Jump Regression

Guilherme Salomé

October 11, 2018

The idea of the jump regression is to study the jump dependence between two processes from high-frequency observations. In this context, we will study whether moves in a stock's price are dependent on jumps in the market index.

1 Setup

The setup we will use for jump regressions uses 2 assets:

$$\begin{aligned} X_1 &: \text{ market index} \\ X_2 &: \text{ stock} \end{aligned}$$

The dynamics are as follows:

$$\begin{aligned} dX_{1,t} &= \sigma_{1,t} dW_{1,t} + \Delta J_{1,t} \\ dX_{2,t} &= \sigma_{2,t} dW_{2,t} + \Delta J_{2,t} + \beta_t^c dX_{1,t}^c + \beta_t^d dX_{1,t}^d \end{aligned}$$

where:

$$\begin{aligned} \sigma_{1,t} &: \text{ market diffusive volatility} \\ \sigma_{2,t} &: \text{ idiosyncratic diffusive volatility} \\ \Delta J_{1,t} &: \text{ market index jumps} \\ \Delta J_{2,t} &: \text{ idiosyncratic jumps} \\ \beta_t^c &: \text{ beta on market diffusive moves} \\ dX_{1,t}^c &: \text{ market diffusive moves } (\sigma_{1,t} dW_{1,t}) \\ \beta_t^d &: \text{ beta on market jumps} \\ dX_{1,t}^d &: \text{ market jump moves } (\Delta J_{1,t}) \end{aligned}$$

Built in the model is the assumption that jumps in the market index (X_1) are independent of jumps in the asset (X_2), which implies that $\Delta J_{1,t} \Delta J_{2,t} = 0$. That is, if at time t there is a jump in $X_{1,t}$, then at that that same instant $X_{2,t}$ cannot have an idiosyncratic jump (jump independent of the market index), but it still can jump (if $\beta_t^d > 0$). Another assumption is that there is a finite number of jumps in the market index and in the stocks.

We can analyze what happens to the stock at times when: the market jumps, when the stock jumps, and when there are no jumps.

Let t_p for $p = 1, 2, \ldots, P$ denote the times when the market jumps. At these instants we know there are no idiosyncratic jumps in the stock, that is, $\Delta J_{2,t_p} = 0$. However, the stock can still jump via a dependence on the market moves as captured by the β_t^d:

$$\Delta X_{2,t_p} = \beta_{t_p}^d \underbrace{\Delta X_{1,t_p}}_{\Delta J_{1,t_p}} \text{ for } p = 1, 2, \ldots, P$$

This equation implies that stock jumps move accordingly to jumps in the market index.

Now, let t_h denote times when there are idiosyncratic jumps in the stock price. At these instants the market index does not jump ($\Delta X_{1,t_h} = 0$), so the jump in the stock price is:

$$\Delta X_{2,t_h} = \Delta J_{2,t_h} \text{ for } t_h \neq t_p$$

Lastly, at times when neither the market nor the stock jump:

$$\Delta X_{1,t} = 0 \text{ and } \Delta X_{2,t} = 0 \text{ for } t \neq t_h \neq t_p$$

Therefore, we can write the jumps in the stock as:

$$\Delta X_{2,t} = \beta_t^d \Delta J_{1,t} + \Delta J_{2,t}, \forall t \in [0, T]$$

and we also have $\Delta J_{1,t} \Delta J_{2,t} = 0$.

If we willing to assume that the jump beta is constant, that is $\beta_{t_p}^d \equiv \beta$, then at times when the market jumps we have the relation:

$$\Delta X_{2,t_p} = \beta \Delta X_{1,t_p}$$

This is an exact linear relationship between jumps in the market index and jumps in the stock.

There are 2 ways in which this constant jump beta assumption can fail:

1. The relation between the jumps is linear but changes over time: β_{t_p}
2. The relationship between jumps is non-linear: $\beta(\Delta X_{1,t_p})$

We will work with the assumption of a constant jump beta, and see whether it makes sense in the data. The relationship we are interested in is of jumps, so next we will discuss more details of how to detect the jumps.

