Modeling the Stock Price Process as a Continuous Time Discrete Jump Process

Rituparna Sen

Monday, Nov 8
1. **Background & Existing Literature**

2. **Overview**

- General discrete jump model
- Introducing rare big jumps
- Bounds on option price
- Real Data Applications
- Birth and Death Process
- Hedging
- Stochastic Intensity
- Updating parameters by Filtering
- Data Applications for Birth-Death model
- Conclusions
3. The General Discrete Jump Model

Stock Price $S_t$. Tick-size $c = \frac{1}{16}$, $N_t = \frac{S_t}{c}$
We suppose that there is a risk free interest rate $\rho_t$.
$N_t$ is modeled as a non-homogeneous (rate of events per individual depends on $t$), linear (rate is proportional to number of individuals present) jump process.

Jumps($Y_t$) of size $i$ with probability $p_i$. Rate of events per individual is $\lambda_t$.

Wait for an exponential time with parameter $N_t \lambda_t$ and then jump $Y_t$.
How to get the correct rate and jump probabilities?
Equate the drift and volatility of this model to those of the Black-Scholes model:

\[
d < S >_t = EY_tN_t\lambda_t = S_t(\sum ip_i)\lambda_t
\]

\[
d < S, S >_t = EY_t^2N_t\lambda_t = S_t^2(\sum \frac{i^2p_i}{N_t})\lambda_t
\]

$\sum p_i = 1$, $\sum ip_i = \frac{r_t}{\lambda_t}$, $\sum i^2p_i = N_t$ → Gives B-S with $\lambda_t = \sigma_t^2$

$\sum p_i = 1$, $\sum ip_i = \frac{r_t}{\lambda_t}$, $\sum i^2p_i = \frac{1}{c}$ → Gives Cox-Ingersoll-Ross model

Rate is uniquely determined.
Jump distribution is not unique → incomplete market.
4.  Estimation

Have to estimate 2 parameters: intensity rate $\lambda$ and expected jump size $\nu$.
Suppose we observe the process from time 0 to $T$ and the jumps occur at times $\tau_1, \tau_2 \ldots \tau_k$.

By IV.4.1.5 of Andersen et al (1993) $\hat{\lambda}_{NPMLE} = k \left[ \sum_{i=1}^{k} \frac{S_{i-1}}{c} (\tau_i - \tau_{i-1}) + \frac{S_k}{c} (T - \tau_k) \right]^{-1}$

Quasi Likelihood estimators for $\nu = r / \lambda$ are obtained as solutions of estimating equations of the form

$$\sum_{i=1}^{n} w_{\nu}(X_{i-1})(X_i - m_{\nu}(X_{i-1})) = 0$$

In our case, the estimator with minimum variance becomes

$$\hat{\nu} = \frac{\sum_{i=1}^{n} \left( \frac{X_i}{X_{i-1}} - 1 \right)}{\sum_{i=1}^{n} \left( \frac{1}{X_{i-1}} \right)}$$
5. **Introducing rare big jumps**

The jumps come from two competing processes:

- Small jumps with rate $= N_t \lambda_t$,
  \[ E(Y_t) = \frac{r_t}{\lambda_t}, E(Y_t^2) = N_t \]

- Big jumps with rate $= \mu_t$,
  \[ E(Y_t) = aN_t, E(Y_t^2) = bN_t^2 \]

In the limit this is analogous to the **Merton Jump-diffusion model**.

Estimation can be done by **E-M algorithm** regarding the data augmented with knowledge of which process it came from as the **complete data**.
6. Upper and Lower bounds on Option price

Eberlein and Jacod Finance and Stochastics(1997)

Let $f$ be the payoff of the option.

$$\gamma(Q) = E_Q[e^{-rT}f(S_T)]$$

Assume $f$ is convex, $0 \leq f(x) \leq x$ for all $x > 0$

$\mathcal{M}_T$ is the class of measures which converge to B-S.

Under each $Q \in \mathcal{M}_T$, $e^{-rt}S_t$ is a martingale.

