Local Volatility, Stochastic Volatility and Jump-Diffusion Models

These notes provide a brief introduction to local and stochastic volatility models as well as jump-diffusion models. These models extend the geometric Brownian motion model and are often used in practice to price exotic derivative securities. It is worth emphasizing that the prices of exotics and other non-liquid securities are generally not available in the market-place and so models are needed in order to both price them and calculate their Greeks. This is in contrast to vanilla options where prices are available and easily seen in the market. For these more liquid options, we only need a model, i.e. Black-Scholes, and the volatility surface to calculate the Greeks and perform other risk-management tasks.

In addition to describing some of these models, we will also provide an introduction to a commonly used fourier transform method for pricing vanilla options when analytic solutions are not available. This transform method is quite general and can also be used in any model where the characteristic function of the log-stock price is available. Transform methods now play a key role in the numerical pricing of derivative securities.

1 Local Volatility Models

The GBM model for stock prices states that

$$dS_t \;=\; \mu S_t \, dt \;+\; \sigma S_t \, dW_t$$

where μ and σ are constants. Moreover, when pricing derivative securities with the cash account as numeraire, we know that $\mu = r - q$ where r is the risk-free interest rate and q is the dividend yield. This means that we have a single free parameter, σ, which we can fit to option prices or, equivalently, the volatility surface. It is not all surprising then that this exercise fails. As we saw before, the volatility surface is never flat so that a constant σ fails to re-produce market prices. This was particularly true after the crash of $'87$ when market participants began to correctly identify that lower strike options should be priced with a higher volatility, i.e. there should be a skew.

After this crash, researchers developed alternative models in an attempt to model the skew. While not the first[1] such model, the local volatility model is probability the simplest extension of Black-Scholes. It assumes that the stock's risk-neutral dynamics satisfy

$$dS_t \;=\; (r-q) S_t \, dt \;+\; \sigma_l(t, S_t) S_t \, dW_t \tag{1}$$

so that the instantaneous volatility, $\sigma_l(t, S_t)$, is now a function of time and stock price. The key result[2] of the local volatility framework is the *Dupire* formula that links the local volatilities, $\sigma_l(t, S_t)$, to the implied volatility surface.

Theorem 1 (The Dupire Formula) *Let $C = C(K, T)$ be the price of a call option as a function of strike and time-to-maturity. Then the local volatility function satisfies*

$$\sigma_l^2(T, K) \;=\; \frac{\frac{\partial C}{\partial T} \;+\; (r-q)K\,\frac{\partial C}{\partial K} \;+\; qC}{\frac{K^2}{2}\,\frac{\partial^2 C}{\partial K^2}}. \tag{2}$$

[1]Earlier models included Merton's jump-diffusion model, the CEV model and Heston's stochastic volatility model. Indeed the first two of these models date from the 1970's.

[2]The local volatility framework was developed by Derman and Kani (1994) and in continuous time by Dupire (1994).

Proof: Recall first Kolmogorov's forward equation for the PDF of the underlying stock. In particular, if $p(y,t)$ is the PDF of the stock price at time t and evaluated at $S_t = y$, then the forward equation is

$$-p_t - (r-q)(yp)_y + \frac{1}{2}(\sigma^2(t,y)y^2p)_{yy} \;=\; 0 \quad \text{for } t > 0 \tag{3}$$

with initial condition $p(y,t) = \delta_{S_0}(y)$ at $t = 0$. We can also write the call option price as

$$C(K,T) \;=\; e^{-rT}\mathsf{E}_0\left[(S_T - K)^+\right] \;=\; e^{-rT}\int_K^\infty (y-K)p(y,T)\,dy \tag{4}$$

and if we differentiate across (4) twice with respect to K we obtain

$$C_{KK}(K,T) \;=\; e^{-rT}p(K,T). \tag{5}$$

Differentiating across (5) with respect to T yields

$$\begin{aligned}
C_{KKT}(K,T) \;+\; rC_{KK}(K,T) &= e^{-rT}p_T(K,T) \\
&= e^{-rT}\left(\frac{1}{2}(\sigma^2(T,K)K^2p)_{KK} - (r-q)(Kp)_K\right) \qquad (6)\\
&= \frac{1}{2}(\sigma^2(T,K)K^2C_{KK})_{KK} - (r-q)(KC_{KK})_K \qquad (7)
\end{aligned}$$

where (6) follows from (3) and (7) follows from substituting C_{KK} for $e^{-rT}p(K,T)$. Integrating across (7) with respect to K yields

$$C_{KT}(K,T) \;+\; rC_K(K,T) \;-\; \frac{1}{2}(\sigma^2(T,K)K^2C_{KK})_K \;+\; (r-q)(KC_{KK}) \;=\; h(T) \tag{8}$$

for some function, $h(T)$. Since $KC_{KK} = (KC_K)_K - C_K$, we can rewrite (8) as

$$C_{KT}(K,T) \;-\; \frac{1}{2}(\sigma^2(T,K)K^2C_{KK})_K \;+\; r(KC_K)_K \;-q((KC_K)_K - C_K) \;=\; h(T). \tag{9}$$

Integrating across (9) again with respect to K yields

$$C_T(K,T) \;-\; \frac{1}{2}\sigma^2(T,K)K^2C_{KK} \;+\; r(KC_K) \;-q(KC_K - C) \;=\; h(T)K \;+\; g(T). \tag{10}$$

for some function, $g(T)$. Now note that as $K \to \infty$ the call option price as well as its derivatives, C_T, C_K and C_{KK} all go to zero. But (10) then implies that $h(T) = g(T) = 0$ for all T after which (10) reduces to (2). ■

Given the implied volatility surface we can easily compute the corresponding *call option price surface* which is the graph of $C(K,T)$ as a function of K and T. It is then clear from (2) that we need to take first and second derivatives of this latter surface with respect to strike and first derivatives with respect to time-to-maturity in order to compute the local volatilities. Calculating the local volatilities from (2) is therefore difficult and can be unstable as computing derivatives numerically can itself be very unstable. As a result, it is necessary to use a sufficiently smooth Black-Scholes implied volatility[3] surface when calculating local volatilities using (2).

Remark 1 *It is worth emphasizing that the local volatility model (1) with $\sigma_l(\cdot,\cdot)$ computed according to (2) is, by construction, a self-consistent model that is capable of producing the implied volatility surface observed in the market place.*

[3]It is also possible to write the Dupire formula in terms of the implied volatilities rather than the call option prices. One can then work directly with the implied volatility surface to compute the local volatilities.

Local volatility is known to suffer from several weaknesses. For example, it leads to unreasonable skew dynamics and underestimates the volatility of volatility or "vol-of-vol". Moreover the Greeks that are calculated from a local volatility model are generally not consistent with what is observed empirically. Nevertheless, it is an interesting model from a theoretical viewpoint and is often used in practice for pricing barrier[4] options for example.

Gyöngy's Theorem

Gyöngy's Theorem is an important theoretical result that links local volatility models to other diffusion models that are also capable of generating the implied volatility surface. Consider a general n-dimensional Itô process, X_t, satisfying

$$dX_t \;=\; \alpha(t,\omega)\,dt \;+\; \beta(t,\omega)\,dW_t$$

where $\alpha(t,\omega)$ and $\beta(t,\omega)$ are $n \times 1$ and $n \times m$ adapted processes, respectively, and ω is a sample path of the m-dimensional Brownian motion, W_t. Then Gyöngy's Theorem states that there is a Markov process, Y_t, satisfying

$$dY_t \;=\; a(t,Y_t)\,dt \;+\; b(t,Y_t)\,dW_t$$

where X_t and Y_t have the same *marginal* distributions, i.e. X_t and Y_t have the same distribution for each t. Moreover, Y_t can be constructed by setting

$$\begin{aligned} a(t,y) &= \mathrm{E}_0\left[\alpha(t,\omega)|X_t = y\right] \quad \text{and} \\ b(t,y)b(t,y)^T &= \mathrm{E}_0\left[\beta(t,\omega)\beta^T(t,\omega)|X_t = y\right]. \end{aligned}$$

In a financial setting, X_t might represent the true risk-neutral dynamics of a particular security. Then $b(t,y)/y$ represents the local volatility function $\sigma_l(t,\cdot)$ in (1). Because X_t and Y_t have the same marginal distributions then we know (why?) that European option prices can be priced correctly if we assume the price dynamics are given by Y_t. In particular Y_t can produce the correct implied volatility surface. Moreover Gyöngy's Theorem therefore implies that the local volatility model of (1) is in some sense the simplest diffusion model capable of doing this, i.e. reproducing the implied volatility surface. Gyöngy's Theorem has been used recently to develop stochastic-local volatility models as well as approximation techniques for pricing various types of basket options.

1.1 The CEV Model

The *constant elasticity of variance* (CEV) model is a particular parametric local volatility model that was introduced by Cox (1975). The risk-neutral dynamics are assumed to follow

$$dS_t \;=\; (r-q)S_t\,dt \;+\; \sigma S_t^{\beta}\,dW_t \tag{11}$$

where as usual q is the dividend yield, r is the risk-free rate and σ and $\beta \in [0,1]$ are the remaining model parameters. Note that the CEV model generalizes GBM which is obtained when we set $\beta = 1$. The popularity of the CEV model is due to its tractability. In particular, analytic[5] expressions for options prices are available in terms of the non-central χ^2 distribution. This tractability also accounts for its use in term structure modeling, often in conjunction with *LIBOR market models*.

