High Frequency Econometrics: The First Lecture

1. Introduction

We consider a very simple setting to illustrate the econometrics of a few commonly used estimators.

2. A toy continuous semimartingale model

Fix a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\).

**Definition 1.** (Brownian Motion, Protter 2004, Page 17) An adapted process \(B = (B_t)_{0 \leq t < \infty}\) taking values in \(\mathbb{R}^n\) is called an \(n\)-dimensional Brownian motion if

(i) for \(0 \leq s < t < \infty\), \(B_t - B_s\) is independent of \(\mathcal{F}_s\) (increments are independent of the past)

(ii) for \(0 < s < t\), \(B_t - B_s\) is a Gaussian random variable with mean zero and variance matrix \((t - s)C\), for a given, non-random matrix \(C\).

It is called “standard” if \(C\) is the identity matrix.

It turns out that one can always find a “version” of Brownian motion with continuous sample path almost surely (i.e. for \(\mathbb{P}\)-a.s. \(\omega \in \Omega, t \mapsto B_t(\omega)\) is a continuous function). By convention, we suppose that Brownian motion has continuous paths. In most of our discussions, we consider 1-dimensional standard Brownian motion, denoted by \(W\).

To keep the technicality at minimum in this first note, let us consider a very simple stochastic integral w.r.t. \(W\). Let \(T\) be a fixed horizon. Let \(\bar{\sigma}_0\) and \(\bar{\sigma}_{T/2}\) be two positive constants. Consider the following process

\[
\sigma_t = \begin{cases} 
\bar{\sigma}_0 & \text{if } t \in [0, T/2) \\
\bar{\sigma}_{T/2} & \text{if } t \geq T/2.
\end{cases}
\]

We can then define a stochastic integral of \(\sigma_t\) w.r.t. \(W_t\): for \(t \in [0, T]\),

\[
\int_0^t \sigma_s dW_s = \begin{cases} 
\bar{\sigma}_0 \cdot (W_t - W_0) & \text{if } t \in [0, T/2) \\
\bar{\sigma}_0 \cdot (W_{T/2} - W_0) + \bar{\sigma}_{T/2} \cdot (W_t - W_{T/2}) & \text{if } t \geq T/2
\end{cases}
\]

Let \(b_t\) be a bounded cadlag (right-continuous with left limits) adapted process. The following integral is defined path-by-path in the usual Riemann sense:

\[
\int_0^t b_s ds.
\]

Now, we can construct a very simple model for stock process

\[
X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s.
\]
3. Integrated Variance

Denote for each \( i = 1, \ldots, n \), where \( n = T/\Delta_n \) is assumed to be an even integer,

\[
\Delta^n_i X = X_{i\Delta_n} - X_{(i-1)\Delta_n}.
\]

Consider the following estimator for the integrated variance:

\[
\hat{IV}_n = \sum_{i=1}^{T/\Delta_n} (\Delta^n_i X)^2.
\]

We set

\[
IV = \int_0^T \sigma_s^2 ds.
\]

In our simple setup above, \( IV = \bar{\sigma}_0^2 T/2 + \bar{\sigma}_{T/2}^2 T/2 \). Below, we derive the consistency and the asymptotic normality of the estimator \( \hat{IV}_n \).

To simplify notations, let \( \lambda^n_i = \int_{(i-1)\Delta_n}^{i\Delta_n} b_s ds \). Then under our simple model, we have

\[
\Delta^n_i X = \begin{cases} 
\lambda^n_i + \bar{\sigma}_0 \Delta^n_i W & \text{if } 1 \leq i \leq n/2 \\
\lambda^n_i + \bar{\sigma}_{T/2} \Delta^n_i W & \text{if } n/2 + 1 \leq i \leq n 
\end{cases}
\]

