1 Example: Realized Beta

The ideas are best explained example. We consider here the realized beta computed via jump truncated returns.

1.1 Description of Problem

Suppose $X_1$ and $X_2$ are two random variables each with mean zero; the theoretical regression of $X_2$ on $X_1$ is

$$X_2 = X_1\beta + \epsilon$$

where by definition

$$\beta = \frac{\text{Cov}(X_1, X_2)}{\text{Var}(X_1)}.$$  \hspace{1cm} (2)

In the theoretical regression, the random variable $\epsilon = X_2 - \beta X_1$ by definition, and by construction $\text{Cov}(\epsilon, X_1) = 0$. Given observations $X_{1i}, X_{2i}, i = 1, 2, \ldots, n$ we can estimate $\beta$ by

$$\hat{\beta} = \frac{\sum_i^n X_{1i}X_{2i}}{\sum_i^n X_{1i}^2}.$$  \hspace{1cm} (3)

Under mild regularity conditions $\hat{\beta}$ is a consistent estimator of $\beta$ with a well defined asymptotic distribution. The theoretical regression is a statistical construction that essentially always exists so long as the moments exist, but it does not always make economic sense.

1.2 The High Frequency Setting

We start with the usual model, where here $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, with $X_1$ being the log of a market price index and $X_2$ is the log price of a particular stock or portfolio. We assume $X_t$ is a semimartingale, which is the usual model for financial prices. Its dynamics are governed by

$$X_t = \int_0^t \sqrt{c_s} \, dW_s + J_t \quad 0 \leq t \leq T.$$  \hspace{1cm} (4)

where $W_t$ is a standard Brownian motion and $J_t$ is a jump process. Here $c_t$ is the symmetric local variance matrix

$$c_t = \begin{bmatrix} c_{11,t} & c_{12,t} \\ c_{21,t} & c_{22,t} \end{bmatrix}.$$  \hspace{1cm} (5)
and we interpret $\sqrt{c_t}$ as the symmetric square root of $c_t$. We decompose $X_t$ into its continuous and discontinuous parts as

$$X_t = X_t^c + X_t^d.$$  \hspace{1cm} (6)

Under the usual sampling scheme we have observations $X_{t\Delta_n}$, $i = 1, 2, \ldots, nT$, at integers (days) $t = 1, 2, \ldots, T$, $\Delta_n = 1/n$. Suppose (for a while) that we directly observe the continuous and discontinuous parts of $X$. We do not, of course, but we can separate out and estimate these two components. Define the returns

$$r_{1t}^c = \Delta_n^{n}_{(t-1)n+i} X_1^c, \quad r_{2t}^c = \Delta_n^{n}_{(t-1)n+i} X_2^c, \quad i = 1, 2, \ldots, n$$  \hspace{1cm} (7)

From our earlier work, we know how to estimate the integrated variance and covariance,

$$\sum_{i=1}^{n} (r_{1t}^c)^2 \rightarrow \int_{t-1}^{t} c_{11,s} \, ds, \quad \sum_{i=1}^{n} r_{1t}^c r_{2t}^c \rightarrow \int_{t-1}^{t} c_{12,s} \, ds, \quad n \rightarrow \infty.$$  \hspace{1cm} (8)

So, by analogy with the theoretical regression ($\text{beta} = \text{Cov}/\text{Var}$) we define the integrated beta as

$$\int_{t-1}^{t} c_{12,s} \, ds \quad \int_{t-1}^{t} c_{11,s} \, ds.$$  \hspace{1cm} (9)

The estimator of the integrated beta is the realized beta given by

$$\hat{R}_t^\beta = \frac{\sum_{i=1}^{n} r_{1t}^c r_{2t}^c}{\sum_{i=1}^{n} (r_{1t}^c)^2}$$  \hspace{1cm} (10)

2 Inference

There are expressions in Ait-Sahalia and Jacod (2014) and the literature for the asymptotic variance of $\hat{R}_t^\beta$, but here we will use a bootstrap approximation. For now, drop the $t$ subscript and consider the $n$ observed $r_i^c = \begin{pmatrix} r_{1,i}^c \\ r_{2,i}^c \end{pmatrix}$. Divide the observations into $M$ intervals of length $k_n$, $M = \lfloor n/K_n \rfloor$. Within interval $j$, $j = 1, \ldots, M$, draw $k_n$ independent $\tilde{r}^c$ by sampling randomly with replacement from $(r_i^c)_{(j-1)k_n+1 \leq i \leq jk_n}$. Compute

$$\tilde{R}_t^\beta = \frac{\sum_{i=1}^{Mk_n} \tilde{r}_{1,i}^c \tilde{r}_{2,i}^c}{\sum_{i=1}^{Mk_n} (r_{1,i}^c)^2}$$  \hspace{1cm} (11)

Repeat 1,000 times yielding $\tilde{R}_1^\beta, \ldots, \tilde{R}_{1000}^\beta$. Compute the 0.025 and 0.975 quantiles which form a 95% confidence interval. Re-do each step over the days $t = 1, 2, \ldots, T$. Some illustrative plots are below.
### 3 Inference for Integrated Variance Functionals

The idea extends in a straightforward way to general integrated variance functionals. Suppose \( X \) is a vector process and we are interested in functions of

\[
V = \begin{bmatrix}
\int_0^1 c_{11,s} \cdots \int_0^1 c_{1p,s} \\
\vdots & \ddots & \vdots \\
\int_0^1 c_{p1,s} \cdots \int_0^1 c_{pp,s}
\end{bmatrix}
\]  

(12)

Say we want to estimate \( h(V) \) for some function \( h \), of which the integrated beta is a special case. We form the vector of truncated (continuous) returns \( r_{ci} = (r_{1i}, \ldots, r_{pi})' \) and take as the estimate of \( V \)

\[
\hat{V} = \sum_{i=1}^n r_{ci}^e r_{ci}^e'
\]

(13)

To get a handle on the sampling fluctuations of \( h(\hat{V}) - h(V) \) we use the bootstrap exactly as above.

Consider the \( n \) observed \( p \times 1 \) vectors \( (r_{ci})_{i=1,\ldots,n} \). Divide the observations into \( M \) intervals of length \( k_n, M = \lfloor n/K \rfloor \). Within interval \( j, j = 1, \ldots, M \), draw \( k_n \) independent \( \tilde{r}_i^e \) by sampling randomly with replacement from \( (r_{ci})_{(j-1)k_n+1 \leq i \leq jk_n} \). Compute

\[
\tilde{V} = \sum_{i=1}^{Mk_n} \tilde{r}_{ci}^e \tilde{r}_{ci}^e', \quad \tilde{h} = h(\tilde{V}).
\]

(14)

Repeat 1,000 times yielding \( \tilde{h}_1, \ldots, \tilde{h}_{1000} \). Compute the 0.025 and 0.975 quantiles of \( \tilde{h}_i \), which form a 95% confidence interval. Re-do each step over the days \( t = 1, 2, \ldots, T \).
Figure 1: 95% Confidence Intervals for Integrated Beta 2007–2014
Figure 2: 95% Confidence Intervals for Integrated Beta

Figure 3: 95% Confidence Intervals for Integrated Beta
References