m3f33chV

Chapter V: MATHEMATICAL FINANCE IN DISCRETE TIME

We follow [BK], Ch. 4 (or see the other sources cited in Ch0).

§1. The Model.

It suffices, to illustrate the ideas, to work with a *finite* probability space $(\Omega, \mathcal{F}, P)$, with a finite number $|\Omega|$ of points ω, each with positive probability: $P(\{\omega\}) > 0$. We will use a finite time-horizon N, which will correspond to the expiry date of the options.

As before, we use a filtration $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_N$: we may (and shall) take $\mathcal{F}_0 = \{\emptyset, \Omega\}$, the trivial σ-field, $\mathcal{F}_N = \mathcal{F} = \mathcal{P}(\Omega)$ (here $\mathcal{P}(\Omega)$ is the *power-set* of Ω, the class of all $2^{|\Omega|}$ subsets of Ω: we need every possible subset, as they all (apart from the empty set) carry positive probability.

The financial market contains $d+1$ financial assets: a riskless asset (bank account) labelled 0, and d risky assets (stocks, say) labelled 1 to d. The prices of the assets at time n are random variables, $S_n^0, S_n^1, \cdots, S_n^d$ say [note that we use superscripts here as labels, *not* powers, and suppress ω for brevity], non-negative and $\mathcal{F}_n$-measurable [at time n, we know the prices S_n^i].

We take $S_0^0 = 1$ (that is, we reckon in units of our initial bank holding). We assume for convenience a constant interest rate $r > 0$ in the bank, so 1 unit in the bank at time 0 grows to $(1+r)^n$ at time n. So $1/(1+r)^n$ is the *discount factor* at time n.

Definition. A *trading strategy* H is a vector stochastic process $H = (H_n)_{n=0}^N = ((H_n^0, H_n^1, \cdots, H_n^d))_{n=0}^N$ which is *predictable* (or *previsible*): each H_n^i is $\mathcal{F}_{n-1}$-measurable for $n \geq 1$.

Here H_n^i denotes the number of shares of asset i held in the portfolio at time n – to be determined on the basis of information available *before* time n; the vector $H_n = (H_n^0, H_n^1, \cdots, H_n^d)$ is the *portfolio* at time n. Writing $S_n = (S_n^0, S_n^1, \cdots, S_n^d)$ for the vector of prices at time n, the *value* of the portfolio at time n is the scalar product

$$V_n(H) = H_n.S_n := \Sigma_{i=0}^d H_n^i S_n^i.$$

The *discounted value* is

$$\tilde{V}_n(H) = \beta_n(H_n.S_n) = H_n.\tilde{S}_n,$$

where $\beta_n := 1/S_n^0$ and $\tilde{S}_n = (1, \beta_n S_n^1, \cdots, \beta_n S_n^d)$ is the vector of discounted prices.
Note. The *previsibility* of H reflects that there is no *insider trading.*

Definition. The strategy H is *self-financing (SF)*, $H \in SF$, if

$$H_n.S_n = H_{n+1}.S_n \qquad (n = 0, 1, \cdots, N-1).$$

Interpretation. When new prices S_n are quoted at time n, the investor adjusts his portfolio from H_n to H_{n+1}, without bringing in or consuming any wealth.

$$\begin{aligned} V_{n+1}(H) - V_n(H) &= H_{n+1}.S_{n+1} - H_n.S_n \\ &= H_{n+1}.(S_{n+1} - S_n) + (H_{n+1}.S_n - H_n.S_n). \end{aligned}$$

For a SF strategy, the second term on the right is zero. Then the LHS, the net increase in the value of the portfolio, is shown as due only to the price changes $S_{n+1} - S_n$. So for $H \in SF$,

$$V_n(H) - V_{n-1}(H) = H_n(S_n - S_{n-1}),$$

$$\Delta V_n(H) = H_n.\Delta S_n, \qquad V_n(H) = V_0(H) + \Sigma_1^n H_j.\Delta S_j$$

and $V_n(H)$ is the *martingale transform* of S by H (IV.6). Similarly with discounting:

$$\Delta\tilde{V}_n(H) = H_n.\Delta\tilde{S}_n, \qquad \tilde{V}_n(H) = V_0(H) + \Sigma_1^n H_j.\Delta\tilde{S}_j$$

$(\Delta\tilde{S}_n := \tilde{S}_n - \tilde{S}_{n-1} = \beta_n S_n - \beta_{n-1} S_{n-1})$.

As in II, we are allowed to borrow (so S_n^0 may be negative) and sell short (so S_n^i may be negative for $i = 1, \cdots, d$). So it is hardly surprising that if we decide what to do about the risky assets, the bank account will take care of itself, in the following sense.

Proposition. If $((H_n^1, \cdots, H_n^d))_{n=0}^N$ is predictable and V_0 is $\mathcal{F}_0$-measurable, there is a unique predictable process $(H_n^0)_{n=0}^N$ such that $H = (H^0, H^1, \cdots, H^d)$ is SF with initial value V_0.

Proof. If H is SF, then as above

$$\tilde{V}_n(H) = H_n.\tilde{S}_n = H_n^0 + H_n^1\tilde{S}_n^1 + \cdots + H_n^d\tilde{S}_n^d,$$

while as $\tilde{V}_n = H.\tilde{S}_n$,

$$\tilde{V}_n(H) = V_0 + \Sigma_1^n(H_j^1\Delta\tilde{S}_j^1 + \cdots + H_j^d\Delta\tilde{S}_j^d)$$

($\tilde{S}_n = (1, \beta_n S_n^1, \cdots, \beta_n S_n^d)$, so $\tilde{S}_n^0 \equiv 1$, $\Delta\tilde{S}_n^0 = 0$). Equate these:

$$H_n^0 = V_0 + \Sigma_1^n(H_j^1\Delta\tilde{S}_j^1 + \cdots + H_j^d\Delta\tilde{S}_j^d) - (H_n^1\tilde{S}_n^1 + \cdots + H_n^d\tilde{S}_n^d),$$

which defines H_n^0 uniquely. The terms in $\tilde{S}_n^i$ are $H_n^i\Delta\tilde{S}_n^i - H_n^i\tilde{S}_n^i = -H_n^i\tilde{S}_{n-1}^i$, which is $\mathcal{F}_{n-1}$-measurable. So

$$H_n^0 = V_0 + \Sigma_1^{n-1}(H_j^1\Delta\tilde{S}_j^1 + \cdots + H_j^d\Delta\tilde{S}_j^d) - (H_n^1 S_{n-1}^1 + \cdots + H_n^d\tilde{S}_{n-1}^d),$$

where as $H^1, \cdots, H^d$ are predictable, all terms on the RHS are $\mathcal{F}_{n-1}$-measurable, so H^0 is predictable. //

Numéraire. What units do we reckon value in? All that is really necessary is that our chosen unit of account should always be *positive* (as we then reckon our holdings by dividing by it, and one cannot divide by zero). Common choices are pounds sterling (UK), dollars (US), euros etc. Gold is also possible (now priced in sterling etc. – but the pound sterling represented an amount of gold, till the UK 'went off the gold standard'). By contrast, risky stocks *can* have value 0 (if the company goes bankrupt). We call such an always-positive asset, used to reckon values in, a *numéraire*.

Of course, one has to be able to change numéraire – e.g. when going from UK to the US or eurozone. As one would expect, this changes nothing important. In particular, we quote (*numéraire invariance theorem* – see e.g. [BK] Prop. 4.1.1) that the set SF of self-financing strategies is invariant under change of numéraire.

Note. 1. This alerts us to what is meant by 'risky'. To the owner of a goldmine, sterling is risky. The danger is not that the UK government might go bankrupt, but that sterling might depreciate against the dollar, or euro, etc.
2. With this understood, we shall feel free to refer to our numéraire as 'bank account'. The point is that we don't trade in it (why would a goldmine owner trade in gold?); it is the other – 'risky' – assets that we trade in.

