Continuous Time Finance Notes, Spring 2004 – Section 10, *April 7, 2004*
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use in connection with the NYU course Continuous Time Finance.

Jump-diffusion models. Merton was the first to explore option pricing when the underlying follows a jump-diffusion model. His 1976 article "Option pricing when underlying stock returns are discontinuous" (reprinted as Chapter 9 of his book *Continuous Time Finance*) is a pleasure to read. My notes follow it – but the article contains much more information than I'm presenting here. Much has happened since 1976 of course; a recent reference is A. Lipton, *Assets with jumps*, *RISK*, Sept. 2002, 149-153.

I will begin with an introduction to jump-diffusions. Then I'll discuss option pricing using such models. This cannot be done using absence of arbitrage alone: when the underlying can jump the market is not complete, since there are two sources of noise (the diffusion and the jumps) but just one tradeable (the underlying). How, then, can we price options? Merton's proposal (still controversial) was to assume that the extra randomness due to jumps is uncorrelated with the market – i.e. its β is zero. This means it can be made negligibile by diversification, and (by the Capital Asset Pricing Model) only the *average* effect of the jumps is important for pricing.

The analogue of the Black-Scholes PDE for a jump-diffusion model is an *integrodifferential* equation. You may wonder how one could ever hope to solve it. In the constant-coefficient setting the Fourier transform is a convenient tool. That's beyond the scope of this course. I've nevertheless included a discussion of the Fourier transform and its use in this setting, as enrichment reading for those who have sufficient background.

****

Jump-diffusion processes. The standard (constant-volatility) Black-Scholes model assumes that the logarithm of an asset price is normally distributed. In practice however the observed distributions are not normal – they have "fat tails," i.e. the probability of a very large positive or negative change is (though small) much larger than permitted by a Gaussian. The jump-diffusion model provides a plausible mechanism for explaining the fat tails and their consequences.

A one-dimensional diffusion solves $dy = \mu\, dt + \sigma\, dw$. (Here μ and σ can be functions of y and t.) A jump-diffusion solves the same stochastic differential equation most of the time, but the solution occasionally jumps.

We need to specify the statistics of the jumps. We suppose the occurrence of a jump is a Poisson process with rate λ. This means the jumps are entirely independent of one another. Some characteristics of Poisson processes:

(a) The probability that a jump occurs during a short time interval of length Δt is $\lambda\Delta t + o(\Delta t)$.

(b) The probability of two or more jumps occuring during a short time interval of length Δt is negligible, i.e. $o(\Delta t)$.

(c) The probability of exactly n jumps occuring in a time interval of length t is $\frac{(\lambda t)^n}{n!} e^{-\lambda t}$.

(d) The mean waiting time for a jump is $1/\lambda$.

We also need to specify what happens when a jump occurs. Our assumption is that a jump takes y to $y + J$. The jump magnitudes are independent, identically distributed random variables. In other words, each time a jump occurs, its size J is selected by drawing from a pre-specified probability distribution.

This model is encapsulated by the equation

$$dy = \mu\, dt + \sigma\, dw + J dN$$

where N counts the number of jumps that have occurred (so it takes integer values, starting at 0) and J represents the random jump magnitude. Ito's Lemma can be extended to this setting: if $v(x,t)$ is smooth enough and y is as above then $v(y(t),t)$ is again a jump-diffusion, with

$$d[v(y(t),t)] = (v_t + \mu v_x + \tfrac{1}{2}\sigma^2 v_{xx})dt + \sigma v_x dw + [v(y(t)+J,t) - v(y(t),t)]dN.$$

All the terms on the right are evaluated at $(y(t),t)$, as usual. In writing the jump term, we're trying to communicate that while the occurence of a jump in $v(y(t),t)$ is determined by N (i.e. by the presence of a jump in y) the size of the jump depends on $y(t)$ and the form of v.