Note that

\[
\hat{IV}_n - IV
\]

\[
= \sum_{i=1}^{n/2} (\lambda^n_i + \bar{\sigma}_0 \Delta^n_i W)^2 + \sum_{i=n/2+1}^n (\lambda^n_i + \bar{\sigma}_{T/2} \Delta^n_i W)^2 - \left( \bar{\sigma}_0^2 T/2 + \bar{\sigma}_{T/2}^2 T/2 \right)
\]

\[
= \sum_{i=1}^n (\lambda^n_i)^2 + 2 \sum_{i=1}^n \sigma(i-1)\Delta_n \lambda^n_i \Delta^n_i W + \sum_{i=1}^{n/2} (\bar{\sigma}_0 \Delta^n_i W)^2 + \sum_{i=n/2+1}^n (\bar{\sigma}_{T/2} \Delta^n_i W)^2
\]

\[
= \sum_{i=1}^n (\lambda^n_i)^2 + 2 \sum_{i=1}^n \sigma(i-1)\Delta_n \lambda^n_i \Delta^n_i W + \sum_{i=1}^{n/2} \bar{\sigma}_0^2 \left( (\Delta^n_i W)^2 - \Delta_n \right) + \sum_{i=n/2+1}^n \bar{\sigma}_{T/2}^2 \left( (\Delta^n_i W)^2 - \Delta_n \right)
\]

where

\[
\zeta^n_i = \begin{cases} 
\bar{\sigma}_0^2 \left( (\Delta^n_i W)^2 - \Delta_n \right) & \text{if } 1 \leq i \leq n/2 \\
\bar{\sigma}_{T/2}^2 \left( (\Delta^n_i W)^2 - \Delta_n \right) & \text{if } n/2 + 1 \leq i \leq n.
\end{cases}
\]
Note that $\zeta^n_i$ forms an array of martingale differences satisfying
\[
\sum_{i=1}^n \mathbb{E} \left[ \left( \Delta_n^{-1/2} \zeta^n_i \right)^2 \right] = \Delta_n \sum_{i=1}^n 2 \sigma^4_{(i-1)\Delta_n} \rightarrow 2 \int_0^T \sigma^4_s ds
\]
\[
\sum_{i=1}^n \mathbb{E} \left[ \left( \Delta_n^{-1/2} \zeta^n_i \right)^4 \right] \rightarrow 0.
\]
Then by a CLT for MD arrays, we have
\[
(0.2) \quad \Delta_n^{-1/2} \sum_{i=1}^n \zeta^n_i \overset{d}{\rightarrow} \mathcal{N} \left(0, 2 \int_0^T \sigma^4_s ds\right).
\]
Since the process $b$ is assumed to be bounded, $|\lambda^n_i| \leq K \Delta_n$. Hence
\[
(0.3) \quad \Delta_n^{-1/2} \sum_{i=1}^n (\lambda^n_i)^2 \leq K \Delta_n^{1/2} \rightarrow 0.
\]
By the triangle inequality and the Cauchy-Schwarz inequality, we then have
\[
(0.4) \quad \mathbb{E} \left| \sum_{i=1}^n \sigma_{(i-1)\Delta_n} \lambda^n_i \Delta^n_{i} W \right| \leq K \sum_{i=1}^n \mathbb{E} |\lambda^n_i \Delta^n_{i} W| \leq K \Delta_n^{1/2} \rightarrow 0.
\]
By (0.2) and (0.3), it is easy to see that the first and the third term on the right-hand side of (0.1) are both $o_p(1)$. Moreover, (0.4) shows that the second term on the right-hand side of (0.1) is $o_p(1)$. Hence, we have
\[
\hat{IV}_n \overset{p}{\rightarrow} IV.
\]
We now show that $\Delta_n^{-1/2} (\hat{IV}_n - IV) \overset{d}{\rightarrow} \mathcal{N}(0, 2 \int_0^T \sigma^4_s ds)$. In view of (0.1), (0.2) and (0.3), we are almost done if we can show that $\sum_{i=1}^n \sigma_{(i-1)\Delta_n} \lambda^n_i \Delta^n_{i} W = o_p(\Delta_n^{1/2})$. But the bound (0.4) is not good enough for this purpose. We need to do some computation more carefully.

