Derivative Securities – Fall 2007– Section 7
Notes by Robert V. Kohn, extended and improved by Steve Allen.
Courant Institute of Mathematical Sciences.

Further discussion of the continuous time framework. Topics in this section: (a) more stochastic calculus; (b) what are the consequences of hedging only at discrete times? (c) the link between risk-neutral expectation and PDE's; and (d) martingales and their importance for option pricing.

****

More stochastic calculus. You'll need the following for HW4. Consider a stochastic integral of the form $\int_a^b g(s)\,dw(s)$ where g is a deterministic function of s. It has mean value zero – we explained this in Section 6. What about its variance? The answer is simple:

$$E\left[\left(\int_a^b g(s)\,dw\right)^2\right] = \int_a^b g^2(s)\,ds.$$

Here is why. Approximating the stochastic integral by a sum, we see that the square of the stochastic integral is approximately

$$\left(\sum_i g(s_i)[w(s_{i+1}) - w(s_i)]\right)\left(\sum_j g(s_j)[w(s_{j+1}) - w(s_j)]\right)$$
$$= \sum_{i,j} g(s_i)g(s_j)[w(s_{i+1}) - w(s_i)][w(s_{j+1}) - w(s_j)] \qquad .$$

For $i \neq j$ the expected value of the i,jth term is 0 since $[w(s_{j+1}) - w(s_j)]$ and $[w(s_{i+1}) - w(s_i)]$ are independent Gaussians, each with mean value 0. For $i = j$ the expected value of the i,jth term is $g^2(s_i)(s_{i+1} - s_i)$. So the expected value of the squared stochastic integral is approximately

$$\sum_i g^2(s_i)(s_{i+1} - s_i),$$

which is a Riemann sum for $\int_a^b g^2(s)\,ds$. By the way: since we are assuming that g is deterministic, $\int_a^b g(s)dw(s)$ is a *Gaussian* random variable. (Proof: recall that a sum of Gaussians is Gaussian; therefore $\sum_i g(s_i)[w(s_{i+1}) - w(s_i)]$ is Gaussian. Now use the fact that a limit of Gaussians is Gaussian.) Since we know its mean and variance, we have completely characterized this random variable.

We are in the habit of focusing on lognormal dynamics, because this is the most basic model for the price of a stock (or the forward price of a stock). Another exactly-solvable SDE is the *Ornstein-Uhlenbeck* process, which solves

$$dy = -cydt + \sigma dw, \quad y(0) = y_0$$

with c and σ constant. (This is *not* a lognormal process, because the coefficient of dw is not proportional to y.) Ito's lemma gives

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

so

$$e^{ct}y(t) - y_0 = \sigma \int_0^t e^{cs}dw,$$

or in other words

$$y(t) = e^{-ct}y_0 + \sigma \int_0^t e^{c(s-t)}dw(s).$$

We see (using the discussion at the beginning of this section) that $y(t)$ is a Gaussian random variable. So it is entirely characterized by its mean and variance. They are easy to compute: the mean is clearly

$$E[y(t)] = e^{-ct}y_0$$

since the "dw" integral has expected value 0, and the variance is

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

The Ornstein-Uhlenbeck process is relevant to finance. 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 $dy = -ay dt + \sigma dw$. Notice that $r(t)$ has a positive probability of being negative (since it is a Gaussian random variable); as a consequence the Vasicek model is not very realistic. Even so, its exact solution formulas provide helpful intuition.

****

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 unfortunately). Making a choice, I'll focus on the hedging of a European option on a non-dividend-paying stock. A parallel discussion can however be given for an option on a forward price.

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)

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

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 $\delta t \to 0$. An 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).

****

The link between risk-neutral expectations and PDE's. We have discussed two apparently different approaches to the valuation of a European option: (i) take the discounted risk-neutral expected payoff, or (ii) solve the Black-Scholes PDE. Let's show now that these two approaches are equivalent.

First, consider options on a forward price. We saw long ago that in the discrete time setting, the forward process satisfies

$$\mathcal{F}_t = E_{\mathrm{RN}}[\mathcal{F}_T]$$

for any $t < T$. (In the terminology we'll introduce soon, $\mathcal{F}_t$ is a *martingale* when we calculate expectations using the risk-neutral measure.) In the continuous-time setting, this means the SDE for $\mathcal{F}$ under the risk-neutral measure has no dt term. If in addition $\mathcal{F}$ is lognormal then its SDE under the risk-neutral measure must be

$$d\mathcal{F} = \sigma \mathcal{F} dw. \tag{1}$$

I claim that if $\mathcal{F}$ satisfies this SDE, and $V(F,t)$ satisfies

$$V_t + \tfrac{1}{2}\sigma^2 F^2 V_{FF} - rV = 0 \quad \text{for } t < T, \text{ with } V(F,T) = \phi(F), \tag{2}$$

then

$$V(\mathcal{F}(0), 0) = e^{-rT} E[\phi(\mathcal{F}_T)]. \tag{3}$$

To see why, let's apply the Ito calculus to the function $H(F,t) = e^{r(T-t)}V(F,t)$. We get

$$dH(\mathcal{F}(t), t) = e^{r(T-t)} \left[V_t + \tfrac{1}{2}\sigma^2 F^2 V_{FF} - rV \right] dt + e^{r(T-t)} \sigma \mathcal{F} V_F \, dw$$

with the understanding that each term on the right is evaluated at $(F,t) = (\mathcal{F}(t), t)$. *According to the PDE (1) the dt term has coefficient zero.* Therefore

$$H(\mathcal{F}(T), T) - H(\mathcal{F}_0, 0) = \int_0^T (\text{stuff}) \, dw.$$

Now take the expected value of both sides. On the right we get 0. The second term on the left is known at time 0. As for the first term: remembering the definition of H and the final-time condition in the definition of V it is is

$$H(\mathcal{F}(T), T) = e^0 V(\mathcal{F}(T), T) = \phi(\mathcal{F}_T).$$

We thus conclude that

$$E[\phi(\mathcal{F}_T)] = H(\mathcal{F}_0, 0) = e^{rT} V(\mathcal{F}_0, 0)$$

which is equivalent to (3).