Now consider the expected final-time payoff

$$u(x,t) = E_{y(t)=x}\left[w(y(T))\right]$$

where $w(x)$ is an arbitrary "payoff" (later it will be the payoff of an option). It is described as usual by a backward Kolmogorov equation

$$u_t + \mathcal{L}u = 0 \text{ for } t < T, \quad \text{with} \quad u(x,T) = w(x) \text{ at } t = T. \tag{1}$$

The operator $\mathcal{L}$ for our jump-diffusion process is

$$\mathcal{L}u = \mu u_x + \tfrac{1}{2}\sigma^2 u_{xx} + \lambda E\left[u(x+J,t) - u(x,t)\right].$$

The expectation in the last term is over the probability distribution of jumps. The proof of (1) follows the standard strategy used for diffusions without jumps (see e.g. Section 1 of my PDE for Finance lecture notes for a review of this topic). Let u solve (1), and apply Ito's formula. This gives

$$\begin{aligned} u(y(T),T) - u(x,t) \;&=\; \int_0^T (\sigma u_x)(y(s),s)\, dw + \int_0^T (u_s + \mu u_x + \tfrac{1}{2}\sigma^2 u_{xx})(y(s),s)\, ds \\ &\quad + \int_0^T [u(y(s)+J,s) - u(y(s),s)]dN. \end{aligned}$$

Now take the expectation. Only the jump term is unfamiliar; since the jump magnitudes are independent of the Poisson jump occurence process, we get

$$E\left([u(y(s)+J,s) - u(y(s),s)]dN\right) = E\left([u(y(s)+J,s) - u(y(s),s)]\right)\lambda ds.$$

Thus when u solves (1) we get

$$E[u(y(T),T)] - u(x,t) = 0.$$

This gives the result, since $u(y(T),T) = w(y(T))$ from the final-time condition on u.

A similar argument shows that the discounted final-time payoff

$$u(x,t) = E_{y(t)=x}\left[e^{-r(T-t)}w(y(T))\right]$$

solves

$$u_t + \mathcal{L}u - ru = 0 \text{ for } t < T, \quad \text{with} \quad u(x,T) = w(x) \text{ at } t = T,$$

using the same operator $\mathcal{L}$.

What about the probability distribution? It solves the forward Kolmogorov equation,

$$p_s - \mathcal{L}^* p = 0 \text{ for } s > 0, \quad \text{with} \quad p(z,0) = p_0(z)$$

where p_0 is the initial probability distribution and $\mathcal{L}^*$ is the adjoint of $\mathcal{L}$. (See Section 1 of my PDE for Finance lecture notes for an explanation why this must be so.) What is the adjoint of the new jump term? For any functions $\xi(z), \eta(z)$ we have

$$\int_{-\infty}^{\infty} E[[\xi(z+J) - \xi(z)]\,\eta(z)\,dz = \int_{-\infty}^{\infty} \xi(z)E[[\eta(z-J) - \xi(z)]\;dz =$$

since $\int_{-\infty}^{\infty} E[\xi(z+J)]\eta(z)\,dz = \int_{-\infty}^{\infty} \xi(z)E[\eta(z-J)]\,dz$. Thus

$$\mathcal{L}^* p = \tfrac{1}{2}(\sigma^2 p)_{zz} - (\mu p)_z + \lambda E\left[p(z-J) - p(z)\right],$$

i.e. the probability distribution satisfies

$$p_s - \tfrac{1}{2}(\sigma^2 p)_{zz} + (\mu p)_z - \lambda E\left[p(z-J,s) - p(z,s)\right] = 0.$$

****

Hedging and the risk-neutral process. I'd like to say that the time-0 value of an option with payoff $w(S)$ should be the discounted final-time payoff under "the risk-neutral dynamics." This is not obvious; indeed, it is a modeling hypothesis not consequence of arbitrage. I now explain the meaning and logic of this hypothesis, following Merton.

It is convenient to focus on the stock price itself not its logarithm. If (as above) J represents a typical jump of $\log S$ then the stock price leaps from S to $e^J S$, i.e. the jump in stock price is $(e^J - 1)S$. So (applying the jump-diffusion version of Ito to $S = e^y$, with $dy = \mu dt + \sigma dw$) the stock dynamics is

$$dS = (\mu + \tfrac{1}{2}\sigma^2)Sdt + \sigma Sdw + (e^J - 1)SdN.$$

The associated risk-neutral process is determined by two considerations:

(a) it has the same volatility and jump statistics – i.e. it differs from the subjective process only by having a different drift; and

(b) under the risk-neutral process $e^{-rt}S$ is a martingale, i.e. $dS - rSdt$ has mean value 0.

