i = e^{\mu_i \delta t}$). The task of calibration is to find the drift parameters a_0, a_1, etc. from the market observables, which are $B(0,1)$, $B(0,2)$, $B(0,3)$, etc.

We proceed inductively. Getting started is easy: $B(0,1) = e^{-r(0)\delta t}$ so $a_0 = r(0)$ is directly observable. To determine a_1 let $Q(1)_u$ be the value at time 0 of the option whose value at time 1 is 1 in the up state and 0 in the down state; let $Q(1)_d$ be the value at time 0 of the option whose value at time 1 is 1 in the down state and 0 in the up state. Their values are evident from the tree:

$$Q(1)_u = \tfrac{1}{2} e^{-r(0)\delta t}, \quad Q(1)_d = \tfrac{1}{2} e^{-r(0)\delta t}.$$

They are useful because examination of the tree gives

$$B(0,2) = Q(1)_u e^{-r(1)_u \delta t} + Q(1)_d e^{-r(1)_d \delta t}.$$

The left hand side is known, while the right hand side depends on a_1; so this equation determines a_1 (it must be found numerically – no analytical solution is available).

Let's do one more step to make the scheme clear. To determine a_2 we define $Q(2)_{uu}$, $Q(2)_{ud}$, and $Q(2)_{dd}$ to be the values at time 0 of options worth 1 and the indicated time-2 node (the

uu node for $Q(2)_{uu}$, etc.) and worth 0 at the other time-2 nodes. Their values are evident from the tree:

$$
\begin{aligned}
Q(2)_{uu} &= \tfrac{1}{2}e^{-r(1)_u \delta t} Q(1)_u \\
Q(2)_{ud} &= \tfrac{1}{2}e^{-r(1)_u \delta t} Q(1)_u + \tfrac{1}{2}e^{-r(1)_d \delta t} Q(1)_d \\
Q(2)_{dd} &= \tfrac{1}{2}e^{-r(1)_d \delta t} Q(1)_d.
\end{aligned}
$$

Notice that a_1 enters this calculation but a_2 does not. Now observe that

$$B(0,3) = Q(2)_{uu} e^{-r(2)_{uu}\delta t} + Q(2)_{ud} e^{-r(2)_{ud}\delta t} + Q(2)_{dd} e^{-r(2)_{dd}\delta t}.$$

The left hand side is known, while the right hand side depends on a_2; so this equation determines a_2 (solving for it numerically).

The general idea should now be clear: at each new timestep j we must find the $j+1$ values of $Q(j)_{xx}$; the tree gives us an formula based on the values of $Q(j-1)_{xx}$ and the time $j-1$ interest rates $r(j-1)_{xx}$. Then $B(0, j+1)$ can be expressed in terms of the various $Q(j)_{xx}$ and the time-j interest rates $r(j)_{xx}$, giving a nonlinear equation to solve for a_j. This procedure can easily be turned into an implementable algorithm. The only thing I haven't explained is how to index the nodes systematically, leading to general formulas for $Q(j)_{xx}$ and $B(j)_{xx}$. This is left as an exercise – or you can find it explained in section 8.4 of Clewlow and Strickland.

Derivative Securities – Fall 2004 – Semester Review
Distributed 12/13/04

1. One-period market models

   - Arbitrage-based pricing: value of a forward contract on a non-dividend-paying stock; value of a forward contract for foreign currency; put-call parity; etc.
   - Binomial market: value any contingent claim by finding a hedge portfolio. Expression for the value as the discounted risk-neutral expectation. The risk-neutral probability is uniquely determined by the condition that it give the right value for a forward.
   - Trinomial and more general markets: we can give upper and lower bounds for the value of a contingent claim by solving a pair of linear programming problems. The dual problem is an optimization over risk-neutral probabilities. A contingent claim is replicatable exactly if the upper and lower bounds coincide.

   Sample exam questions: value a forward contract by identifying a portfolio of known value with the same payoff at time T. Or take advantage of a mispriced forward – how much can you gain through this arbitrage?.

2. Multiperiod binomial trees

   - Valuing contingent claims: working backward in the tree.
   - Hedging: dynamic replication of a contingent claim.
   - The valuation formula: for a European option, value = discounted risk-neutral expectation of payoff at maturity.
   - American options: check at each node for the possibility of early exercise.
   - Continuous dividend yield, foreign currency, options on futures: similar framework, but the formula for the risk-neutral probability is different
   - Passage to the continuous-time limit: application of the central limit theorem. (Don't confuse the subjective and risk-neutral processes.)
   - Futures prices. (Note: the futures price is a martingale under the risk-neutral probability; it is not the price of a tradeable.)

   Sample exam questions: Value a contingent claim by working backward in a tree. Specify the hedging strategy, including rebalancing What if the option is American, i.e. permits early exercise? When considering an option on foreign currency, what is the proper choice of the risk-neutral probability and why? Consider an N-step multiplicative tree with $u = \exp(\mu\delta t + \sigma\sqrt{\delta t})$ and $d = \exp(\mu\delta t - \sigma\sqrt{\delta t})$; if the probability of going up is p and the probability of going down is $1-p$, find the mean and variance of $\log s(N\delta t)$.

3. Derivation and use of the Black-Scholes formulas

- I'll give you the formula for $E[e^{aX}$ restricted to $X \geq k]$ if you need it. But you'll need to know how to choose the mean and variance of X, what to use for a, etc. for a specific application.
- Option values: deriving the BS formula for valuing a put or a call; interpretations of the terms; analogous formulas for other options such as a powered call.
- Hedging: deriving the formulas for Delta, Vega, etc.
- Qualitative properties, for puts and calls. Early exercise can be optimal for an American put.
- Continuous dividend yield, options on futures: Black's formula

Sample exam questions: for the risk-neutral price process, find the probability that its value at time T is greater than K. Derive a formula for the value or the Delta of a particular option (like HW3 problem 6, valuing an option with payoff s_T^n; or HW4 problem 1, concerning an option with payoff $(s_T - K)_+^2$).

4. The basic continuous-time theory

- Stochastic differential equations. The lognormal stock process as a special case.
- Applications of Ito's formula. Consequences of the fact that dw integrals have mean value 0 (are martingales).
- Derivation of the Black-Scholes PDE based on hedging and rebalancing.
- Interpretation of the Black-Scholes PDE: $e^{-rt}V(s(t), t)$ is a martingale for the risk-neutral price process, which solves $ds = rsdt + \sigma sdw$.
- Equivalence of pricing based on the solution formula (value = discounted expected payoff, using the risk-neutral probabilities) and based on solving the Black-Scholes PDE (value = $V(s_0, 0)$ where V solves the PDE with the option payoff as final-time at $t = T$.)

Sample exam questions: show that if $ds = rsdt + \sigma sdw$ then $\log s(t)$ is Gaussian with mean $r - (1/2)\sigma^2$ and variance $\sigma^2 t$. Show that the solution of $ds = rsdt + \sigma sdw$ has the property that $s(t)e^{-rt}$ is a martingale, i.e. $E[s(t)e^{-rt}]$ is independent of time; show further that if $V(s, t)$ solves the Black-Scholes differential equation then $V(s(t), t)e^{-rt}$ is a martingale, i.e. $E[V(s(t), t)e^{-rt}]$ is independent of time; use this to connect our two methods for finding the value of an option (by solving the Black-Scholes PDE, and by evaluating the discounted expected payoff using the risk-neutral process.) Derive the Black-Scholes PDE by considering a suitable hedging strategy.

5. Further continuous-time theory

- Equivalence of the Black-Scholes PDE and the linear heat equation
- American options

Sample exam questions: show that early exercise is never optimal, for an American call on a non-dividend-paying stock. Show that it can be optimal for an American put. Show that it can be optimal on an American call, if the underlying has continuous dividend yield D with $D > r$. Value a perpetual call on an asset with nonzero dividend yield [I would not ask you to reproduce from memory the change of variables that transforms Black-Scholes to the linear heat equation.]

6. Stochastic interest rates

   - Various representations: discount rates, term rates, etc.
   - The forward rate $F_0(t,T) = B(0,T)/B(0,t)$ and its interpretation.
   - Value of a forward rate agreement.
   - Value of a swap.
   - Binomial interest-rate trees: using a tree to evaluate $B(0,T)$ for various T, to value options, and to hedge options.

   Sample exam questions: Value a specific forward rate agreement. Value a specific swap, or find the par swap rate. Given a tree, find the associated discount rates; use it to value an option; hedge the option, e.g. by a suitable portfolio of bonds.

7. Caps, floors, and swaptions

   - Black's formula applied to options on zero-coupon bonds; to caps (viewed as sums of caplets) and floors (viewed as sums of floorlets); and to swaptions.
   - Justification of Black's formula by change of numeraire.

   Sample exam questions: For a specific caplet, floorlet, or swaption, explain which version of Black's formula should be used (give an expression for the instrument's value, without actually doing any arithmetic). Explain, on a tree, why if f and g are tradeables then there's a measure that makes f/g a martingale.

8. Credit risk

   - Default probabilities and their relationship with defaultable bond prices.
   - Credit default swaps; the CDS spread.
   - Merton's framework for relating equity prices with default probabilities.

   Sample exam equations: extract default probabilities from bond prices; evaluate a CDS spread. [New material covered on 12/13, like the last bullet above, will not be on the exam.]

Derivative Securities – Homework 1 – distributed 9/13/04, due 9/27/04

(1) We used arbitrage to value a forward contract on a non-dividend-paying asset. Similar principles can be used to value a forward contract on an asset with a dividend yield, or a forward contract for foreign currency (where the foreign interest rate is like a dividend yield).

(a) Suppose the underlying asset pays cash dividends continuously at constant rate q. (This is a good approximation for a stock index fund.) Show that a forward contract with delivery price K and maturity T has present value $S_0 e^{-qT} - Ke^{-rT}$, where S_0 is the spot price and r is the risk-free interest rate. What is the forward rate (the choice of K for which the contract has present value 0)?

(b) Now consider a forward contract to buy francs for K dollars/franc at time T. Show that its present value is $S_0 e^{-qT} - Ke^{-rT}$ where S_0 is the present exchange rate, r is the risk-free interest rate for dollar investments, and q is the risk-free rate for franc investments. What is the forward exchange rate (the choice of K for which the contract has present value 0)?

(2) [like Jarrow-Turnbull 2.1] The present exchange rate between US dollars and Euros is 1.22 $/Euro. The price of a domestic 180-day Treasury bill is $99.48 per $100 face value. The price of the analogous Euro instrument is 99.46 Euros per 100 Euro face value.

(a) What is the theoretical 180-day forward exchange rate?

(b) Suppose the 180-day forward exchange rate available in the marketplace is 1.21 $/Euro. This is less than the theoretical forward exchange rate, so an arbitrage is possible. Describe a risk-free strategy for making money in this market. How much does it gain, for a contract size of 100 Euro?

(3) Let $B(t, T)$ be the cost at time t of a risk-free dollar at time T.

(a) Suppose $B(0, 1)$, $B(0, 2)$ and $B(1, 2)$ are all known at time 0 (i.e. interest rates are deterministic). Show that the absence of arbitrage requires $B(0, 1)B(1, 2) = B(0, 2)$.

(b) Now suppose $B(0, 1)$ and $B(0, 2)$ are known at time 0 but $B(1, 2)$ will not be known till time 1. What goes wrong with your argument for (a)? Show that if we know with certainty that $m \leq B(1, 2) \leq M$ then we can still conclude $mB(0, 1) \leq B(0, 2) \leq MB(0, 1)$.

(4) Which functions $\phi(S_T)$ can be the value-at-maturity of a portfolio of calls? Such a portfolio consisting of a_i call options with strike price price K_i, $1 \leq i \leq N$, all having the same maturity date T. (We permit short as well as long positions, i.e. a_i can be positive or negative. We may suppose $0 < K_1 < \ldots < K_N$. The value of this portfolio at maturity is $\phi(S_T) = \sum_{i=1}^{N} a_i(S_T - K_i)_+$.)

(a) Show that ϕ is a continuous, piecewise linear function of S_T, with $\phi(S_T) = 0$ for S_T near 0, and $\phi(S_T) = a_\infty S_T + b_\infty$ when S_T is sufficiently large.

(b) Show that any such ϕ can be realized by a suitable portfolio, and the portfolio is uniquely determined by ϕ. (Hint: think about the graph of ϕ. How does it determine K_i and a_i?)

(c) Show that $a_\infty = \sum_{i=1}^{N} a_i$ and $b_\infty = -\sum_{i=1}^{N} a_i K_i$.

(5) An investor holds a European call with strike K_c and maturity T on a non-dividend-paying asset whose current price is S_0. Suppose the investor can write a put with any strike strike price K_p, write a forward with any delivery price K_f, and can borrow any amount B at the risk-free rate (if B is negative this is a loan). What are the conditions on K_p, K_f, and B that make this combination of positions a constructive sale (i.e. that have the same effect as selling the call)?

(6) [like Jarrow-Turnbull 3.10] The present price of a stock is 50. The market value of a European call with strike 47.5 and maturity 180 days is 4.375. The cost of a risk-free dollar 180 days hence is $B(0, 180) = .9948$.

(a) For a European put with a strike price of 47.5 you are quoted a price of 1.450. Show this is inconsistent with put-call parity.

(b) Describe how you can take advantage of this situation, by finding a combination of purchases and sales which provides an instant profit with no liability 180 days from now.

(7) Show, using the absence of arbitrage, that price of a call must be a decreasing function of its strike price. In other words, if the price is $c[S_0, K, T]$ as a function of spot price S_0, strike price K, and maturity T, show that $c[S_0, K_2, T] < c[S_0, K_1, T]$ when $K_1 < K_2$.

Derivative Securities – Homework 1 Solutions
Distributed 9/27/04

(1a) Consider **Portfolio 1** = e^{-qT} units stock (with all dividends reinvested) + a bond worth K at time T, and **Portfolio 2** = the forward contract, with delivery price K. Both have value $S_T - K$ at maturity (note that portfolio 1 contains 1 unit of stock at maturity, due to the reinvested dividends). So both have the same value now. But the present value of portfolio 1 is clearly $e^{-qT}S_0 - e^{-rT}K$.

The forward rate K_* is the value satisfying $S_0e^{-qT} - K_*e^{-rT} = 0$. This gives $K_* = S_0e^{(r-q)T}$.

Note: the assertion of the problem is correct only if we assume dividends are paid at a constant rate *per share*. This is the interpretation taken above.

(1b) This is really the same calculation as 1a, with interest on foreign currency holdings taking the place of dividends.

Consider **Portfolio 1** = e^{-qT} francs + dollar debt of $e^{-rT}K$, and **Portfolio 2** = the forward contract as stated. Both have value $S_T - K$ at maturity, where S_T is now the spot exchange rate (in dollars/franc) at time T. So their present value is the same. But the present value of portfolio 1 is clearly $e^{-qT}S_0 - e^{-rT}K$, where S_0 is the spot exchange rate now, in dollars/franc. As in part (a), the forward exchange rate is $K_* = S_0e^{(r-q)T}$.

(2a) Evidently $e^{-rT} = .9948$ and $e^{-qT} = .9946$. So the theoretical forward exchange rate is $1.22 \times \frac{.9946}{.9948} = 1.2198$ dollars/euro.

(2b) The price of the forward is too low, so the arbitrage involves *buying* forwards.

- go long on a forward contract for X Euros with delivery price 1.21 dollars/euro
- borrow $e^{-qT}X$ Euros now, convert to dollars at 1.22 dollars/euro and invest at the dollar rate.

At maturity: Fulfill the contract, paying $1.21X$ dollars for X euros, and clear your cash positions. Euros cancel, and you have $e^{-qT} \times 1.22 \times e^{rT} \times X$ dollars. You pocket

$$(1.22e^{(r-q)T} - 1.21)X = (1.2198 - 1.21)X = 0.0098X \text{ dollars}$$

If $X = 100$, you make .98 cents risk free at maturity.