Consider the following decomposition:
\[
\lambda^n_i = b_{(i-1)\Delta_n} \Delta_n + \lambda'^n_i, \quad \text{where} \quad \lambda'^n_i = \int_{(i-1)\Delta_n}^{i\Delta_n} (b_s - b_{(i-1)\Delta_n}) ds.
\]
Hence,
\[
(0.5) \quad \Delta_n^{-1/2} \sum_{i=1}^n \sigma_{(i-1)\Delta_n} \lambda^n_i \Delta^n_{i} W = \Delta_n^{-1/2} \sum_{i=1}^n \sigma_{(i-1)\Delta_n} b_{(i-1)\Delta_n} \Delta_n \Delta^n_{i} W + \Delta_n^{-1/2} \sum_{i=1}^n \sigma_{(i-1)\Delta_n} \lambda'^n_i \Delta^n_{i} W.
\]
Note that
\[
(0.6) \quad \text{the first term on the right-hand side of the above display is } o_p(1),
\]
because
\[ \mathbb{E} \left| \Delta_n^{-1/2} \sum_{i=1}^{n} \sigma_{(i-1)\Delta_n} b_{(i-1)\Delta_n} \Delta_n^{n} W_i \right|^2 = \Delta_n \sum_{i=1}^{n} \mathbb{E} \left| \sigma_{(i-1)\Delta_n} b_{(i-1)\Delta_n} \Delta_n^{n} W_i \right|^2 \leq K \Delta_n \to 0. \]

Now, let us impose some smoothness on the drift coefficient \( b_i \) by assuming that for some (arbitrarily small) \( r > 0 \), \( \mathbb{E} |b_u - b_v|^r \leq K |u - v|^r \) for any \(|u - v| \leq 1\). Then, we have
\[ \mathbb{E} |\lambda_i^n|^2 = \mathbb{E} \int_{(i-1)\Delta_n}^{i\Delta_n} (b_s - b_{(i-1)\Delta_n}) ds \leq K \Delta_n \int_{(i-1)\Delta_n}^{i\Delta_n} \mathbb{E} \left[ (b_s - b_{(i-1)\Delta_n})^2 \right] ds \leq K \Delta_n^{2+r}. \]

Then by the triangle inequality and the Cauchy-Schwarz inequality,
\[
\mathbb{E} \left| \Delta_n^{-1/2} \sum_{i=1}^{n} \sigma_{(i-1)\Delta_n} \lambda_i^n \Delta_n^{n} W_i \right| \leq \sum_{i=1}^{n} \mathbb{E} \left| \sigma_{(i-1)\Delta_n} \lambda_i^n \Delta_n^{n} W_i \right| \leq K \sum_{i=1}^{n} \left( \mathbb{E} |\lambda_i^n|^2 \right)^{1/2} \left( \mathbb{E} \left| \Delta_n^{n} W_i \right|^2 \right)^{1/2} \leq K \Delta_n^{r/2} \to 0.
\]

(0.7)

Combining (0.5)-(0.7), we have \( \Delta_n^{-1/2} \sum_{i=1}^{n} \sigma_{(i-1)\Delta_n} \lambda_i^n \Delta_n^{n} W_i = O_p(1) \), as we wanted. Hence,
\[
\Delta_n^{-1/2} \left( \hat{I}_V - IV \right) \overset{d}{\to} \mathcal{N} \left( 0, 2 \int_0^T \sigma_s^2 ds \right).
\]