§2. Viability (NA): Existence of Equivalent Martingale Measures.

Although we are allowed to borrow (from the bank), and sell (stocks) short, we are – naturally – required to stay solvent (recall that trading while insolvent is an offence under the Companies Act!).

Definition. A strategy H is *admissible* if it is self-financing (SF), and $V_n(H) \geq 0$ for each time $n = 0, 1, \cdots, N$.

Recall that arbitrage is riskless profit – making 'something out of nothing'. Formally:

Definition. An *arbitrage strategy* is an admissible strategy with zero initial value and positive probability of a positive final value.
Definition. A market is *viable* if no arbitrage is possible, i.e. if the market is arbitrage-free (no-arbitrage, NA).

This leads to the first of two fundamental results:

Theorem (No-Arbitrage Theorem: NA iff EMMs exist). The market is viable (is arbitrage-free, is NA) iff there exists a probability measure P^* equivalent to P (i.e., having the same null sets) under which the discounted asset prices are P^*-martingales – that is, iff there exists an equivalent martingale measure (EMM).

Proof. $\Leftarrow$. Assume such a P^* exists. For any self-financing strategy H, we have as before

$$\tilde{V}_n(H) = V_0(H) + \Sigma_1^n H_j.\Delta\tilde{S}_j.$$

By the Martingale Transform Lemma, $\tilde{S}_j$ a (vector) P^*-martingale implies $\tilde{V}_n(H)$ is a P^*-martingale. So the initial and final P^*-expectations are the same: using E^* for P^*-expectation,

$$E^*[\tilde{V}_N(H)] = E^*[\tilde{V}_0(H)].$$

If the strategy is admissible and its initial value – the RHS above – is zero, the LHS $E^*[\tilde{V}_N(H)]$ is zero, but $\tilde{V}_N(H) \geq 0$ (by admissibility). Since each $P(\{\omega\}) > 0$ (by assumption), each $P^*(\{\omega\}) > 0$ (by equivalence). This and $\tilde{V}_N(H) \geq 0$ force $\tilde{V}_N(H) = 0$ (sum of non-negatives can only be 0 if each term is 0). So no arbitrage is possible. //

The converse is true, but harder, and needs a preparatory result – which is interesting and important in its own right.

Separating Hyperplane Theorem (SHT).

In a vector space V, a *hyperplane* is a translate of a (vector) subspace U of codimension 1 – that is, U and some one-dimensional subspace, say $\mathbb{R}$, together span V: V is the direct sum $V = U \oplus \mathbb{R}$ (e.g., $\mathbb{R}^3 = \mathbb{R}^2 \oplus \mathbb{R}$). Then

$$H = [f, \alpha] := \{x : f(x) = \alpha\}$$

for some α and linear functional f. In the finite-dimensional case, of dimension n, say, one can think of $f(x)$ as an inner product,

$$f(x) = f.x = f_1 x_1 + \ldots + f_n x_n.$$

The hyperplane $H = [f, \alpha]$ *separates* sets $A, B \subset V$ if

$$f(x) \geq \alpha \qquad \forall\, x \in A, \qquad f(x) \leq \alpha \qquad \forall\, x \in B$$

(or the same inequalities with A, B, or $\geq, \leq$, interchanged).

Call a set A in a vector space V *convex* if

$$x, y \in A, \quad 0 \leq \lambda \leq 1 \quad \Rightarrow \quad \lambda x + (1 - \lambda) y \in A$$

– that is, A contains the line-segment joining any pair of its points.

We can now state (without proof) the SHT (see e,g, [BK] App. C).

SHT. Any two non-empty disjoint convex sets in a vector space can be separated by a hyperplane.

A *cone* is a subset of a vector space closed under vector addition and multiplication by *positive* constants (so: like a vector subspace, but with a sign-restriction in scalar multiplication).

We turn now to the proof of the converse.

Proof of the converse (not examinable). $\Rightarrow$: Write Γ for the cone of strictly positive random variables. Viability (NA) says that for any admissible strategy H,

$$V_0(H) = 0 \qquad \Rightarrow \qquad \tilde{V}_N(H) \notin \Gamma. \qquad (*)$$

To any admissible process $(H_n^1, \cdots, H_n^d)$, we associate its discounted cumulative *gain* process

$$\tilde{G}_n(H) := \Sigma_1^n (H_j^1 \Delta \tilde{S}_j^1 + \cdots + H_j^d \Delta \tilde{S}_j^d).$$

By the Proposition, we can extend $(H_1, \cdots, H_d)$ to a unique predictable process (H_n^0) such that the strategy $H = ((H_n^0, H_n^1, \cdots, H_n^d))$ is self-financing with initial value zero. By NA, $\tilde{G}_N(H) = 0$ – that is, $\tilde{G}_N(H) \notin \Gamma$.

We now form the set $\mathcal{V}$ of random variables $\tilde{G}_N(H)$, with $H = (H^1, \cdots, H^d)$ a previsible process. This is a vector subspace of the vector space $\mathbb{R}^\Omega$ of random variables on Ω, by linearity of the gain process $G(H)$ in H. By $(*)$, this subspace $\mathcal{V}$ does not meet Γ. So $\mathcal{V}$ does not meet the subset

$$K := \{X \in \Gamma : \Sigma_\omega X(\omega) = 1\}.$$

Now K is a convex set not meeting the origin. By the Separating Hyperplane Theorem, there is a vector $\lambda = (\lambda(\omega) : \omega \in \Omega)$ such that for all $X \in K$

$$\lambda.X := \Sigma_\omega \lambda(\omega) X(\omega) > 0, \tag{1}$$

but for all $\tilde{G}_N(H)$ in $\mathcal{V}$,

$$\lambda.\tilde{G}_N(H) = \Sigma_\omega \lambda(\omega) \tilde{G}_N(H)(\omega) = 0. \tag{2}$$

Choosing each $\omega \in \Omega$ successively and taking X to be 1 on this ω and zero elsewhere, (1) tells us that each $\lambda(\omega) > 0$. So

$$P^*(\{\omega\}) := \lambda(\omega)/(\Sigma_{\omega' \in \Omega} \lambda(\omega'))$$

defines a probability measure equivalent to P (no non-empty null sets). With E^* as P^*-expectation, (2) says that

$$E^*[\tilde{G}_N(H)] = 0 : \qquad E^*[\Sigma_1^N H_j.\Delta \tilde{S}_j] = 0.$$

In particular, choosing for each i to hold only stock i,

$$E^*[\Sigma_1^N H_j^i \Delta \tilde{S}_j^i] = 0 \qquad (i = 1, \cdots, d).$$

By the Martingale Transform Lemma, this says that the discounted price processes $(\tilde{S}_n^i)$ are P^*-martingales. //

§3. Complete Markets: Uniqueness of EMMs.

A contingent claim (option, etc.) can be defined by its *payoff* function, h say, which should be non-negative (options confer rights, not obligations, so negative values are avoided by not exercising the option), and $\mathcal{F}_N$-measurable (so that we know how to evaluate h at the terminal time N).

Definition. A contingent claim defined by the payoff function h is *attainable* if there is an admissible strategy worth (i.e., replicating) h at time N. A market is *complete* if every contingent claim is attainable.

Theorem (Completeness Theorem: complete iff EMM unique). A viable market is complete iff there exists a unique probability measure P^* equivalent to P under which discounted asset prices are martingales – that is, iff equivalent martingale measures are unique.