(3a) Compare **Portfolio 1** = a bond worth 1 dollar at time 2, and **Portfolio 2** = a bond worth $B(1,2)$ at time 1, to be reinvested at time 1. Both have the same value at time 2, namely 1 dollar. So their present values are equal. But the present value of Portfolio 1 is $B(0,2)$, while that of Portfolio 2 is $B(0,1)B(1,2)$.

(3b) The argument fails because $B(1,2)$ is not known at time 0, so we don't know how to choose Portfolio 2. But we know $B(1,2) \geq m$ then we can use **Portfolio 2** = a bond worth m at time 1, to be reinvested at time 1. By hypothesis Portfolio 1 is worth a least as much as Portfolio 2 at time 2. So present value of Portfolio 1 $\geq$ present value of Portfolio 2, which gives $B(0,2) \geq mB(0,1)$. The other inequality $B(0,2) \leq MB(0,1)$ is shown similarly, using Portfolio 2 = a bond worth M at time 1, to be reinvested at time 1.

(4a) The ith term has piecewise linear payoff and value 0 near $S_T = 0$. These properties are preserved under summation. The slope discontinuities are clearly at the strike prices. Above the largest strike K_N the function ϕ is linear, so it has the form $a_\infty S_T + b_\infty$ for some a_∞ and b_∞.

(4b) Given such ϕ, we can pick out the associated options from the discontinuities of ϕ': K_i is the ith discontinuity of ϕ', and $a_i = \phi'(K_i + 0) - \phi'(K_i - 0)$ is the value of the slope discontinuity at K_i.

(4c) For $S_T > K_N$ all the calls in the sum are in the money, so

$$\phi = \sum_{i=1}^{N} a_i(S_T - K_i) = \left(\sum_{i=1}^{N} a_i\right) S_T - \sum_{i=1}^{N} a_i K_i.$$

(5) The essence of put-call parity is $(S_T - K)_+ - (K - S_T)_+ = S_T - K$. Extension to forwards with delivery prices:

$$(S_T - K_c)_+ - (K_c - S_T)_+ = (S_T - K_f) + (K_f - K_c).$$

So if you

- write a forward with delivery price K_f
- write a put with strike K_c
- borrow $(K_f - K_c)e^{-rT}$

your liability at maturity will be $-(K_c - S_T)_+ - (S_T - K_f) - (K_f - K_c) = -(S_T - K_c)_+$, exactly offsetting the call. Thus the conditions for a constructive sale are $K_p = K_c$ and $B = (K_f - K_c)e^{-rT}$.

(6a) Put-call parity gives $p = c - f = 4.375 - (50 - 47.5 * .9948) = 1.628$. Thus the market value of the put is too high.

(6b) Puts are underpriced, so you can profit by *buying* them:

- Buy the offered put at market value.
- Sell an 'artifical' put which you construct from a forward and a call.

That is:

- buy a put from the market (pay 1.45)
- write a call (receive 4.375)
- buy a forward with delivery price 47.5 (pay $50 - 47.5 * .9948 = 2.747$).

The excess of receipts over payments is $.1780 = 1.628 - 1.450$. This is pure profit, since the positions above are self-cancelling at maturity.

(7) This problem is very similar to those above, so we rephrase it in more general terms: let P_T denote your portfolio at time T, and let $V(P_T)$ denote the value of this portfolio. The principle of no-arbitrage indicates that if $V(P_T) > 0$ then $V(P_0) > 0$. Now, let $K_1 < K_2$ and consider the portfolio which has

- a long call, strike K_1
- a short call, strike K_2.

This portfolio has positive value (examine its graph vs. S_T) and therefore it must have a positive value at time 0. But at time zero its value is $c[S_0, K_1, T] - c[S_0, K_2, T]$. So, we have shown that

$$c[S_0, K_1, T] - c[S_0, K_2, T] > 0$$

Another way to see this: construct an arbitrage portfolio if the assertion fails. If $c[S_0, K_1, T] \leq c[S_0, K_2, T]$, then consider the portfolio with

- a long call, strike K_1 – pay $c[S_0, K_1, T]$
- a short call, strike K_2 – receive $c[S_0, K_2, T]$

Your balance will be $c[S_0, K_2, T] - c[S_0, K_1, T]$, which is greater than or equal to zero by assumption. Now note:

- if at maturity $S_T < K_1$ then both options are worthless, but you still can make a profit from the above transaction:

$$(c[S_0, K_2, T] - c[S_0, K_1, T])e^{rT}$$

- if at maturity $K_1 < S_T < K_2$, then your long call is in the money, but the short call is worthless to its holder, so you make a profit of

$$(S_T - K_1) + (c[S_0, K_2, T] - c[S_0, K_1, T])e^{rT}$$

- if at maturity $K_2 < S_T$, then both calls are in the money, and you receive $(S_T - K_1)$ but pay out $(S_T - K_2)$, for a net profit of $K_2 - K_1$, in addition to your original gain of $c[S_0, K_2, T] - c[S_0, K_1, T]$. So, your total profit is

$$(c[S_0, K_2, T] - c[S_0, K_1, T])e^{rT} + (K_2 - K_1)$$

In all circumstances we have realized a profit (except perhaps the first, in which case we may come out even), even though we started with no money. Thus, $c[S_0, K_1, T] \leq c[S_0, K_2, T]$ creates an arbitrage opportunity and cannot hold.

Derivative Securities – Homework 2 – distributed 9/27/04, due 10/11/04

Class mailing list: a new mailing list has been set up for this class. To sign up, go to http://cs.nyu.edu/mailman/listinfo/g63_2791_001_fa04.

Problems 1 and 2 reinforce our discussion of one-period markets (the Section 2 notes). Problems 3-5 reinforce our discussion of multiperiod binomial trees (the Section 3 notes). The classic article on binomial trees, still well worth reading, is J. Cox, S. Ross, and M. Rubinstein, *Option pricing: a simplified approach*, J. Financial Economics 7 (1979) 229-263. It is available online through www.sciencedirect.com (to gain access using a non-NYU computer use the NYU proxy server; for instructions see http://library.nyu.edu/help/proxy.html).

(1) Consider the one-period trinomial model with

- asset 1 = risk-free, interest rate $r > 0$
- asset 2 = risky, initial unit price s_0, final unit prices $s_0 d$, s_0, $s_0 u$

with $d < 1 < u$. Assume that $d < e^{rT} < u$ so the market admits no arbitrage. You want to buy a call option on the risky asset with strike price K. Let's find the largest and smallest prices you should consider paying for it, based on considerations of arbitrage.

(a) Let π_1, π_2, π_3 be the risk-neutral probabilities associated to the down, no-change, and up states respectively (these are the $\hat{\pi}$'s of the Section 2 notes). They must satisfy

$$\pi_1 + \pi_2 + \pi_3 = 1, \quad \pi_1 d + \pi_2 + \pi_3 u = e^{rT}, \quad \pi_i \geq 0 \text{ for each } i.$$

These relations restrict (π_1, π_2, π_3) to a line segment. What are its endpoints?

(b) Any contingent claim in this market is described by a vector $f = (f_1, f_2, f_3)$ giving the payoffs if the final-time stock price is $s_0 d$, s_0, and $s_0 u$ respectively. Let $V_-(f)$ and $V_+(f)$ be the smallest and largest prices permitted for payoff f. Using part (a), give simple formulas for $V_-(f)$ and $V_+(f)$. (Your formulas should involve the min and max of two expressions.) What choice of f corresponds to a call with strike K?

(c) The payoff of a call is a *convex* function of the stock price, i..e $F(x) = (s_0 x - K)_+$ is a convex function of x. Use Jensen's inequality to show that the min in part (b) is at the endpoint with $\pi_1 = 0$ and the max is at the endpoint with $\pi_2 = 0$. Conclude that for a call with strike K in this 3-period market the smallest and largest prices consistent with the absence of arbitrage are

$$V_- = \frac{ue^{-rT} - 1}{u - 1}(s_0 - K)_+ + \frac{1 - e^{-rT}}{u - 1}(s_0 u - K)_+$$

and

$$V_+ = \frac{ue^{-rT} - 1}{u - d}(s_0 d - K)_+ + \frac{1 - e^{-rT} d}{u - d}(s_0 u - K)_+ \ .$$

(2) Consider the following one-period market with 3 assets and 4 states:

- Asset 1 is a riskless bond, paying no interest.
- Asset 2 is a stock with initial price 1 dollar/share; its possible final prices are d and u, with $d < 1 < u$.
- Asset 3 is another stock with initial price 1 dollar/share and possible final prices d and u (same d and u).
- To keep the arithmetic simple, let's assume that $u = 1 + \epsilon$ and $d = 1 - \epsilon$ for some $\epsilon > 0$. To avoid confusion, let's number the states: 1 = both stocks go up; 2 = asset 2 goes up, asset 3 goes down; 3 = asset 2 goes down, asset 3 goes up; 4 = both stocks go down.

(a) What system of equations and inequalities characterizes the associated risk-neutral probabilities?

(b) Show that the general solution to (a) is $\pi = (t, \frac{1}{2} - t, \frac{1}{2} - t, t)$ for $0 \leq t \leq 1/2$.

(c) Consider the contingent claim with payoff $f = (f_1, f_2, f_3, f_4)$. Let $V_-(f)$ and $V_+(f)$ be the smallest and largest prices permitted by the absence of arbitrage. Using part (b), give simple formulas for $V_-(f)$ and $V_+(f)$, expressing each as the min or max of two expressions.

(d) Does $f_\alpha \geq 0$ for all α and $V_-(f) = 0$ imply $f = 0$? Explain.

(e) Which f's are replicatable?

Problems 3-5 are from the book by Jarrow and Turnbull. For the binomial trees in these problems please use constant risk-free rate r, and $s_{\text{up}} = us_{\text{now}}$and $s_{\text{down}} = ds_{\text{now}}$ with

$$u = e^{[(r-\sigma^2/2)\delta t + \sigma\sqrt{\delta t}]}, \quad d = e^{[(r-\sigma^2/2)\delta t - \sigma\sqrt{\delta t}]}$$

where r is the risk-free rate, σ is the volatility of the underlying, and δt is the time interval. (We'll discuss the logic behind this choice soon; if you'd like to read ahead, see e.g. Chapter 5 of Jarrow-Turnbull, or my old Section 4 notes.) Convention concerning units: if r and σ are given "per year" then $\delta t = 1/2$ for a time period of 6 months, $\delta t = 1/4$ for a time period of 3 months, etc.

(3) A European put option with strike price 45 dollars matures in one year. The underlying asset has volatility 20 percent per annum, and the current spot price is 50 dollars. The risk-free interest rate is 5.60 percent per annum. Divide the one-year interval into two six-month intervals, and use the recombinant tree with u and d as given above.

(a) Show that $u = 1.172832$ and $d = .883891$. Evaluate the risk-neutral probabilities.

(b) Determine the put price by working backward through the tree.

(c) Determine the put price by using the formula which gives it as an average over all final-time payoffs. Of course your answer should be the same as for (b).

(d) Describe the associated trading strategy. In other words, specify how many units of stock and how much debt you should hold at each node after rebalancing.

(4) The current stock price is 100, and the volatility is 30 percent per annum. The risk-free interest rate is 6 percent per annum. Consider a one-year European call option on this stock with strike price 100.

(a) Divide the one-year period into two six-month intervals, and use the recombinant tree with u and d as given above. Calculate the risk-neutral probabilities. Show that the option value is 13.65.

(b) Suppose the market price of the option is 14 7/8. Assuming the market is truly described by the tree of part (a), there must be an arbitrage. Explain in detail (specifying all trades) how you can take advantage of the "incorrect" market price to earn a risk-free profit.

(4) A special kind of one-year put option is written on a stock. The current stock price is 40 and the current strike price is 40. At month 6, if the stock price is below 35 the strike price is lowered to 35; otherwise it remains unchanged. The risk-free interest rate is 5 percent per annum; the volatility of the stock is 35 percent per annum.

(a) Use a 2-period binomial tree with the usual choices of u and d to value the option.

(b) Now use a 4-period binomial tree to value the option.

(c) What is the difficulty with valuing this type of option?

Derivative Securities – Homework 2 Solutions
Distributed 10/11/04

Problem 1

(a) Let π_1, π_2, π_3 be the risk-neutral probabilities. They satisfy (by definition – these are the $\hat{\pi}$'s of the Section 2 notes)

$$\pi_1 + \pi_2 + \pi_3 = 1, \quad \pi_1 d + \pi_2 + \pi_3 u = e^{rT}, \quad \pi_i \geq 0 \text{ for each } i.$$

We have three equations and two equality constraints, so the solutions should form a line segment; let's find it. Solving for π_1 and π_3 in terms of π_2 gives

$$\pi_1 = \frac{(u - e^{rT}) - \pi_2(u-1)}{u-d} \quad \text{and} \quad \pi_3 = \frac{(e^{rT} - d) - \pi_2(1-d)}{u-d}$$

and the constraints $\pi_i \geq 0$ become

$$\pi_2 \leq \frac{u - e^{rT}}{u-1}, \quad \pi_2 \leq \frac{e^{rT} - d}{1-d}, \quad \pi_2 \geq 0.$$

One verifies that $u > d$, $r > 0$ imply $\frac{u-e^{rT}}{u-1} < \frac{e^{rT}-d}{1-d}$ since $(u-e^{rT})(1-d)-(e^{rT}-d)(u-1) = (u-d)(1-e^{rT}) < 0$. So the range of possible values for π_2 is $0 \leq \pi_2 \leq \frac{u-e^{rT}}{u-1}$, and the endpoints of the segment of risk-neutral probabilities are

$$\pi = \left(\frac{u - e^{rT}}{u-d}, 0, \frac{e^{rT} - d}{u-d}\right) \quad \text{and} \quad \pi = \left(0, \frac{u - e^{rT}}{u-1}, 0, \frac{e^{rT} - 1}{u-1}\right).$$

(b) We conclude that for any contingent claim $f = (f_1, f_2, f_3)$, the smallest and largest prices permited by arbitrage are $V_-(f)$ and $V_+(f)$ respectively, where

$$V_-(f) = e^{-rT} \cdot \min\left\{\frac{u - e^{rT}}{u-d} f_1 + \frac{e^{rT} - d}{u-d} f_3, \frac{u - e^{rT}}{u-1} f_2 + \frac{e^{rT} - 1}{u-1} f_3\right\}$$

and

$$V_+(f) = e^{-rT} \cdot \max\left\{\frac{u - e^{rT}}{u-d} f_1 + \frac{e^{rT} - d}{u-d} f_3, \frac{u - e^{rT}}{u-1} f_2 + \frac{e^{rT} - 1}{u-1} f_3\right\}.$$

This applies to any contingent claim, so in particular it applies to the call, which has payoff $(s_T - K)_+$. We have only to substitute

$$f_1 = (s_0 d - K)_+ \quad f_2 = (s_0 - K)_+ \quad f_3 = (s_0 u - K)_+$$

into the preceding formulas for $V_\pm$.

