Jump Inference

1 Jump Detection and Estimation

Let’s think in terms of typical day with continuous $t \in [0, 1]$. The dynamics are

$$X = \int_0^1 \sqrt{c_s} dw_s + \sum_{0 \leq s \leq 1} \Delta X_s$$

and discrete observations:

$$\Delta^n_i X \quad i = 1, 2, \ldots, n$$

Truncation Indicator

$$T_i = \begin{cases} 
1 & \text{if } |\Delta^n_i X| \leq \alpha \sqrt{\tau_i BV}\Delta^n_\sigma n \\
0 & \text{otherwise}
\end{cases}$$

Suppose we detect a jump at time $t = \tau$ somewhere in the day; actually, we only know that

for some $\tilde{j}$ that $T_{\tilde{j}} = 1$ where $(\tilde{j} - 1)\Delta_n < \tau < \tilde{j}\Delta_n$. (Our lack of knowledge about where jumps lies within the interval turns out to matter.) Interestingly, we can think of the return over the jump interval as an estimator of the jump itself. Since

$$\Delta^n_{\tilde{j}} X = \Delta^n_{\tilde{j}} X^c + \Delta^\tau X$$

then

$$\Delta^n_{\tilde{j}} X - \Delta^\tau X = \Delta^n_{\tilde{j}} X^c$$

In terms of our other notation we would write

$$r^c_{\tilde{j}} = r^d_{\tilde{j}}$$

The “error” in using the observed return to estimate the jump size is exactly the diffusive return over the interval, and it $\to 0$ as $n \to \infty$.

2 Inference about Jump Size: Continuous Volatility

For a while we will assume the volatility function $c_t$ is continuous. We will relax that assumption and see that it matters, a lot. From (5) or (6) we know that asymptotically

$$\Delta^n_{\tilde{j}} X - \Delta^\tau X \approx N(0, \Delta_n c_\tau)$$

and

$$\sqrt{\text{Var}(\Delta^n_{\tilde{j}} X - \Delta^\tau X)} \approx \Delta_n^{1/2} c_\tau^{1/2}$$

i.e., the the error is approximately Gaussian and depends upon just the local continuous part of the model. The statement makes a lot of sense on reflection, especially if we think in terms of estimation error.
\[
\Delta_n^{-1/2}(\Delta_j^n X - \Delta_r X) \approx \sqrt{c_r} Z \tag{9}
\]

where \( Z \sim N(0,1) \). We can set a confidence interval for the jump size if we know the local variance \( \sqrt{c_r} \), but that is exactly the estimation problem we have been studying. Another way to write this is

\[
\Delta_n^{-1/2}(\hat{J}_r - J_r) = \Delta_n^{-1/2} \int_{(i-1)\Delta_n}^{i\Delta_n} \sqrt{c_s} dw_s \approx \sqrt{c_r} Z, \quad Z \sim N(0,1) \tag{10}
\]

so long as the local variance \( c_s \) is continuous across the jump at time \( \tau \). The feasible version is

\[
\frac{\Delta_n^{-1/2}(\hat{J}_t - J_t)}{\sqrt{\hat{c}_\tau}} \overset{p}{\approx} N(0,1) \tag{11}
\]

To implement the feasible version, we need to get an estimate of \( c_\tau \) for the jump time \( \tau \). But we know how to do this. A natural thing (given continuity of \( c \)) is to estimate it from the right and left and then average the two:

\[
\hat{c}_\tau^+ = \frac{1}{k_n \Delta_n} \sum_{i=1}^{k_n} r_{j+i}^c, \quad \hat{c}_\tau^- = \frac{1}{k_n \Delta_n} \sum_{i=1}^{k_n} r_{j-i}^c \tag{12}
\]

Again, if we think of \( \hat{J}_r = \Delta_j^n X \) as the “estimate” of the jump, and \( J_r = \Delta_r X \) as the unobserved jump itself, then we can form a confidence interval for the jump at time \( \tau \) as

\[
J_r = \hat{J}_r \pm z_{0.025} \Delta_n^{1/2} \sqrt{\hat{c}_\tau} \tag{13}
\]

at level 95 percent.