Proof. $\Rightarrow$: Assume viability and completeness. Then for any $\mathcal{F}_N$-measurable random variable $h \geq 0$, there exists an admissible (so SF) strategy H replicating h: $h = V_N(H)$. As H is SF, by §1

$$h/S_N^0 = \tilde{V}_N(H) = V_0(H) + \Sigma_1^N H_j.\Delta\tilde{S}_j.$$

We know by the Theorem of §2 that an equivalent martingale measure P^* exists; we have to prove uniqueness. So, let P_1, P_2 be two such equivalent martingale measures. For $i = 1, 2$, $(\tilde{V}_n(H))_{n=0}^N$ is a P_i-martingale. So,

$$E_i[\tilde{V}_N(H)] = E_i[V_0(H)] = V_0(H),$$

since the value at time zero is non-random ($\mathcal{F}_0 = \{\emptyset, \Omega\}$). So

$$E_1[h/S_N^0] = E_2[h/S_N^0].$$

Since h is arbitrary, E_1, E_2 have to agree on integrating all non-negative integrands. Taking negatives and using linearity: they have to agree on non-positive integrands also. Splitting an arbitrary integrand into its positive and negative parts: they have to agree on all integrands. Now E_i is expectation (i.e., integration) with respect to the measure P_i, and measures that agree on integrating all integrands must coincide. So $P_1 = P_2$. //

Before proving the converse, we prove a lemma. Recall that an admissible strategy is a SF strategy with all values non-negative. The Lemma shows that the non-negativity of contingent claims extends to all values of any SF strategy replicating it – in other words, this gives equivalence of admissible and SF replicating strategies. [SF: isolated from external wealth; admissible:

actually worth something. These sound similar; the Lemma shows they are the same here. So we only need one term; we use SF as it is shorter, but w.l.o.g. this means admissible also.]

Lemma. In a viable market, any attainable h (i.e., any h that can be replicated by a SF strategy H) can also be replicated by an admissible strategy H.

Proof. If H is SF and P^* is an equivalent martingale measure under which discounted prices $\tilde{S}$ are P^*-martingales (such P^* exist by viability and the Theorem of §2), $\tilde{V}_n(H)$ is also a P^*-martingale, being the martingale transform of $\tilde{S}$ by H (see §1). So

$$\tilde{V}_n(H) = E^*[\tilde{V}_N(H)|\mathcal{F}_n] \qquad (n = 0, 1, \cdots, N).$$

If H replicates h, $V_N(H) = h \geq 0$, so discounting, $\tilde{V}_N(H) \geq 0$, so the above equation gives $\tilde{V}_n(H) \geq 0$ for each n. Thus *all* the values at each time n are non-negative – not just the final value at time N – so H is admissible. //

Proof of the Theorem (continued). $\Leftarrow$ (*not examinable*): Assume the market is viable but incomplete: then there exists a non-attainable $h \geq 0$. By the Proposition of §1, we may confine attention to the risky assets $S^1, \cdots, S^d$, as these suffice to tell us how to handle the bank account S^0.

Call $\tilde{\mathcal{V}}$ the set of random variables of the form

$$U_0 + \Sigma_1^N H_n.\Delta\tilde{S}_n$$

with U_0 $\mathcal{F}_0$-measurable (i.e. deterministic) and $((H_n^1, \cdots, H_n^d))_{n=0}^N$ predictable; this is a vector space. (Here $(H^1, \ldots, H^d)$ extends to $H := (H^0, H^1, \ldots, H^d)$, by the Proposition of §1, and H can be any strategy here.) Then as h is not attainable, the discounted value h/S_N^0 does not belong to $\tilde{\mathcal{V}}$, so $\tilde{\mathcal{V}}$ is a *proper* subspace of the vector space $\mathbb{R}^\Omega$ of all random variables on Ω. Let P^* be a probability measure equivalent to P under which discounted prices are martingales (such P^* exist by viability, by the Theorem of §2). Define the scalar product

$$(X, Y) \to E^*[XY]$$

on random variables on Ω. Since $\tilde{\mathcal{V}}$ is a proper subspace, by Gram-Schmidt orthogonalisation there exists a non-zero random variable X orthogonal to $\tilde{\mathcal{V}}$. That is,

$$E^*[X] = 0.$$

Write $\|X\|_\infty := \max\{|X(\omega)| : \omega \in \Omega\}$, and define P^{**} by

$$P^{**}(\{\omega\}) = \Big(1 + \frac{X(\omega)}{2\|X\|_\infty}\Big) P^*(\{\omega\}).$$

By construction, P^{**} is equivalent to P^* (same null-sets - actually, as $P^* \sim P$ and P has no non-empty null-sets, neither do P^*, P^{**}). As X is non-zero, P^{**} and P^* are *different*. Now

$$\begin{aligned} E^{**}[\Sigma_1^N H_n.\Delta\tilde{S}_n] &= \Sigma_\omega P^{**}(\omega)\Big(\Sigma_1^N H_n.\Delta\tilde{S}_n\Big)(\omega) \\ &= \Sigma_\omega \Big(1 + \frac{X(\omega)}{2\|X\|_\infty}\Big) P^*(\omega)\Big(\Sigma_1^N H_n.\Delta\tilde{S}_n\Big)(\omega). \end{aligned}$$

The '1' term on the right gives $E^*[\Sigma_1^N H_n.\Delta\tilde{S}_n]$, which is zero since this is a martingale transform of the E^*-martingale $\tilde{S}_n$. The 'X' term gives a multiple of the inner product

$$(X, \Sigma_1^N H_n.\Delta\tilde{S}_n),$$

which is zero as X is orthogonal to $\tilde{\mathcal{V}}$ and $\Sigma_1^N H_n.\Delta\tilde{S}_n \in \tilde{\mathcal{V}}$. By the Martingale Transform Lemma, $\tilde{S}_n$ is a P^{**}-martingale since H (previsible) is arbitrary. Thus P^{**} is a second equivalent martingale measure, different from P^*. So incompleteness implies non-uniqueness of equivalent martingale measures. //

Martingale Representation. To say that every contingent claim can be replicated means that every P^*-martingale (where P^* is the risk-neutral measure, which is unique) can be written, or *represented*, as a martingale transform (of the discounted prices) by the replicating (perfect-hedge) trading strategy H. In stochastic-process language, this says that all P^*-martingales can be *represented* as martingale transforms of discounted prices. Such Martingale Representation Theorems hold much more generally, and are very important. For the Brownian case, see VII and [RY], Ch. V.
Note. In the example of Chapter II, we saw that the simple option there could be replicated. More generally, in our market set-up, *all* options can be replicated – our market is *complete*. Similarly for the Black-Scholes theory below.

§4. The Fundamental Theorem of Asset Pricing; Risk-neutral valuation.

We summarise what we have learned so far. We call a measure P^* under which discounted prices $\tilde{S}_n$ are P^*-martingales a *martingale measure*. Such

a P^* equivalent to the true probability measure P is called an *equivalent martingale measure*. Then
1 (**No-Arbitrage Theorem**: §2). If the market is *viable* (arbitrage-free), equivalent martingale measures P^* *exist*.
2 (**Completeness Theorem**: §3). If the market is *complete* (all contingent claims can be replicated), equivalent martingale measures are *unique*. Combining:

Theorem (Fundamental Theorem of Asset Pricing, FTAP). In a complete viable market, there exists a unique equivalent martingale measure P^* (or Q).