(c) The preceding answer is satisfactory, however we can improve it by observing that *for a call* the max is achieved at the endpoint with $\pi_2 = 0$ and the min is achieved at the endpoint with $\pi_1 = 0$. To see this, observe that

$$\begin{aligned}
\sum_{i=1}^{3} \pi_i f_i &= \left(\frac{u-e^{rT}}{u-d} - \frac{u-1}{u-d}\pi_2\right) f_1 + \pi_2 f_2 + \left(\frac{e^{rT}-d}{u-d} - \frac{1-d}{u-d}\pi_2\right) f_3 \\
&= (\text{term independent of } \pi_2) + \left(f_2 - \frac{u-1}{u-d} f_1 - \frac{1-d}{u-d} f_3\right)\pi_2.
\end{aligned}$$

To justify our assertion we must show that for a call, the coefficient of π_2 is *negative*. This follows from Jensen's inequality and the convexity of the function $F(x) = (s_0 x - K)_+$. In fact:

$$\begin{aligned}
\frac{u-1}{u-d} f_1 + \frac{1-d}{u-d} f_3 &= \frac{u-1}{u-d} F(d) + \frac{1-d}{u-d} F(u) \\
&\geq F\left(\frac{u-1}{u-d} d + \frac{1-d}{u-d} u\right) \quad \text{by Jensen} \\
&= F(1) = f_2
\end{aligned}$$

so $f_2 - \frac{u-1}{u-d} f_1 - \frac{1-d}{u-d} f_3 \leq 0$. Thus our final conclusion is

$$\begin{aligned}
V_-(f) &= e^{-rT}\left[\frac{u-e^{rT}}{u-1} f_2 + \frac{e^{rT}-1}{u-1} f_3\right] \\
&= \frac{ue^{-rT}-1}{u-1}(s_0 - K)_+ + \frac{1-e^{-rT}}{u-1}(s_0 u - K)_+
\end{aligned}$$

and similarly

$$V_+(f) = \frac{ue^{-rT}-1}{u-d}(s_0 d - K)_+ + \frac{1-e^{-rT} d}{u-d}(s_0 u - K)_+.$$

Problem 2

(a) The cash flow matrix is

$$D_{i\alpha} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ u & u & d & d \\ u & d & u & d \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1+\epsilon & 1+\epsilon & 1-\epsilon & 1-\epsilon \\ 1+\epsilon & 1-\epsilon & 1+\epsilon & 1-\epsilon \end{bmatrix}$$

The equality constraints on the risk-neutral probabilities are $\sum_\alpha D_{i\alpha}\pi_\alpha = 1$, $i = 1, 2, 3$. This amounts to the linear system

$$\begin{aligned}
\pi_1 + \pi_2 + \pi_3 + \pi_4 &= 1 \\
u\pi_1 + u\pi_2 + d\pi_3 + d\pi_4 &= 1 \\
u\pi_1 + d\pi_2 + u\pi_3 + d\pi_4 &= 1
\end{aligned}$$

(b) The above equations reduce easily to $\pi_2 = \pi_3$, $\pi_1 = \pi_4$, $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$. The general solution is thus

$$\pi = (t, \tfrac{1}{2} - t, \tfrac{1}{2} - t, t).$$

We also have the positivity constraints $\pi_\alpha \geq 0$; these hold exactly for $0 \leq t \leq \frac{1}{2}$. (Note that the market admits no arbitrage, since choosing $0 < t < 1$ gives a risk-neutral probability with each $\pi_\alpha > 0$.)

Final answer to part b: the risk-neutral probabilities form a line segment in R^4 with endpoints $(0, \frac{1}{2}, \frac{1}{2}, 0)$ and $(\frac{1}{2}, 0, 0, \frac{1}{2})$.

(c) The general theory gives

$$V_-(f) = \min_\pi f \cdot \pi \text{ and } V_+(f) = \max_\pi f \cdot \pi$$

for the smallest and largest values permitted by arbitrage. But the restriction of a linear function to a line segment always assumes its minimum and maximum values at the endpoints, and the endpoint values are $f \cdot (0, \frac{1}{2}, \frac{1}{2}, 0) = (f_2 + f_3)/2$ and $f \cdot (\frac{1}{2}, 0, 0, \frac{1}{2}) = (f_1 + f_4)/2)$. So we have

$$V_-(f) = \tfrac{1}{2} \min \{f_2 + f_3, f_1 + f_4\} \text{ and } V_+(f) = \tfrac{1}{2} \max \{f_2 + f_3, f_1 + f_4\}.$$

(d) No: if $f_2 = f_3 = 0$ and $f_1 > 0$, $f_4 > 0$ then $V_-(f) = 0$ but $f \neq 0$. Such a claim has nonnegative and sometimes positive payoff, yet arbitrage considerations alone don't force it to have a positive value. (Of course such an f should have a positive value. But to explain why, one must go beyond arbitrage considerations, using arguments based on the optimization of utility.)

(e) f is replicatable exactly if $f \cdot \pi$ is constant along the segment of risk-neutral probabilities, i.e. if $f \perp (1, -1, -1, 1)$. In other words f is replicatable exactly if $f_1 + f_4 = f_2 + f_3$.

[That was the easy way to do this part. The hard way, also acceptable, is to look for a replicating portfolio by solving four linear equations (one for each state) in three unknowns (representing the portfolio) – and to show it has a solution exactly when $f_1 + f_4 = f_2 + f_3$.]

Problem 3

(a) $u = \exp[(r - \frac{1}{2}\sigma^2)\delta t + \sigma\sqrt{\delta t}] = 1.172832$ and $d = \exp[(r - \frac{1}{2}\sigma^2)\delta t - \sigma\sqrt{\delta t}] = .883891$ using $r = .056$, $\sigma = 0.2$, $\delta t = 0.5$. These give $q = (e^{r\delta t} - d)/(u - d) = .500118$.

(b) The tree of stock prices and option values is shown in the figure. It is obtained by working backward in the tree, using final value $f = (45 - s_T)_+$ and the rule $f_{\text{now}} = e^{-r\delta t}[q f_{\text{up}} + (1 - q) f_{\text{down}}]$. In particular the option price is \$1.402715.

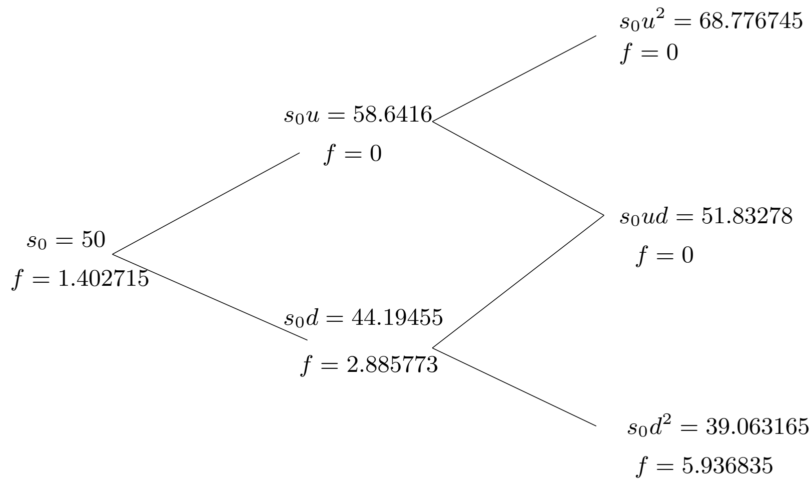


Figure 1: Problem 3 — Option valuation on a tree.

(c) Option price $= e^{-rT}[q^2 \cdot 0 + 2q(1-q) \cdot 0 + (1-q)^2 \cdot 5.936835] = 1.402715$

(d) The initial replicating portfolio has $\phi_0 = (f_{\text{up}} - f_{\text{down}})/(s_{\text{up}} - s_{\text{down}}) = -.199748$ units of stock, and $f_0 - \phi_0 s_0 = 11.390115$ dollars of cash invested in a riskless bond.

If the stock goes up to s_0u, then $\phi = 0$, so you clear your short stock position by buying .199748 units of stock. This costs 11.713542 dollars, which is exactly the value of the bond (with interest). Your portfolio is empty, since at the next stage the payoff is sure to be 0.

If the stock goes down to s_0d, then the new ϕ is -.464919, so you short another .265171 units of stock, investing the proceeds in the riskless bond. This brings your total bond holding (with interest) to value $2.885773 + .464919 s_0 d = 23.432659$. At the final time this bond is worth 24.098045 due to interest. If the final stock price is s_0ud then your stock position is worth $(s_0ud))(-.464919) = -24.098044$, i.e. your net position is 0 (modulo roundoff). If the stock's final price is s_0d^2 then your stock position is worth $(s_0d^2)(-.464919) = -18.1612076$ so your net position is worth 5.936837 dollars, exactly the payoff of the option. Either way, the replicating portfolio has final value equal to that of the option.

Problem 4

(a) $e^{r\delta t} = e^{.03} = 1.030$, $u = 1.245$, $d = 0.815$, $q = 0.500$, all given to three decimal places.

The tree of price and option values is shown in the figure. Evidently the value assigned to the option is 13.718.

(b) An arbitrageur makes money by "buying cheap and selling dear." So: if an investment bank writing the option can sell it for 14 7/8 then it is assured a profit of 14 7/8 - 13.65 = 1.225 by investing in the replicating portfolio and trading as necessary. That's a profit of 1.225 today, equivalent to $1.225e^{.06} = 1.3$ a year from now.

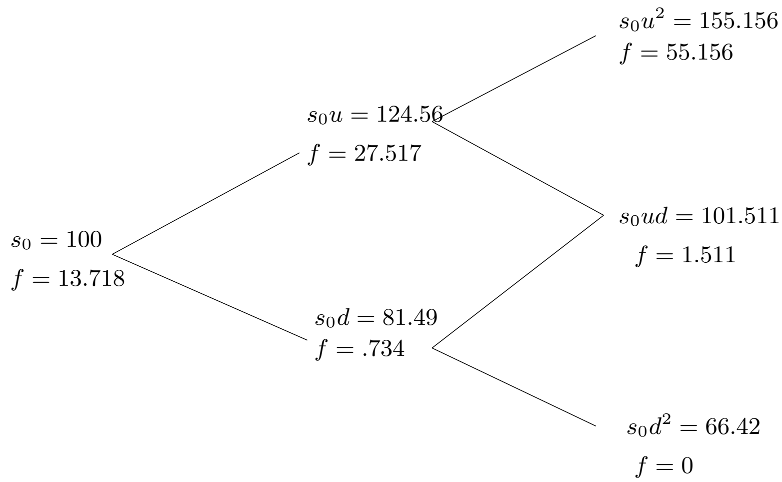


Figure 2: Problem 4 — Option valuation using a tree.

Problem 5

(a) $e^{r\delta t} = 1.0253$, $u = 1.2737$, $d = .7764$, $q = 0.500$ to adequate accuracy. The payoff is history-dependent, so we must not "recombine" the tree. The price tree is shown in the top part of the figure. The option value is $e^{-r}[(1/4)(.444)+(1/4)(10.888)] = 2.695$.

(b) We could similarly use a fully-not-recombined tree, but it saves arithmetic to recombine where the character of the option permits. Now $e^{r\delta t} = 1.0126$, $u = 1.1879$, $d = 0.8371$, and $q = .500$ to adequate accuracy. The price tree is shown in the bottom part of the figure. The option value is the discounted expected payoff using the risk-neutral probability:

$$\text{value} = e^{-r}\left[\frac{5}{16}(.45) + \frac{1}{8}(12.13) + \frac{1}{8}(7.13) + \frac{1}{16}(15.36)\right] = 3.337.$$

(c) The problem is that the most obvious trees to use (nonrecombinant) are computationally infeasible. This option can be prices using partially-recombinant trees, but this requires an ad-hoc analysis.

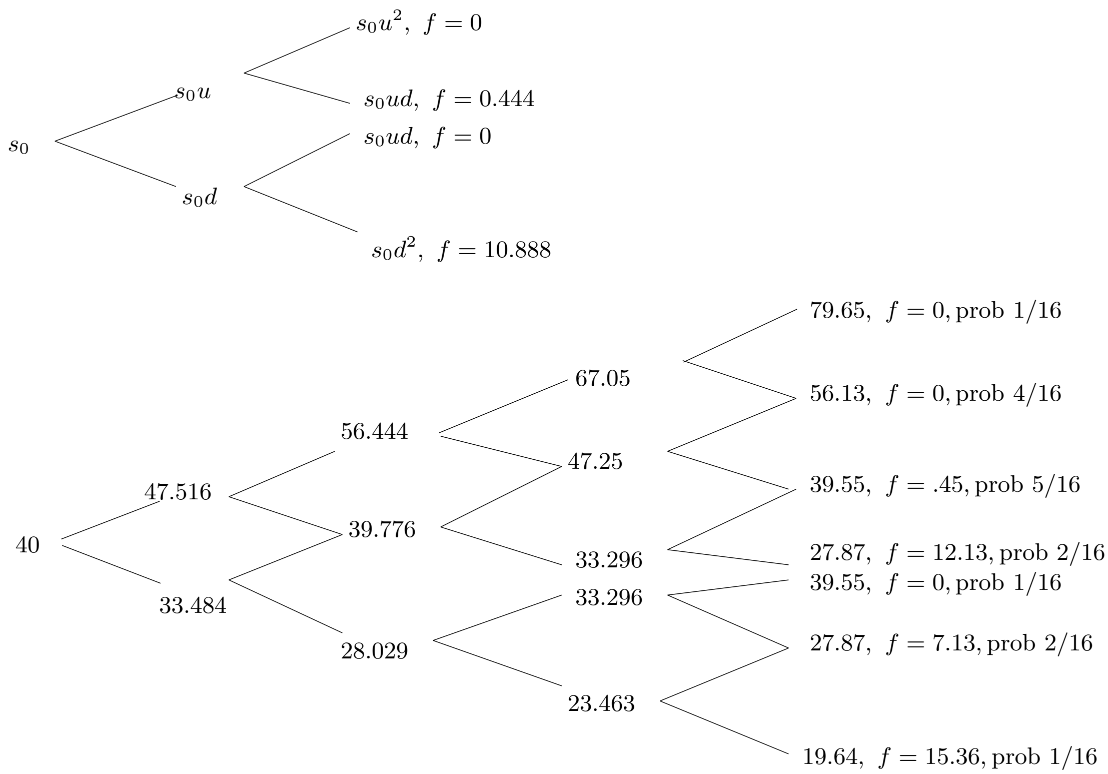


Figure 3: Problem 5 — Valuation of "history dependent" option.

Derivative Securities – Homework 3 – distributed 10/11/04, due 10/25/04

Problem 1 provides practice with lognormal statistics. Problems 2-4 explore the consequences of our formula for the value of an option, as the discounted risk-neutral expected payoff. Problem 5 makes sure you have access to a numerical tool for playing with the Black-Scholes formula and the associated "Greeks." Problem 6 examines the question: exactly which binomial trees are consistent, in the continuous-time limit, with our continuous-time valuation formula.

Convention: when we say a risky asset has "lognormal dynamics with drift μ and volatility σ" we mean $\log s(t_2) - \log s(t_1)$ has mean $\mu(t_2 - t_1)$ and variance $\sigma^2(t_2 - t_1)$; here μ and σ are constant. When pricing options, we assume the underlying has lognormal dynamics and pays no dividend, and the risk-free rate is (constant) r.

(1) Consider a stock whose price has lognormal dynamics with drift μ and volatility σ (as defined above). Suppose the stock price now is s_0.

(a) Give a 95% confidence interval for the price at time T, using the fact that with 95% confidence, a Gaussian random variable lies within 1.96 standard deviations of its mean.

(b) Give the mean and variance of the price at time T.

(c) Give a formula for the likelihood that an option with strike price K and maturity T will be in-the-money at maturity.

(d) If the mean return is 16% per annum and the volatility is 30% per annum, what do (a) and (b) tell you about tomorrow's closing price in terms of today's closing price?

(e) What is the probability that $s_T > E[s_T]$? (Note: the answer is not 1/2.)

(2) Consider a derivative with payoff s_T^n at maturity. Show that its value at time t is

$$s_t^n e^{[\frac{1}{2}\sigma^2 n(n-1) + r(n-1)](T-t)}$$

where r is the risk-free rate and σ is the volatility of the underlying asset. (Hint: use the option valuation formula $e^{-rT}E_{\text{RN}}[\text{payoff}]$.)

(3) Consider a squared call with strike K and maturity T, i.e. an option whose payoff at maturity is $(s_T - K)_+^2$.

(a) Evaluate its hedge ratio (its "Delta") by differentiating under the integral, then evaluating the resulting expression.

(b) Give a formula for the value of the squared call at time 0, analogous to the standard formula $s_0 N(d_1) - Ke^{-rT}N(d_2)$ for an ordinary call.