We easily deduce that the risk-neutral process is

$$dS = (r - \lambda E[e^J - 1])Sdt + \sigma Sdw + (e^J - 1)SdN. \tag{2}$$

Applying Ito once more, we see that under the risk-neutral dynamics $y = \log S$ satisfies

$$dy = (r - \tfrac{1}{2}\sigma^2 - \lambda E[e^J - 1])dt + \sigma dw + JdN$$

Thus the formalism developed in the preceding subsection can be used to price options; we need only set $\mu = r - \frac{1}{2}\sigma^2 - \lambda E[e^J - 1]$.

But is this right? And what are its implications for hedging? To explain, let's examine what becomes of the standard demonstration of the Black-Scholes PDE in the presence of jumps. Assume the option has a well-defined value $u(S(t), t)$ at time t. Suppose you try to hedge it by holding a long position in the option and a short position of Δ units of stock. Then over a short time interval the value of your position changes by

$$\begin{aligned} d[u(S(t),t)] - \Delta dS \;\; &= \;\; u_t dt + u_S([\mu + \tfrac{1}{2}\sigma^2]Sdt + \sigma Sdw) + \tfrac{1}{2}u_{SS}\sigma^2 S^2 dt \\ &\quad\; + [u(e^J S(t), t) - u(S(t), t)]dN \\ &\quad\; - \Delta([\mu + \tfrac{1}{2}\sigma^2]Sdt + \sigma Sdw) - \Delta(e^J - 1)SdN. \end{aligned}$$

There are two sources of randomness here – the Brownian motion dw and the jump process dN – but only one tradeable. So the market is incomplete, and there is no choice of Δ that makes this portfolio risk-free.

But consider the choice $\Delta = u_S(S(t), t)$. With this choice the randomness due to dw cancels, leaving only the uncertainty due to jumps:

$$\text{change in portfolio value} = (u_t + \tfrac{1}{2}\sigma^2 S^2 u_{SS})dt + \{[u(e^J S(t), t) - u(S(t), t)] - u_S(e^J S - S)\}dN.$$

To make progress, we must assume something about the statistics of the jumps. Merton's suggestion (still controversial) was to assume they are uncorrelated with the marketplace. The impact of such randomness can be eliminated by diversification. Put differently: according the the Capital Asset Pricing Model, for such an investment (whose β is zero) only the mean return is relevant to pricing. So the mean return on our hedge portfolio should be the risk-free rate:

$$(u_t + \tfrac{1}{2}\sigma^2 S^2 u_{SS})dt + \lambda E[u(e^J S(t), t) - u(S(t), t) - (e^J S - S)u_S]dt = r(u - Su_S)dt. \tag{3}$$

After rearrangement, this is precisely the backward Kolmogorov equation describing the discounted final-time payoff under the risk-neutral dynamics (2):

$$u_t + (r - \lambda E[e^J - 1])Su_S + \tfrac{1}{2}\sigma^2 S^2 u_{SS} - ru + \lambda E[u(e^J S, t) - u(S, t)] = 0.$$

A final remark about the experience of the investor who follows this hedge rule. If the option value is convex in S (as we expect for a call or a put) then the term in (3) associated with the jumps is positive:

$$E[u(e^J S(t),t) - u(S(t),t) - (e^J S - S)u_S] \geq 0.$$

So in the absence of jumps the value of the hedge portfolio (long the option, short $\Delta = u_S$ units of stock) rises a little slower than the risk-free rate. Without accounting for jumps, the investor who follows this hedge appears to be falling behind (relative to a cash investment at the risk-free rate). But due to convexity the net effect of the jumps is favorable – exactly favorable enough that the investor's long-term (mean) experience is risk-free.

****

Let's review why we're doing this. Certainly not for pricing European options, which are relatively liquid – their prices are visible in the marketplace. Interpreting their prices using standard Black-Scholes (deducing an implied volatility from the option price) one obtains a result in contradiction to the model: the implied volatility depends on maturity and strike price (the dependence on strike price is often called the "volatility smile"). The introduction of jumps provides a plausible family of models that's better able to fit the market data. But it introduces headaches of modeling and calibration (e.g. how to choose the distribution of jumps?). If the goal were merely to price Europeans, there would be no reason to bother – their prices are visible in the marketplace.