Let h (≥ 0, $\mathcal{F}_N$-measurable) be any contingent claim, H an admissible strategy replicating it:

$$V_N(H) = h.$$

As $\tilde{V}_n$ is the martingale transform of the P^*-martingale $\tilde{S}_n$ (by H_n), $\tilde{V}_n$ is a P^*-martingale. So $V_0(H)(= \tilde{V}_0(H)) = E^*[\tilde{V}_N(H)]$. Writing this out in full:

$$V_0(H) = E^*[h/S_N^0].$$

More generally, the same argument gives $\tilde{V}_n(H) = V_n(H)/S_n^0 = E^*[(h/S_N^0)|\mathcal{F}_n]$:

$$V_n(H) = S_n^0 E^*[\frac{h}{S_N^0}|\mathcal{F}_n] \qquad (n = 0, 1, \cdots, N).$$

It is natural to call $V_0(H)$ above the *value* of the contingent claim h at time 0, and $V_n(H)$ above the value of h at time n. For, if an investor *sells* the claim h at time n for $V_n(H)$, he can follow strategy H to replicate h at time N and clear the claim. To sell the claim for *any other amount* would provide an arbitrage opportunity (as with the argument for put-call parity). So this value $V_n(H)$ is the *arbitrage price* (or more exactly, *arbitrage-free price* or *no-arbitrage price*); an investor selling for this value is *perfectly hedged*.

We note that, to calculate prices as above, we need to know only
(i) Ω, the set of all possible states,
(ii) the σ-field $\mathcal{F}$ and the filtration (or information flow) $(\mathcal{F}_n)$,
(iii) the EMM P^* (or Q).
We do **NOT** need to know the underlying probability measure P – only its null sets, to know what 'equivalent to P' means (actually, in this model, only

the empty set is null).

Now option pricing is our central task, and for pricing purposes P^* is vital and P itself irrelevant. We thus may – and shall – focus attention on P^*, which is called the *risk-neutral* probability measure. *Risk-neutrality* is the central concept of the subject. The concept of risk-neutrality is due in its modern form to Harrison and Pliska [HP] in 1981 – though the idea can be traced back to actuarial practice much earlier. Harrison and Pliska call P^* the *reference measure*; other names are *risk-adjusted* or *martingale measure*. The term 'risk-neutral' reflects the P^*-martingale property of the risky assets, since martingales model fair games.

To summarise, we have the Risk-Neutral Valuation (or Pricing) Formula:

Theorem (Risk-Neutral Valuation Formula). In a complete viable market, arbitrage-free prices of assets are their discounted expected values under the risk-neutral (equivalent martingale) measure P^* (or Q). With payoff h,

$$V_n(H) = (1+r)^{-(N-n)} E^*[V_N(H)|\mathcal{F}_n] = (1+r)^{-(N-n)} E^*[h|\mathcal{F}_n].$$

§5. European Options. The Discrete Black-Scholes Formula.

We consider the simplest case, the Cox-Ross-Rubinstein *binomial model* of 1979; see [CR], [BK]. We take $d = 1$ for simplicity (one risky asset, one bank account); the price vector is (S_n^0, S_n^1), or $((1+r)^n, S_n)$, where

$$S_{n+1} = \begin{cases} S_n(1+a) & \text{with probability } p, \\ S_n(1+b) & \text{with probability } 1-p \end{cases}$$

with $-1 < a < b$, $S_0 > 0$. So writing N for the expiry time,

$$\Omega = \{1+a, 1+b\}^N,$$

each $\omega \in \Omega$ representing the successive values of $T_{n+1} := S_{n+1}/S_n$. The filtration is $\mathcal{F}_0 = \{\emptyset, \Omega\}$ (trivial σ-field), $\mathcal{F}_T = \mathcal{F} = 2^\Omega$ (power-set of Ω: all subsets of Ω), $\mathcal{F}_n = \sigma(S_1, \cdots, S_n) = \sigma(T_1, \cdots, T_n)$. For $\omega = (\omega_1, \cdots, \omega_N) \in \Omega$, $P(\{\omega_1, \cdots, \omega_N\}) = P(T_1 = \omega_1, \cdots, T_N = \omega_N)$, so knowing the pr. measure P (i.e. knowing p) means we know the distribution of $(T_1, \cdots, T_N)$.

For $p^* \in (0,1)$ to be determined, let P^* correspond to p^* as P does to p. Then the discounted price $(\tilde{S}_n)$ is a P^*-martingale iff

$$E^*[\tilde{S}_{n+1}|\mathcal{F}_n] = \tilde{S}_n, \qquad E^*[(\tilde{S}_{n+1}/\tilde{S}_n)|\mathcal{F}_n] = 1, \qquad E^*[T_{n+1}|\mathcal{F}_n] = 1+r,$$

since $S_n = \tilde{S}_n(1+r)^n$, $T_{n+1} = S_{n+1}/S_n = (\tilde{S}_{n+1}/\tilde{S}_n)(1+r)$. But

$$E^*[T_{n+1}|\mathcal{F}_n] = (1+a).p^* + (1+b).(1-p^*)$$

is a weighted average of $1+a$ and $1+b$; this can be $1+r$ iff $r \in [a,b]$. As P^* is to be *equivalent* to P and P has no non-empty null-sets, $r = a, b$ are excluded. Thus by §2:

Lemma. The market is viable (arbitrage-free) iff $r \in (a,b)$.

Next, $1+r = (1+a)p^* + (1+b)(1-p^*)$, $r = ap^* + b(1-p^*)$: $r - b = p^*(a-b)$:

Lemma. The equivalent mg measure exists, is unique, and is given by

$$p^* = (b-r)/(b-a).$$

Corollary. The market is complete.

Now $S_N = S_n \Pi_{n+1}^N T_i$. By the Fundamental Theorem of Asset Pricing, the price C_n of a call option with strike-price K at time n is

$$\begin{aligned} C_n &= (1+r)^{-(N-n)} E^*[(S_N - K)_+|\mathcal{F}_n] \\ &= (1+r)^{-(N-n)} E^*[(S_n \Pi_{n+1}^N T_i - K)_+|\mathcal{F}_n]. \end{aligned}$$

Now the conditioning on $\mathcal{F}_n$ has no effect – on S_n as this is $\mathcal{F}_n$-measurable (known at time n), and on the T_i as these are independent of $\mathcal{F}_n$. So

$$\begin{aligned} C_n &= (1+r)^{-(N-n)} E^*[(S_n \Pi_{n+1}^N T_i - K)_+] \\ &= (1+r)^{-(N-n)} \Sigma_{j=0}^{N-n} \binom{N-n}{j} p^{*j}(1-p^*)^{N-n-j}(S_n(1+a)^j(1+b)^{N-n-j} - K)_+; \end{aligned}$$

here j, $N-n-j$ are the numbers of times T_i takes the two possible values $1+a, 1+b$. This is the *discrete Black-Scholes formula* of Cox, Ross & Rubinstein (1979) for pricing a European call option in the binomial model. The European put is similar – or use put-call parity (II.3).

To find the (perfect-hedge) strategy for replicating this explicitly: write

$$c(n,x) := \Sigma_{j=0}^{N-n} \binom{N-n}{j} p^{*j}(1-p^*)^{N-n-j}(x(1+a)^j(1+b)^{N-n-j} - K)_+.$$

Then $c(n, x)$ is the undiscounted P^*-expectation of the call at time n given that $S_n = x$. This must be the value of the portfolio at time n if the strategy $H = (H_n)$ replicates the claim:

$$H_n^0(1+r)^n + H_n S_n = c(n, S_n)$$

(here by previsibility H_n^0 and H_n are both functions of $S_0, \cdots, S_{n-1}$ only). Now $S_n = S_{n-1}T_n = S_{n-1}(1+a)$ or $S_{n-1}(1+b)$, so:

$$\begin{aligned} H_n^0(1+r)^n + H_n S_{n-1}(1+a) &= c(n, S_{n-1}(1+a)) \\ H_n^0(1+r)^n + H_n S_{n-1}(1+b) &= c(n, S_{n-1}(1+b)). \end{aligned}$$

Subtract:

$$H_n S_{n-1}(b-a) = c(n, S_{n-1}(1+b)) - c(n, S_{n-1}(1+a)).$$

So H_n in fact depends only on $S_{n-1}, H_n = H_n(S_{n-1})$ (by previsibility), and

Proposition. The perfect hedging strategy H_n replicating the European call option above is given by

$$H_n = H_n(S_{n-1}) = \frac{c(n, S_{n-1}(1+b)) - c(n, S_{n-1}(1+a))}{S_{n-1}(b-a)}.$$

Notice that the numerator is the difference of two values of $c(n, x)$ with the larger value of x in the first term (recall $b > a$). When the payoff function $c(n, x)$ is an increasing function of x, as for the European call option considered here, this is non-negative. In this case, the Proposition gives $H_n \geq 0$: the replicating strategy does not involve short-selling. We record this as:

Corollary. When the payoff function is a non-decreasing function of the final asset price S_N, the perfect-hedging strategy replicating the claim does not involve short-selling of the risky asset.