By writing (11) as

$$\frac{dS_t}{S_t} \;=\; (r-q)\,dt \;+\; \sigma S_t^{\beta-1}\,dW_t$$

we see that there is a negative relationship between price level and instantaneous volatility when $\beta < 1$. The CEV model is therefore able to capture some of the skew that is observed empirically in practice. It is also worth noting that when $\beta < 1/2$, there is a strictly positive probability that the CEV process will hit zero.

[4]While the Black-Scholes GBM framework can be used to barrier options analytically, it is well known that the Black-Scholes model is in fact a truly awful model for barrier options and that it should never be used in practice. As a result more sophisticated models and numerical methods such as PDE or Monte-Carlo methods are used. We will return to barrier options at a later stage to highlight the danger of using just one model to price exotic options.

[5]See Cox (1975) but also Schroder (1989).

1.2 How Well Do Local Volatility Models Work in Practice

A local volatility model that is calibrated using the Dupire formula and he market's implied volatility surface will of course perfectly capture the observed skew at any maturity. A more interesting question is whether or not the dynamics implied by (1) are an accurate representation of the true dynamics. Leaving aside the inability of any diffusion model to capture jumps, the answer must be "no". Gyöngy's Theorem provides some justification for this answer in that it suggests that the local volatility model is in some sense only a projection of a richer and possibly more accurate diffusion model. This projection property also suggests that local volatility models underestimate the volatility of volatility, also known as the *vol-of-vol*. Moreover this observation has been borne out in practice where it has been found that local volatility models often significantly underprice structured products. This is also related to a phenomenon whereby local volatility models generate *forward skews* that are too flat. As such, they are unsuited for pricing securities such as *cliquet options* that are sensitive to the forward skew.

Nonetheless, local volatility models are often used in practice to price various types of barrier options, for example. In order to account for the possibility of jumps and the inability to hedge these jumps, various ad-hoc adjustments are often made to the local volatility prices. There has also been some controversy[6] regarding the accuracy of *deltas* that are obtained from local volatility models. Indeed, users of local volatility models will often persist in using Black-Scholes deltas to hedge rather than the deltas implied by their local volatility model.

---

2 Stochastic Volatility Models

The most well-known and important stochastic volatility model is due to Heston (1989) and in this section we will concentrate exclusively on this model. It is a two-factor model and assumes separate dynamics for both the stock price and instantaneous volatility. In particular, it assumes

$$dS_t = (r-q)S_t\,dt + \sqrt{\sigma_t}S_t\,dW_t^{(s)} \tag{12}$$

$$d\sigma_t = \kappa\,(\theta-\sigma_t)\,dt + \gamma\sqrt{\sigma_t}\,dW_t^{(vol)} \tag{13}$$

where $W_t^{(s)}$ and $W_t^{(vol)}$ are standard Q-Brownian motions with constant correlation coefficient, ρ. In practice, ρ is generally (why?) negative.

Remark 2 *The volatility process in (13) is commonly used in interest rate modeling where it is known as the CIR[7] model. It has the property that the process will remain non-negative with probability one. For certain parameter combinations, it will always be strictly positive with probability one.*

Whereas the local volatility model is a complete model, Heston's stochastic volatility model is an incomplete[8] model. This should not be too surprising as there are two sources of uncertainty in the Heston model, $W_t^{(s)}$ and $W_t^{(vol)}$, but only one risky security and so not every security is replicable. Put another way, while the drift in (12) must be $r-q$ under any EMM with the cash account as numeraire, we could use Girsanov's Theorem to change the drift in (13) in infinitely many different ways without changing the drift in (12).

To see this let us first suppose that the P-dynamics of S_t and σ_t satisfy

$$dS_t = \mu_t S_t\,dt + \sqrt{\sigma_t}S_t\,dW_t^{(1)} \tag{14}$$

$$d\sigma_t = \nu_t\,dt + \gamma\sqrt{\sigma_t}\,\left(\rho dW_t^{(1)} + \sqrt{1-\rho^2}dW_t^{(2)}\right) \tag{15}$$

where μ_t and ν_t are some $\mathcal{F}_t$-adapted processes, and $W_t = (W_t^{(1)}, W_t^{(2)})$ is a standard 2-dimensional P-Brownian motion. Let us now define

$$L_t := \exp\left(-\int_0^t \eta_s'\,dW_s - \frac{1}{2}\int_0^t \eta_s'\,\eta_s\,ds\right)$$

[6] The SABR model was introduced to overcome the problems associated with the deltas of local volatility models.
[7] After Cox, Ingersoll and Ross (1985) who used this model for modeling the dynamics of the short interest rate.
[8] Assuming as usual that the stock and the cash-account are the only traded securities.

for $t \in [0, T]$ and where $\eta_t = (\eta_t^{(1)}, \eta_t^{(2)})$ is a 2-dimensional adapted process. Then[9] Girsanov's Theorem implies $\widehat{W}_t := W_t + \int_0^t \eta_s ds$ is a standard 2-dimensional Q^η-Brownian motion where $dQ^\eta/dP = L_T$. In particular the Q^η-dynamics of S_t and σ_t satisfy

$$dS_t = \left(\mu_t - \sqrt{\sigma_t}\eta_t^{(1)}\right) S_t\, dt + \sqrt{\sigma_t} S_t\, d\widehat{W}_t^{(1)} \tag{16}$$

$$d\sigma_t = \left(\nu_t - \gamma\sqrt{\sigma_t}\rho\eta_t^{(1)} - \gamma\sqrt{\sigma_t}\sqrt{1-\rho^2}\eta_t^{(2)}\right) dt + \gamma\sqrt{\sigma_t}\left(\rho d\widehat{W}_t^{(1)} + \sqrt{1-\rho^2} d\widehat{W}_t^{(2)}\right). \tag{17}$$

In order for Q^η to be an EMM it is only necessary (why?) that

$$\mu_t - \sqrt{\sigma_t}\eta_t^{(1)} = r - q$$

and we are free to choose $\eta_t^{(2)}$ subject only to the usual integrability constraints. This means that there are infinitely many EMM's and so[10] the model is incomplete. In the Heston model we choose $\eta_t^{(2)}$ so that

$$\nu_t - \gamma\sqrt{\sigma_t}\rho\eta_t^{(1)} - \gamma\sqrt{\sigma_t}\sqrt{1-\rho^2}\eta_t^{(2)} = \kappa\,(\theta - \sigma_t)$$

is satisfied. We therefore recover (12) and (13) once we identify $\widehat{W}_t^{(1)}$ with $W_t^{(s)}$ and $W_t^{(vol)}$ with (via Levy's Theorem) $\rho\widehat{W}_t^{(1)} + \sqrt{1-\rho^2}\widehat{W}_t^{(2)}$. Note that we still have several free parameters which in practice we would determine by calibrating the model to the market prices of European options. This is the typical method of choosing an EMM in incomplete market models.

2.1 The Pricing PDE

Because Heston's model is incomplete it is not possible to price options using the replication arguments that apply to complete market models. Instead, one assumes Q-dynamics as in (12) and (13) and then prices all securities using this EMM. (The model parameters are chosen by calibrating model prices to observable market prices.) The pricing PDE that the price, $C(t, S_t, \sigma_t)$, of any derivative security must satisfy in Heston's model is given by

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma\gamma S \frac{\partial^2 C}{\partial S \partial \sigma} + \frac{1}{2}\gamma^2\sigma\frac{\partial^2 C}{\partial \sigma^2} + (r-q)S\frac{\partial C}{\partial S} + \kappa\,(\theta - \sigma)\frac{\partial C}{\partial \sigma} = rC. \tag{18}$$

This PDE follows of course via Feynman-Kac and the fact that

$$M(t, S_t, \sigma_t) := e^{-rt}\,\mathrm{E}_t^{\mathcal{Q}}\left[e^{-r(T-t)}\text{ Option Payoff}\right] = e^{-rt}\, C(t, S_t, \sigma_t) \tag{19}$$

must be (why?) a $\mathcal{Q}$-martingale. Derivative prices can then be obtained by solving (18) subject to the relevant boundary conditions or by using Monte-Carlo methods to estimate (19).

Heston succeeded in solving (18) in the case of European call options (and therefore put options via put-call parity) by conjecturing a solution of the form

$$C(t, S_t, \sigma_t) = S_t P_1(t, S_t, \sigma_t) - K e^{-r(T-t)} P_2(t, S_t, \sigma_t) \tag{20}$$

where K is the option strike, $T > t$ is the option maturity, and P_1 and P_2 are functions to be determined. Each of the two terms on the right-hand-side of (20) must (why?) also satisfy (18). Substituting each of them in turn into (18) leads to a corresponding PDE and terminal condition for P_j, for $j = 1, 2$. It is not possible to solve these PDE's for P_1 and P_2 in closed form but Heston was able to compute their Fourier transforms by guessing their functional forms and then reducing each PDE to a series of two ODEs which could be solved analytically. These transforms could then be inverted numerically to obtain the price of the call option via (20). Heston was also able to interpret P_1 and P_2 as risk-neutral probabilities (with respect to different EMMs) of the option expiring in the money. This observation should not[11] be surprising and indeed it holds more generally.

[9] We are assuming the necessary conditions. e.g. Novikov's condition, to ensure that L_t is a martingale.
[10] But we could make the model complete by introducing into the model another security whose price process depends on σ_t.
[11] See also Exercise 16.

Table 1: Call Option Price Estimates Using an Euler Scheme for Heston's Stochastic Volatility Model. The true option price is 13.085.