Since $f$ in convex, the process $A_t = f(e^{r(T-t)}S_t)$ is a $Q$-submartingale. So

$$\gamma(Q) = e^{-rT}E_Q[A_T] \geq e^{-rT}f(e^{rT}S_0)$$

We have $e^{-rT}f(S_T) < e^{-rT}S_T$. So

$$\gamma(Q) < E_Q[e^{-rT}S_T] = S_0$$
7. An Example

\[
\begin{align*}
S_0 &= 466.77 \\
K &= 450.00 \\
r &= 5\% \\
T &= 19 \text{ days} \\
\text{B-S Price} &= 18.93
\end{align*}
\]

Upper bounds

\[
\begin{align*}
\text{E-J} &= 466.77 \\
\text{Cont time} &= 20.38 \\
\text{Discrete time} &= 21.06
\end{align*}
\]
8. The algorithm to compute bounds

\( M_t \)=Number of jumps in time 0 to \( t \)
For each value \( m \) of \( M_T \)
For each \( i \leq m \)
For each value \( x \) of \( N_{T_{i-1}} \)

\[
f_i(x) = \begin{cases} 
(x - \frac{K}{c}) + & \text{for } i = m \\
\max E(f_{i+1}(x + Y_{i+1}) \mid M_T = m, N_{T_i} = x) & \text{for } 0 \leq i \leq m - 1 
\end{cases}
\]

The maximum value is \( f_0(N_0) \)
The problem reduces to:

- Get the distribution of \( M_T \)
- Maximize \( E(f_{i+1}(x + Y_{i+1}) \mid M_T = m, N_{T_i} = x) \) over the distribution on \( Y_{i+1} \)
9. Distribution of jump size

Let \( p_y = P(Y_t = y \mid N_{T_{i-1}}) \)

We have the constraints:
\[
\sum p_y = 1 \quad \sum y p_y = \frac{r}{\lambda} \quad \sum y^2 p_y = N_{T_{i-1}}
\]

\[
E(f_i(l + Y_t) \mid M_T = m, N_{T_{i-1}} = l) = \frac{\sum_y f_i(l + Y_t) p_y P(M_T = m \mid N_{T_{i-1}} = l, Y_t = y)}{\sum_y p_y P(M_T = m \mid N_{T_{i-1}} = l, Y_t = y)}
\]

\[
P(M_T = m \mid l, y) = \int_q q_{i,l,y}(t) Q_{T-t, l+y}(m - i) dt
\]

where \( Q_{t,l}(j) = P(M_t = k \mid N_0 = l) \) and

\( q_i(t) \) is the conditional density of \( T_i \) given \( N_{T_{i-1}} = l, Y_t = y \).

We have to maximize the ratio of 2 linear functions of \( p_y \) under three linear constraints.

So we have a four point distribution where the maximum is attained.

Can argue further that this is indeed a three-point distribution.
Date: June 1, 2001. Tick = $\frac{1}{100}$. $a =$ Distance from bid-ask midpoint to closest point of prediction interval. $b =$ Length of interval. $n = 92$. Units of tick. Range of option prices $0$-$37$. Avg bid ask spread = 10.
Figure 2: Estimation: ABMD Data

Date: Dec 2, 2000. Tick = \frac{1}{100}. n = 12. Range of option prices $0-$1.60. Avg bid ask spread = 50.
Figure 3: Estimation: Ford Data

Date: June 1, 2001. Tick = \( \frac{1}{16} \). n=34. Range of option prices $0-$11.25. Avg bid ask spread = 2.
Figure 4: Prediction: Ford Data

Date: Dec, 2000. Tick = $\frac{1}{16}$. n = 578. Range of option prices $0-$16. Avg bid ask spread = 2. $a = 4.33$, $b = 17.14$ In learning sample, $a = 3.27$, $b = 19.21$
Date: June, 2001. Tick = \frac{1}{100}. n = 1823. Range of option prices $0$-$80. Avg bid ask spread = 10. a = 55.68, b = 273.17 In Learning Sample a = 21.23 b = 345.76
10. **Birth & Death Model**

*Perrakis(1988), Korn et. al.(1998)*

**Jump size** $\pm c$.

$N_t = c \times S_t$ is a birth and death process.