§6. Continuous-Time Limit of the Binomial Model.

Suppose now that we wish to price an option in continuous time with initial stock price S_0, strike price K and expiry T. We can use the work above to give a discrete-time approximation, where $N \to \infty$. We write (temporarily) $\rho \geq 0$ for the instantaneous interest rate in continuous time, and define (again temporarily) r by

$$r := \rho T/N: \qquad e^{\rho T} = \lim_{N\to\infty}(1 + \frac{\rho T}{N})^N = \lim_{N\to\infty}(1+r)^N.$$

Here r, which tends to zero as $N \to \infty$, represents the interest rate in discrete time for the approximating binomial model.
For $\sigma > 0$ fixed (σ^2 plays the role of a variance, corresponding in continuous time to the *volatility* of the stock – below), define a, b ($\to 0$ as $N \to \infty$) by

$$\log((1+a)/(1+r)) = -\sigma/\sqrt{N}, \qquad \log((1+b)/(1+r)) = \sigma/\sqrt{N}.$$

We now have a sequence of binomial models, for each of which we can price options as in §5. We shall show that the pricing formula converges as $N \to \infty$ to a limit. This is the famous *Black-Scholes formula*, the central result of the course. We shall meet it later, and re-derive it, in *continuous time*, its natural setting, in Ch. VII; see also e.g. [BK], 4.6.2. Fortunately, the continuous Black-Scholes formula is much neater than its discrete counterpart, which involves the unwieldy binomial sum above.

Lemma. Let $(X_j^N)_{j=1}^N$ be iid with mean μ_N satisfying

$$N\mu_N \to \mu \qquad (N \to \infty)$$

and variance $\sigma^2(1+o(1))/N$. If $Y_N := \Sigma_1^N X_j^N$, then Y_N converges in distribution to normality:

$$Y_N \to Y = N(\mu, \sigma) \qquad (N \to \infty).$$

Proof. Use characteristic functions (CFs), I.4: since Y_N has mean and variance as given, it also has second moment $\sigma^2(1+o(1))/N$, so has CF

$$\phi_N(u) := E\exp\{iuY_N\} = \Pi_1^N E\exp\{iuX_j^N\} = [E\exp\{iuX_1^N\}]^N$$

$$= (1 + \frac{iu\mu}{N} - \frac{1}{2}\frac{\sigma^2 u^2}{N} + o(\frac{1}{N}))^N \to \exp\{iu\mu - \frac{1}{2}\sigma^2 u^2\} \qquad (N \to \infty),$$

the CF of the normal law $N(\mu, \sigma)$. Convergence of CFs implies convergence in distribution by Lévy's continuity theorem for CFs ([W], §18.1). //

We can apply this to pricing the call option above (the details of the calculation below, which are messy, are not hard, and not crucial; the point is that this can be done):

$$C_0^{(N)} = (1+\frac{\rho T}{N})^{-N} E^*[(S_0 \Pi_1^N T_n - K)_+]$$

$$= E^*[(S_0 \exp\{Y_N\} - (1+\frac{\rho T}{N})^{-N} K)_+], \qquad (1)$$

where

$$Y_N := \sum\nolimits_1^N \log(T_n/(1+r)).$$

Since $T_n = T_n^N$ above takes values $1+b, 1+a$, $X_n^N := \log(T_n^N/(1+r))$ takes values $\log((1+b)/(1+r))$, $\log((1+a)/(1+r)) = \pm\sigma/\sqrt{N}$ (so has second moment σ^2/N). Its mean is

$$\mu_N := \log\Big(\frac{1+b}{1+r}\Big)(1-p^*) + \log\Big(\frac{1+a}{1+r}\Big)p^* = \frac{\sigma}{\sqrt{N}}(1-p^*) - \frac{\sigma}{\sqrt{N}}p^* = (1-2p^*)\sigma/\sqrt{N}$$

(we shall see below that $1-2p^* = O(1/\sqrt{N})$, so the Lemma will apply). Now (recall $r = \rho T/N = O(1/N)$)

$$a = (1+r)e^{-\sigma/\sqrt{N}} - 1, \qquad b = (1+r)e^{\sigma/\sqrt{N}} - 1,$$

so $a, b, r \to 0$ as $N \to \infty$, and

$$1-2p^* = 1 - 2\frac{(b-r)}{(b-a)} = 1 - 2\frac{[(1+r)e^{\sigma/\sqrt{N}} - 1 - r]}{[(1+r)(e^{\sigma/\sqrt{N}} - e^{-\sigma/\sqrt{N}})]}$$

$$= 1 - 2\frac{[e^{\sigma/\sqrt{N}} - 1]}{[e^{\sigma/\sqrt{N}} - e^{-\sigma/\sqrt{N}}]}.$$

Now expand the two $[\cdots]$ terms above by Taylor's theorem: they give

$$\frac{\sigma}{\sqrt{N}}(1 + \frac{1}{2}\frac{\sigma}{\sqrt{N}} + \cdots), \qquad \frac{2\sigma}{\sqrt{N}}(1 + \frac{\sigma^2}{6N} + \cdots).$$

So, cancelling $\sigma/\sqrt{N}$,

$$1-2p^* = 1 - \frac{2(1+\frac{1}{2}\frac{\sigma}{\sqrt{N}} + \cdots)}{2(1+\frac{\sigma^2}{6N} + \cdots)} = -\frac{1}{2}\frac{\sigma}{\sqrt{N}} + O(1/N):$$

$$N\mu_N = N.\frac{\sigma}{\sqrt{N}}.(-\frac{1}{2}\frac{\sigma}{\sqrt{N}} + O(1/N)) \to \mu := -\frac{1}{2}\sigma^2 \qquad (N \to \infty).$$

We now need to change notation:
(i) We replace the variance σ^2 above by $\sigma^2 T$. So σ^2 is the *variance per unit time* (which is more suited to the work of Ch. VI, VII in continuous time); the standard deviation (SD) σ is called the *volatility*. It measures the variability of the stock, so its riskiness, or its sensitivity to new information.
(ii) We replace ρ in the above by r. This is the standard notation for the riskless interest rate in continuous time, to which we are now moving.

As usual, we write the standard normal density function as ϕ and distribution function as Φ:

$$\phi(x) := \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}, \qquad \Phi(x) := \int_{-\infty}^{x} \phi(u)du = \int_{-\infty}^{x} \frac{e^{-\frac{1}{2}u^2}}{\sqrt{2\pi}} du.$$

Note that as ϕ is even, the left and right tails of Φ are equal:

$$\phi(x) = \phi(-x), \qquad \text{so} \qquad \int_{-\infty}^{-x} \phi(u)du = \int_{x}^{\infty} \phi(u)du : \qquad \Phi(-x) = 1 - \Phi(x).$$

Theorem (Black-Scholes formula (for calls), 1973). The price of the European call option is

$$c_t = S_t\Phi(d_+) - Ke^{-r(T-t)}\Phi(d_-), \tag{BS}$$

where S_t is the stock price at time $t \in [0, T]$, K is the strike price, r is the riskless interest rate, σ is the volatility and

$$d_\pm := [\log(S/K) + (r \pm \frac{1}{2}\sigma^2)(T-t)]/\sigma\sqrt{T-t} : \qquad d_+ = d_- + \sigma\sqrt{T-t}.$$

For completeness, we state the corresponding Black-Scholes formula for puts. The proofs of the two results are closely analogous, and one can derive either from the other by put-call parity.