2 Detecting Market Jump Returns

The idea of the jump regression is to use the data at times when the market jumped to learn how the moves on the stock depend on the jumps in the market index. We have seen previously that we can detect jumps by analyzing the size of returns in comparison to a local variance estimate. Indeed, we defined jump returns as:

$$r_{t,i}^d \equiv r_{t,i} 1_{\{|r_{t,i}| > \alpha cut_{t,i}\}}$$

We will use a high α to select only returns where we are quite confident the market jumped.

We can define the set of times when the market price actually jumps:

$$\mathcal{T} \equiv \{t \in [0, T] : \Delta X_{1,t} \neq 0\}$$

The times collected in this set are the true jump times for the market index. Because there are only a finite number of jumps, we can say the set $\mathcal{T}$ contains P elements.

For Δ_n small enough, each jump time $t \in \mathcal{T}$ will lie in a distinct discrete interval of size Δ_n. That is:

$$\exists \Delta_n > 0 : \forall t_p \in \mathcal{T}, \exists i_p \in \mathbb{N} : t_p \in ((i_p - 1)\Delta_n, i_p \Delta_n)$$

We can collect these indices i_p's in a set:

$$\mathcal{I} \equiv \{i_p\}_{p=1,\ldots,P}$$

If we sampled the market index every Δ_n, then the intervals defined by the i_p's would each contain a single jump in the market index.

We can use the jump threshold to detect the indices of the intervals where we identify jumps. There are two sets of interest. First, is the set of indices i_p where our jump threshold technique identifies jumps in intervals where jumps happened:

$$\mathcal{I}_n \equiv \{i \in \mathcal{I} : |\Delta_i^n X_{1,t}| > \alpha cut_{t,i}\}$$

The indices in this set are indices of the intervals where two things happen: the market index actually jumped and the jump threshold detects a jump the market index. It must be the case that $\mathcal{I}_n \subset \mathcal{I}$. We do not have an equality between the sets because the detector might miss some of the jumps.

Second, is the set of indices where our jump threshold identifies jumps:

$$\mathcal{I}'_n \equiv \{i \in \{1, 2, \ldots, nT\} : |\Delta_i^n X_{1,t}| > \alpha cut_{t,i}\}$$

The indices in this set are indices where the jump threshold detected a jump. In general, $\mathcal{I}'_n \neq \mathcal{I}$. There are two reasons for this:

1. The jump threshold can miss an actual jump: $i \in \mathcal{I}$ but $i \notin \mathcal{I}'_n$
2. The jump threshold declares that an interval contains a jump when it does not: $i \in \mathcal{I}'_n$ but $i \notin \mathcal{I}$

The set $\mathcal{I}_n$ contains only the jumps that actually happened and we correctly identified. In practice, however, we do not know this set. And the set $\mathcal{I}'_n$ contains jumps that we identified, but could miss jumps or incorrectly identify moves as jumps. In practice, we can compute the indices of this set. These two sets are of importance because of Li, Todorov, and Tauchen (2017). These authors show that as $\Delta_n \to 0$ we have:

$$\mathcal{I}_n \to \mathcal{I}$$
$$\mathcal{I}'_n \to \mathcal{I}$$

That is, the jump detector is correct in the limit. The speed (rate) of convergence is high, so we can basically act as if $\mathcal{I}'_n$ coincides with $\mathcal{I}$.

3 Jump Beta

The theory becomes simpler when we make two additional assumptions: the variance processes for the market index and stock are continuous, and that the diffusive beta equals the jump beta across the intervals where the market jumps. The theory can handle some deviations from these assumptions, but at the expense of extra theoretical derivations.

In the classical case of ordinary linear regression we have:

$$y = X\beta + u \text{ and } \mathrm{Var}[u] = \Omega \neq \sigma^2 I$$

And we know β can be estimated by:

$$\hat{\beta} \equiv (X'X)^{-1}X'y \to \beta$$

It is also possible to show that:

$$\begin{aligned}\hat{\beta} - \beta &= (X'X)^{-1}X'u \\ \mathrm{Var}\left[\hat{\beta}\right] &= (X'X)^{-1}X'\Omega X(X'X)^{-1}\end{aligned}$$

Very similar expressions show up in the jump regression setting, but the core theory is entirely different.