So why are we doing this? Three reasons. One is the desire for a consistent theoretical framework. The second, more practical, is the need to hedge (not simply to price) options – the Delta predicted by a jump-diffusion model is different from that of the Black-Scholes framework. A third reason is the need to price and hedge exotic options (e.g. barriers) which are less liquid. The basic idea: calibrate your jump-diffusion model using the market prices of European options, then use it to price and hedge barrier options.

****

Solution via Fourier transform. To assess what a jump-diffusion model says about "fat tails" we need a scheme for solving the forward Kolmogorov equation. And to value options (e.g. to calibrate the model to market prices) we need a scheme for solving the backward Kolmogorov equation. In the constant-coefficient setting these integrodifferential equations can be solved using the Fourier transform. The rest of these notes explain how. (This material is offered for enrichment purposes only; it will not be presented in lecture, and will not be required for homeworks or the exam.)

We henceforth focus on

$$dy = \mu dt + \sigma dw + J dN$$

with μ and σ *constant*. To be specific let's focus on the *forward* Kolmogorov equation, which is now

$$p_s - \tfrac{1}{2}\sigma^2 p_{zz} + \mu p_z - \lambda E\left[p(z-J,s) - p(z,s)\right] = 0 \tag{4}$$

and let us solve it with initial condition $p(0) = \delta_{z=0}$. (This will give us the fundamental solution, i.e. $p(z, s)$ = probability of being at z at time s given that you started at 0 at time 0.)

Why is the Fourier transform useful in this setting? Basically, because the forward equation is a mess – nonlocal in space (due to the jump term) and involving derivatives too (the familiar terms). But when we take its Fourier transform in space we get a simple, easy-to-solve ODE. The result is a simple expression not for the probability distribution itself, but rather for its Fourier transform – what a probabilist would call the *characteristic function* of the distribution.

Most students in this class will be relatively unfamiliar with the Fourier transform. Here's a brief summary of what we'll use:

(a) Given a function $f(x)$, its Fourier transform $\hat{f}(\xi) = \mathcal{F}[f](\xi)$ is defined by

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{ix\xi}\,dx.$$

Notice that even if f is real-valued, $\hat{f}$ is typically complex-valued.

(b) Elementary manipulation reveals that the translated function $f(x - a)$ has Fourier transform

$$\mathcal{F}[f(x-a)](\xi) = e^{i\xi a}\hat{f}(\xi)$$

(c) Integration by parts reveals that the Fourier transform takes differentiation to multiplication:

$$\mathcal{F}[f_x](\xi) = -i\xi\hat{f}(\xi)$$

(d) It is less elementary to prove that the Fourier transform takes convolution to multiplication: if $h(x) = \int_{-\infty}^{\infty} f(x-y)g(y)\,dy$ then

$$\hat{h}(\xi) = \hat{f}(\xi)\hat{g}(\xi).$$

(e) Another less elementary fact is Plancherel's formula:

$$\int_{-\infty}^{\infty} \overline{f} g\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\mathcal{F}[f]}\mathcal{F}[g]\,d\xi$$

where $\overline{f}$ is the complex conjugate of f.

(f) The Fourier transform is invertible, and

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ix\xi}\hat{f}(\xi)\,d\xi$$

(g) The Fourier transform of a Gaussian is again a Gaussian. More precisely, for a centered Gaussian with variance σ^2,

$$\mathcal{F}\left[\frac{1}{\sqrt{2\pi}\sigma}e^{-x^2/2\sigma^2}\right] = e^{-\sigma^2\xi^2/2}.$$

OK, let's use this tool to solve the forward equation (4). We apply the Fourier transform (in space) to each term of the equation. The first three terms are easy to handle using property (c). For the jump term, let the jump J have probability density ω, that

$$E\left[p(z-J)-p(z)\right]=\int [p(z-J)-p(z)]\omega(J)\,dJ=\int p(z-J)\omega(J)\,dJ-p(z).$$

By property (d) this the Fourier transform of this term is $(\hat{\omega}-1)\hat{p}$. Thus in the Fourier domain the equation becomes

$$\hat{p}_s+\tfrac{1}{2}\sigma^2\xi^2\hat{p}-i\xi\mu\hat{p}-\lambda(\hat{\omega}-1)\hat{p}=0$$

with initial condition $\hat{p}(0)=\int e^{i\xi z}\delta_{z=0}\,dz=1$. Writing the equation in the form

$$\hat{p}_s=K(\xi)\hat{p}$$

with

$$K(\xi)=-\tfrac{1}{2}\sigma^2\xi^2+i\xi\mu+\lambda(\hat{\omega}(\xi)-1)$$

we recognize immediately that the solution is

$$\hat{p}(\xi,s)=e^{sK(\xi)}. \tag{5}$$

The probability density itself can of course be obtained by taking the inverse Fourier transform:

$$p(z,s)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iz\xi+sK(\xi)}\,d\xi. \tag{6}$$