**Probability** that jump size $Y_t = 1$ is $p_t$.

Such a process can be considered as a discretized version of the Black-Scholes model if the **intensity of jumps** is proportional to $N_t^2$.

Consider processes with intensity $\lambda_t N_t^2$.

- $\lambda_t$ is a **constant**.

- $\lambda_t$ a **stochastic process** and $N_t$ is a birth and death process conditional on the $\lambda_t$ process.

Let the measure associated with the process $N_t$ be $\mathcal{P}$.

Let $\xi(t)$ be the underlying process of event times.

So $d\xi(t) = 1$ if there is a jump at time $t$.

$dS(t) = cY(t)d\xi(t)$
11. Edgeworth expansion for Option Prices

Let us define \( X_t^{(n)} = \ln \left( \frac{N_t^{(n)}}{n} \right) \)

\[
X_t^{*(n)} = X_t^{(n)} - X_0^{(n)} - \int_0^t \left[ p_{u,N_u} \log(1 + \frac{1}{N_u}) \right. \\
\left. + (1 - p_{u,N_u}) \log(1 - \frac{1}{N_u}) \right] N_u^2 \sigma_u^2 \, du
\]

where \( p_{t,N_t} = \frac{1}{2} \left( 1 + \frac{\rho_t}{\sqrt{N_t \sigma_t^2}} \right) \).

Let \( C \) be the class of functions \( g \) that satisfy the following:

(i) \( \int |\hat{g}(x)| \, dx < \infty \), uniformly in \( C \), and \( \left\{ \sum_u x_u^2 \hat{g}(x), g \in C \right\} \) is uniformly integrable (here, \( \hat{g} \) is the Fourier transform of \( g \), which must exist for each \( g \in C \)); or

(ii) \( g \) nd \( g'' \) bounded, uniformly in \( C \), and with \( g'' \) equicontinuous almost everywhere (under Lebesgue measure).

Under assumptions (I1) and (I2), for any \( g \in C \),

\[
E_g(X_T^{*(n)}) = E_g(N(0, \lambda T)) + o(1/n)
\]

(I1) There are \( \underline{k}, \bar{k}, \underline{k} < \lambda T < \bar{k} \) so that

\[
n \left( l_T^{(n)} - \lambda T \right) \mathbf{1}(\underline{k} \leq l_T^{(n)} \leq \bar{k}) \text{ is uniformly integrable, where } l_T^{(n)} = (X_t^{*(n)}, X_t^{*(n)})_T
\]

(I2) For the same \( \underline{k}, \bar{k} \),

\[
P(\underline{k} \leq (X_t^{*(n)}, X_t^{*(n)})_T \leq \bar{k}) = 1 - o(1/n)
\]
12. Hedging

The market is complete when we add a market traded derivative security. We can hedge an option by trading the stock, the bond and another option. Let $F_2(x, t), F_3(x, t)$ be the prices of two options at time $t$ when price of stock is $cx$. Let $F_1(x, t) = cx$ be the price of the stock and $F_0(x, t)B_0 \exp\{ - \int_0^t \rho_s ds \}$ be price of the bond. Assume $F_i$ are continuous in both arguments.

We shall construct a self financing risk-less portfolio

$$V(t) = \sum_{i=0}^{3} \phi^i(t) F_i(x, t)$$

Let $u^i(t) = \frac{\phi^i(t)F_i(x,t)}{V(t)}$ be the proportion of wealth invested in asset $i$. $\sum u^i = 1$

Since $V_t$ is self financing,

$$\frac{dV(t)}{V(t)} = \sum_{i=0}^{3} u^i(t) \frac{dF(x, t)}{F(x, t)}$$

$$= u^{(0)}(t) \rho_t dt + u^{(1)}(t) \frac{1}{cx} (dN_{1t} - dN_{2t})$$

$$+ \sum_{i=2}^{3} u^i(t) (\alpha_{Fi}(x, t) dt + \beta_{Fi}(x, t) dN_{1t} + \gamma_{Fi}(x, t) dN_{2t})$$
\( V_t \) is risk-less \( \implies \)
\[
u^{(1)}(t) \frac{1}{c_x} + \sum_{i=2}^{3} u^{(i)}(t) \beta_{F_i}(x, t) = 0,
- u^{(1)}(t) \frac{1}{c_x} + \sum_{i=2}^{3} u^{(i)}(t) \gamma_{F_i}(x, t) = 0
\]