Theorem (Black-Scholes formula for puts, 1973). The price of the corresponding put option is

$$p_t = Ke^{-r(T-t)}\Phi(-d_-) - S_t\Phi(-d_+). \tag{BS - p}$$

The Black-Scholes formula is not perfect – indeed, Fischer Black himself famously wrote a paper called *The holes in Black-Scholes*. But it is very useful, as a benchmark and first approximation.

Proof of the Black-Scholes formula.

It suffices to take $t = 0$ – so T is the *remaining* time to expiry.

We use the Lemma, with $\mu = -\frac{1}{2}\sigma^2 T$ (in our new notation). In (1), we have $Y_N \to Y$ in distribution and (replacing R in the Lemma by r, as above) $(1 + \frac{rT}{N})^{-N} \to e^{-rT}$ as $N \to \infty$. This suggests

$$C_0^{(N)} \to C_0 := E_Y[(S_0 e^Y - e^{-rT}K)_+] = e^{-rT} E_Y[(S_0 e^{rT+Y} - K)_+],$$

where E_Y is the expectation for the distribution of Y, which is $N(-\frac{1}{2}\sigma^2 T, \sigma^2 T)$ (in our current notation). This can be justified, by standard properties of convergence in distribution (see e.g. [W], Ch. 18). So if $Z := (Y + \frac{1}{2}\sigma^2 T)/(\sigma\sqrt{T})$, $Z \sim N(0,1)$, $Y = -\frac{1}{2}\sigma^2 T + \sigma\sqrt{T} Z$, and

$$C_0 = e^{-rT} E_Z[(S_0 e^{(r-\frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z} - K)_+]$$

$$= e^{-rT} \int_{-\infty}^{\infty} [S_0 \exp\{(r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}x\} - K]_+ \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} dx.$$

Similarly, with payoff h, the time-0 price of the claim, or option is

$$e^{-rT} \int_{-\infty}^{\infty} h(S_0 \exp\{(r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}x\}) \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} dx. \qquad (*)$$

To evaluate the integral, note first that $[...] > 0$ where

$$S_0 \exp\{(r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}x\} > K : \quad x > [\log(K/S_0) - (r - \frac{1}{2}\sigma^2)T]/\sigma\sqrt{T} = c, \quad \text{say. So}$$

$$C_0 = S_0 \int_c^{\infty} e^{-\frac{1}{2}\sigma^2 T}. \exp\{-\frac{1}{2}x^2 + \sigma\sqrt{T}x\} dx/\sqrt{2\pi} - K e^{-rT}[1 - \Phi(c)],$$

and the last term is $Ke^{-rT}\Phi(-c) = Ke^{-rT}\Phi(d_-)$ ($-c = [\log(S/K) + (r - \frac{1}{2}\sigma^2)T]/\sigma\sqrt{T} = d_-$, when $t = 0$). The remaining integral is

$$\int_c^{\infty} \exp\{-\frac{1}{2}(x - \sigma\sqrt{T})^2\} dx/\sqrt{2\pi} = \int_{c-\sigma\sqrt{T}}^{\infty} \exp\{-\frac{1}{2}u^2\} du/\sqrt{2\pi}$$

$$= 1 - \Phi(c - \sigma\sqrt{T}) = \Phi(-c + \sigma\sqrt{T}) = \Phi(d_+),$$

as $-c + \sigma\sqrt{T} = d_- + \sigma\sqrt{T} = d_+$ when $t = 0$. So the option price is given in terms of the initial price S_0, strike price K, expiry T, interest rate r and volatility σ by

$$C_0 = S_0\Phi(d_+) - Ke^{-rT}\Phi(d_-), \qquad d_\pm := [\log(S/K) + (r \pm \frac{1}{2}\sigma^2)T]/\sigma\sqrt{T}. \ //$$

Note. 1. *Normal approximation to binomial.* The proof above starts from a binomial distribution and ends with a normal distribution. The binomial distribution is that of a sum of independent Bernoulli random variables. That sums (or averages) of independent random variables with finite means and variances gives a normal limit is the content of the *Central Limit Theorem* or CLT (the *Law of Errors*, as physicists would say). This form of the CLT is the *de Moivre-Laplace limit theorem.*

The picture for this is familiar. The Binomial distribution $B(n, p)$ has a histogram with $n + 1$ bars, whose heights peak at the mode and decrease to either side. For large n, one can draw a smooth curve through the histogram. This curve is the relevant approximating normal density.

2. The Cox-Ross-Rubinstein binomial model above goes over in the passage to the limit to the geometric Brownian motion model of VII.1. We will later re-derive the continuous Black-Scholes formula in Ch. VII, using continuous-time methods (Itô calculus), rather than, as above, deriving the discrete BS formula and going to the limit on the *formula*, rather than the *model.*

3. For similar derivations of the discrete Black-Scholes formula and the passage to the limit to the continuous Black-Scholes formula, see e.g. [CR], §5.6.

4. One of the most striking features of the Black-Scholes formula is that it does **not** involve the mean rate of return μ of the stock – only the riskless interest-rate r and the volatility of the stock σ. Mathematically, this reflects the fact that the *change of measure* involved in the passage to the risk-neutral measure involves a *change of drift.* This eliminates the μ term; see VII.

5. *Volatility.* The volatility σ can be estimated in two ways:

a. *Historic volatility.* Directly from the movement of a stock price in time, using Time Series methods in discrete time [see Ch. VI for continuous time].

b. *Implied volatility.* From the observed market prices of options: if we know everything in the Black-Scholes formula (including the price at which the option is traded) *except* the volatility σ, we can solve for σ. Since σ appears inside the argument of the normal distribution function Φ as well as outside it, this is a transcendental equation for σ and has to be solved numerically by iteration (Newton-Raphson method). We quote (see 'The Greeks' below, and

Problems 7) that the Black-Scholes price is a monotone (increasing) function of the volatility (more volatility doesn't make us 'more likely to win', but when we do win, we 'win bigger'), so there is a unique root of the equation.

In practice, one sees discrepancies between historic and implied volatility, which show limitations to the accuracy of the Black-Scholes model. But it is the standard 'benchmark model', and useful as a first approximation.

The classical view of volatility is that it is caused by future uncertainty, and shows the market's reaction to the stream of new information. However, studies taking into account periods when the markets are open and closed [there are only about 250 trading days in the year] have shown that the volatility is less when markets are closed than when they are open. This suggests that *trading itself is one of the main causes of volatility.*

Note. This observation has deep implications for the macro-prudential and regulatory issues discussed in Ch. II. The real economy cannot afford too much volatility. Volatility is (at least partly) caused by trading. Conclusion: there is too much trading. Policy question: how can we reduce the volume of trading (much of it speculative, designed to enrich traders, and not serving a more widely useful economic purpose)? One answer is the so-called *Tobin tax* (also known as the "Robin Hood tax") (James Tobin (1918-2002), American economist; Nobel Prize for Economics, 1981). This would levy a small charge (e.g. 0.01%) on *all* financial transactions. This would both provide a major and useful source of tax revenue, and – more importantly – would discourage a lot of speculative trading, thereby (shrinking the size of the financial services industry, but) diminishing volatility, to the benefit of the general economy (Problems 7 again).

If the Black-Scholes model were perfect, historic and implied volatility estimates would agree (to within sampling error). But discrepancies can be observed, which shows the imperfections of our model.