A similar calculation applies to options on a non-dividend-paying stock. We learned in Section 4 that if $s(t)$ is lognormal, then under the risk-neutral measure $s(t) = s_0 e^X$ where X is Gaussian with mean $(r - \frac{1}{2}\sigma^2)t$ and variance $\sigma^2 t$. Put differently: $s_t =$

$s_0 \exp\left[(r - \frac{1}{2}\sigma^2)t + \sigma w(t)\right]$. We now know (from Section 6) an equivalent statement using SDE's: under the risk-neutral measure, s solves the SDE

$$ds = rs\,dt + \sigma s\,dw. \tag{4}$$

I claim that if s satisfies this SDE, and $V(F,t)$ satisfies

$$V_t + rsV_s + \tfrac{1}{2}\sigma^2 s^2 V_{ss} - rV = 0 \quad \text{for } t < T, \text{ with } V(s,T) = \phi(s), \tag{5}$$

then

$$V(s_0, 0) = e^{-rT} E[\phi(s_T)]. \tag{6}$$

The argument is just as before: we apply the Ito calculus to $H(F,t) = e^{r(T-t)}V(s,t)$. We get

$$dH(s(t),t) = e^{r(T-t)}\left[V_t + rsV_s + \tfrac{1}{2}\sigma^2 s^2 V_{ss} - rV\right]dt + e^{r(T-t)}\sigma s V_s\,dw$$

with the understanding that each term on the right is evaluated at $(s,t) = (s(t),t)$. The PDE assures us that the dt term has coefficient zero, so that

$$H(s(T),T) - H(s_0,0) = \int_0^T (\text{stuff})\,dw.$$

Taking the expectation of both sides, we find

$$E[\phi(s_T)] = H(s_0,0) = e^{rT}V(s_0,0)$$

which is equivalent to (6).

****

Martingales and their importance for option pricing. The SDE's (1) and (4) are clearly fundamental. The argument we gave above for the former was hopefully convincing. The argument we gave for the latter was perhaps less so. We'll do better now, as we reformulate what we've been doing in terms of *martingales*. This viewpoint has many advantages; in particular, it explains the origin of our SDE's, and it extends easily to stochastic interest rates (which we'll begin addressing very soon).

The basic prescription for working backward in a binomial tree was this: if V is the value of a tradeable non-dividend-paying security (such as an option) then

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

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 of the "up" state. (I wrote "futures" rather than "forward" on purpose. When the interest rate is deterministic, futures prices and forward prices are the same. But when the interest rate is random, it is the futures price not the forward price that satisfies the preceding equation.)

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

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

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 $0 \le t < t' \le T$:

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

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 (for a non-dividend-paying stock), or equivalently by the fact that

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

where $\mathcal{F}(t)$ is a futures price. Options are tradeables, so the value V of any option is determined by the condition that

- $V(t)/A(t)$ 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 $dy = fdt + gdw$ is a martingale exactly if $f = 0$.

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 $A(t) = e^{rt}$, and we focus (just to be specific) on options on a non-dividend-paying stock (rather than on a futures price).

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.

The preceding questions and answers are, of course, simply convenient reorganizations our prior calculations connecting risk-neutral expectation to the Black-Scholes PDE.

Why, exactly, must $V(t)/A(t)$ be a martingale, if V is the price of a tradeable? In discrete time, this is true because $V(0)/A(0) = V(0)$ is the intial cost of a self-financing trading strategy that replicates the value of V at time T. In continuous time the same assertion holds. In general it is a consequence of the *martingale representation theorem*, which lies beyond the scope of this class. But for an option on a lognormal stock in a constant interest rate environment the argument reduces to our second explanation of the Black-Scholes PDE (bottom of page 7, Section 6). In fact, in that setting $V(t)/A(t) = V(t)e^{-rt}$ is a martingale because

$$d(e^{-rt}V(s(t),t)) = e^{-rt}\sigma s V_s \, dw$$

by Ito combined with the Black-Scholes PDE. This is equivalent (another application of Ito) to

$$dV(s(t),t) = \sigma s V_s \, dw + rV \, dt = V_s \, ds + r(V - sV_s) \, dt.$$

This equation is familiar: we used it in Section 6 to know that our trading strategy (holding V_s units of stock and a bond worth $V - sV_s$ at each time) was self-financing. In summary: if we knew nothing about the Black-Scholes PDE, but we knew that $V(t)e^{-rt}$ was a martingale (and a little more: we would need the coeffient of dw in the SDE satisfied by Ve^{-rt}), we could identify – by arguing as above – a trading strategy with initial cost $V(0)$ that replicates the option.