Time Steps	Sticky Zero	Reflection
100	28.3	45.1
200	27.1	41.3
500	25.6	37.1
1000	24.8	34.6

2.2 Simulating the Heston Model

Suppose we need to simulate a multi-dimensional SDE of the form

$$dX_t = \mu(t, X_t)\ dt + \sigma(t, X_t)\ dW_t \tag{21}$$

where W_t is an n-dimensional Brownian motion. Then the simplest discretization scheme for doing so is the so-called *Euler* scheme which satisfies

$$\widehat{X}_{kh} \ = \ \widehat{X}_{(k-1)h} + \mu\left((k-1)h,\ \widehat{X}_{(k-1)h}\right)\ h\ +\ \sigma\left((k-1)h,\ \widehat{X}_{(k-1)h}\right)\ \sqrt{h}\ Z_k \tag{22}$$

where the Z_k's are $\text{MVN}_n(0, I_n)$. We can then simulate $\{\widehat{X}_h, \widehat{X}_{2h},\ \ldots\ , \widehat{X}_{mh}\}$, where m is the number of time steps, h is a constant and $mh = T$. The smaller the value of h, the closer the discretized path in (22) will be to the continuous-time path of (21) that we wish to simulate. Of course this will be at the expense of greater computational effort.

Special care must be taken, however, if we wish to simulate the dynamics in (12) and (13). In fact the Euler scheme of (22) does not always converge well when the dynamics are as in (13), even when we take m very large. For example, Andersen[12] (2007) considers the problem of pricing an at-the-money 10-year call option when $r = q = 0$. He takes $\kappa = .5$, $\sigma_0 = \theta = .04$, $\gamma = 1$, $S_0 = 100$ and $\rho := \text{Corr}(W_t^{(s)}, W_t^{(vol)}) = -0.9$. The true value of the option is 13.085. Using one million sample paths and a "sticky zero" or "reflection" assumption[13], he obtains the estimates displayed in the table above for the option price as a function of m, the number of discretization points. It is clear that neither scheme has converged even when taking as many as $1,000$ discretization points.

One therefore needs to be very careful when applying an Euler scheme to the Heston SDE and indeed to jump-diffusion models where the diffusion component has the dynamics of (13). In fact when simulating Heston dynamics it is much better to use an alternative scheme such as that proposed by Andersen (2007). More generally, it is a good idea to price vanilla securities in the Monte-Carlo alongside the more exotic securities that are of direct interest. If the estimated vanilla security prices are not comparable to their analytic prices, then we know the scheme has not converged.

2.3 The Characteristic Function of the log-Stock Price

We will see in Section 4.1 that the characteristic function, i.e. the Fourier transform, of the log-stock price is very useful for pricing options. For the Heston model of (12) and (13), the characteristic function satisfies

$$\begin{aligned}
\phi_T(u) &= \text{E}[\exp(iulog(S_T) \mid S_0, \sigma_0)] \\
&= \exp(iu(log(S0) + (r-q)T)) \\
&\quad \times\ \exp(\theta\kappa\gamma^{-2}((\kappa - \rho\gamma ui - d)T - 2\log((1 - g\exp(-dT))/(1-g)))) \\
&\quad \times \exp(\sigma_0^2\gamma^{-2}(\kappa - \rho\gamma ui - d)(1 - \exp(-dT))/(1 - g\exp(-dT)))
\end{aligned} \tag{23}$$

[12] "Simple, Efficient, and Robust Simulation of Stochastic Volatility Models, ICBI (2007) conference presentation.

[13] The sticky zero assumption simply means that anytime the variance process, V_t, goes negative in the Monte-Carlo it is replaced by 0. The reflection assumption replaces V_t with $|V_t|$. In the limit as $m \to \infty$, the variance will stay non-negative with probability 1 so both assumptions are unnecessary in the limit and the option prices should be identical.

where $i = 0 + 1i$ is imaginary and

$$\begin{aligned} d &= \sqrt{(\rho\gamma ui - \kappa)^2 + \gamma^2(iu + u^2)} \\ g &= (\kappa - \rho\gamma ui - d)/(\kappa - \rho\gamma ui + d). \end{aligned}$$

It is worth noting that there is another equivalent[14] representation, $\widehat{\phi}_T(u)$ say, of the characteristic function that is very similar to $\phi_T(u)$ in (23). While this representation is correct, using the *principal value* of the $\log$ function in $\widehat{\phi}_T(u)$ causes $\widehat{\phi}_T(u)$ to jump discontinuously when the imaginary component of the argument of the $\log$ function crosses the negative real axis. As a result, when $\widehat{\phi}_T(u)$ is used in a numerical integration scheme such as (62) to price options, wildly inaccurate prices can be obtained. The representation in (23) avoids this problem.

2.4 How Well Does Heston's Model Capture the Skew?

An interesting question that arises is whether or not Heston's model accurately represents the dynamics of stock prices. This question is often reduced in practice to the less demanding question of how well the Heston model captures the volatility skew. By "capturing" the skew we have in mind the following: once the Heston model has been calibrated, then European option prices can be computed using numerical techniques such as Monte-Carlo, PDE or transform methods. The resulting option prices can then be used to determine the corresponding Black-Scholes implied volatilities. These volatilities can then be graphed to create the *model's* implied volatility surface which can then be compared to the *market's* implied volatility surface.

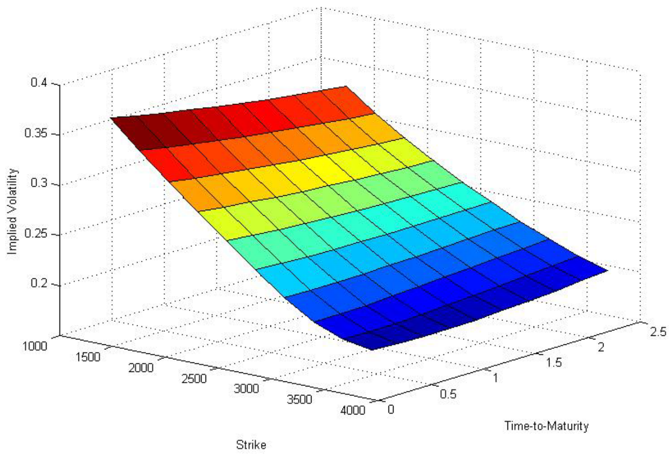

Figure 1: An Implied Volatility Surface under Heston's Stochastic Volatility Model

Figure 1 displays the implied volatility surface for the following choice of parameters: $r = .03$, $q = 0$, $\sigma_0 = \sqrt{.0654}$, $\gamma = .2928$, $\rho = -.7571$, $\kappa = .6067$ and $\theta = .0707$. Perhaps the most noticeable feature of this surface is the persistence of the skew for long-dated options. Indeed the Heston model generally captures longer-dated skew quite well but it typically struggles to capture the near term skew, particularly when the latter is very steep. The problem with a steep short-term skew is that any diffusion model will struggle to capture it as there is not enough time available for the stock price to diffuse sufficiently far from its current level. In order to solve this problem jumps are needed.

Note that some instruments can be priced analytically in Heston's model. For example, the price of the continuous-time version of a variance swap has a closed-form solution as the following example demonstrates.

[14] Heston's original paper as well as most papers in the literature report this alternative representation. See Section 4.6 of *Applications of Fourier Transform to Smile Modeling* (2010), Springer, by J.Zhu for a discussion of this problem and other possible remedies.

Example 1 (Variance-swaps in Heston's model)
Let $V_c(0,T)$ denote the total continuous realized variance in the Heston model from $t = 0$ to $t = T$. Then we know from (13) that (see Exercise 3)

$$\begin{aligned}
\mathrm{E}_0^{\mathbb{Q}}\left[V_c(0,T)\right] &= \mathrm{E}_0^{\mathbb{Q}}\left[\int_0^T \sigma_t \, dt\right] \\
&= \int_0^T \mathrm{E}_0^{\mathbb{Q}}\left[\sigma_t\right] \, dt \qquad (24) \\
&= \int_0^T \left(e^{-\kappa t}(\sigma_0 - \theta) + \theta\right) \, dt \qquad (25) \\
&= \frac{1 - e^{-\kappa T}}{\kappa}(\sigma_0 - \theta) \; + \; \theta T
\end{aligned}$$

so that the annualized variance is given by

$$\frac{\mathrm{E}_0^{\mathbb{Q}}\left[V_c(0,T)\right]}{T} = \frac{1 - e^{-\kappa T}}{\kappa T}(\sigma_0 - \theta) \; + \; \theta. \qquad (26)$$

▮

Note that the fair price of a variance-swap in Heston's model does not depend on γ or ρ. This should not be too surprising.

---

3 An Introduction to Jump-Diffusion Models

We now give a brief introduction[15] to jump-diffusion models.

Definition 1 *We say that J_t is a* **pure jump process** *if it is constant between jumps and is adapted and right-continuous.*

Typical examples of pure jump processes are Poisson processes, compound Poisson processes as described in Example 2 and *point processes* more generally.

Example 2 (A Compound Poisson Process)
Let N_t be a Poisson process with intensity λ. Then

$$X_t := \sum_{i=0}^{N_t} Y_i$$

where the Y_i's are IID random variables is a compound Poisson process. It is easy to check that $\mathrm{E}[X_t] = \lambda t \mu_y$ where $\mu_y := \mathrm{E}[Y]$ and that $\mathrm{Var}(X_t) = \lambda t \mathrm{E}[Y^2]$. It is also easy to check that $M_t := X_t - \lambda \mu_y t$ is a martingale and that the moment generating function (MGF) of X_t is given by

$$\phi_{X_t}(u) := e^{\lambda t(\phi_Y(u) - 1)} \qquad (27)$$

where $\phi_Y(u)$ is the MGF of Y. ▮

We now state[16] an important result that aids our understanding of compound processes and jump-diffusions more generally. It effectively states that if we have a compound Poisson process with a finite number of possible jump sizes, then we can view this process equivalently as a sum of independent Poisson processes in which the size one jumps are replaced by jumps of a fixed size.