(Hint: For part (b) use the fact that $(e^x - K)^2 = e^{2x} - 2Ke^x + K^2$. You could of course differentiate your answer to (b) to find Delta, but that's the hard way.)

(4) Consider a "cash-or-nothing" option with strike price K, i.e. an option whose payoff at maturity is

$$f(s_T) = \begin{cases} 1 & \text{if } s_T \geq K \\ 0 & \text{if } s_T < K \end{cases}$$

It can be interpreted as a bet that the stock will be worth at least K at time T.

(a) Give a formula for its value at time t, in terms of the spot price s_t.

(b) Give a formula for its Delta (i.e. its hedge ratio). How does the Delta behave as t gets close to T?

(c) Why is it difficult, in practice, to hedge such an instrument?

[Comment: Such options are rarely found "naked" but they often arise in "structured products" calling for a fixed payment to be made if an asset price is above a certain value on a certain date. In view of (c) it is not entirely clear that the Black-Scholes valuation formula is valid for such an option. What do you think?)

(5) Suppose r is 5 percent per annum and σ is 20 percent per annum. Let's consider standard put and call options with strike price $K = 50$. Do this problem using the Black-Scholes formulas (not a binomial tree).

(a) Suppose the spot price is $s_0 = 50$ and the maturity is one year. Find the value, Delta, and Vega of the put. Same request for the call.

(b) Graph the value of a European call as a function of the spot price s_0, for several maturities. Display all the graphs on a single set of axes, and comment on the trends they reveal.

(c) Same as (b) but for a European put.

(d) Your answer to (c) should show that the value of the put is lower than $(K - s_0)_+$ for $s_0 < s_*$ and higher for $s_0 > s_*$. Estimate the critical value s_* when the maturity T is 2 years.

[Comment: Use whatever means (matlab, mathematica, spreadsheet) is most convenient, but say briefly what you used. One point of this problem is to visualize the behavior of the Black-Scholes pricing formulas. Another is to be sure you have a convenient tool for exploring further on your own.]

(6) We saw in Section 4 that different binomial trees (associated with different values of μ) can give the same values for options in the continuum limit. So it makes sense to ask: for a given risk-free rate r and volatility σ, which binomial trees give the correct continuum limit? Let us refine this question a bit. We consider only recombinant trees of the form $s_{\text{up}} = us_{\text{now}}$, $s_{\text{down}} = ds_{\text{now}}$. The continuum limit corresponds to $n \to \infty$ time steps of

size $\delta t = T/n$. We expect u and d to depend on n, i.e. $u = u_n$, $d = d_n$. For any fixed n the value of the option is $e^{-rT} E_{\text{RN}}[f(s_T)]$; this is the value obtained by working backward through the tree, using the risk-neutral probability $q = q_n = (e^{r\delta t} - d)/(u - d)$. In the continuum limit we know the value should be $e^{-rT} E[f(s_0 e^X)]$ where X is Gaussian with mean $(r - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$. Our task is to find conditions on u_n and d_n such that

$$E_{\text{RN}}[f(s_T)] \to E[f(s_0 e^X)] \quad \text{as } n \to \infty. \tag{1}$$

The main point of this problem is to show that (1) holds if $u = u_n$ and $d = d_n$ are chosen so that

$$qu + (1-q)d = e^{r\delta t}, \quad qu^2 + (1-q)d^2 = e^{(2r+\sigma^2)\delta t}. \tag{2}$$

Of course the first relation is equivalent to the definition of the risk-neutral probability $q = q_n$, so only the second relation is new. Notice that (2) gives two equations in three unknowns (u, d, q), so there is one remaining degree of freedom.

(a) Define a_n and b_n by $u_n = e^{a_n}$ and $d_n = e^{b_n}$. Show, by arguing as in the Section 4 notes, that (1) holds if

$$n(q_n a_n + (1-q_n)b_n) \to (r - \tfrac{1}{2}\sigma^2)T \quad \text{and} \quad n q_n (1-q_n)(a_n - b_n)^2 \to \sigma^2 T \tag{3}$$

as $n \to \infty$.

(b) Show, by algebraic manipulation, that (2) is equivalent to

$$u = e^{r\delta t}\left(1 + \sqrt{\frac{1-q}{q}(e^{\sigma^2\delta t} - 1)}\right) \qquad d = e^{r\delta t}\left(1 - \sqrt{\frac{q}{1-q}(e^{\sigma^2\delta t} - 1)}\right)$$

so that

$$a_n = r\delta t + \log\left[1 + \sqrt{\frac{1-q_n}{q_n}\omega_n}\right] \qquad b_n = r\delta t + \log\left[1 - \sqrt{\frac{q_n}{1-q_n}\omega_n}\right]$$

with $\omega_n = e^{\sigma^2\delta t} - 1$.

(c) Use the Taylor expansion of $\log(1+x)$ near $x = 0$ to verify the limits (3) as $n \to \infty$.

(d) How should we choose u_n and d_n, if we want $q_n = 1/2$ *exactly* for each n?

Derivative Securities – Homework 3 Solutions
Distributed 11/25/04

Problem 1.

(a) $s(T) = s_0 e^X$ where X is normal with mean μT and variance $\sigma^2 T$. We have

$$\text{Prob}\left\{\mu T - 1.96\sigma\sqrt{T} \leq X \leq \mu T + 1.96\sigma\sqrt{T}\right\} = 0.95$$

so

$$\text{Prob}\left\{s_0 e^{\mu T - 1.96\sigma\sqrt{T}} \leq s(T) \leq s_0 e^{\mu T + 1.96\sigma\sqrt{T}}\right\} = 0.95.$$

Expressed more informally:

$$s_0 e^{\mu T - 1.96\sigma\sqrt{T}} \leq s(T) \leq s_0 e^{\mu T + 1.96\sigma\sqrt{T}} \quad \text{with confidence 95\%.}$$

(b) The mean is $s_0 e^{\mu T + \frac{1}{2}\sigma^2 T}$ by the lemma at the end of the Section 4 notes. To get the variance, observe that $s^2(T) = s_0^2 e^{2X}$, and $2X$ is normal with mean $2\mu T$ and variance $4\sigma^2 T$. So $E[s^2(T)] = s_0^2 e^{2\mu T + 2\sigma^2 T}$ and the variance is

$$E[s^2(T)] - (E[s(T)])^2 = s_0^2 e^{(2\mu + \sigma^2)T}\left(e^{\sigma^2 T} - 1\right).$$

(c) $s(T) = s_0 e^X \geq K \iff X \geq \log(K/s_0)$. The probability that this occurs is

$$\int_{\log(K/s_0)}^{\infty} \frac{1}{\sigma\sqrt{2\pi T}} e^{-(x-\mu T)^2/2\sigma^2 T}\, dx = \int_{y_0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy$$

with $y = (x - \mu T)/\sigma\sqrt{T}$, $y_0 = [\log(K/s_0) - \mu T]/\sigma\sqrt{T}$. In terms of the cumulative normal $N(y)$ this is

$$\text{Prob}\{s(T) \geq K\} = 1 - N\left(\frac{\log(K/s_0) - \mu T}{\sigma\sqrt{T}}\right).$$

(d) Working in years, we're given $\mu = .16$, $\sigma = .30$, and $T = 1/365$. So $\mu T = 4.3836 \times 10^{-4}$, $\sigma^2 T = 2.4658 \times 10^{-5}$, and $1.96\sigma\sqrt{T} = 3.0777 \times 10^{-2}$. Using part (a),

$$.9701 \leq s(T)/s_0 \leq 1.0317 \quad \text{with confidence 95\%.}$$

Using part (b), $s(T)$ has mean $1.00056 s_0$ and variance $.000247 s_0^2$.

(e) Using the result of part (c), with $K = E(s(T))$, we find that

$$\text{Prob}\{s(T) > E(s(T))\} = 1 - N\left(\frac{\log(E(s(T))/s_0) - \mu T}{\sigma\sqrt{T}}\right)$$

The mean $E(s(T))$ is given by part (b), and we find that

$$\frac{\log(E(s(T))/s_0) - \mu T}{\sigma\sqrt{T}} = \frac{\sigma\sqrt{T}}{2}$$

$$\Longrightarrow \text{Prob}\{s(T) > E(s(T))\} = 1 - N\left(\frac{\sigma\sqrt{T}}{2}\right)$$

So, for example, with $T = 1$ (one year) and $\sigma = .30$ as in (d), we get that

$$\text{Prob}\{\ s(T)\ >\ E(s(T))\ \}\ =\ .4404$$

.

Problem 2. There's no loss of generality taking $t = 0$. Recall th at with respect to the risk-neutral probability $s(T) = s_0 e^X$ with X Gaussian, having mean $(r - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$. So $s^n(T) = s_0 e^{nX}$, and nX is Gaussian with mean $n(r - \frac{1}{2}\sigma^2)T$ and variance $n^2\sigma^2 T$. Thus

$$E_{\text{RN}}[s^n(T)] = s_0 e^{[n(r-\frac{1}{2}\sigma^2)T + \frac{1}{2}n^2\sigma^2 T]}$$

and the value of the option is $e^{-rT}E_{\text{RN}}[s^n(T)]$, which simplifies to

$$s_0^n \exp\left[\left(r(n-1) + \tfrac{1}{2}\sigma^2 n(n-1)\right)T\right]$$

If $t \neq 0$, we have need only replace s_0 by $s(t)$ and T by $T - t$ in this formula to get the answer to the question.

Problem 3. The solution formula gives

$$V = e^{-rT}E\left[(s_0 e^X - K)_+^2\right]$$

for the value of the option, where X is Gaussian with mean $(r - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$. Differentiating inside the expectation gives

$$\Delta = \frac{\partial V}{\partial s_0} = e^{-rT}E\left[2(s_0 e^X - K)_+ e^X\right].$$

(This calculation is legitimate because the payoff is continuous and piecewise differentiable. We're using a result from advanced calculus about differentiating under integrals.) Preparing for the task of evaluating these expectations: the derivation of the Black-Scholes formula shows that

$$E\left[e^X \text{ restricted to } X \geq \ln(K/s_0)\right] = e^{rT}N(d_1)$$

and

$$E\left[1 \text{ restricted to } X \geq \ln(K/s_0)\right] = N(d_2)$$

with the usual definitions

$$d_1 = \frac{\ln(s_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

and

$$d_2 = \frac{\ln(s_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}.$$

Another application of the usual lemma shows that

$$E\left[e^{2X} \text{ restricted to } X \geq \ln(K/s_0)\right] = e^{(2r+\sigma^2)T}N(d_0)$$

with

$$d_0 = \frac{\ln(s_0/K) + (r + \frac{3}{2}\sigma^2)T}{\sigma\sqrt{T}} = d_1 + \sigma\sqrt{T}.$$

Now let's use these facts:

(a) The hedge ratio Δ is

$$e^{-rT}\left[2s_0 e^{(2r+\sigma^2)T}N(d_0) - 2Ke^{rT}N(d_1)\right],$$

or equivalently

$$2s_0 e^{(r+\sigma^2)T}N(d_0) - 2KN(d_1).$$

(b) The value of the option is

$$e^{-rT}\left[s_0^2 e^{(2r+\sigma^2)T}N(d_0) - 2Ks_0e^{rT}N(d_1) + K^2N(d_2)\right]$$

or equivalently

$$s_0^2 e^{(r+\sigma^2)T}N(d_0) - 2Ks_0N(d_1) + K^2e^{-rT}N(d_2)$$

Problem 4.

(a) The solution formula specifies the value of this option at time 0 as

$$e^{-rT}E\left[1 \text{ restricted to } X \geq \ln(K/s_0)\right]$$

where X is Gaussian with mean $(r - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$. This is, of course, e^{-rT} times the probability (with respect to the risk-neutral process) that the option is in the money at maturity. We know from deriving the Black-Scholes solution formula that this is precisely $e^{-rT}N(d_2)$. The value at any time t is obtained by replacing s_0 by s_t and T by $T - t$:

$$\text{value} = e^{-r(T-t)}N(d_2), \quad d_2 = \frac{\ln(s_t/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

(b) We cannot differentiate under the expectation this time, because the payoff is discontinuous. Instead we must differentiate the solution formula obtained in (a). This gives

$$\begin{aligned}\text{Delta} &= e^{-r(T-t)}N'(d_2)\frac{\partial d_2}{\partial s_t} \\ &= e^{-r(T-t)}N'(d_2)\frac{1}{s_t\sigma\sqrt{T-t}}.\end{aligned}$$

Remembering that $N'(d) = \frac{1}{\sqrt{2\pi}}e^{-d^2/2}$ is the density of a standard Gaussian, we conclude that

$$\text{Delta} = e^{-r(T-t)}\frac{1}{s_t}\left\{\frac{1}{\sigma\sqrt{2\pi(T-t)}}e^{-\frac{\left[\ln(K/s_t)-(r-\sigma^2/2)(T-t)\right]^2}{2\sigma^2(T-t)}}\right\}.$$

Notice that the expression in curly brackets is the density of a Gaussian random variable with mean $(r - \frac{1}{2}\sigma^2)(T-t)$ and standard deviation $\sigma\sqrt{T-t}$, evaluated at $\ln(K/s_t)$.

Close to maturity, the option's value is nearly 0 if s_t is significantly less than K (specifically: if $(s_t/K) - 1 \ll \sigma\sqrt{T-t}$). Similarly its value is nearly 1 if s_t is significantly greater than K (specifically: if $(s_t/K) - 1 \gg \sigma\sqrt{T-t}$). Correspondingly, Δ is very large if $s_t \approx K$ and $t \approx T$; it has order of magnitude $\frac{1}{K\sigma\sqrt{2\pi(T-t)}}$ for $s_t = K \pm K\sigma\sqrt{T-t}$.

(c) Hedging is impractical because it requires taking a hugely leveraged position if $s_t \approx K$ near maturity. (For large Delta, the hedge portfolio has a large stock position, purchased with mostly borrowed funds.) Such a position would be extremely sensitive to inaccuracies of the model (failing to trade continuously in time, getting the volatility wrong, etc). No sensible person would put so much faith in a model that is, after all, at best an approximate description of the real world.

Problem 5. I used Matlab. First I created an mfile that computes the value, delta, and vega of a call, as functions of the spot price, strike price, risk-free rate, volatility, and time-to-maturity; then I created an mfile that computes the values for several different maturities and graphs the results. I did this separately for calls and puts.

(a) For the put: Value = 2.7868, Delta = -.3632, Vega = 18.762. For the call: Value = 5.2253, Delta = .6368, Vega = 18.762. (Note that, by put-call parity, the put and the call have the same vega, and Delta(call) - Delta(put) = 1.)

(b) See the figure, which graphs the value of the call when the time-to-maturity is .1, .6, 1.1, 1.6, 2.1, and 2.6 years. As t increases (time-to-maturity decreases) the value decreases and approaches the payoff. As s increases, the slope (Delta) increases monotonically from nearly 0 to nearly 1. The payoff is already smoothed considerably when when $T = .1$.

(c) See the figure, which graphs the value of the put when the time-to-maturity is .1, .6, 1.1, 1.6, 2.1, and 2.6 years. As t increases (time-to-maturity decreases) the value increases when s is much less than the strike, and it decreases when s is much larger than the strike. As s increases, the slope (Delta) decreases from nearly -1 to nearly 0. The graph of the value always crosses the graph of the payout, at a location $s_*(t) < K$; this location is nearly (but not exactly) independent of t.

(d) When the time-to-maturity is 2 years, value$+s-50$ is $-.0207$ for $s = 44.6$ and $+.0331$ for $s = 44.7$, so s_* is between 44.6 and 44.7.

The Matlab mfiles I used for valuing and graphing the call are given at the end of this solution sheet.