Volatility estimation is a major topic, both theoretically and in practice. We return to this in V.7.3-4 below and VII.7.5-8. But we note here:
(i) trading is itself one of the major causes of volatility, as above;
(ii) options like volatility [i.e., option prices go up with volatility].
Recalling Ch. II, this shows that volatility is a 'bad thing' from the point of view of the real economy (uncertainty about, e.g., future material costs is nothing but a nuisance to manufacturers), but a 'good thing' for financial markets (trading increases volatility, which increases option prices, which generates more trade ...) – at the cost of increased instability.

§7. More on European Options

1. *Bounds.* We use the notation above. We also write c, p for the values of European calls and puts, C, P for the values of the American counterparts.

Obvious upper bounds are $c \leq S, C \leq S$, where S is the stock price (we can buy for S on the market without worrying about options, so would not pay more than this for the right to buy). For puts, one has correspondingly the obvious upper bounds $p \leq K, P \leq K$, where K is the strike price: one would not pay more than K for the right to sell at price K, as one would not spend more than one's maximum return. For lower bounds:

$c_0 \geq \max(S_0 - Ke^{-rT}, 0)$.

Proof. Consider the following two portfolios:

I: one European call plus Ke^{-rT} in cash; II: one share. Show "I $\geq$ II".

$p_0 \geq \max(Ke^{-rT} - S_0, 0)$ (proof: by above and put-call parity).

2. *Dependence of the Black-Scholes price on the parameters.*

Recall the Black-Scholes formulae for the values c_t, p_t for the European call and put: with

$$d_\pm := [\log(S_t/K) + (r \pm \frac{1}{2}\sigma^2)(T-t)]/\sigma\sqrt{(}T-t),$$

$$c_t = S_t\Phi(d_+) - Ke^{-r(T-t)}\Phi(d_-), \qquad p_t = Ke^{-r(T-t)}\Phi(-d_-) - S_t\Phi(-d_+),$$

(a). S. As the stock price S increases, the call option becomes more likely to be exercised. As $S \to \infty$, $d_\pm \to \infty$, $\Phi(d_\pm) \to 1$, so $c_t \sim S_t - Ke^{-r(T-t)}$. This has a natural economic interpretation: as the value of a *forward contract* with *delivery price* K (Hull [H1] Ch. 3, [H2] Ch. 3).

(b). σ. When the volatility $\sigma \to 0$, the stock becomes riskless, and behaves like money in the bank. Again, $d_\pm \to \infty$, as above, with the same economic interpretation.

3. *The Greeks.*

These are the partial derivatives of the option price with respect to the input parameters. They have the interpretation of *sensitivities*.

(i) For a call, say, $\partial c/\partial S$ is called the *delta*, Δ. Adjusting our holdings of stock to eliminate our portfolio's dependence on S is called *delta-hedging*.

(ii) Second-order effects involve *gamma* $:= \partial(\Delta)/\partial S$.

(iii) Time-dependence is given by *Theta* is $\partial c/\partial t$.

(iv) Volatility dependence is given by *vega* $:= \partial c/\partial \sigma$.[1]

[1] Of course, vega is not a letter of the Greek alphabet! (it is the Spanish word for 'meadow', as in Las Vegas) – presumably so named for "v for volatility, variance and vega", and because vega sounds quite like beta, etc.

Vega.

From the Black-Scholes formula (which gives the price explicitly as a function of σ), one can check by calculus (Problems 7) that

$$\partial c/\partial \sigma > 0,$$

and similarly for puts (or, use the result for calls and put-call parity). In sum: *options like volatility.* This fits our intuition. The more uncertain things are (the higher the volatility), the more valuable protection against adversity – or insurance against it – becomes (the higher the option price).
(v) *rho* is $\partial c/\partial r$, the sensitivity to interest rates.

§8. American Options.

We now consider an American call option (value C), in the simplest case of a stock paying no dividends. The following result goes back (at least) to R. C. MERTON in 1973.

Theorem (Merton's theorem). It is never optimal to exercise an American call option early. That is, the American call option is equivalent to the European call, so has the same value:

$$C = c.$$

First Proof. Consider the following two portfolios:
I: one American call option plus cash Ke^{-rT}; II: one share.
The value of the cash in I is K at time T, $Ke^{-r(T-t)}$ at time t. If the call option is exercised early at $t < T$, the value of Portfolio I is then $S_t - K$ from the call, $Ke^{-r(T-t)}$ from the cash, total

$$S_t - K + Ke^{-r(T-t)}.$$

Since $r > 0$ and $t < T$, this is $< S_t$, the value of Portfolio II at t. So Portfolio I is *always* worth less than Portfolio II if exercised *early.*

If however the option is exercised instead at expiry, T, the American call option is then the same as a European call option. Then at time T, Portfolio I is worth $\max(S_T, K)$ and Portfolio II is worth S_T. So:

$$\begin{array}{rcl} \text{before} \quad T, & & I < II, \\ \text{at} \quad T, & & I \geq II \text{ always, and} > \text{sometimes.} \end{array}$$

This direct comparison with the underlying [the share in Portfolio II] shows that early exercise is never optimal. Since an American option at expiry is the same as a European one, this completes the proof. //
Second Proof. One can instead use the bounds of §7.1. For details, see e.g. [BK, Th. 4.7.1].

Financial Interpretation.

There are two reasons why an American call should not be exercised early:
1. *Insurance.* Consider an investor choosing to hold a call option instead of the underlying stock. He does not care if the share price falls below the strike price (as he can then just discard his option) – but if he held the stock, he would. Thus the option *insures* the investor against such a fall in stock price, and if he exercises early, he loses this insurance.
2. *Interest on the strike price.* When the holder exercises the option, he buys the stock and pays the strike price, K. Early exercise at $t < T$ loses the interest on K between times t and T: the later he pays out K, the better.
Economic Note. Despite Merton's theorem, and the interpretation above, there are plenty of real-life situations where early exercise of an American call might be sensible, and indeed it is done routinely. Consider, for example, a manufacturer of electrical goods, in bulk. He needs a regular supply of large amounts of copper. The danger is future price increases; the obvious precaution is to hedge against this by buying call options. If the expiry is a year but copper stocks are running low after six months, he would exercise his American call early, to keep an adequate inventory of copper, his crucial raw material. This ensures that his main business activity – manufacturing – can continue unobstructed. Neither of the reasons above applies here:
Insurance. He doesn't care if the price of copper falls: he isn't going to sell his copper stocks, but use them.
Interest. He doesn't care about losing interest on cash over the remaining six months. He is in manufacturing to use his money to make things, and then sell them, rather than put it in the bank.

This neatly illustrates the contrast between *finance* (money, options etc.) and *economics* (the real economy – goods and services).

Put-Call Symmetry.

The BS formulae for puts and calls resemble each other, with stock price S and discounted strike K interchanged. Results of this type are called *put-call symmetry.*

American Puts.

Recall the put-call parity of Ch. II (valid only for European options): $c - p = S - Ke^{-rT}$. A partial analogue for American options is given by the inequalities below:

$$S - K < C - P < S - Ke^{-rT}.$$

For proof (as above) and background, see e.g. Ch. 8 (p. 216) of [H1].

We now consider how to evaluate an American put option, European and American call options having been treated already. First, we will need to work in discrete time. We do this by dividing the time-interval $[0, T]$ into N equal subintervals of length Δt say. Next, we take the values of the underlying stock to be discrete: we use the binomial model of §5, with a slight change of notation: we write u, d ('up', 'down') for $(1 + b), (1 + a)$: thus stock with initial value S is worth $Su^i d^j$ after i steps up and j steps down. Consequently, after N steps, there are $N + 1$ possible prices, $Su^i d^{N-i}$ $(i = 0, \cdots, N)$. It is convenient to display the possible paths followed by the stock price as a binomial tree, with time going left to right and two paths, up and down, leaving each node in the tree, until we reach the $N + 1$ terminal nodes at expiry. There are 2^N possible paths through the tree. It is common to take N of the order of 30, for two reasons:
(i) typical lengths of time to expiry are measured in months (9 months, say); this gives a time-step around the corresponding number of days,
(ii) 2^{30} paths is about the order of magnitude that can be easily handled by computers (recall that $2^{10} = 1,024$, so 2^{30} is somewhat over a billion).