---

[15] This introduction is taken from Chapter 11 of *Stochastic Finance for Finance II: Continuous-Time Models* by Steve Shreve. This chapter is an excellent introduction and contains far more material than we can cover in these notes.
[16] This is Corollary 11.3.4 in Shreve which also contains an outline proof of the result.

Theorem 2 *Let $y_1, \ldots, y_M$ be a finite set of nonzero numbers and let $p_1, \ldots, p_M$ be positive numbers that sum to 1. Let $N(t)$ be a Poisson process with intensity λ and define the compound Poisson process*

$$X_t := \sum_{i=1}^{N(t)} Y_i$$

where the Y_i's are an IID sequence of random variables with $P(Y_i = y_j) = p_j$ for all i, j. For $m = 1, \ldots, M$, let $N_m(t)$ be the number of jumps of size y_m up to and including time t. Then

$$N(t) = \sum_{m=1}^{M} N_m(t) \quad \text{and} \quad X_t = \sum_{m=1}^{M} y_m N_m(t).$$

The processes $N_1, \ldots, N_M$ are independent Poisson processes and each N_m has intensity λp_m.

Definition 2 *A* **jump-diffusion** *or* **jump** *process is a process of the form*

$$\begin{aligned} X_t &= X_0 + \int_0^t \gamma_s \, dW_s + \int_0^t \theta_s \, ds + J_t \\ &=: X_t^c + J_t \end{aligned} \tag{28}$$

where J_t is a pure jump process and X_t^c is the continuous *part of X_t. Of course, γ_s and θ_s are adapted processes and W_s is a standard Brownian motion.*

Note that definition 1 implies that jump-diffusion processes as defined in (28) are *right-continuous*. This means that $X_t = \lim_{s \downarrow t} X_s$. We therefore use X_{t-} to denote the left-continuous version of the process so that

$$X_{t-} := \lim_{s \uparrow t} X_s.$$

The difference between X_t and X_{t-} is then the jump-size, $\Delta X_t = \Delta J_t$, at time t. In the case of the jump-diffusion of (28) we see that

$$X_{t-} = X_0 + \int_0^t \gamma_s \, dW_s + \int_0^t \theta_s \, ds + J_{t-}$$

as the Riemann (or Lebesgue) and stochastic integral components of X_t are both continuous in t.

3.1 Stochastic Integrals with Respect to a Jump-Diffusion

We can now define the stochastic integral of an adapted process with respect to X_t.

Definition 3 *The* **stochastic integral** *of the process Φ_t with respect to a jump-diffusion, X_t, is*

$$\int_0^t \Phi_s \, dX_s = \int_0^t \Phi_s \gamma_s \, dW_s + \int_0^t \Phi_s \theta_s \, ds + \sum_{0 < s \le t} \Phi_s \Delta J_s. \tag{29}$$

In differential notation we write (29) as

$$\Phi_t \, dX_t = \Phi_t \gamma_t \, dW_t + \Phi_t \theta_t \, dt + \Phi_t \, \Delta J_t.$$

Example 3 **(Shreve E.G.** 11.4.4)
Let $X_t := N_t - \lambda t$ where N_t is a Poisson process with intensity λ. Now let $\Phi_s = \Delta N_s$ so that Φ_s is 1 if there is jump at time s and 0 otherwise. Using Definition 3 we find

$$\begin{aligned} \int_0^t \Phi_s \, dX_s &= \int_0^t \Phi_s \, dN_s \; - \; \lambda \int_0^t \Phi_s \, ds \\ &= N_t - 0 = N_t. \end{aligned} \tag{30}$$

▮

It is perhaps surprising that the stochastic integral in Example 3 fails to be a martingale despite the fact that the integrator, X_t, is a martingale. This occurs because the integrand, Φ_s, is not left-continuous. Indeed we have the following[17] theorem.

Theorem 3 *If the jump-diffusion process, X_t, of (28) is a martingale and Φ_t is left-continuous and adapted, and $\mathrm{E}[\int_0^t \gamma_s^2 \Phi_s^2 \, ds] < \infty$ for all $t \geq 0$, then the stochastic integral $\int_0^t \Phi_s \, dX_s$ is a martingale.*

When we work with jump processes we will often want to insist that Φ_s be left-continuous or (almost equivalently), predictable. This is particularly true for financial applications where Φ_s can then be interpreted as a trading strategy. However, it is worth emphasizing that (29) is still defined if we only assume Φ_s is adapted.

Example 4 **(Shreve E.G.** 11.4.6)
Let X_t be as in Example 3 and define $\Phi_s := I_{[0,S_1]}(s)$ where S_1 is the time of the first jump. Note that Φ_s is left-continuous and we now obtain

$$\begin{aligned} \int_0^t \Phi_s \, dX_s &= \begin{cases} -\lambda t, & 0 \leq t < S_1 \\ 1 - \lambda S_1 & t \geq S_1 \end{cases} \\ &= I_{[S_1,\infty)}(t) - \lambda(t \wedge S_1). \end{aligned} \tag{31}$$

where $x \wedge y := \min(x, y)$. We can verify by direct computation that $I_{[S_1,\infty)}(t) - \lambda(t \wedge S_1)$ is a martingale as implied by Theorem 3. It can also be checked that if we had taken $\Phi_s = I_{[0,S_1)}(s)$ which is right-continuous but not left-continuous, then we would have found that

$$\int_0^t \Phi_s \, dX_s \;=\; -\lambda(t \wedge S_1).$$

Since $\mathrm{E}_0[-\lambda(t \wedge S_1)] = e^{-\lambda t} - 1$ it follows that in this case $\int_0^t \Phi_s \, dX_s$ is not a martingale. ▮

3.2 Quadratic Variation and Cross Covariation

Consider a partition of the time interval $[0, T]$ given by $\Pi := 0 = t_0 < t_1 < t_2 < \ldots < t_n := T$. Let X_t be a jump-process as in (28) and consider the sum of squared changes

$$Q_\Pi(X) := \sum_{i=1}^n \left(X_{t_i} - X_{t_{i-1}}\right)^2 \tag{32}$$

so that the *quadratic variation* of X on $[0, T]$ is defined as

$$[X, X](T) := \lim_{||\Pi|| \to 0} Q_\Pi(X).$$

[17]This is Theorem 11.4.5 in Shreve. He points out that the result holds more generally when the integrand is *predictable*. Given a given probability space $(\Omega, \mathcal{F}, P)$, we say that a process Y_t is predictable if $Y : \Omega \times [0, \infty) \to \mathcal{R}$ is measurable with respect to the σ-algebra generated by the set of all left-continuous adapted processes.

If for example, $X_t = \int_0^t \gamma_s \, dW_s$ then we know from our results on diffusions that $[X,\ X](T) = \int_0^T \gamma_s^2 \, ds$ which of course is stochastic in general as γ_s itself is stochastic (and adapted) in general.

Now let X_1 and X_2 be two jump processes so that for $i = 1, 2$ we have

$$X_i(t) \;=\; X_i(0) \;+\; \int_0^t \gamma_i(s)\; dW_s \;+\; \int_0^t \theta_i(s)\; ds \;+\; J_i(t) \tag{33}$$

where $J_i(t)$ is right-continuous. We then define the cross-variation process

$$[X_1,\ X_2](T) := \lim_{||\Pi||\to 0} C_\Pi(X_1, X_2)$$

where

$$C_\Pi(X_1, X_2) \;:=\; \sum_{i=1}^{n} \left(X_1(t_i) - X_1(t_{i-1})\right)\left(X_2(t_i) - X_2(t_{i-1})\right).$$

The following result[18] is straightforward to prove given the corresponding results for diffusion processes.

Theorem 4 *Let X_1 and X_2 be two jump processes as in (33). Then*

$$[X_1,\ X_1](T) \;=\; \int_0^T \gamma_1^2(s)\; ds \;+\; \sum_{0<s\le T} (\Delta J_1(s))^2$$

and

$$\begin{aligned} [X_1,\ X_2](T) \;&=\; \int_0^T \gamma_1(s)\gamma_2(s)\, ds \;+\; \sum_{0<s\le T} \Delta J_1(s)\, \Delta J_2(s) \\ &=\; [X_1^c,\ X_2^c](T) \;+\; [J_1,\ J_2](T). \end{aligned} \tag{34}$$

Note that in differential notation, the identity (34) states that

$$dX_1(t)\, dX_2(t) \;=\; dX_1^c(t)\, dX_2^c(t) \;+\; dJ_1(t)\, dJ_2(t)$$

so that in particular

$$dX_1^c(t)\, dJ_2(t) \;=\; dX_2^c(t)\, dJ_1(t) \;=\; 0$$

where $X_i^c(t)$ is the continuous part of $X_i(t)$. More generally, we can see from (34) that we need both processes to have a diffusion component or both processes to have simultaneous jumps in order for the cross-variation process to be non-zero.

3.3 Examples of Popular Jump-Diffusion Models

Merton (1975) was the first to propose a jump-diffusion model for pricing options. While the assumptions of Merton's model are not very realistic, it is easy to analyze and worth investigating in some detail.