No arbitrage \( \implies u^{(0)}(t) \rho_t dt + \sum_{i=2}^{3} u^{(i)}(t) \alpha_{F_i}(x, t) = \rho_t \)

The hedge ratios are:

\[
u^{(2)} = \left[ \frac{1 - \alpha_{F_2}}{\rho} - x \beta_{F_2} \right] \left( 1 - \frac{\gamma_{F_2} + \beta_{F_2}}{\gamma_{F_3} + \beta_{F_3}} \right)^{-1}
\]

\[
u^{(3)} = \left[ \frac{1 - \alpha_{F_3}}{\rho} - x \beta_{F_3} \right] \left( 1 - \frac{\gamma_{F_3} + \beta_{F_3}}{\gamma_{F_2} + \beta_{F_2}} \right)^{-1}
\]

\[
u^{(0)} = - \frac{1}{\rho_t} (\nu^{(2)} \alpha_{F_2} + \nu^{(3)} \alpha_{F_3})
\]

\[
u^{(1)} = -x (\nu^{(2)} \beta_{F_2} + \nu^{(3)} \beta_{F_3})
\]
13. Stochastic Intensity

Now we consider the case where the unobserved intensity $\lambda_t$ is a stochastic process. We first assume a two state Markov model for $\lambda_t$. \cite{Naik1993}

Later we describe how we can have similar results for other models on $\lambda_t$. \cite{HullWhite1987}

Suppose there is an unobserved state process $\theta_t$ which takes 2 values, say 0 and 1. The transition matrix is $Q$.

When $\theta_t = i$, $\lambda_t = \lambda_i$.

Counting process associated with $\theta_t$ is $\zeta_t$.

Let us denote by $\{G_t\}$ the complete filtration $\sigma(S_u, \lambda_u, 0 \leq u \leq t)$ and by $P$ the probability measure on $\{G_t\}$ associated with the process $(S_t, \lambda_t)$. 
14. Risk-neutral distribution

Let us assume that the risk-neutral measure is the measure associated with a birth-death process with event rate $\lambda^*_t N^2_t$ and probability of birth $p^*_t = \frac{1}{2} \left(1 + \frac{\rho_t}{\lambda^*_t N_t}\right)$ where $\lambda^*_t$ is a Markov process with state space $\{\lambda_1, \lambda_2\}$.
We get two different values of the expected price under the two values of $\theta(0)$.
The $\theta$ process is unobserved.
We cannot invert an option to get $\theta(0)$ because it takes two discrete values.
Need to introduce $\pi_i(t) = P(\theta_t = i \mid \mathcal{F}_t)$ where $\mathcal{F}_t = \sigma(S_u, 0 \leq u \leq t)$
As shown in Snyder(1973), under any $\hat{P} \in \mathcal{P}$ the $\pi_{it}$ process evolves as:
$d\pi_{1t} = a(t)dt + b(t, 1)dN_{1t} + b(t, 2)dN_{2t}$ where $a(t)$ and $b(t, i)$ are $\mathcal{F}_t$ adapted processes.
We still need one market traded option and the stock to hedge an option. But to get the hedge ratios, we need $\pi^*(t), a(t), b(t)$.

- The hedge ratios involve $a(t), b(t)$. So we need them to be predictable. But if we have to invert an option to get them, then we need to observe the price at time $t$ to infer $\pi_t$ and from there to get $a_t$ and $b_t$. So they are no more predictable.
- Given a value of $\pi(t_0)$, the whole process $\pi(t), t > t_0$ is determined by the conditional distribution of the $\theta$ process given the observed process $S_t$. $\pi(t)$ is completely determined by historical data. We do not have the freedom of determining $\pi(t)$ by inverting options. Thus, inferring $\pi(t)$ at each time point $t$ independently will give incorrect prices and lead to arbitrage.
15. Bayesian Framework