We now return to the binomial model in §§5,6, with a slight change of notation. Recall that in §5 (discrete time) we used $1 + r$ for the discount factor. Now call this $1 + \rho$ instead, freeing r for its usual use as the short rate of interest in continuous time. Thus $1 + \rho = e^{r\Delta t}$, and the risk-neutrality condition $p^* = (b - r)/(b - a)$ of §5 becomes

$$p^* = (u - e^{r\Delta t})/(u - d).$$

Now recall (§7) $(1+a)/(1+r) = \exp(-\sigma/\sqrt{N})$, $(1+b)/(1+r) = \exp(\sigma/\sqrt{N})$. We replaced σ^2 by $\sigma^2 T$ (to make σ the volatility per unit time), and $T = N.\Delta t$, so $\sigma/\sqrt{N}$ becomes $\sigma\sqrt{T}/\sqrt{N} = \sigma\sqrt{\Delta t}$. So now

$$u/e^{r\Delta t} = e^{\sigma/\sqrt{\Delta t}}, \qquad d/e^{r\Delta t} = e^{-\sigma\sqrt{\Delta t}} : \qquad ud = e^{2r\Delta t}.$$

Since $\sqrt{\Delta t}$ is small, its square Δt is a second-order term. So to first order, $ud = 1$, which simplifies filling in the terminal values in the binary tree.

We begin again: define our up and down factors u, d so that

$$ud = 1; \qquad (*)$$

define the risk-neutral probability p^* so as to have

$$p^* = (u - e^{r\Delta t})/(u - d)$$

(to get the mean return from the risky stock the same as that from the riskless bank account), and the volatility σ to get the variance of the stock price S' after one time-step when it is worth S initially as $S^2\sigma^2\Delta t$:

$$S^2\sigma^2\Delta t = p^* S^2 u^2 + (1 - p^*)S^2 d^2 - S^2[p^* u + (1 - p^*)d]^2$$

(using $var S' = E(S'^2) - [ES']^2$). Then to first order in $\sqrt{\Delta t}$ (which is all the accuracy we shall need), one can check that we have as before

$$u = \exp(\sigma\sqrt{\Delta t}), \qquad d = \exp(-\sigma\sqrt{\Delta t}). \qquad (**)$$

We can now calculate both the value of an American put option and the optimal exercise strategy by working backwards through the tree (this method of backward recursion in time is a form of the Dynamic Programming [DP] technique (Richard Bellman (1920-84) in 1953, book, 1957), which is important in many areas of optimization and Operational Research (OR)).

1. Draw a binary tree showing the initial stock value and having the right number, N, of time-intervals.

2. Fill in the stock prices: after one time interval, these are Su (upper) and Sd (lower); after two time-intervals, Su^2, S and $Sd^2 = S/u^2$; after i time-intervals, these are $Su^j d^{i-j} = Su^{2j-i}$ at the node with j 'up' steps and $i - j$ 'down' steps (the '(i, j)' node).

3. Using the strike price K and the prices at the *terminal nodes*, fill in the payoffs ($f_{N,j} = \max[K - Su^j d^{N-j}, 0]$) from the option at the terminal nodes (where, at expiry, the values of the European and American options coincide) underneath the terminal prices.

4. Work back down the tree one time-step. Fill in the 'European' value at the penultimate nodes as the discounted values of the upper and lower right (terminal node) values, under the risk-neutral measure – 'p^* times lower right

plus $1-p^*$ times upper right' [notation of V.6]. Fill in the 'intrinsic' (or early-exercise) value - the value if the option is exercised. Fill in the American put value as the higher of these.
5. Treat these values as 'terminal node values', and fill in the values one time-step earlier by repeating Step 4 for this 'reduced tree'.
6. Iterate. The value of the American put at time 0 is the value at the root - the last node to be filled in. The 'early-exercise region' is the node set where the early-exercise value is the higher; the rest is the 'continuation region'.
Note. The above procedure is simple to describe and understand, and simple to programme. It is laborious to implement numerically by hand, on examples big enough to be non-trivial. Numerical examples are worked through in detail in [H1], 359-360 and [CR], 241-242.

Mathematically, the task remains of describing the *continuation region* - the part of the tree where early exercise is not optimal. This is a classical *optimal stopping problem*. No explicit solution is known (and presumably there isn't one). We will, however, connect the work above with that of IV.7 on the Snell envelope. Consider the pricing of an American put, strike price K, expiry N, in discrete time, with discount factor $1+r$ per unit time as earlier. Let $Z = (Z_n)_{n=0}^N$ be the payoff on exercising at time n. We want to price Z_n, by U_n say (to conform to our earlier notation), so as to avoid arbitrage; again, we work backwards in time. The recursive step is

$$U_{n-1} = \max(Z_{n-1}, \frac{1}{1+r} E^*[U_n | \mathcal{F}_{n-1}]),$$

the first alternative on the right corresponding to early exercise, the second to the discounted expectation under P^*, as usual. Let $\tilde{U}_n = U_n/(1+r)^n$ be the discounted price of the American option. Then

$$\tilde{U}_{n-1} = \max(\tilde{Z}_{n-1}, E^*[\tilde{U}_n | \mathcal{F}_{n-1}]) :$$

$(\tilde{U}_n)$ is the *Snell envelope* (III.7) of the discounted payoff process $(\tilde{Z}_n)$. So:
(i) a P^*-supermartingale,
(ii) the smallest supermartingale dominating $(\tilde{Z}_n)$,
(iii) the solution of the optimal stopping problem for $\tilde{Z}$.

Note. One can use the Snell envelope to prove Merton's theorem (equivalence of American and European calls) without using arbitrage arguments. For details see e.g. [BK, Th. 4.7.1 and Cor. 4.7.1].

P-measure and P^*- (or $Q-$) measure.

We use P and P^* in the above, as E and E^* are convenient, but P and Q when the emphasis is on Q, for brevity.

The measure P, the *real* (or real-world) probability measure, models the uncertainty driving prices, which are indeed uncertain, thus allowing us to bring mathematics to bear on financial problems. But P is difficult to get at directly. By contrast, Q is more accessible: the *market* tells us about Q, or more specifically, *trading* does. In addition, trading also tells us about the *volatility* σ, via implied volatility, which we can infer from observing the prices at which options are traded. So Q is certainly more accessible than P. There is thus a sense in which it is Q, rather than P, which is the more real.

It is as well to bear all this in mind when looking at specific problems, particularly numerical ones. Now that we know the CRR binomial-tree model, which gives us the Black-Scholes formula in discrete time (and hence also, by the limiting argument above, the Black- Scholes formula in continuous time, the main result of the course), we can recognise the 'one-period, up or down' model ($/SFr in II.8, price of gold in Problems 5), though clearly artificial and stylised, as a workable 'building block' of the whole theory. Because P itself does not occur in the Black-Scholes formula(e), from a purely financial point of view there is little need to try to construct more realistic, and so more complicated, models of P. Instead, one can exploit what one can infer about Q, which does occur in Black-Scholes, from seeing the prices at which options trade.

From the economic point of view, it is the real world, the real economy, and so the real probability measure P, that matters. The 'Q-measure-eye view of the world' has a degree of artificiality, in so far as options do. One can eat food, and needs to. One can't eat options.

A fuller discussion of Q-measure involves *Arrow-Debreu* prices, equilibria etc., but we omit this for lack of time, and because it would take us too far into Economics.