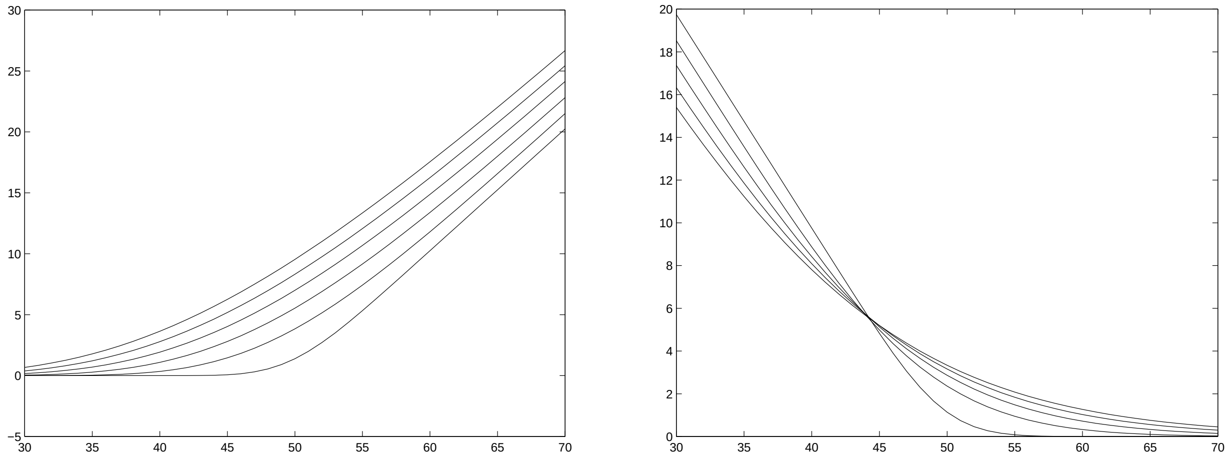


Figure 1: *The value of the call (left) and the put (right) as a function of spot price, at various times-to-maturity, for Problem 3.*

Problem 6.

(a) Arguing as in the notes,

$$s(T) = s_o e^{X_n a_n + (n - X_n) b_n}$$

at the end of a binomial process with n steps ($T = n\delta t$), where

$$X_n = \text{number of heads in } n \text{ flips of a biased coin } (q_n = \text{prob. of heads}).$$

For valuing an option q_n shold be the risk-neutral probability of the "up" state.

The exponent $X_n a_n + (n - X_n) b_n$ is the sum of n independent, identically distributed random variables, each taking value a_n with probability q_n and b_n with probability $1 - q_n$. Each of these random variables has

$$\text{mean} = q_n a_n + (1 - q_n) b_n$$

and

$$\text{variance} = q_n a_n^2 + (1 - q_n) b_n^2 - (q_n a_n + (1 - q_n) b_n)^2 = q_n (1 - q_n)(a_n - b_n)^2$$

so

$X_n a_n + (n - X_n) b_n$ has mean $n(q_n a_n + (1 - q_n) b_n)$ and variance $n q_n (1 - q_n)(a_n - b_n)^2$.

A variant of the central limit theorem (strictly speaking: Lindeberg's theorem – the point is that for different choices of n we're using slightly different basic random variables) says that the limit of $X_n a_n + (n - X_n) b_n$ is Gaussian. So to check whether we get the desired (risk-neutral) distribution we need only match the mean and variance, i.e. we need only require

$$n(q_n a_n + (1 - q_n) b_n) \to (r - \tfrac{1}{2}\sigma^2)T \quad \text{and} \quad n q_n (1 - q_n)(a_n - b_n)^2 \to \sigma^2 T.$$

(b) Set $\alpha = e^{r\delta t}$ and $\beta = e^{(2r+\sigma^2)\delta t}$. Solving the first equation of (2) for u, then substituting the result into the second equation of (2) and simplifying, we get

$$d^2 - 2\alpha d + \frac{\alpha^2 - q\beta}{1-q} = 0.$$

The roots are $d = \alpha \pm \sqrt{\frac{q}{1-q}(\beta - \alpha^2)}$, and the associated values of u are $\alpha \mp \sqrt{\frac{1-q}{q}(\beta - \alpha^2)}$. Since $u > d$ by definition we conclude that

$$u = \alpha + \sqrt{\frac{1-q}{q}(\beta - \alpha^2)}, \quad d = \alpha - \sqrt{\frac{q}{1-q}(\beta - \alpha^2)}$$

which simplifies to the desired formula.

(c) Since $\log(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \dots$ (and since $\omega_n \to 0$ as $n \to \infty$) we have

$$\begin{aligned} a_n &= r\delta t + \sqrt{\frac{(1-q_n)}{q_n}\omega_n} - \frac{1-q_n}{2q_n}\omega_n + \text{error of order } |\delta t|^{3/2} \\ b_n &= r\delta t - \sqrt{\frac{q_n}{(1-q_n)}\omega_n} - \frac{q_n}{2(1-q_n)}\omega_n + \text{error of order } |\delta t|^{3/2}. \end{aligned}$$

Checking conditions (3) we find:

$$\begin{aligned} n(q_n a_n + (1-q_n)b_n) &= n[r\delta t - \tfrac{1}{2}\omega_n + O(|\delta t|^{3/2}) \\ &= rT - \tfrac{1}{2}\sigma^2 T + \text{error of order } T/\sqrt{n} \\ &\to (r - \tfrac{1}{2}\sigma^2)T \quad \text{as } n \to \infty \end{aligned}$$

and similarly

$$\begin{aligned} nq_n(1-q_n)(a_n - b_n)^2 &= nq_n(1-q_n)\left(\left[\sqrt{\frac{1-q_n}{q_n}} + \sqrt{\frac{q_n}{1-q_n}}\right]\sqrt{\omega_n} + O(\delta t)\right)^2 \\ &= nq_n(1-q_n)\left[\frac{1}{q_n(1-q_n)}\omega_n + O(|\delta t|^{3/2})\right] \\ &= n\sigma^2\delta t + \text{error of order } T/\sqrt{n} \\ &\to \sigma^2 T \quad \text{as } n \to \infty. \end{aligned}$$

(d) Substitute $q = 1/2$ into the assertion of part (b): to get $q = 1/2$ exactly, at each n, we should use

$$u = e^{r\delta t}\left(1 + \left[e^{\sigma^2\delta t} - 1\right]^{1/2}\right) \quad \text{and} \quad d = e^{r\delta t}\left(1 - \left[e^{\sigma^2\delta t} - 1\right]^{1/2}\right).$$

Derivative Securities – Homework 3 Solution Addendum
Matlab m-files associated with Problem 5.

For brevity, I give only the files associated with the call.

File call.m

```
function value=call(r, sigma, T, K, s0)
% This is mfile call.m; it computes the value and Greeks for a call.
% r = risk-free rate (example: 5% per year means r=.05)
% sigma = volatility (example: 20% per year means sigma=.2)
% T = maturity (example: one year means T=1)
% K = strike price
% s0 = spot price
% output is call(1)=value, call(2)=delta, call(3)=gamma, call(4)=vega
% generate parameters d1, d2
d1 = log(s0/K) + (r+ (sigma^2)/2)*T;
d1 = d1/(sigma*sqrt(T));
d2 = d1- (sigma*sqrt(T));
% N1 = N(d1), N2 = N(d2) using relation between cumulative normal
% and the error function erf
N1 = ( erf(d1/sqrt(2)) + 1 )/2;
N2 = ( erf(d2/sqrt(2)) + 1 )/2;
% f = normal distribution evaluated at d1
f= exp(-d1^2/2)/sqrt(2*pi);
% evaluate using the Black-Scholes formula
value(1) = s0 * N1 - K * N2 * exp(-r*T);
value(2)=N1;
value(3)=f/(s0*sigma*sqrt(T));
value(4)=s0*sqrt(T)*f;
```

File bscalls.m

```
% This is mfile bscalls.m; it visualizes the graphs of some calls
% In array value(i,j), each value of i corresponds to a different time-to-maturity:
% i=1,2,...,6 correspond to T=.1, T=.6; T=1.1, T=1.6, T=2.1, and T=2.6.
value=[];
for i=.1:.5:2.6
valuei=[];
  for j=30:1:70
  x=call(.05,.2,i,50,j);
  valuei = [valuei,x(1)];
  end
value=[value;valuei];
end
hold off
for k=1:1:5
plot(30:1:70,value(k,:))
hold on
end
```

Derivative Securities – Homework 4 – distributed 10/25/04, due 11/08/04

These problems provide some practice with the Black-Scholes PDE (Problem 1) and stochastic differential equations (Problems 2-5).

1) We considered, in HW3, a derivative whose payoff was $s^n(T)$ at maturity, where $s(t)$ has lognormal dynamics with constant volatility σ, and the risk-free rate is r (also constant). We showed there that the derivative has value

$$s^n(t) \exp\left([\tfrac{1}{2}\sigma^2 n(n-1) + r(n-1)](T-t)\right)$$

at time t. Let's give a different derivation of the same result, using the Black-Scholes PDE.

(a) Substitute $V(s,t) = h(t)s^n$ into the Black-Scholes PDE. What ODE must $h(t)$ solve? What is the appropriate final-time condition?

(b) Verify that $h(t) = \exp\left([\tfrac{1}{2}\sigma^2 n(n-1) + r(n-1)](T-t)\right)$ solves the ODE you found in (a), with the appropriate final-time condition.

(2) Consider the solution of

$$ds = r(t)s\,dt + \sigma(t)s\,dw, \quad s(0) = s_0. \tag{1}$$

where $r(t)$ and $\sigma(t)$ are deterministic functions of time.

(a) Show that $\log s(t)$ is a Gaussian random variable, with mean $\int_0^t [r(s) - \frac{1}{2}\sigma^2(s)]\,ds$ and variance $\int_0^t \sigma^2(s)\,ds$.

(b) Show that $s(T) = s_0 \exp\left([\overline{r} - \tfrac{1}{2}\overline{\sigma}^2]T + \overline{\sigma}\sqrt{T}Z\right)$ where Z is a standard Gaussian,

$$\overline{r} = \frac{1}{T}\int_0^T r(s)\,ds \quad \text{and} \quad \overline{\sigma}^2 = \frac{1}{T}\int_0^T \sigma^2(s)\,ds.$$

[Comment: we'll show soon that (1) is the "risk-neutral" stock price process when the risk-free rate and volatility are deterministic functions of t. This problem shows that options can be valued in that setting using the standard Black-Scholes formula, with r replaced by $\overline{r}$ and σ replaced by $\overline{\sigma}$.]

(3) We showed in class using Ito's formula that if $s(t) = s(0)e^{\mu t + \sigma w(t)}$ then $ds = (\mu + \frac{1}{2}\sigma^2)sdt + \sigma s dw$.

(a) Conclude that $E[s(t)] - E[s(0)] = (\mu + \frac{1}{2}\sigma^2)\int_0^t E[s(\tau)]d\tau$, where E denotes expected value.

(b) Conclude that $E[s(t)] = s(0)e^{(\mu + \frac{1}{2}\sigma^2)t}$.

[Comment: taking $t = 1$, this gives a new proof of the lemma, stated at the end of the Section 4 notes, that if X is Gaussian with mean μ and standard deviation σ then $E[e^X] = e^{\mu+\sigma^2/2}$.]

(4) This problem should help you understand Ito's formula. If w is Brownian motion, then Ito's formula tells us that $z = w^2$ satisfies the stochastic differential equation $dz = 2wdw+dt$. Let's see this directly:

(a) Suppose $a = t_0 < t_1 < \ldots < t_{N-1} < t_N = b$. Show that $w^2(t_{i+1}) - w^2(t_i) = 2w(t_i)(w(t_{i+1}) - w(t_i)) + (w(t_{i+1}) - w(t_i))^2$, whence

$$w^2(b) - w^2(a) = 2\sum_{i=0}^{N-1} w(t_i)(w(t_{i+1}) - w(t_i)) + \sum_{i=0}^{N-1}(w(t_{i+1}) - w(t_i))^2$$

(b) Let's assume for simplicity that $t_{i+1} - t_i = (b-a)/N$. Find the mean and variance of $S = \sum_{i=0}^{N-1}(w(t_{i+1}) - w(t_i))^2$.

(c) Conclude by taking $N \to \infty$ that

$$w^2(b) - w^2(a) = 2\int_a^b w\,dw + (b-a).$$

[Comment: we did parts of this calculation in the notes and in class, but because it's so enlightening I'm asking you to go through it carefully here.]

(5) Here's a cute application of the Ito calculus. Let

$$\beta_k(t) = E[w^k(t)]$$

where $w(t)$ is Brownian motion (with $w(0) = 0$). Show using Ito's formula that for $k = 2, 3, \ldots$,

$$\beta_k(t) = \frac{1}{2}k(k-1)\int_0^t \beta_{k-2}(s)\,ds.$$

Deduce that $E[w^4(t)] = 3t^2$. What is $E[w^6(t)]$?

[Comment: the moments of w can also be calculated from its distribution function, since $w(t)$ is Gaussian with mean 0 and variance 1. But the method in this problem is easier, and good practice with Ito's lemma.)

Derivative Securities – Homework 4 Solutions
Distributed 11/08/04

Problem 1. Substitution of $V(s,t) = h(t)s^n$ into the Black-Scholes PDE yields

$$[h_t + rnh + \tfrac{1}{2}\sigma^2 n(n-1)h - rh]s^n = 0$$

which holds (for all s) only if

$$h_t = -ch, \quad c = rn + \tfrac{1}{2}\sigma^2 n(n-1) - r = (n-1)(r + \tfrac{1}{2}\sigma^2 n).$$

The final-time condition $V(s,T) = s^n$ gives $h(T) = 1$. The general solution of this ODE for h is

$$h(t) = h_0 e^{-ct}$$

and the final-time condition gives $h_0 e^{-cT} = 1$ whence

$$h_0 = e^{cT}.$$

Thus

$$h(t) = e^{c(T-t)} = e^{(n-1)(r+\frac{1}{2}\sigma^2 n)(T-t)}$$

as asserted.

Problem 2

(a) We note that $d(\log(s(t)))$ is given by Ito's lemma:

$$d(\log(s(t))) = \frac{1}{s}ds - \frac{1}{2s^2}ds \cdot ds$$

$ds = r(t)sdt + \sigma(t)sdw$, so $ds \cdot ds = \sigma(t)^2 s^2 dt$ which gives

$$\begin{aligned} d(\log(s(t))) &= \frac{1}{s}\left(r(t)sdt + \sigma(t)sdw\right) - \frac{1}{2s^2}\sigma(t)^2 s^2 dt \\ &= \left(r(t) - \frac{1}{2}\sigma(t)^2\right)dt + \sigma(t)dw \end{aligned}$$

If we write the above equation as an integral we get

$$\log(s(t)) - \log(s(0)) = \int_0^t \left(r(\tau) - \frac{1}{2}\sigma(\tau)^2\right)d\tau + \int_0^t \sigma(\tau)dw(\tau).$$

Now, the first term on the right hand side has no randomness to it, so its expected value is itself. The second term is an integral with respect to brownian motion, so its expected value is zero (see section 7 notes). So,

$$E(\log(s(t))) = \log(s(0)) + \int_0^t \left(r(\tau) - \frac{1}{2}\sigma(\tau)^2\right)d\tau.$$

To find the variance, we use the fact that for a random variable X, $\text{var}(X) = E(X - m)^2$ where m is the mean of the random variable. So,

$$\begin{aligned}
\text{var}(\log(s(t))) &= E\left(\log(s(t)) - \log(s(0)) - \int_0^t \left(r(\tau) - \frac{1}{2}\sigma(\tau)^2\right) d\tau\right)^2 \\
&= E\left(\int_0^t \sigma(\tau) dw\right)^2 \\
&= E\left(\int_0^t \sigma(\tau)^2 d\tau\right) \\
&= \int_0^t \sigma(\tau)^2 \, d\tau
\end{aligned}$$

Finally, to see $\log(s(t))$ is a Gaussian random variable, we just note that the term $\int_0^t \sigma(\tau)dw$ must be Gaussian. To prove this:

$$\int_0^t \sigma(\tau)dw = \lim \sum_{0=\tau_1<\ldots<\tau_k=t} \sigma(\tau_j)\left(w(\tau_{j+1}) - w(\tau_j)\right)$$

is a limit of sums of Gaussian random variables (we use here the hypothesis that $\sigma(\tau)$ is deterministic, so each term $\sigma(\tau_j)[w(\tau_{j+1}) - w(\tau_j)]$ is Gaussian). Now, sums of Gaussians are Gaussian, and a limit of Gaussians is again Gaussian (provided the means and variances converge, which is the case here). Thus the stochastic integral $\int_0^t \sigma(\tau)dw$ is Gaussian. So therefore is

$$\log(s(t)) = \log(s(0)) + \int_0^t \left(r(t) - \frac{1}{2}\sigma(t)^2\right) dt + \int_0^t \sigma(\tau)dw$$

since the first two terms are deterministic. In summary: $\log(s(t))$ is Gaussian. We computed its mean and variance above, and these completely specify its distribution.