Merton's Jump-Diffusion Model

Merton's jump diffusion model assumes that the time t stock price, S_t, satisfies

$$S_t \;=\; S_0\; e^{(\mu-\sigma^2/2)t+\sigma W_t} \prod_{i=1}^{N_t} Y_i \tag{35}$$

where N_t is a Poisson process with mean arrival rate λ, and the Y_i's are IID log-normal random variables with $\mu_y := \mathrm{E}[Y_i]$ for all i. The Poisson process and Brownian motions are independent processes and between jumps

[18] This is Theorem 11.4.7 in Shreve.

the stock price behaves like a regular GBM. If the dynamics in (35) are under an EMM, $\mathcal{Q}$, then the model parameters are constrained in such a way that the $\mathcal{Q}$-expected rate of return must equal $r-q$ where q is the stock's dividend yield. To be specific, note that

$$\begin{aligned}
\mathrm{E}_0^{\mathcal{Q}}[S_t] &= S_0 e^{\mu t}\, \mathrm{E}_0^{\mathcal{Q}}\left[\prod_{i=1}^{N_t} Y_i\right] \\
&= S_0 e^{\mu t}\, \mathrm{E}_0^{\mathcal{Q}}\left[\mathrm{E}_0^{\mathcal{Q}}\left[\prod_{i=1}^{N_t} Y_i \mid N_t\right]\right] \\
&= S_0 e^{\mu t}\, \mathrm{E}_0^{\mathcal{Q}}\left[\mu_y^{N_t}\right] \\
&= S_0 e^{\mu t} \sum_{i=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^i}{i!}\mu_y^i \\
&= S_0 e^{\mu t + \lambda t(\mu_y - 1)}. \qquad (36)
\end{aligned}$$

If $\mathcal{Q}$ is an EMM (with the cash account as numeraire) then the expected growth rate under $\mathcal{Q}$ must be $r-q$ and so (36) implies that we must have

$$\mu + \lambda(\mu_y - 1) = r - q. \qquad (37)$$

This is an equation in three unknowns and so it has infinitely many solutions. We can therefore conclude from the Second Fundamental Theorem of Asset Pricing that Merton's model is incomplete. Indeed this is true of almost all jump-diffusion models.

We would like to be able to price European options in Merton's model and there are several ways to do this including Monte-Carlo simulation and Laplace or Fourier transform methods. We can also price these options, however, by expressing them as an infinitely weighted sum of Black-Scholes options prices. To see this, note that *conditional* on $N_T = n$ we can write

$$\begin{aligned}
S_T &= S_0 \exp\left((\mu - \sigma^2/2)T + \sigma W_T + \sum_{i=1}^{n} Z_i\right) \\
&=^{\text{dist}} S_0 \exp\left((\mu - \sigma^2/2)T + n\mu_z + \sqrt{\sigma^2 + n\sigma_z^2/T}\, W_T\right) \\
&= S_0 \exp\left(\left(\mu + \frac{2n\mu_z + n\sigma_z^2}{2T} - \widehat{\sigma}_n^2/2\right) T + \widehat{\sigma}_n W_T\right) \qquad (38)
\end{aligned}$$

where $Z_i := \log(Y_i) \sim N(\mu_z, \sigma_z^2)$ are IID, "$=^{\text{dist}}$" denotes "equal in distribution" and $\widehat{\sigma}_n := \sqrt{\sigma^2 + n\sigma_z^2/T}$. Conditional on $N_T = n$, we therefore see that the risk-neutral drift of S_t in (38) is given by

$$\begin{aligned}
\mu + \frac{2n\mu_z + n\sigma_z^2}{2T} &= r - q - \lambda(\mu_y - 1) + \frac{2n\mu_z + n\sigma_z^2}{2T} \qquad (39) \\
&= r - q - \lambda(e^{\mu_z + \sigma_z^2/2} - 1) + n(\mu_z + \sigma_z^2/2)/T \qquad (40) \\
&= r - \widehat{q}_n \qquad (41)
\end{aligned}$$

where $\widehat{q}_n := q + \lambda(e^{\mu_z + \sigma_z^2/2} - 1) - n(\mu_z + \sigma_z^2/2)/T$. Note that (39) follows from (37) and (40) follows because $\mu_y = \exp(\mu_z + \sigma_z^2/2)$. We are now in a position to derive an expression for European call options in Merton's jump-diffusion model. We obtain

$$\begin{aligned}
\mathrm{E}_0^{\mathcal{Q}}[e^{-rT}(S_T - K)^+] &= \sum_{n=0}^{\infty} e^{-\lambda T}\frac{(\lambda T)^n}{n!}\mathrm{E}_0^{\mathcal{Q}}[e^{-rT}(S_T - K)^+ | N_T = n] \\
&= \sum_{n=0}^{\infty} e^{-\lambda T}\frac{(\lambda T)^n}{n!} C_{bs}(S_0, K, r, \widehat{q}_n, \widehat{\sigma}_n, T) \qquad (42)
\end{aligned}$$

where C_{bs} is the usual Black-Scholes call option price and (42) follows from (41) and the log-normality of S_T conditional on $N_T = n$.

An interesting question to consider is how well Merton's jump-diffusion model can replicate the implied volatility surfaces that are typically observed in the market. Note that in contrast to the geometric Brownian motion

(GBM), we have five parameters[19], σ, λ, μ, μ_y and σ_y^2 and just one equation to satisfy, namely (37). We therefore have much more flexibility than GBM which can only achieve constant implied volatility surfaces.

Figure 2 displays the implied volatility surface under a Merton jump-diffusion model when $\sigma = 20\%$, $r = 2\%$, $q = 1\%$, $\lambda = 10\%$, $\mu_z = -.05$ and $\sigma_z = \sqrt{0.1}$. While other shapes are also possible by varying these

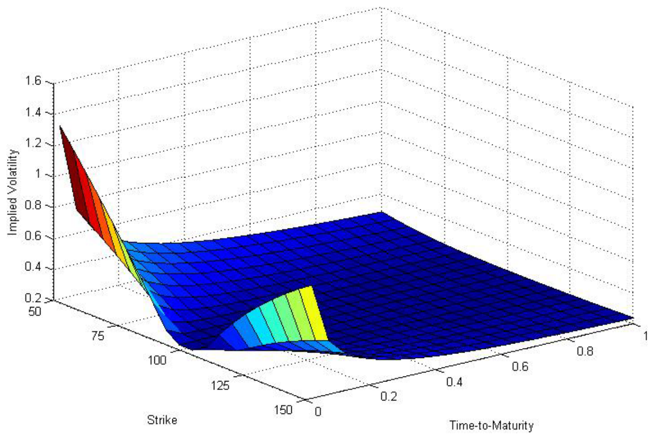

Figure 2: An Implied Volatility Surface under Merton's Jump Diffusion Model

parameters, Figure 2 demonstrates one of the principal weaknesses of Merton's jump-diffusion model, namely the rapid flattening of the volatility surface as time-to-maturity increases. For very short time-to-maturities, however, the model has no difficulty with producing a steep volatility skew. This is in contrast to stochastic volatility models which do not allow jumps.

Kou's Model

Kou (2002) developed the double-exponential jump-diffusion model where the jump-sizes have a double exponential distribution. In particular, the stock price process, S_t, has Q-dynamics that satisfy

$$\frac{dS_t}{S_{t-}} \;=\; (r - \lambda\xi)\, dt \;+\; \sigma\, dW_t \;+\; d\left(\sum_{i=1}^{N_t}(V_i - 1)\right)$$

where W_t and N_t are a Q-Brownian motion and Poisson process with intensity, λ, respectively. A simple application of Itô's Lemma (see Section 3.4) implies that the log-stock price process, $X_t := \log(S_t/S_0)$, then satisfies

$$X_t \;=\; \left(r - \sigma^2/2 - \lambda\xi\right) t \;+\; \sigma W_t \;+\; \sum_{i=1}^{N_t} Y_i\,, \qquad X_0 = 0$$

where $Y_i = \log(V_i)$. The Y_i's are an IID sequence of double-exponential random variables with density

$$f_Y(y) \;=\; p\eta_1 e^{-\eta_1 y} 1_{\{y \geq 0\}} \;+\; q\eta_2 e^{\eta_2 y} 1_{\{y<0\}}$$

where $p, q \geq 0$ with $p + q = 1$, $\eta_1 > 1$ and $\eta_2 > 0$. In order to ensure that Q is indeed an EMM, it must be the case that

$$\xi \;:=\; \mathsf{E}[V] - 1 \;=\; \frac{p\eta_1}{\eta_1 - 1} \;+\; \frac{q\eta_2}{\eta_2 + 1} \;-\; 1. \tag{43}$$

The double-exponential jump-diffusion model is quite tractable due to the memoryless property of exponential random variables. This enables us, for example, to compute analytical expressions for expectations involving first passage times.

[19] We could use σ_z and σ_z^2 in place of μ_y and σ_y^2.

The Bates Model

The Bates model assumes the stock price, S_t, has risk-neutral dynamics given by

$$\begin{aligned} dS_t/S_t &= (r - q - \lambda(\bar{\alpha} - 1))\,dt \;+\; \sqrt{\sigma_t}\,dW_t^{(s)} \;+\; (\alpha_t - 1)\,dN_t && (44)\\ d\sigma_t &= \kappa\,(\theta - \sigma_t)\,dt \;+\; \gamma\sqrt{\sigma_t}\,dW_t^{(vol)} && (45)\end{aligned}$$

where N_t is a Poisson process with intensity λ and α_t is log-normally distributed with $\mathsf{E}^Q_{t-}[\alpha_t] = \bar{\alpha}$. It is clear that the Bates model is simply the Heston model with independent jumps added to the security price dynamics. This independence implies that we can also easily compute the characteristic function of the $\log$-stock price in this model given that we already know how to do so for the Heston model.