Option pricing. What use is this? Well, one can explore what the tails of the distribution look like, as you vary the jump rate λ and jump distribution ω. Or one can use it to price options. Recall that the value of an option should be the discounted expected payoff relative to the risk-neutral probability distribution. Assume for the moment that the under the risk-neutral distribution the stock price is e^y where y is a jump-diffusion as above. Then the time-0 value of an option with payoff w is

$$e^{-rT}E_{y(0)=\ln S_0}[w(e^y)]$$

if the time-0 stock price is S_0. If $w(S)$ vanishes for S near 0 and ∞ (corresponding to y very large negative or positive) then the option price is easy to express. Let $v(y)=w(e^y)$ be the payoff as a function of y, and let $x=\ln S_0$. By translation invariance, the probability density of y is $p(z-x,T)$ where p is given by (6). Therefore

$$\begin{aligned}E_{y(0)=\ln S_0}[w(e^y)] &= \int p(z-x,T)v(z)\,dz\\ &= \frac{1}{2\pi}\int\overline{\mathcal{F}[p(z-x,T)]}\mathcal{F}[v]\,d\xi\\ &= \frac{1}{2\pi}\int e^{-i\xi x}\hat{p}(-\xi,T)\hat{v}(\xi)d\xi.\end{aligned} \tag{7}$$

In the last step we used that $\mathcal{F}[p(z-x,T)] = e^{i\xi x}\hat{p}(\xi,T)$, and the fact that the complex conjugate of $\hat{p}(\xi,T) = \int e^{i\xi z}p(z,T)\,dz$ is $\int e^{-i\xi z}p(z,T)\,dz = \hat{p}(-\xi,T)$ since p is real. Equation (7) reduces the task of option pricing to the calculation of two Fourier transforms (those of ω and v) followed by a single integration (7).

The hypothesis that the payoff $w(S)$ vanishes near $S = 0$ and $S = \infty$ is inconvenient, because neither a put nor a call satisfies it. Fortunately this hypothesis can be dispensed with. Consider for example the call $w(S) = (S-K)_+$, for which $v(y) = (e^y - K)_+$. Its Fourier transform is not defined on the real axis, because the defining integral diverges as $y \to \infty$. But $e^{-\alpha y}v(y)$ decays near ∞ for $\alpha > 1$. So its Fourier transform in y is well-defined. This amounts to examining the Fourier transform of v along the line $\Im\xi = \alpha$ (here $\Im\xi$ is the imaginary part of ξ) since

$$\mathcal{F}[e^{-\alpha y}v(y)](\xi_1) = \int_{-\infty}^{\infty} e^{i(\xi_1+i\alpha)y}v(y)\,dy.$$

Fortunately, the Plancherel formula isn't restricted to integrating along the real axis in Fourier space; one can show that

$$\int_{-\infty}^{\infty} \overline{f}g\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\mathcal{F}[f]}(\xi_1+i\alpha)\mathcal{F}[g](\xi_1+i\alpha)\,d\xi_1$$

if the Fourier transforms of f and g exist (and are analytic) at $\Im\xi = \alpha$. Using this, an argument similar to that given above shows that

$$E_{y(0)=\ln S_0}[w(e^y)] = \int p(z-x,T)v(z)\,dz = \frac{1}{2\pi}\int e^{i(-\xi_1+i\alpha)x}\hat{p}(-\xi_1+i\alpha,T)\hat{v}(\xi_1+i\alpha)d\xi_1.$$

By the way, in the case of the call the Fourier transform of v is explicit and easy:

$$\hat{v}(\xi) = \int_{\ln K}^{\infty} e^{iy\xi}(e^y - K)\,dy = -\frac{K^{1+i\xi}}{\xi^2 - i\xi}$$

by elementary integration. Here $\xi = \xi_1 + i\alpha$ is any point in the complex plane such that the integral converges (this requires $\alpha > 1$).