3 Jump Estimation - Scalar Case, Discontinuous Volatility

3.1 A single jump

In general, the volatility function is discontinuous; and in particular there is ample evidence that the variance usually jumps at the same time price jumps:

\[
c_\tau^- = \lim_{u \uparrow \tau} c_u, \quad c_\tau^+ = \lim_{r \downarrow \tau} c_r, \quad c_\tau^- \neq c_\tau^+ \tag{14}
\]
As usual, however, \( c \) is a right continuous function so \( c_\tau = c_\tau^+ \). The variance jump complicates the asymptotic theory of how the observed return behaves across the jump interval,

\[
[(\tilde{j} - 1)\Delta_n, \tilde{j}\Delta_n)]
\]

It is the uncertainty about where the jump time \( \tau \) lies in the above interval that causes the complications. The key result is

\[
\Delta_n^{-1/2}(\tilde{J}_\tau - J_\tau) \overset{D}{\approx} \sqrt{\kappa c^-_\tau} Z_1 + \sqrt{(1 - \kappa)c^+_{\tau}} Z_2, \quad \kappa \sim \text{Uniform}(0, 1), \quad Z_1, Z_2 \sim N(0, 1).
\] (15)

The above represents the gist of (Jacod and Protter, 2012, p. 121). Equation (15) is crucial for much of this course. The limiting distribution is mixed normal with mixing variable \( \kappa \), and \( \kappa \) reflects the fact that we do not know where the jump occurred in the interval that the detected jump lies in.

### 3.2 Multiple Jumps

Suppose now our data are \( X_{i\Delta_n}, i = 0, 1, 2, \ldots, nT \); that is \( n \) observations per day for \( T \) days. Suppose we have detected jumps at intervals \( (i_p)_{p=1,2,\ldots,P} \). These jumps occurred at times \( t_p \in [(i_p - 1)\Delta_n, i_p\Delta_n] \). The collection of jump times is \( (t_p)_{p=1,2,\ldots,P} \), but the complication is that we do not know where each jump occurred within the interval.

Our key results are these:

\[
\Delta_n^{-1/2}(\Delta^n_{i_p} X - \Delta X_{t_p}) \overset{D}{\approx} \zeta_p
\] (16)

where

\[
\zeta_p = \sqrt{(1 - \kappa_i)c^-_{i_p}} Z_{1p} + \sqrt{\kappa_i c^+_{i_p}} Z_{2p}
\] (17)

for \( p = 1, 2, \ldots, P \), where the \( \kappa_i \) are iid uniform \((0, 1)\) and the \( Z_{1p}, Z_{2p} \) are iid \( N(0, 1) \) random variables.

The asymptotic results of equations (16) and (17) are due to Jacod and Protter (2012), and they are fundamental for much of the material to follow in this course.

Unfortunately, the asymptotic distribution is not pivotal (it depends upon unknown parameters) and it can only be made feasible via simulation. First estimate the right and left variances:

\[
\hat{c}^{+}_{i_p} = \frac{1}{k_n\Delta_n} \sum_{i=1}^{k_n} (\bar{r}^c_{i_p+i})^2
\] (18)

\[
\hat{c}^{-}_{i_p} = \frac{1}{k_n\Delta_n} \sum_{i=1}^{k_n} (\bar{r}^c_{i_p-i})^2
\]

Now simulate \( \tilde{\kappa}_p, \tilde{Z}_{1p}, \tilde{Z}_{2p} \) and form

\[
\tilde{\zeta}_p = \sqrt{(1 - \tilde{\kappa}_i)\hat{c}^-_{i_p}} \tilde{Z}_{1p} + \sqrt{\tilde{\kappa}_i \hat{c}^+_{i_p}} \tilde{Z}_{2p}.
\] (19)
Repeat the simulation for 1000 or more replications. The distribution of the simulated $\tilde{\zeta}_p$ will capture the asymptotic distribution of $\Delta_n^{-1/2}(\Delta_{t_p}^n X - \Delta X_{t_p})$ from equation (16).
References