3.4 Itô's Lemma for Jump-Diffusions

The stochastic calculus we have studied for diffusion processes can be extended to jump-diffusions and indeed other stochastic processes, including Levy processes and more generally, semimartingales. In this section we will state Itô's Lemma for jump-diffusions.

Theorem 5 (Itô's Lemma for 1-Dimensional Jump-Diffusions)
Let X_t be a jump process and $f(x)$ a function for which $f'(x)$ and $f''(x)$ are defined and continuous. Then

$$f(X_t) \;=\; f(X_0) \;+\; \int_0^t f'(X_s)\;dX_s^c \;+\; \frac{1}{2}\int_0^t f''(X_s)\;dX_s^c dX_s^c \;+\; \sum_{0<s\leq t}\left[f(X_s) - f(X_{s-})\right] \qquad (46)$$

where X_t^c denotes the continuous, i.e. non-jump, component of X_t and the summation in (46) is over the jump times of the process.

The statement[20] of Theorem 5 should not be surprising given our knowledge of Itô's Lemma for diffusions. In particular, the jump-diffusion behaves as a diffusion in between jumps so it is only necessary to add the finite number of changes in $f(\cdot)$ due to jumps to the usual terms that come from the diffusive component of the process.

Example 5 (Merton's Jump Diffusion Model)
Let $X_t := \log(S_t)$ where S_t is given by (35). Then $X_t = \log(S_0) + (\mu - \sigma^2/2)t + \sigma W_t + \sum_{i=1}^{N_t}\log(Y_i)$ and so

$$dX_t \;=\; (\mu - \sigma^2/2)\;dt \;+\; \sigma\;dW_t \;+\; \log(Y_i)\;dN_t.$$

Applying Itô's Lemma to $S_t = e^{X_t}$ we recover the dynamics for S_t and obtain[21]

$$\begin{aligned} dS_t &= S_t(\mu - \sigma^2/2)\,dt \;+\; \sigma S_t\,dW_t \;+\; \frac{1}{2}S_t\sigma^2\,dt \;+\; S_{t-}(Y_t - 1)\,dN_t\\ &= S_t(\mu + (\mu_y - 1)\lambda)\,dt \;+\; \sigma S_t\,dW_t \;+\; \left[S_{t-}(Y_t - 1)\,dN_t \;-\; S_t(\mu_y - 1)\lambda\,dt\right]. && (47)\end{aligned}$$

Referring to (47) note that the dW_t term and the term in the square brackets are both martingales. Therefore if (47) describes the risk-neutral dynamics of S_t is must be the case that the drift term equals $S_t(r - q)\;dt$ and so we obtain (37) once again. ∎

[20] This is Theorem 11.5.1 in Shreve where a proof of the result may also be found.
[21] We could of course have written the dynamics for S_t directly using (35) but then we wouldn't have been able to practice using Itô's Lemma.

Example 6 (Shreve E.G. 11.5.2: Geometric Poisson Process)
Suppose S_t satisfies

$$S_t \ := \ S_0 \exp\left(N_t \log(\sigma+1) - \lambda\sigma t\right) \ = \ S_0 e^{-\lambda\sigma t} \left(\sigma+1\right)^{N_t}$$

where $\sigma > -1$ and N_t is again a Poisson process. Note that if $\sigma > 0$ then this process only jumps up and drifts down between jumps. If $-1 < \sigma < 0$ then the process only jumps down and drifts up between jumps. Let's apply Itô's Lemma to show that S_t is a martingale.

Define $X_t = N_t \log(\sigma+1) - \lambda\sigma t$ so that $S_t = S_0 f(X_t)$ where $f(x) = \exp(x)$. Itô's Lemma now implies

$$\begin{aligned}
S_t &= S_0 f(X_t) \\
&= S_0 f(X_0) \ - \ \lambda\sigma S_0 \int_0^t f'(X_u)\, du \ + \ S_0 \sum_{0<u\leq t} \left[f(X_u) - f(X_{u-})\right] \\
&= S_0 \ - \ \lambda\sigma \int_0^t S_u\, du \ + \ \sum_{0<u\leq t} \left[S_u - S_{u-}\right] \\
&= S_0 \ - \ \lambda\sigma \int_0^t S_u\, du \ + \ \sum_{0<u\leq t} \sigma S_{u-} && (48) \\
&= S_0 \ - \ \lambda\sigma \int_0^t S_{u-}\, du \ + \ \int_0^t \sigma S_{u-}\, dN_u && (49) \\
&= S_0 \ + \ \sigma \int_0^t S_{u-}\, (dN_u - \lambda\, du).
\end{aligned}$$

where we have used the fact that $S_u = (\sigma+1)S_{u-}$ if a jump takes place at time u and where S_{u-} is the value of S immediately before that jump. Note that because we were able to write the jump in S at time u in terms of S_{u-}, we can write the SDE in differential form

$$dS_t \ = \ -\lambda\sigma S_t\ dt + \sigma S_{t-}\ dN_t = \ \sigma S_{t-}\ dX_t$$

where $X_t := N_t - \lambda t$ is clearly a martingale. Note also that Theorem 3 implies (why?) that S_t is a martingale. ▮

Remark 3 *Note that in going from (48) to (49) we replaced $\int_0^t S_u\, du$ with $\int_0^t S_{u-}\, du$ and this presents no problem as the two expressions are equal whenever there are only countably many jumps. This is indeed the case for all jump-diffusion processes.*

Example 7 (Independence of Brownian Motion and a Poisson Process)
Let W_t and N_t be a Brownian motion and Poisson process with intensity λ, respectively, where both processes are defined on the same probability space $(\Omega, \mathcal{F}, P)$ and relative to the same filtration, $\mathcal{F}_t$, $t \geq 0$. Using Itô's Lemma we will show that W_t and N_t are independent.

Towards this end, let $Y_t := \exp(X_t)$ where

$$X_t \ := \ u_1 W_t + u_2 N_t - \frac{1}{2}u_1^2 t - \lambda(e^{u_2} - 1)t$$

where u_1 and u_2 are fixed real numbers. By applying Itô's Lemma to the function $f(x) = e^x$ we obtain after some calculations

$$Y_t \ = \ 1 \ + \ u_1 \int_0^t Y_s\, dW_s \ + \ (e^{u_2} - 1) \int_0^t Y_{s-}\, dM_s \tag{50}$$

where $M_s := N_s - \lambda s$. From (50) we can easily conclude (why?) that Y_t is a martingale from which it follows that $\mathsf{E}[Y_t] = 1$ for all t. But this implies that the joint MGF of W_t and N_t can be written as

$$\mathsf{E}_0\left[e^{u_1 W_t + u_2 N_t}\right] \ = \ e^{\frac{1}{2}u_1^2 t}\ e^{\lambda(e^{u_2}-1)t}. \tag{51}$$

We can therefore conclude (why?) that W_t and N_t are independent. Given this key step, how would you prove the full result? ∎

Example 7 is the main step in proving a more general result. This result[22] states that a Brownian motion and a Poisson process, both defined on a common probability space and relative to the same filtration, must be independent. We now state Itô's Lemma for multi-dimensional jump-diffusions.

Theorem 6 (Itô's Lemma for Multi-Dimensional Jump-Diffusions)
Let $X_t := (X_1(t), \ldots, X_n(t))$ where $X_i(t)$ is a jump process for $i = 1, \ldots, n$ and let $f(\cdot)$ be a C^2 function of n variables. Then

$$f(X_t) \;=\; f(X_0) + \sum_{i=1}^{n} \int_0^t f_i(X_s)\, dX_i^c(s) + \frac{1}{2} \sum_{i,j=1}^{n} \int_0^t f_{ij}(X_s)\, dX_i^c(s) dX_j^c(s) + \sum_{0<s\leq t} [f(X_s) - f(X_{s-})] \tag{52}$$

where f_i and f_{ij} denote the first and second partial derivatives, respectively, with respect to the appropriate arguments of f.

The following special case of Theorem 6 arises frequently enough as to warrant special attention.

Theorem 7 (Itô's Product Rule for Jump-Diffusions) *Let $X_1(t)$ and $X_2(t)$ be jump processes. Then*

$$\begin{aligned} X_1(t)X_2(t) \;&=\; X_1(0)X_2(0) \;+\; \int_0^t X_2(s)\, dX_1^c(s) \;+\; \int_0^t X_1(s)\, dX_2^c(s) \;+\; [X_1^c,\, X_2^c](t) \\ &\quad + \sum_{0<s\leq t} [X_1(s)X_2(s) \;-\; X_1(s-)X_2(s-)] \qquad (53) \\ &=\; X_1(0)X_2(0) \;+\; \int_0^t X_2(s-)\, dX_1(s) \;+\; \int_0^t X_1(s-)\, dX_2(s) \;+\; [X_1,\, X_2](t) \qquad (54) \end{aligned}$$

While (53) follows almost immediately from an application of Theorem 6, proving (54) requires some additional work. In particular, we can use (34) and the fact that jumps in $X_i(t)$ are the same as jumps in $J_i(t) := X_i(t) - X_i^c(t)$ to go directly from (54) to (53).