The setup with constant jump beta implies the model:

$$\Delta X_{2,t} = \beta \Delta J_{1,t} + \Delta J_{2,t} \text{ and } \Delta J_{1,t}\Delta J_{2,t} = 0, \forall t \in [0, T]$$

The jump beta estimator is given by:

$$\hat{\beta} \equiv \frac{\sum_{p=1}^{P_n}(\Delta_{i_p}^n X_1)(\Delta_{i_p}^n X_2)}{\sum_{p=1}^{P_n}(\Delta_{i_p}^n X_1)^2} \to \beta$$

where P_n is the number of detected jumps in the set $\mathcal{I}'_n$. Notice that this estimator is an analogue of the OLS estimator, it is the same estimator we would obtain from a regression of jumps in the stock on the jumps in the market index. However, the underlying theory to prove the estimator's convergence and asymptotic distribution is entirely different.

It is possible to show that the difference between the jump beta estimator and the actual jump beta is:

$$\hat{\beta} - \beta = \frac{\sum_{p=1}^{P_n} \Delta_{i_p}^n X_1 \Delta_{i_p}^n E}{\sum_{p=1}^{P_n}(\Delta_{i_p}^n X_1)^2}$$

where $E_t \equiv X_{2,t} - \beta X_{1,t}$ is the error process, which itself has a variance process denoted by $c_{e,t}$.

Let's see what $\Delta_{i_p}^n E$ is (this is the "return" of the error process across intervals where the market jumped):

$$\begin{aligned}\Delta_{i_p}^n E &= \Delta_{i_p}^n X_2 - \beta \Delta_{i_p}^n X_1 \\ &= \Delta_{i_p}^n (X_2^c + X_2^d) - \beta \Delta_{i_p}^n (X_1^c + X_1^d) \\ &= \underbrace{\left[\Delta_{i_p}^n X_2^c - \beta \Delta_{i_p}^n X_1^c\right]}_{\approx \sqrt{\Delta_n c_{e,t_p}} Z_p} + \underbrace{\left[\Delta_{i_p}^n X_2^d - \beta \Delta_{i_p}^n X_1^d\right]}_{= 0 \text{ at the } i_p \text{ intervals}}\end{aligned}$$

where the approximation is shown in Jacod and Protter (2012). Therefore,

$$\hat{\beta} - \beta \approx \frac{\sum_{p=1}^{P_n} \Delta_{i_p}^n X_1 \sqrt{\Delta_n c_{e,t_p}} Z_p}{\sum_{p=1}^{P_n} (\Delta_{i_p}^n X_1)^2}$$

The asymptotic distribution of the jump beta estimator is:

$$\Delta_n^{-\frac{1}{2}} (\hat{\beta} - \beta) \to \mathcal{N}(0, V_\beta)$$
$$V_\beta \equiv \frac{\sum_{p=1}^{P_n} (\Delta_{i_p}^n X_1)^2 c_{e,t_p}}{(\sum_{p=1}^{P_n} (\Delta_{i_p}^n X_1))^2}$$

If we plug in an estimator for c_{e,t_p} and assume n is large enough, we get the following applied form:

$$\hat{\beta} \approx \mathcal{N}\left(\beta, \Delta_n \frac{\sum_{p=1}^{P_n} (\Delta_{i_p}^n X_1)^2 \hat{c}_{e,t_p}}{(\sum_{p=1}^{P_n} (\Delta_{i_p}^n X_1))^2}\right)$$

We can estimate the variance of E_t using the local variance estimator:

$$\hat{c}_{e,t_p} \equiv \frac{1}{(2k_n + 1)\Delta_n} \sum_{j=-k_n}^{k_n} \left(\Delta_{i_p+j}^n X_{2,t} - \hat{\beta} \Delta_{i_p+j}^n X_{1,t}\right)^2 1_{\left\{\left|\Delta_{i_p+j}^n X_{1,t}\right| \leq \alpha cut_{t,i_p+j}\right\}}$$

And set $k_n = 7$ (depends on how smooth we think the variance is across the jump). Notice that we would need to adjust the local variance estimator if the time is close to the beginning or the end of the market hours. The jump threshold used in the equation above is the one for the market index.

References

Jacod, Jean and Philip Protter (2012). *Discretization of Processes*. Springer-Verlag.

Li, Jia, Viktor Todorov, and George Tauchen (2017). "Jump Regressions". In: **Econometrica** 85, pp. 173–195.