(b) We've shown that

$$\log(s(t)) \sim N\left(\log(s(0)) + \int_0^t \left(r(\tau) - \frac{1}{2}\sigma(\tau)^2\right) d\tau, \int_0^t \sigma(\tau)^2 d\tau\right)$$

Any a random variable $X \sim N(m, v^2)$ can be written $X = m + vZ$ where Z is a standard Gaussian, $Z \sim N(0,1)$. Thus, letting

$$\begin{aligned}
\bar{r} &= \frac{1}{T}\int_0^T r(t)dt \\
\bar{\sigma}^2 &= \frac{1}{T}\int_0^T \sigma(t)^2 dt,
\end{aligned}$$

we have that

$$\log(s(T)) \sim N\left(\log(s(0)) + \left(\bar{r} - \frac{1}{2}\bar{\sigma}^2\right) T, \bar{\sigma}^2 T\right)$$

So finally putting together all of what's above, and exponentiating $\log(s(T))$ we get:

$$s(T) = s(0)\exp\left(\left[\bar{r} - \frac{1}{2}\bar{\sigma}^2\right] T + \bar{\sigma}\sqrt{T}Z\right)$$

where $Z \sim N(0,1)$ as before.

Problem 3 We use the fact that every stochastic differential equation has an equivalent integral form, and the fact that stochastic integrals of the form $\int_a^b g dw$ have expected value 0 (here w is Brownian motion).

(a) Integrating the SDE gives $s(t) - s(0) = \int_0^t (\mu + \frac{1}{2}\sigma^2) s(\tau)\, d\tau + \int_0^t \sigma s(\tau)\, dw(\tau)$. Taking the expected value of both sides we get the desired assertion, since the expected value of the stochastic integral is 0 and $\mu + \frac{1}{2}\sigma^2$ is constant.

(b) Let $f(t) = E[s(t)]$. The assertion of (a) is that $f(t) - f(0) = (\mu + \frac{1}{2}\sigma^2) \int_0^t f(\tau)\, d\tau$. Differentiation gives $df/dt = (\mu + \frac{1}{2}\sigma^2) f$, so $f(t) = f(0) e^{(\mu + \frac{1}{2}\sigma^2)t}$. Finally $f(0) = E[s(0)] = s_0$ since $s(0)$ was specified as data (i.e. it is deterministic).

Problem 4

(a) This is a simple algebraic manipulation.

(b) $\Delta_i w = w(t_{i+1}) - w(t_i)$ is Gaussian, with mean 0 and standard deviation $\sqrt{\Delta t} = \sqrt{(b-a)/N}$. So $(\Delta_i w)^2 = \frac{b-a}{N} Z_i^2$ where $\{Z_i^2\}_{i=1}^N$ are independent random variables, each distributed as the square of a standard Gaussian. Evidently the mean of Z_i is 1, and its variance is finite (in fact it is 2, as one easily shows using Problem 6). So the mean of S is $b - a$ and the variance of S is proportional to $1/N$ (in fact it is $2(b-a)^2/N$).

(c) Part (b) shows that the $S \to (b-a)$ in distribution. Actually we can say more, since the law of large numbers is directly applicable:

$$S_N = (b-a)\frac{Z_1^2 + \cdots + Z_N^2}{N} \to (b-a)E[Z_i^2] = (b-a) \quad \text{almost surely.}$$

The sum $2\sum_{i=0}^{N-1} w(t_i)(w(t_{i+1}) - w(t_i))$ converges to the stochastic integral $\int_a^b w dw$, by the very definition of a stochastic integral.

Problem 5 Ito's formula gives

$$d(w^k) = k w^{k-1} dw + \tfrac{1}{2}k(k-1) w^{k-2} dt.$$

Integrating then taking the expected value of each side (as in Problem 4) gives the desired assertion

$$\beta_k(t) = \tfrac{1}{2}k(k-1)\int_0^t \beta_{k-2}(s) ds$$

since $w(0) = 0$ and the stochastic integral $\int w^{k-1} dw$ has expected value 0. From the definition of Brownian motion, $\beta_2(t) = E[w^2(t)] = t$. So

$$\beta_4(t) = E[w^4(t)] = 6\int_0^t s ds = 3t^2$$

and

$$\beta_6(t) = E[w^6(t)] = 15\int_0^t 3s^2 ds = 15t^3.$$

Derivative Securities – Homework 5 – distributed 11/15/04, due 11/29/04

(1) Consider options on an underlying with continuous dividend yield D (assumed constant and positive). Show that the value of a European call with strike K is smaller than $s_0 - K$ when s_0 is sufficiently large. Conclude that it can be optimal to exercise an American call prior to maturity.

(2) Let's value a *perpetual American put* with strike K, written on a non-dividend-paying stock with lognormal dynamics. By definition this instrument never matures, and it can be exercised at any time t yielding payoff $(K - s(t))_+$. Since it never matures, we expect its value to be a function of $s(t)$ alone, i.e. to have the form $V(s(t))$ where

$$\tfrac{1}{2}V_{ss}\sigma^2 s^2 + rsV_s - rV \leq 0 \quad \text{and} \quad V(s) \geq (K - s)_+$$

for all s. Moreover we expect that there is an exercise threshold s_* such that

$$\tfrac{1}{2}V_{ss}\sigma^2 s^2 + rsV_s - rV = 0 \quad \text{for } s \geq s_*,$$

and

$$V(s) = (K - s)_+ \quad \text{for } s \leq s_*,$$

and we expect that V and V_s are continuous at s_*. Finally we expect that $V(s) \to 0$ as $s \to \infty$.

(a) Show that if $k = 2r/\sigma^2$ then $f(s) = As + Bs^{-k}$ solves the PDE $\frac{1}{2}f''\sigma^2 s^2 + rsf' - rf = 0$ for any choice of the constants A and B. (If you know some PDE then you'll recognize that this is the most general possible solution.)

(b) Show that to have $f(s_*) = (K - s_*)$ and $f(\infty) = 0$ we must set $A = 0$ and $B = s_*^k(K - s_*)$.

(c) Show that the choice $s_* = \frac{k}{1+k}K$ gives $f'(s_*) = -1$, and that the resulting function $f(s)$ has all the properties listed above for $V(s)$. This is the desired function $V(s)$ which values the perpetual American put.

(3) [Jarrow-Turnbull, chapter 7, problem 2]. An American put option with a maturity of one year and a strike price of 60 is written on a non-dividend-paying stock. Assume the current stock price s_0 is 60, the volatility σ is 35 percent per year, and the risk-free rate r is 6 percent per year. To keep things simple, let's use a two-period binomial tree to value the option.

(a) Construct an appropriate two-period recombining price tree using $u = \exp[(r - \frac{1}{2}\sigma^2)\delta t + \sigma\sqrt{\delta t}]$ and $d = \exp[(r - \frac{1}{2}\sigma^2)\delta t - \sigma\sqrt{\delta t}]$. Notice that when δt is 6 months, $u = 1.2800$, $d = .7803$, and the risk-neutral probability of the up state is $q = .5007$. (You may round this to $q = .5$.) Value the option by working backward through the tree. Don't forget to check whether early exercise is optimal.

(b) Describe the replicating portfolio at each node. Verify that the associated trading strategy is self-financing, and that it replicates the payoff.

(4) [like Jarrow-Turnbull, chapter 11, problem 5]. Consider a call option written on Euros with a maturity of one year. The spot exchange rate is 1.30 dollars/Euro, its volatility is 12 percent per year, and the strike price is 1.1818. The contract size is $250,000$ Euros. To keep things simple, let's use a two-period binomial tree to value the option. Assume that if you invest one dollar for six months at the risk-free rate it will be worth 1.0151 dollars, and if you invest one Euro for six months at the risk-free rate it will be worth 1.0305 Euros.

(a) Construct an appropriate two-period recombining price tree using $u = \exp[(r - D - \frac{1}{2}\sigma^2)\delta t + \sigma\sqrt{\delta t}] = 1.068436$ and $d = \exp[(r - D - \frac{1}{2}\sigma^2)\delta t - \sigma\sqrt{\delta t}] = .901667$.

(b) Explain briefly how we should use this tree to value the option. How does the Euro interest rate affect the calculation?

(c) If the option is European (exercisable only at maturity) what is its value? (You may round off the risk-neutral probability to $q = .5$).

(d) If the option is American (exercisable at six months or at maturity) what is its value?

(e) Describe the hedging strategy at each node. (The hedge portfolio consists of a Euro bond holding and a dollar bond holding. Notice that each bond earns interest at the appropriate risk-free rate. I am asking for the *dollar-investor's* hedge portfolio; equivalently, I'm asking for the trading strategy a dollar investor can use to replicate the option.)

5) The Section 9 notes discuss use of Black's formula to price options on futures. Focusing for simplicity on calls: if the futures price F_t is lognormal with volatility σ_F then the value of a European call with maturity T (whose payoff is $(F_T - K)_+$ at time T)is $e^{-rT}[F_0 N(d_1) - KN(d_2)]$, where $d_1 = [\log(F_0/K) + \frac{1}{2}\sigma^2 T]/\sigma\sqrt{T}$ and $d_2 = [\log(F_0/K) - \frac{1}{2}\sigma^2 T]/\sigma\sqrt{T}$.

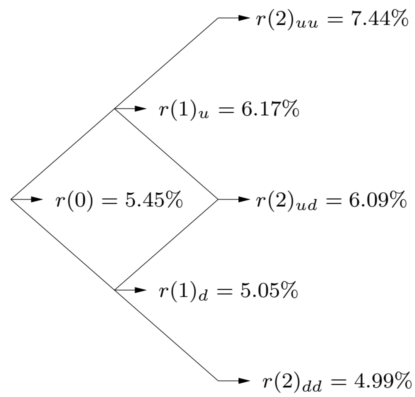
(a) Suppose the underlying is a non-dividend paying stock with volatility σ, and the futures prices under consideration are for delivery at time T_F. What value do you use for σ_F in Black's formula?

(b) This call can be hedged by continuously trading futures contracts (with delivery time T_F) and the risk-free asset. Describe the appropriate trading strategy.

(c) The call can alternatively be hedged by continuously trading the underlying and the risk-free asset. Describe the appropriate trading strategy.

6) [Jarrow & Turnbull, Chapter 15, problem 4.] Consider the binomial tree of interest rates shown in the figure (each time interval is one year, and the rates shown are per annum with continuous compounding). Assume the risk-neutral probabilities are 1/2 for each branch.

(a) Find the values of $B(0,1)$, $B(0,2)$, and $B(0,3)$.



(b) Consider the following European call option written on a one year Treasury bill: its maturity is $T = 2$, and its strike is 0.945, so the payoff at time 2 is $(B(2,3) - 0.945)_+$. Find the value of this option at time 0.

(c) Suppose you wish to hedge this option using two-year and three-year treasury bills. Find the hedge portfolio at time 0. [Hint: to find the hedge portfolio at a given node, solve an appropriate system of two linear equations in two unknowns.]

Derivative Securities – Homework 5 Solutions
Distributed 12/06/04

Problem 1. We want to show that

$$S_0e^{-DT}N(d_1) - Ke^{-rT}N(d_2) < S_0 - K$$

If S_0 is very large with respect to K (and all other parameters, for convenience), then $S_0 - K \simeq S_0$, and $d_1, d_2 >> 1$. So, $N(d_1) \simeq N(d_2) \simeq 1$ and hence

$$S_0 - K \simeq S_0$$

$$S_0e^{-DT}N(d_1) - Ke^{-rT}N(d_2) \simeq S_0e^{-DT}N(d_1)$$

and since $S_0e^{-DT}N(d_1) < S_0$ for all d_1 if $D > 0$ we get the result that

$$S_0e^{-DT}N(d_1) - Ke^{-rT}N(d_2) < S_0 - K.$$

Problem 2. If V is independent of t then the Black-Scholes PDE becomes $\frac{1}{2}\sigma^2s^2V_{ss} + rsV_s - rV = 0$. Trying $V(s) = s^{-a}$ we find that this is a solution if

$$\tfrac{1}{2}\sigma^2a(a+1) - ra - r = (\tfrac{1}{2}\sigma^2a - r)(a+1) = 0$$

which holds when $a = -1$ and when $a = k = 2r/\sigma^2$. Since the equation is linear we conclude that $V(s) = As + Bs^{-k}$ is a solution for any choice of the constants A and B. (Since the equation is of second-order, the solution space is two-dimensional, so this is the most general possible solution.) The condition $V(s) \to 0$ as $s \to \infty$ forces $A = 0$. We thus expect that

$$V(s) = \begin{cases} (K-s)_+ & \text{for } s < s_* \\ Bs^{-k} & \text{for } s > s_* \end{cases}$$

for some choice of B and s_*. The values of B and s_* are determined by the condition that both V and V_s be continuous at $s = s_*$. If $s_* < K$ then continuity of V requires $K - s_* = Bs_*^{-k}$, whence

$$B = s_*^k(K - s_*).$$

Still assuming $s_* < K$, we see that continuity of V_s requires $-1 = -kBs_*^{-k-1}$; substituting the preceding formula for B then solving for s_* gives

$$s_* = \frac{k}{k+1}K.$$

This value is less than K, so our calculation is self-consistent. It remains to show that the function V defined this way satisfies

$$\tfrac{1}{2}\sigma^2s^2V_{ss} + rsV_s - rV \le 0 \quad \text{and} \quad V \ge (K-s)_+$$

for all s. The first assertion is easy, since $\frac{1}{2}\sigma^2s^2V_{ss} + rsV_s - rV = 0$ for $s > s_*$ and $\frac{1}{2}\sigma^2s^2V_{ss} + rsV_s - rV = -rK < 0$ for $s < s_*$. The second assertion is only slightly harder: for $s < s_*$ we have $V - (K-s)_+ = 0$; for $s_* < s < K$ we see that the function $g(s) = V(s) - (K-s)$ has $g' > 0$ and $g(s_*) = 0$, so $g > 0$; for $s > K$ the assertion is clear since $V > 0$.

Problem 3.

(a) At the end of one period the stock price is either 76.8 or 46.818. At the end of two periods it can be 98.304, 59.927, or 36.5321. The corresponding values of the option at time two are 0, 0.073, and 23.4679 respectively. At time 1, if the stock price is 76.8 then the option value is (1/2)(.073)/1.030455 = .0354. At time 1, if the stock price is 46.818, the value of the option without early exercise would be (1/2)(.073+23.4679)/1.030455=11.422575, but early exercise yields 13.182 which is better. So early exercise is optimal at this node, and the value is 13.182. At time 0 the value of the option is (1/2)(.0354+13.182)/1.030455=6.4139.

(b) The replicating portfolio at time 0 has (0.0354-13.182)/(76.8-46.82) = -.4385 units of stock, worth -26.31, and a bond position worth 6.4139 + 26.31 = 32.72. The value of the replicating portfolio is equal to that of the option.