Definition 4 (Doleans-Dade Exponential) *Let X_t be a jump-diffusion. Then the Doleans-Dade exponential of X is defined to be*

$$Z_t^X \;:=\; \exp\left(X_t^c - \frac{1}{2}[X^c,\, X^c](t)\right) \prod_{0<s\leq t} (1 + \Delta X_s). \tag{55}$$

Theorem 8 *The Doleans-Dade exponential, Z_t^X, is the solution to the stochastic differential equation*

$$Z_t^X \;=\; 1 \;+\; \int_0^t Z_{s-}^X\; dX_s. \tag{56}$$

Theorem 8 can be proven by applying Itô's product rule to $Z_t^X = Y_t V_t$ where $Y_t := \exp\left(X_t^c - \frac{1}{2}[X^c,\, X^c](t)\right)$ and $V_t := \prod_{0<s\leq t}(1 + \Delta X_s)$. The Doleans-Dade exponential plays an important role in stochastic calculus, particularly in the context of Girsanov's Theorem. This follows because we can see from (55) and (56) that Z_t^X will be[23] a positive martingale with $\mathsf{E}[Z_t^X] = 1$ if X_t is a martingale. In that case it can be used to define a change of probability measure.

[22]The steps of Example 7 are presented in Corollary 11.5.3 of Shreve which states the result concerning the independence of a Brownian motion and Poisson process. See also Exercises 11.4 to 11.6 in Shreve for related results on (i) the independence of multiple Poisson processes that do not have simultaneous jumps and (ii) the independence of a compound Poisson process and a Brownian motion.

[23]Subject to technical restrictions on X_t. Otherwise Z_t^X will be a *local martingale* which we will not define here.

4 Fourier Transform Methods

Fourier and Laplace transform methods have proved particularly useful for pricing derivative securities when closed-form solutions are not available. As mentioned earlier, they were first used in a financial engineering context by Heston (1993). Transform methods can be used to solve the pricing PDE or to compute the risk-neutral expected value directly. In section 4.1 below we briefly describe the Carr-Madan approach[24] which uses Fourier transforms to compute option prices directly.

A popular alternative approach is the Lewis method[25] which computes the *generalized* Fourier transform of the derivative price with respect to the log-stock price. The Lewis approach is more general than the Carr-Madan approach below as it does not require the concept of a "strike" but it is also somewhat trickier to implement.

4.1 The Carr-Madan Approach

The Carr-Madan approach requires the characteristic function or the Fourier transform of the log-stock price. This transform can be computed for a wide array of models including, for example, the Heston model. In fact the characteristic function of the log-stock price in the Heston model is given in (23). The Carr-Madan approach can be used to price European call options as we will now show.

First note that we can write the call option price as

$$C_T(k) = \int_k^\infty e^{-rT}(e^s - e^k)q_T(s)\,ds \tag{57}$$

where q_T is the risk-neutral density of the log-stock price, $s := \log(S_T)$, T is the time-to-maturity and $k = \log(K)$ is the log-strike. Because $C_T(k) \to S_0$ as $k \to -\infty$, $C_T(k)$ will not be square-integrable. We overcome this problem by defining

$$c_T(k) := \exp(\alpha k)C_T(k) \tag{58}$$

for some $\alpha > 0$. (Values of $\alpha = .75$ have been recommended in the literature but depending on the application at hand, a different value may be required. Note that (62) is valid for any positive α so any difficulties that might arise with α are due to the difficulties that arise with the numerical inversion of the right-hand-side of (62).) Consider now the Fourier transform of $c_T(k)$ which is defined as

$$\psi_T(v) := \int_{-\infty}^\infty e^{ivk}c_T(k)\,dk. \tag{59}$$

In Exercise 17 you are asked to show that

$$\begin{aligned}\psi_T(v) &= \frac{e^{-rT}\,\mathrm{E}_0\left[e^{i(v-(\alpha+1)i)\log(S_T)}\right]}{\alpha^2+\alpha-v^2+i(2\alpha+1)v} \\ &= \frac{e^{-rT}\,\phi_T(v-(\alpha+1)i)}{\alpha^2+\alpha-v^2+i(2\alpha+1)v}\end{aligned} \tag{60}$$

where $i = 0 + i$ is imaginary and $\phi_T(\cdot)$ is the characteristic function of the log-stock price, $\log(S_T)$. The Fourier inversion formula then implies that the option price is given by

$$C_T(k) = \frac{e^{-\alpha k}}{2\pi}\int_{-\infty}^\infty e^{-ivk}\psi_T(v)\,dv \tag{61}$$

$$= \frac{e^{-\alpha k}}{\pi}\int_0^\infty \mathrm{Re}\left(e^{-ivk}\psi_T(v)\right)\,dv \tag{62}$$

[24] This approach was developed in "Option Valuation Using the Fast Fourier Transform" by Carr and Madan in the *Journal of Computational Finance* (1998). Other popular transform methods include the Lewis method

[25] "A Simple Option Formula for General Jump-Diffusion and Other Exponential Levy Processes" (2001), available from http:www.optioncity.net.

where Exercise 17 again asks you to justify going from (61) to (62). The option price in (62) can be found using standard Fourier inversion techniques. If many options with different strikes but the same time-to-maturity need to be priced, then the Fast Fourier transform can be used. Indeed this was the approach proposed originally by Carr and Madan and for this reason we review the FFT in Appendix A. Alternatively, if just a single option price is required a standard[26] numerical integration of the right-hand-side of (62) can be performed. In fact, even if we are pricing multiple options with the same time-to-maturity it is not clear[27] that using the FFT is the most efficient method for doing so.

Finally, we mention here the implied volatility surface in Figure 1 was constructed by using (62) to compute call option prices in the Heston model.

Remark 4 *As pointed out by Carr and Madan, for very short maturities the option price approaches its intrinsic value which is non-analytic. This causes the integrand in (62) to be very oscillatory and therefore very difficult to integrate numerically. They circumvent this problem by developing an expression for $z_T(k)$ where $z_T(k)$ is the* time *value of the option. This expression is again in terms of the characteristic function of the* log*-stock terminal price and is obtained in a similar manner to our derivation of (62).*

Appendix A: The Fast Fourier Transform

Let X be a random variable with PDF, $f(\cdot)$, and let

$$\widehat{f}_x(s) \;:=\; \mathrm{E}\left[e^{-isx}\right] \;=\; \int_{-\infty}^{\infty} e^{-isx} f(x)\,dx$$

be the Fourier transform of X, i.e. the Fourier transform of $f(\cdot)$. Then the standard Fourier inversion formula states that

$$f(x) \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{f}_x(s) e^{isx} ds. \tag{63}$$

Now suppose we know $\widehat{f}_x(\cdot)$ and we wish to compute $f(\cdot)$ numerically using the inverse FFT. Then using (63) we can approximate $f(\cdot)$ as

$$f(x) \;\approx\; \frac{1}{2\pi}\sum_{j=1}^{N} \widehat{f}_x(s_j) e^{is_j x}\,\eta \tag{64}$$

where $\eta =$ "ds" and

$$s_j := -T + \eta(j-1) \;\; \text{for } j = 1,\ldots,N \tag{65}$$

so that the lower and upper limits of integrations are $-T$ and $-T+\eta(N-1)$, respectively. Let $\eta\lambda = 2\pi/N$. The FFT returns then $f(\cdot)$ at the values

$$x_n \;:=\; -b + \lambda(n-1) \quad \text{for } n = 1,\ldots,N. \tag{66}$$

In particular substituting (66) and then (65) into (64) we obtain

$$\begin{aligned} f(x_n) \;&\approx\; \frac{1}{2\pi}\sum_{j=1}^{N} \widehat{f}_x(s_j) e^{is_j(-b+\lambda(n-1))}\,\eta \\ &=\; \frac{\eta}{2\pi}\sum_{j=1}^{N} \widehat{f}_x(s_j) e^{-ibs_j} e^{i\eta\lambda(n-1)(j-1) - i\lambda T(n-1)} \\ &=\; \frac{\eta e^{-i\lambda T(n-1)}}{2\pi}\sum_{j=1}^{N} \widehat{f}_x(s_j) e^{-ibs_j} e^{\frac{2\pi i}{N}(n-1)(j-1)} \end{aligned} \tag{67}$$

---

[26]Note that some software packages / languages such as *Matlab* can handle complex calculations whereas others such as *VBA* require additional functions.

[27]There is an ongoing debate as to what are the most efficient methods for evaluating (62).

since $\eta\lambda = 2\pi/N$. Rearranging (67) we obtain

$$\frac{2\pi}{\eta N} f(x_n) e^{i\lambda T(n-1)} = \frac{1}{N} \sum_{j=1}^{N} \widehat{f}_x(s_j) e^{-ibs_j} e^{\frac{2\pi i}{N}(n-1)(j-1)}$$

or equivalently,

$$v_n = \frac{1}{N} \sum_{j=1}^{N} u_j \, e^{\frac{2\pi i}{N}(n-1)(j-1)} \tag{68}$$

where

$$\begin{aligned} v_n &:= \frac{2\pi}{\eta N} f(x_n) e^{i\lambda T(n-1)} \quad \text{and} \\ u_j &:= \widehat{f}_x(s_j) e^{-ibs_j}. \end{aligned} \tag{69}$$

(68) is now in the form required by most software packages. That is we first compute the u_j's and then pass them through the inverse FFT to obtain the v_n's. Finally, we compute the $f(x_n)$'s using (69). Note that for a fixed N, there is a tradeoff between the accuracy of the numerical integration and the fineness of the grid where we compute $f(x)$. This is because we must have $\eta\lambda = 2\pi/N$.