As shown in Yashin (1970) and Elliott et. al. (1995), the posterior of $\theta_j(t)$ is given by:

$$\pi_j(t) = \pi_j(0) + \int_0^t \sum_i q_{ij} \pi_i(u)du$$

$$+ \int_0^t \pi_j(u)(\bar{\lambda}(u) - \lambda_j)N_u^2 du$$

$$+ \sum_{0<u<t} b_j(u)$$

where $\bar{\lambda}(t) = \sum_i \pi_i(t)\lambda_i$ and $b_j(u) = \pi_j(u-)
\left(\frac{\lambda_j p_{\lambda_j}(S_{u-\rightarrow S_u})}{\sum_i \pi_i(u)\lambda_i p_{\lambda_i}(S_{u-\rightarrow S_u})} - 1\right)$

Thus, $a_j(u) = \sum_i q_{ij} \pi_i(u) + \int_0^t \pi_j(u)(\bar{\lambda}(u) - \lambda_j)N_u^2$

Now we can hedge as in the constant intensity case with modified hedge ratios.
16. Pricing

How to get $E(e^{-\int_0^T \rho_t ds} X \mid \mathcal{F}_0)$ under fixed values of $Q, \lambda_0, \lambda_1, \pi_0(0)$?
Fix $\theta_0 = i$ Generate the $\theta$ process. Generate the $\xi$ waiting times as non-homogeneous Poisson process. At each event time $S_t$ jumps by $\pm c$ with probability $p_t$ and $1 - p_t$. Get the expectation under $\theta_0 = i$. Now take the average of these with respect to $\pi_0$.
Generating $\xi$: Generate $T_0$ from $\text{Exp}(\lambda_{\theta_0} N^2_0)$
Let $\tau_0 = \inf\{t : \theta_t \neq \theta_0\}$
If $T_0 < \tau_0$, jump at time $T_0$.
Otherwise, generate $T_1$ from $\text{Exp}(\lambda_{\theta_{\tau_0}} N^2_{\tau_0})$
$\tau_1 = \inf\{t : \theta_t \neq \theta_{\tau_0}\}$
If $T_1 < \tau_1$, jump at time $\tau_0 + T_1$
Continue. This is justified by memorylessness.
We need to infer 5 parameters $q_{01}, q_{10}, \lambda_0, \lambda_1, \pi_0(0) \Rightarrow$ Invert the prices of 5 options with different maturities at time 0. Using the filtering equations it is possible to infer the required quantities by using one option at 5 points of time close to 0.
$\pi_0(t_j) - \pi_0(t_0) = [q_{00}\pi_0(t_0) + q_{10}\pi_1(t_0) + \pi_0(t_0)](t_j - t_0) + [\pi_0(t_0) - \pi_1(t_0)]\lambda_2 \ldots (\ast)$
Another possibility of reducing the number of parameters to be estimated is to assume that the $\theta$ process is in equilibrium when we start observing. Then we need to estimate (or invert for) 4 parameters($q_{01}, q_{10}, \lambda_0, \lambda_1$).
The chain is irreducible and $T_{ii} \sim \text{Exp}(q_{10} + q_{01})$ that is non-lattice and finite mean.
$\pi_0(0) = \frac{q_{10}}{q_{10} + q_{01}}$ (e.g. Ross Pg214)
17. Generalizations

According to Snyder(1973), the $\pi_i(t)$ process evolves as:

$$d\pi(t) = a(t)dt + b(t, 1)dN_{1t} + b(t, 2)dN_{2t}$$

where $a(t)$ and $b(t, i)$ are $\mathcal{F}_t$ adapted processes. Even if we do not want to specify the model for the $\lambda$ process, it is still of that form.