- At time 1, the bond is worth 32.72*1.030455=33.72. If the stock price is 76.8 then the position of -.4385 units is worth -33.68. Rebalancing must occur to produce a new portfolio with (0-.073)/(98.304-59.927)=-.0019 units of stock, worth -0.1460, and a bond worth .0354 + .1460 = .1814. No new investment is required. If instead the stock price is 46.82 then then the stock position is worth -20.53. The holder of the option should exercise it, and the replicating portfolio provides a liquidated value exactly equal to the payoff. If the holder of the option foolishly fails to exercise it, then the holder of the replicating portfolio can put (13.182-11.422575) in his pocket and rebalance to create the replicating portfolio worth 11.422575 appropriate for the European option.
- At time 2, the value of the replicating portfolio is in each case the value of the option's payoff. (If the stock price reached 46.82 but the holder of the option failed to exercise it, then the holder of the replicating portfolio achieves an arbitrage – i.e. a gain obtained with no risk of loss.)

Problem 4.

(a) The appropriate tree is below, and lists the possible exchange rates at each time.

(b) The tree can be used just as with stocks to compute the value of the option: we take the discounted, risk-neutral expectation of the payout on the tree. The Euro interest rate enters only in the calculation of the risk neutral probability, which is roughly 1/2.

(c) For the European option, the valuation is given in the tree below. We work backward in the tree, taking discounted risk neutral expectations. The final value of the contract is $250,000$ Euros $\times$ {dollar value}/euro = $\$250,000 \times 0.1075 = \$26,875$. See Figure 2.

(d) Valuing the American option is a little more difficult since we have to check whether early exercise is better after 1 period. At the up node, the value of the contract exercised early would be $1.3889 - 1.1818 = 0.2071 > 0.18363$, so early exercise is optimal. At the down node, early exercise would give $1.1721 - 1.1818 < 0$ and therefore is not optimal. The resulting tree has a different up node, and thus the American

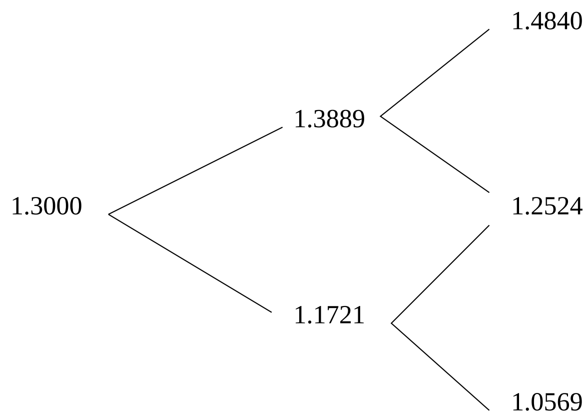


Figure 1: Problem 4a: Tree of exchange rates ($/Euro).

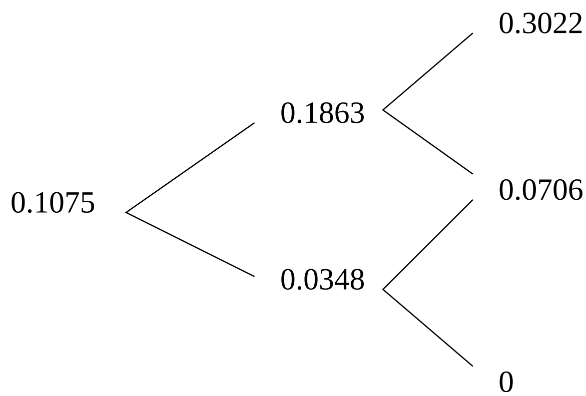


Figure 2: Problem 4c: Tree of European option values ($/Euro).

option has a different value. The final value of the contract is $\$250,000 \times 0.1191 = \$29,775$. (See Figure 3.)

(e) We'll do the replication for a contract on 1 Euro; the extension to 250,000 euros will be clear (just multiply by 250,000).

The strategy is to hold ϕ dollars and ψ Euros, each invested at their risk-free rates. The value of this portfolio, in dollars, is $\phi + 1.3\psi$. At the end of one period, we want ϕ and ψ to satisfy:

$$\begin{aligned} \phi e^{rT} + s_1 \psi e^{DT} &= \text{option up} \\ \phi e^{rT} + s_2 \psi e^{DT} &= \text{option down} \end{aligned}$$

where s_i is the spot exchange rate at 6 months. So,

$$\begin{aligned} 1.0151\phi + 1.3889 \times 1.0305\psi &= 0.2071 \\ 1.0151\phi + 1.1721 \times 1.0305\psi &= 0.0348 \end{aligned}$$

Solving these equations gives $\phi = -0.8834$ and $\psi = 0.7712$. As a consistency check we also note that $\phi + 1.3\psi = 0.1191$, which is the value of the option.

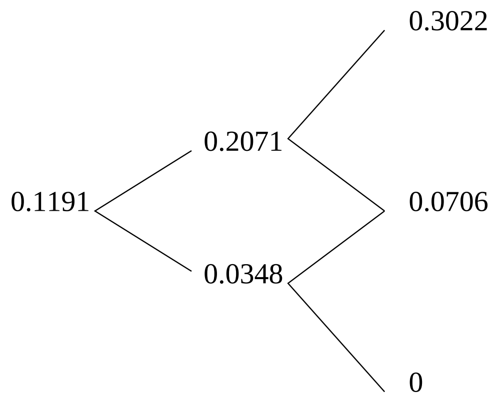


Figure 3: Problem 4d: Tree of American option values (\$/Euro).

At the other nodes, the argument is the same: determine ϕ and ψ so that the value of the portfolio will match the payoff at the up and down nodes. The only subtlety is that at the 6 month up node, we have replicated the American option, and there it is optimal for the holder to exercise the option. If the holder foolishly does not exercise, then we only need the European value at that node to replicate the proceeding up/down nodes. That is, we only need \$0.1863 to replicate the possibilites of 0.3022 and 0.0706. However, we will have \$0.2071, and therefore can pocket \$0.0208. In this event we gain a risk-free profit of $\$0.0208 \times 1.0151 = \$0.0211$ on each Euro in the contract.

Problem 5.

(a) The two volatilities should be the same. To see this, note that when interest rates are constant (as they are here), the futures price is the forward price, and so $F_t = e^{r(T-t)} S_t$. Taking d of this expression to see what SDE F_t satisfies gives:

$$\begin{aligned} d(F_t) &= d\left(e^{r(T-t)} S_t\right) \\ &= -re^{r(T-t)} S_t dt + e^{r(T-t)} dS_t \\ &= -rF_t dt + e^{r(T-t)} \left(rS_t dt + \sigma S_t dB_t\right) \\ &= -rF_t dt + rF_t dt + \sigma F_t dB_t \\ &= \sigma F_t dB_t \end{aligned}$$

(b) On intuitive grounds the 'delta' for this contract should just be $\frac{\partial V}{\partial F}$. (Since we're dealing with a call option, we have a very explicit formula for this, namely $\partial V/\partial F = N(d_1)$. To see why the hedge portfolio holds $\partial V/\partial F$ futures, consider a replication strategy (our argument is identical to that in Section 6 of the notes): we hold a portfolio of ϕ shares of futures contracts and ψ units of money market. We want to choose ϕ so that the growth of the portfolio which has a long option and short ϕ futures will grow deterministically, so that the addition of the money market account can replicate the option. So, letting $\Pi = V(F_t, t) - \phi F_t$, we take d and hold ϕ fixed, giving:

$$\begin{aligned} d\Pi &= dV_t - \phi dF_t \\ &= \left(\frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial F}dF + \frac{1}{2}\frac{\partial^2 V}{\partial F^2}dFdF\right) - \phi\left(\sigma F_t dB_t\right) \\ &= \{\text{terms}\}dt + \left(\frac{\partial V}{\partial F}\sigma F_t - \phi\sigma F_t\right)dB_t \end{aligned}$$

which is deterministic iff $\phi = \frac{\partial V}{\partial F}$.

(c) We invoke the same argument, with $\Pi = V(F_t, t) - \phi S_t$ and get

$$\begin{aligned} d\Pi &= \{\text{terms}\}dt + \left(\frac{\partial V}{\partial F}\sigma F_t - \phi\sigma S_t\right)dB_t \\ &= \{\text{terms}\}dt + \left(\frac{\partial V}{\partial F}\sigma e^{r(T-t)}S_t - \phi\sigma S_t\right)dB_t \\ \Longrightarrow \phi &= \frac{\partial V}{\partial F}e^{r(T-t)} \end{aligned}$$

Problem 6.

(a) For the bond with maturity 1:

$$B(0,1) = e^{-.0545} = .94696.$$

For the bond with maturity 2: its value at time 1 is

$$B(1,2) = \begin{cases} e^{-.0617} = .94016 & \text{at node } u \\ e^{-.0505} = .95075 & \text{at node } d \end{cases}$$

so its value at time 0 is

$$B(0,2) = .94696 \cdot (1/2)[.94016 + .95075] = .89531.$$

For the bond with maturity 3: its value at time 2 is

$$B(2,3) = \begin{cases} e^{-.0744} = .92830 & \text{at node } uu \\ e^{-.0609} = .94092 & \text{at node } ud \\ e^{-.0499} = .95132 & \text{at node } dd \end{cases}$$

so its value at time 1 is

$$B(1,3) = \begin{cases} .94016 \cdot (1/2)[.92830 + .94092] = .87868 & \text{at node } u \\ .95075 \cdot (1/2)[.94092 + .95132] = .89952 & \text{at node } d \end{cases}$$

and its value at time 0 is

$$B(0,3) = .94696 \cdot (1/2)[.87868 + ..89952] = .84194.$$

(b) At time 2, the option's value is

$$V_2 = \begin{cases} (.92830 - .945)_+ = 0 & \text{at node } uu \\ (.94092 - .945)_+ = 0 & \text{at node } ud \\ (.95132 - .945)_+ = .00632 & \text{at node } dd. \end{cases}$$

Working backward in the tree, we find that its value at time 1 is

$$V_1 = \begin{cases} .94016 \cdot (1/2)[0 + 0] = 0 & \text{at node } u \\ .95075 \cdot (1/2)[0 + .00632] = .00300 & \text{at node } d \end{cases}$$

and its value at time 0 is

$$V_0 = .94696 \cdot (1/2)[0 + .00300] = .00142$$

(c) We wish to build a hedge portfolio consisting of a zero-coupon bond worth ϕ_2 dollars at time 2, and a zero-coupon bond worth ϕ_3 dollars at time 3. Their present values are of course $\phi_2 B(0,2)$ and $\phi_3 B(0,3)$ respectively. To provide a perfect hedge, the portfolio's value must match that of the option at time 1, at both nodes u and d. This gives two equations in the two unknowns ϕ_2 and ϕ_3:

$$\begin{aligned} \phi_2 \times .94016 + \phi_3 \times .87868 &= 0 \\ \phi_2 \times .95075 + \phi_3 \times .89952 &= .00300 \end{aligned}$$

which has the unique solution

$$\phi_2 = -0.25623, \quad \phi_3 = 0.27416.$$

Of course the value of the hedge portfolio at time 0 should match the value of the option, and indeed it does:

$$\phi_2 \times .89531 + \phi_3 \times .84194 = .00142.$$

Derivative Securities – Homework 6 (complete version) – distributed 12/6/04, due 12/13/04

Note 1: As previously announced, the final exam will be Monday December 20, in the normal class hour and location. You may bring two pages of notes (8.5 × 11, both sides, any font). The exam questions will focus on fundamental ideas and examples covered in the lectures and homework.

Note 2: The "first installment" of HW6 posted 12/2/04 had just 4 problems, corresponding to material covered in lecture on 11/29/04. This "complete version" consists of those 4 problems plus two more on material covered 12/6/04.

1) [Jarrow-Turnbull chapter 14, problems 1 and 2, somewhat modified.] Suppose the LIBOR discount rate $B(0,t)$ are given by the table below. Consider a 3-year swap whose floating payments are at the then-current LIBOR rate, and whose fixed payments are at the term rate of $R_{\rm fix}$ per annum.

payment date t_i	$B(0,t_i)$
0.5	.9748
1.0	.9492
1.5	.9227
2.0	.8960
2.5	.8687
3.0	.8413

(a) Suppose $R_{\rm fix}$ is 6.5 percent per annum and the notional principal is 1 million dollars. What is the value of the swap?

(b) What is the par swap rate? In other words: what value of $R_{\rm fix}$ sets the value of the swap to 0?

2) There are two ways to think about the value of a swap:

(i) One approach (presented in class on 11/29/04) views the swap as a collection of forward rate agreements. If the payment dates are $0 < t_1 < \ldots < t_N$ and L is the notional principal, this approach gives

$$\text{swap value} = \sum_{i=1}^{N} B(0,t_i)[R_{\rm fix} - f_0(t_{i-1},t_i)](t_i - t_{i-1})L$$

where $f_0(t_{i-1},t_i)$ is the forward term rate for lending from t_{i-1} to t_i, defined by

$$f_0(t_{i-1},t_i)(t_i - t_{i-1}) = \frac{B(0,t_{i-1})}{B(0,t_i)} - 1.$$

(ii) The other approach (implicit but not explicit in the Section 10 notes) views the swap as a long position in a coupon bond paying $R_{\rm fix}$ plus a short position in a floating rate bond. This approach gives the formula

$$\text{swap value} = \sum_{i=1}^{N} B(0,t_i)R_{\rm fix}(t_i - t_{i-1})L - (1 - B(0,t_N))L.$$

Show that these two approaches are consistent, i.e. the swap values given in (i) and (ii) above are equal.

3) [Hull, Chapter 22, problem 28, slightly modified.] Calculate the price of a cap on the three-month LIBOR rate in nine months' time when the principal amount is $1000. Use Black's model and the following information:

- The nine-month Eurodollar futures price is 92 (ignore the difference between forwards and futures).
- The interest rate volatility implied by a nine-month Eurodollar option is 15 percent per annum.
- The current 12-month interest rate with continuous compounding is 7.5 percent per annum.
- The cap rate is 8 percent per annum.

[See the Section 11 notes for help interpreting this jargon.]

4) [Hull, Chapter 22, problem 29] Suppose the LIBOR yield curve is flat at 8% with annual compounding. Consider a swaption that gives its holder the right to receive 7.6% in a five-year swap starting in four years. Payments are made annually. The volatility for the swap rate is 25% per annum and the principal is $1 million. Use Black's model to price the swaption.

5) [Hull, Chapter 26, problem 16, slightly modified]. Suppose the risk-free yield curve is flat at 6% with annual compounding. One-year, two-year, and three-year bonds yield 7.2%, 7.4%, and 7.6% with annual compounding. All pay 6% coupons. Assume that in case of default the recovery is 40% of principal, with no payment of accrued interest. Find the risk-neutral probability of default during each year.

6) [Hull, Chapter 27, problem 20, slightly modified.] Suppose the risk-free yield curve is flat at 6% per annum with continuous compounding, and defaults can occur at times 1 year, 2 years, 3 years, and 4 years in a four-year plain vanilla credit default swap with semiannual payments. Suppose the recovery rate is 20% and the probabilities of default at times 1 yr, 2yrs, 3yrs, and 4yrs are .01, .015, .02, and .025 respectively. The reference obligation is a bond paying a coupon semiannually of 8% per year. Assume any default takes place immediately before a coupon date, and the recovery does not include any accrued interest. What is the credit default swap spread?

Derivative Securities – Homework 6 Solutions
Distributed 12/13/04

Problem 1.

(a) The value of the fixed-rate payments is

$$R_{\text{fix}} L\Delta t(B(0,.5) + \cdots + B(0,3)) = R_{\text{fix}} L(.5)(.9748 + \cdots + .8413) = 2.72635 R_{\text{fix}} L.$$

Since $R_{\text{fix}} = .065$ this is $.1772128L$. The value of the floating payments is

$$[1 - B(0,3)]L = .1587L.$$

Therefore the value of the swap is $(.1772128 - .1587)L = .0185128L$. Since $L = 10^6$ the value is 18,512.8 dollars.

(b) Using the calculations in part (a), we see that the par swap rate solves

$$2.72635 R_{\text{fix}} = .1587.$$

This gives $R_{\text{fix}} = .0582097$, in other words (approximately) 5.82 percent.