Exercises

1. This question refers to the exercise from the *Black-Scholes and the Volatility Surface* where an implied volatility surface was fitted to European call and put prices and then (in a later exercise) used to compute the price of a digital option. We will fit an implied volatility surface to the option data and then use this surface to compute a local volatility surface using Dupire's formula. (Note that there are better ways to do the various steps below but if done properly they should produce a reasonable local volatility surface. Feel free to try your own way if you prefer.)

   (a) Use a spline function to fit the implied volatility skew for *each* option maturity. (As there are four maturities you should have four fitted splines. Note also that in practice you should know (and be satisfied with) how the splines extrapolate beyond the observed implied volatilities.)

   (b) Now write a function that takes each of the fitted splines from part (a) and uses them to compute an implied volatility for any strike-maturity pair, (K, T). A good way to do this is to interpolate using *total variance*. That is, suppose we want $\sigma_{bs}(K, T)$ where $K = 75$ and $T = .8$. We then use the fitted splines at $T = .5$ and $T = 1$ to compute $\sigma_{bs}(75, .5)$ and $\sigma_{bs}(75, 1)$. We could then estimate $\sigma_{bs}(75, .8)$ as

$$\sigma_{bs}(75, .8) = \sqrt{\frac{\frac{2}{5} \times .5\, \sigma_{bs}^2(75, .5) + \frac{3}{5} \times \sigma_{bs}^2(75, 1)}{.8}}. \tag{70}$$

   Note that interpolating total variance as we do in (70) is much better than interpolating total volatility. Note also that you need to make some assumption regarding how to extrapolate to maturities less than $.25$ years and greater than 1.5 years. For the purpose of this question you can assume that the term structure of implied volatility is constant up until $t = .25$ and constant beyond $t = 1.5$.

   (c) Now plot your fitted implied volatility surface using a grid of strike-maturity pairs that is larger than the original grid.

   (d) Use the Dupire formula to write a function that computes the local volatility at a given strike-maturity pair, (K, T). You should use your function from part (b) to estimate the various partial derivatives numerically. In anticipation of simulating many paths of the local volatility process simultaneously, it would actually be a good idea if your function was *vectorized* along the strike dimension so that one call to the function could return $\sigma_l(K_1, T), \ldots, \sigma_l(K_n, T)$ where n is an arbitrary number of strikes.

(e) Now plot your local volatility surface. What do you notice?

(f) Simulate your local volatility model (using an Euler scheme) to price the original options that you used to construct the volatility surface. Note that you can use the same Monte-Carlo to price all of the options. (If necessary the Brownian bridge construction together with stratified sampling can be used to speed up your Monte-Carlo.) How do your Monte-Carlo prices compare to the original prices? (Hint: they should be almost identical with any differences being due to statistical error (from the Monte-Carlo) and numerical error (from the Euler scheme and estimation of the derivatives).)

2. Consider the CEV model with Q-dynamics given by (11). Find the $\tilde{Q}$-dynamics of the deflated cash account when we deflate by the risky asset and $\tilde{Q}$ is the corresponding EMM. (You may assume that $\beta > 1/2$ so that there is no possibility of the risky asset price reaching zero.)

3. Justify the step where we went from (24) to (25) in Example 1. (*Hint*: Take expectations in (13) and then use the martingale property of stochastic integrals to eliminate the last term. You can then obtain a simple ODE for $\mathrm{E}_0^Q[\sigma_t]$.)

4. Confirm that (18) is indeed the pricing PDE that corresponds to the Heston model of (12) and (13).

5. **(Volatility and Variance Swaps in Heston's Model)**
A standard result states that the square-root function can be expressed as

$$\sqrt{x} \;=\; \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1-e^{-sx}}{s^{3/2}}\, ds. \tag{71}$$

(a) Use (71) to obtain an expression for the fair value of a volatility swap under the Heston model. You may assume that the payoff of a volatility swap is given by $\sqrt{V_c(0,T)} - K^*_{vol}$ where K^*_{vol} is the strike (typically chosen at inception so that the initial value of the swap is 0) and

$$V_c(0,T) = \frac{1}{T}\int_0^T \sigma_t\, dt$$

is the continuous realized variance with σ_t having the dynamics of (13).

Hint: The Laplace transform of $V_c(0,T)$ can be computed analytically. In particular,

$$\mathsf{E}_0\left[e^{-sV_c(0,T)} \,|\, \sigma_0\right] \;=\; e^{A(T,s)-B(T,s)\sigma_0} \tag{72}$$

where $A(T,s)$ and $B(T,s)$ are known deterministic functions of T. (See Cairns (2000) for details.)

(b) Use (72) to compute the fair value of a variance swap in Heston's model. Your answer should of course agree with the expression we found in (26).

6. Consider a two-dimensional Markov process for the foreign price of a stock, S_t, and the exchange rate, X_t, where both processes follow local volatility models:

$$dS_t \;=\; (r_f - q)S_t\, dt \;+\; \sigma_s(S_t,t)S_t\, dW_s(t) \tag{73}$$

$$dX_t \;=\; (r - r_f)X_t\, dt \;+\; \sigma_s(X_t,t)X_t\, dW_x(t) \tag{74}$$

where $dW_s dW_x = \rho(S_t, X_t, t)\, dt$, r_f is the foreign risk-free rate, r is the domestic risk-free rate and q is the dividend yield. In addition, W_s and W_x are $\tilde{Q}$- and Q-Brownian motions, respectively, where $\tilde{Q}$ is the EMM for a foreign investor with the foreign cash account as numeraire, and Q is the EMM of a domestic investor with the domestic cash account as numeraire. Find an expression for the Q-dynamics in a local volatility model of S_t.

7. Show that the MGF of a compound Poisson process, X_t, is given by $\phi_{X_t}(u) = e^{\lambda t(\phi_Y(u)-1)}$ as in equation (27) of Example 2.

8. Write a program to construct the implied volatility surface of Figure 2. To do this you need to write a program to compute option prices in Merton's model. This can be done, for example, by either evaluating the expression in (42) or by using Monte-Carlo or by computing the characteristic function of the log-stock price and then using Fourier transform techniques.

9. Referring to Example 4, show that the process $M_t := I_{[S_1,\infty)}(t) - \lambda(t \wedge S_1)$ in (31) is indeed a martingale.

10. Use Itô's Lemma to confirm that (43) must hold if the dynamics specified for S_t in Kou's jump-diffusion model are risk-neutral or Q-dynamics.

11. **(Exercise 11.3 in Shreve)** Let N_t be a Poisson process with intensity $\lambda > 0$, and let S_0 and $\sigma > -1$ be given. Use the stationary and independent increments property of a Poisson process rather than Itô's Lemma to show directly that

$$S_t = \exp\left(N_t \log(\sigma + 1) - \lambda\sigma t\right) \;=\; (\sigma + 1)^{N_t}\, e^{-\lambda\sigma t}$$

is a martingale.

12. **(Exercise 11.4 in Shreve)** Suppose $N_1(t)$ and $N_2(t)$ are independent Poisson processes with intensities λ_1 and λ_2, respectively, both defined on the same probability space $(\Omega, \mathcal{F}, P)$ and relative to the same filtration $\mathcal{F}_t$, $t \geq 0$. Show that almost surely $N_1(t)$ and $N_2(t)$ can have no simultaneous jump. (Hint: Define the compensated Poisson processes $M_1(t) = N_1(t) - \lambda_1 t$ and $M_2(t) = N_2(t) - \lambda_2 t$, which like N_1 and N_2, are independent. Use Itô's product rule for jump processes to compute $M_1(t)M_2(t)$ and take expectations.)

13. Referring to Example 7, use Itô's Lemma to derive (50).

14. Prove Itô's Product Rule for jump-diffusions. That is, prove Theorem 7.

15. Use Itô's product rule to prove Theorem 8.

16. **(A Standard Fourier Transform Approach to Option Pricing)**
Let $C_0 = \mathsf{E}_0^Q[e^{-rT}\,(S_T - K)^+$ be the time $t = 0$ price of a call option with strike K, maturity T, risk-free-rate r and underlying security price process S_t. Note that Q is the EMM corresponding to the cash account as numeraire.

    (a) Show that the option price may be expressed as

$$C_0 \;=\; S_0 Q_1\left(X_T > \ln K\right) \;-\; e^{-rT} K Q_2\left(X_T > \ln K\right) \tag{75}$$

    where $X_T := \ln(S_T)$ and Q_1 and Q_2 are EMM's corresponding to specific changes of numeraire. Be sure to identify these numeraires. Does the expression in (75) agree with the interpretation of P_1 and P_2 that we gave at the end of the paragraph immediately following (20)?

    (b) Let f_1 and f_2 be the characteristic functions under Q_1 and Q_2, respectively, of X_T so that

$$f_j(u) \;:=\; \mathsf{E}_0^{Q_j}\left[e^{iuX_T}\right], \quad j = 1, 2.$$

    Give an expression for C_0 in terms of f_1 and f_2. Be sure to simplify it as much as possible.

17. In this exercise we justify the Carr-Madan call option-pricing formula of (62).

    (a) By substituting for $c_T(k)$ in (59) using (57) and (58), show that $\psi_T(v)$ satisfies (60).

    (b) Justify going from (61) to (62).

18. Using the Carr-Madan approach, write a program to recreate the implied volatility surface in Figure 1 corresponding to the Heston model with parameters as in Section 2.4. (Note that characteristic function of the $\log$-stock price is given in (23).)