Let us see the form of the posterior in the general case:

Let $c_t(v | N_{t_0,t})$ be the posterior characteristic function for $\lambda_t$ given an observed path

$$dc_t(v | N_{t_0,t}) = \lambda_{N_t,N_{t+dt}}^{-1} \times E(\exp(iv\lambda_t)(\lambda_{N_t,N_{t+dt}} - \lambda_{N_t,N_{t+dt}}^*) | N_{t_0,t})d\xi_t$$

$$+ E(\exp(iv\lambda_t)\Psi_t(v | N_{t_0,t}, \lambda_t) | N_{t_0,t}) dt - E(\exp(iv\lambda_t)g(N_t)(\lambda_t - \hat{\lambda}_t) | N_{t_0,t}) dt$$

where $\lambda_{N_t,\zeta_t} = g(N_t)\lambda_t p_{\lambda_t,N_t}$ if $\zeta_t = N_t + 1$

$$g(N_t)\lambda_t(1 - p_{\lambda_t,N_t})$$ if $\zeta_t = N_t - 1$

$$0 \text{ o. w.}$$

$$\lambda_{N_t,\zeta_t}^* = E(\lambda_{N_t,\zeta_t}(t, N_{t_0,t}, \lambda_t) | N_{t_0,t})$$

$$\Psi_t(v | N_{t_0,t}, \lambda_t) = E(\exp(iv\Delta \lambda_t) | N_{t_0,t}, \lambda_t)$$

For example, let $d\lambda_t = f_t(\lambda_t)dt + G_t(\lambda_t)dW_t$, where $W_t$ is Brownian motion. In this case

$$\Psi_t(v | N_{t_0,t}, \lambda_t) = ivf_t(\lambda_t) - \frac{1}{2}v^2G_t^2(\lambda_t).$$

As long as the observed process is Markov jump process, the second and third terms are same as in the two state Markov case and hence yield the same function on taking inverse Fourier transform. The first term, on taking inverse Fourier transform yields $L$, the Kolmogorov-Fokker-Plank differential operator associated with $\lambda$.

$$d\pi_t = L + \int_0^t \pi_j(u)(\lambda_j - \lambda_j)g(S(u))du + \sum_{0 < u < t} \pi_j(u-)
\left(\frac{\lambda_j p_{\lambda_j}(S_u \to S_u)}{\sum_i \pi_i(u)\lambda_i p_{\lambda_i}(S_u \to S_u)} - 1\right)$$
Another possible direction is to consider jumps of size $> 1$. But then we no longer have the distribution of jump size from simple martingale considerations. We have to either assume the jump distribution, or estimate it, or impose some optimization criterion to get a unique price. Also, if the jump magnitude can take $k$ values, we need $k - 1$ market traded options and the stock to hedge an option.
Figure 6: Estimation: ABMD Data, general model

Figure 7: Estimation: ABMD Data, Birth Death model with Stochastic volatility
Black Scholes with constant volatility
Figure 8: Error in CALL price for training sample of IBM data
Figure 9: Error in CALL price for test sample of IBM data
18. **Stochastic Intensity Rate**

- To estimate 5 parameters: $\lambda_0, \lambda_1, q_{01}, q_{10},$ and $\pi$.

- The objective is to find the parameter set that minimizes the root mean square error between the bid-ask-midpoint and the daily average of the predicted option price, for all options in the training sample.

- We followed a diagonally scaled steepest descent algorithm with central difference approximation to the differential.

- The starting values of $\lambda_0, \lambda_1$ are taken to be equal to the value of the estimator $\hat{\lambda}$ obtained in the constant intensity model.

- The starting values of $q_{01}, q_{10}$ are obtained by a hidden Markov model approach using an iterative method (Ref Elliott 1995).

- We do a finite search on the parameter $\pi$.

- For the ABMD and Ford datasets, the RMSE of prediction obtained from the constant intensity method is less than the bid-ask spread.

- For IBM data, $q_{01}, q_{10}$ and $\pi$ are $8.64e-02, 1.2126$ and $\lambda_0$ and $\lambda_1$ are $1.042039e-06$ and $8.326884e-08$. 

• The RMSE is 29.6799. Compare this to the birth death model with constant intensity (RMSE=52.7729) or Black-Scholes model with constant volatility (RMSE=46.7688).

• This is an ill-posed problem.
19. Conclusions

- Quadratic Variation is not observable
- General model is Nonparametric
- Bounds are small enough to be of practical use
- New Statistical problem of predicting an interval
- General jump models are not amenable to hedging, Birth-Death process is
- Stochastic intensity
- Combining risk neutral estimation with updating by historical data
- Need extra derivative to hedge an option