Problem 2. We just need to show that the different expressions for the floating sides are equal. Using the fact that $f_0(t_{i-1}, t_i)(t_i - t_{i-1}) = \frac{B(0,t_{i-1})}{B(0,t_i)} - 1$ we get

$$\begin{aligned}
V_{float} &= \sum_{i=1}^{N} B(0,t_i) f_0(t_{i-1}, t_i)(t_i - t_{i-1})L \\
&= L \sum_{i=1}^{N} B(0,t_i) \left[\frac{B(0,t_{i-1})}{B(0,t_i)} - 1 \right] \\
&= L \sum_{i=1}^{N} \left[B(0,t_{i-1}) - B(0,t_i) \right] \\
&= L \left[B(0,0) - B(0,t_N) \right] \\
&= L \left[1 - B(0,t_N) \right].
\end{aligned}$$

Thus the two expressions for the floating side are equal, and the two swap values are consistent.

Problem 3. There is only one payment date, so this cap is really a caplet. According to Black's model its value is

$$B(0,T')L\Delta T \left[f_0 N(d_1) - R_K N(d_2) \right]$$

with

$$\begin{aligned}
T' &= \text{time when funds are received} = 1 \text{ year} \\
B(0,T') &= e^{-.075} = .9277 \\
L &= \text{principal} = 1000 \text{ dollars} \\
\Delta T &= \text{term of loan} = .25 \text{ year} \\
R_K &= \text{cap rate} = .08 \text{ per annum} \\
f_0 &= \text{forward term rate} = .08 \text{ per annum.}
\end{aligned}$$

Brief explanation about f_0: we are told that the Eurodollar futures price is 92. We suppose the Eurodollar futures price refers to a 3-month contract (this is the market convention). Ignoring the difference between forwards and futures means viewing the futures price as a forward term rate; this means we can secure, at no cost now, the right to a 3 month Eurodollar contract starting 9 months from now at 100-92=8 percent per annum (another market convention). Therefore $f_0 = .08$.

In the definition of d_1 and d_2 we are given $\sigma = .15$ per annum, and the option's maturity date is $T = .75$ years. Since $f_0 = R_K$ we have $\log(f_0/R_K) = 0$ and

$$d_1 = \tfrac{1}{2}\sigma\sqrt{T} = (.5)(.15)\sqrt{.75} = .0650, \quad d_2 = d_1 - \sigma\sqrt{T} = -\tfrac{1}{2}\sigma\sqrt{T} = -.0650.$$

Since $N(.0650) = .5259$ and $N(-.0650) = .4741$ we conclude that

$$\text{value } = (.9277)(1000)(.25)[.08 \cdot .5259 - .08 \cdot .4741] = 0.96.$$

The present value of the cap is 96 cents.

Problem 4. The swap gives its holder the right to *receive* the fixed rate. So it is in the money if $R_K > R_{\text{swap}}$ when the option matures. Therefore it is like a *put* on R_{swap}. This is different from the example in the Section 11 notes – we must use the Black-Scholes formula for a put rather than a call. Black's formula becomes is

$$\text{swaption value } = LA\left[R_K N(-d_2) - F_{\text{swap}} N(-d_1)\right]$$

with

$$\begin{aligned}
A &= \sum B(0,t_i)\Delta_i t = (1.08)^{-5} + \cdots + (1.08)^{-9} = 2.9348 \\
L &= 10^6 \text{ dollars} \\
R_K &= .076 \\
d_1 &= \frac{\log(F_{\text{swap}}/R_K) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} \\
d_2 &= d_1 - \sigma\sqrt{T}.
\end{aligned}$$

In the calculation of d_1 and d_2 we are given $\sigma = .25$, and $T = 4$ is the maturity date of the option. It remains to find F_{swap}. Since the yield curve is flat, the forward term rate for any future period is 8 percent per annum (compounded annually), so it is intuitively clear that $F_{\text{swap}} = .08$. Of course our formula gives the same result: it determines F_{swap} by solving

$$F_{\text{swap}}\left[(1.08)^{-1} + \cdots (1.08)^{-5}\right] = 1 - (1.08)^{-5}.$$

Applying the identity $\rho + \cdots + \rho^5 = \frac{\rho(1-\rho^5)}{1-\rho}$ with $\rho = (1.08)^{-1}$ we get $F_{\text{swap}} = \frac{1}{\rho} - 1 = .08$ as expected. Now

$$d_1 = \frac{\log(.08/.076) + \frac{1}{2}(.25)^2(4)}{(.25)(2)} = .3526, \quad d_2 = d_1 - (.25)(2) = -.1474,$$

and using $N(.1474) = .5586$ and $N(-.3526) = .3622$ we get

$$\text{value of swaption } = (2.9348)L[(.076)(.5586) - (.08)(.3622)] = .039554L.$$

The swaption is worth $39,554.

Problem 5.

We assume the coupons are paid annually. Moreover if a bond defaults during a certain year, we assume the coupon due at the end of the year is not paid, and the recovery (40% of principal) is paid at the end of the year.

To begin, let's find the risky bond prices from the yield information in the problem. We use the standard relationship between a bond's yield and its price (ignoring the possibility of default); for example the yield y of a two-year bond with principal P and coupon payment c is related to the price of the bond by

$$\text{price} = c/(1+y) + c/(1+y)^2 + P/(1+y)^2.$$

Using the data in the problem, for bonds with principal one dollar, we find:

$$\begin{aligned}
\text{price of 1-yr bond} &= .06/1.072 + 1/1.072 = .9888 \\
\text{price of 2-yr bond} &= .06/1.074 + .06/(1.074)^2 + 1/(1.074)^2 = .9748 \\
\text{price of 3-yr bond} &= .06/1.076 + .06/(1.076)^2 + .06/(1.076)^3 + 1/(1.076)^3 = .9585.
\end{aligned}$$

Now we find the default probabilities from the fact that these bond prices must be the RN-expected discounted values of the associated cash flows. Let p_i be the probability that a default occurs in year i.

Focusing on the one-year bond first, we have

$$.9888 = (1 - p_1)(.06/1.06 + 1/1.06) + p_1(.4/1.06) = 1 - p_1 + .3774p_1$$

so $p_1 = (.9888 - 1)/(.3774 - 1) = .0180$.

Considering the two-year bond next, we have

$$.9748 = p_1(.4/1.06) + p_2[.06/1.06 + .4/(1.06)^2] + (1 - p_1 - p_2)[1]$$

using for the last term the fact that $.06/1.06 + .06/(1.06)^2 + 1/(1.06)^2 = 1$. Since we know p_1, this reduces to a linear equation for p_2, namely

$$.9748 = .0068 + .4126p_2 + (.9820 - p_2).$$

It follows that $p_2 = (.9748 - .0068 - .9820)/(.4126 - 1) = .0238$.

Finally we use the three-year bond price to find p_3. We have

$$\begin{aligned} .9585 &= p_1(.4/1.06) + p_2[.06/1.06 + .4/(1.06)^2] \\ &\quad + p_3[.06/1.06 + .06/(1.06)^2 + .4/(1.06)^3] + (1 - p_1 - p_2 - p_3)[1]. \end{aligned}$$

Using the known values of p_1 and p_2 this becomes

$$.9585 = .0166 + .4459p_3 + (.9582 - p_3)$$

so $p_3 = (.9585 - .0166 - .9582)/(.4459 - 1) = .0294$.

Notice that the probabilities p_i used above are actual probabilities, not conditional probabilities. They can be written in terms of the conditional probabilities of default (also called hazard rates)

$$q_i = \text{prob of default in year } i\text{, given that the bond survived to year } i$$

by

$$\begin{aligned} p_1 = \text{prob of default in year 1} &= q_1 \\ p_2 = \text{prob of default in year 2} &= (1 - q_1)q_2 \\ p_3 = \text{prob of default in year 3} &= (1 - q_1)(1 - q_2)q_3. \end{aligned}$$

It was most natural for this problem to solve directly for the unconditional probabilities p_i; but the conditional probabilities q_i can easily be found, if desired, from the formulas just above.

Problem 6.

As in the notes, the CDS spread s is given by

$$s = \frac{(1-R)\sum_{i=1}^{n} p_i B_*(0, t_i)}{\pi u(T) + \sum_{i=1}^{n} p_i u(t_i)}$$

where R is the recovery rate, p_i are the risk-neutral probabilites of default (extracted from bond prices presumably), B_* is the risk-free bond price, π is the probability of no default, and $u(t_i)$ is the present value of payments per payment dollar. With semi-annual payments, this becomes

$$u(t_i) = \sum_{n=1}^{i} \left(\frac{1}{2} B_*(0, n - 1/2) + \frac{1}{2} B_*(0, n) \right)$$

Plugging in $i = 1, \ldots, 4$ and using 6% continuous compounding we get $u(1) = 0.9561$, $u(2) = 1.8565$, $u(3) = 2.7045$, and $u(4) = 3.5031$. Then

$$\begin{aligned} s &= \frac{(1-R)\sum_{i=1}^{4} p_i B_*(0, i)}{\pi u(4) + \sum_{i=1}^{4} p_i u(i)} \\ &= \frac{0.0473}{3.4370} \\ &= 0.0138 \end{aligned}$$

So, since payments are semiannual, the terms of the CDS under the CDS spread s should be that the buyer pays \$0.0138/2 per dollar face value of the underlying defaultable bond every six months.

Derivative Securities Final Exam

Fall 2004 – G63.2791 – Professor Kohn

- You may bring two 8.5×11 pages of notes [both sides] to this exam.
- Put your answers on the exam paper; use the back of the page if you need more space, and attach additional sheets if necessary. I will grade *only* your exam paper, not your scratch paper.
- Part A consists of 3 "longer-answer" problems, worth 20 points each. Part B consists of 10 "shorter-answer" problems, worth 10 each. The total possible score is thus 160.
- Show your work, and explain all answers (at least briefly). Partial credit will be given for correct ideas.

NAME: ____________________________

A) Longer-answer problems: 20 points each

A1) _______
A2) _______
A3) _______

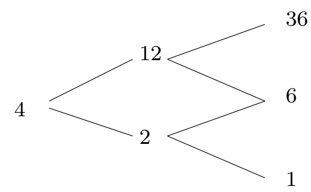
B) Shorter-answer problems: 10 points each

B1) _______	B6) _______
B2) _______	B7) _______
B3) _______	B8) _______
B4) _______	B9) _______
B5) _______	B10) _______

Total: _____________

Part A: Longer-answer questions. Each of the 3 problems in Part A has several parts, worth a total of 20 points.

1. (20 points) Suppose the price of a non-dividend-paying stock is restricted to the multiplicative binomial tree shown below. Notice that $u = 3$, $d = 1/2$. Assume the risk-free



rate satisfies $e^{r\delta t} = 2$. Consider a European put option with strike price $K = 3$.

(a) Find the value of this option by working backward in the tree.

(b) Specify the replicating (hedge) portfolio at time $t = 0$.

(c) Suppose the stock goes up to 12 in the first time period. How should the replicating portfolio be changed?

(d) Now consider the associated American put – with the same strike and underlying, but permitting early exercise. Is its value at time 0 different? Explain briefly.

2. (20 points) Consider a non-dividend-paying stock whose price at time t is $s(t) = e^{y(t)}$ where y solves the stochastic differential equation $dy = \mu y\, dt + \sigma y\, dw$. Notice that $y = \log s$.

   (a) What stochastic differential equation does s solve?

   (b) Consider an option on this stock, and assume its value at time t has the form $V(s(t), t)$. Find the associated hedge, by determining the choice of ϕ that makes $dV - \phi ds$ have no "dw" term.

   (c) What partial differential equation does V solve?

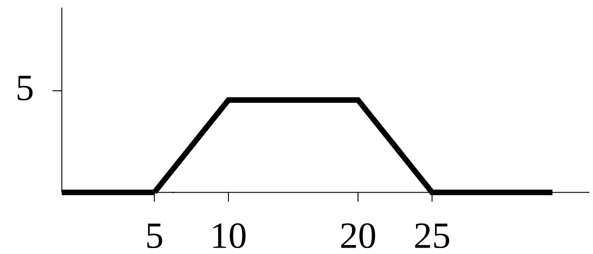
3. Suppose y solves the stochastic differential equation $dy = \mu y\, dt + \sigma y dw$ with $y(0) = y_0$.

(a) Show that if $V(y,t)$ solves $V_t + \mu y V_y + \frac{1}{2}\sigma^2 y^2 V_{yy} - rV = 0$ then $e^{-rt}V(y,t)$ is a martingale.

(b) Now suppose that in addition to the PDE, V satisfies the final-time condition $V(y,T) = f(y)$. Show that $V(y_0, 0) = e^{-rT} E_{y(0)=y_0}[f(y(T))]$.

Part B: Shorter-answer questions. The problems in Part B can be answered relatively briefly; they are worth 10 points each.

1. (10 points) Consider the payoff $f(s_T)$ sketched below. How can it be achieved by a combination of call options?



[Note: the payoff is 0 for $s_T < 5$ and for $s_T > 25$; it equals 5 for $10 < s_T < 20$; its slope is 1 for $5 < s_T < 10$; and its slope is -1 for $20 < s_T < 25$.]

2. (10 points) Let $B(0, T)$ be the price in dollars of a risk-free bond worth one dollar at time T, and let $D(0, T)$ be the price in Euros of a risk-free bond worth one Euro at time T. Consider a forward contract, which obligates the holder to buy 100 Euros at K dollars per Euro, at time T. Express its value today (time 0) in terms of the current exchange rate S_0 (in dollars per Euro). Briefly justify your answer.

3. (10 points) Consider the trading strategy that replicates the payoff of a European call on a non-dividend-paying stock with lognormal dynamics. Does it ever require you to take a short position in the underlying? Explain briefly.

4. (10 points) We learned that to hedge an option with value $V(s(t), t)$ you should hold $-\Delta$ units of the underlying, where $\Delta = \partial V / \partial s$. Suppose rather than the underlying, you wish to hedge using futures on the underlying. What should your futures position be at time t? (Assume the risk-free rate is constant r.)

5. (10 points) Consider a non-dividend-paying stock with lognormal dynamics, $ds = \mu s\,dt + \sigma s\,dw$. The risk-free rate is r (assumed constant). Consider the digital option that's worth 1 at time T if $s_T > K$ and 0 otherwise. What is its value at time 0?

6. (10 points) Consider a lognormal stock with continuous dividend yield d. (This means the stock price satisfies $ds = \mu s\,dt + \sigma s dw$, and if you start with one share at time 0 and reinvest all dividends in stock you'll have e^{dt} shares at time t.) Show that early exercise can be optimal for an American call on such a stock.

7. (10 points) Suppose the yield curve is flat at 5%. Consider the following two-year swap: it starts at the end of year 1; the holder pays 6% per annum and receives the floating rate for years 2 and 3; payments are annual. What is the present value of this swap?

8. (10 points) Consider a floorlet with fixed rate R_K for lending one year from now with maturity two years from now. Explain how Black's formula specifies its value now (at time 0) in terms of the notional principal L, the current prices of zero-coupon bonds $B(0, T)$, and a suitable volatility parameter.

9. (10 points) When we use Black's formula to value options with maturity T on an underlying V, we are asserting (or assuming) the existence of a probability measure such that (a) the underlying is lognormal, and (b) the option with payoff $f(V_T)$ has value $B(0,T)E[f(V_T)]$. Show that if these hypotheses hold then the $E[V_T]$ must be the forward price of V for delivery at time T.

10. (10 points) This problem concerns credit risk. Let $B(0,T)$ be the price of a risk-free zero-coupon bond worth T at maturity. Consider two-year corporate bond with principal L, paying annual coupons at 5 percent per annum. Let p_1 be the risk-free probability of default in year 1, and p_2 the risk-free probability of default in year 2. Assume that after default there is no recovery of principal or interest. What is the present value of the corporate bond?