4. The Realized Laplace Transform

The IV estimator can be casted in the form of \( \hat{I}_V = \Delta_n \sum_{i=1}^{T/\Delta_n} f(\Delta_{i}^n X/\Delta_n^{1/2}) \) with the function \( f(x) = x^2 \). Let us now consider the other function \( f(x) = \cos (\sqrt{2u} x) \) for some \( u > 0 \). Define
\[
\hat{L}_n(u) = \Delta_n \sum_{i=1}^{T/\Delta_n} \cos \left( \frac{\sqrt{2u} \Delta_{i}^n X}{\Delta_n^{1/2}} \right)
= \Delta_n \sum_{i=1}^{T/\Delta_n} \left( \cos \left( \frac{\sqrt{2u} \Delta_{i}^n X}{\Delta_n^{1/2}} \right) - \cos \left( \frac{\sqrt{2u} \sigma_{(i-1)\Delta_n} \Delta_n^{n} W_i}{\Delta_n^{1/2}} \right) \right)
+ \Delta_n \sum_{i=1}^{T/\Delta_n} \left( \cos \left( \frac{\sqrt{2u} \sigma_{(i-1)\Delta_n} \Delta_n^{n} W_i}{\Delta_n^{1/2}} \right) - \mathbb{E} \left[ \cos \left( \frac{\sqrt{2u} \sigma_{(i-1)\Delta_n} \Delta_n^{n} W_i}{\Delta_n^{1/2}} \right) \right] \mathcal{F}_{(i-1)\Delta_n} \right)
= \Delta_n \sum_{i=1}^{T/\Delta_n} \mathbb{E} \left[ \cos \left( \frac{\sqrt{2u} \sigma_{(i-1)\Delta_n} \Delta_n^{n} W_i}{\Delta_n^{1/2}} \right) \right] \mathcal{F}_{(i-1)\Delta_n}.
\]

(0.8)

Note that \( \Delta_{i}^n W_i / \Delta_n^{1/2} \sim \mathcal{N}(0, 1) \) and for any \( a \in \mathbb{R} \), \( \mathbb{E} \left[ \cos \left( a \cdot \mathcal{N}(0, 1) \right) \right] = \exp \left( -a^2/2 \right) \). Hence
\[
\mathbb{E} \left[ \cos \left( \frac{\sqrt{2u} \sigma_{(i-1)\Delta_n} \Delta_n^{n} W_i}{\Delta_n^{1/2}} \right) \right] \mathcal{F}_{(i-1)\Delta_n} = \exp \left( -u \sigma_{(i-1)\Delta_n}^2 \right).
\]
In our simple model, the volatility process is assumed piecewise constant. It is then easy to see that
\[
\Delta_n \sum_{i=1}^{T/\Delta_n} \mathbb{E} \left[ \cos \left( \frac{\sqrt{2u} \sigma_{(i-1)\Delta_n} \Delta_i^W}{\Delta_{n}^{1/2}} \right) \bigg| \mathcal{F}_{(i-1)\Delta_n} \right] = \Delta_n \sum_{i=1}^{T/\Delta_n} \exp \left( -u \sigma_{(i-1)\Delta_n}^2 \right)
\]
(0.9)

Now, note that the second term on the right-hand side of (0.8) is a sum of MDS by construction. To simplify notations, denote
\[
\zeta_i^n = \Delta_n^{1/2} \left( \cos \left( \frac{\sqrt{2u} \sigma_{(i-1)\Delta_n} \Delta_i^W}{\Delta_{n}^{1/2}} \right) \right) - \mathbb{E} \left[ \cos \left( \frac{\sqrt{2u} \sigma_{(i-1)\Delta_n} \Delta_i^W}{\Delta_{n}^{1/2}} \right) \bigg| \mathcal{F}_{(i-1)\Delta_n} \right].
\]

As said, \( \mathbb{E} [\zeta_i^n | \mathcal{F}_{(i-1)\Delta_n}] = 0 \). Moreover,
\[
\sum_{i=1}^{T/\Delta_n} \mathbb{E} \left[ (\zeta_i^n)^2 \bigg| \mathcal{F}_{(i-1)\Delta_n} \right] = \Delta_n \sum_{i=1}^{T/\Delta_n} \mathbb{E} \left[ \left( \cos \left( \frac{\sqrt{2u} \sigma_{(i-1)\Delta_n} \Delta_i^W}{\Delta_{n}^{1/2}} \right) \right)^2 \bigg| \mathcal{F}_{(i-1)\Delta_n} \right]
\]
\[
- \Delta_n \sum_{i=1}^{T/\Delta_n} \left( \mathbb{E} \left[ \cos \left( \frac{\sqrt{2u} \sigma_{(i-1)\Delta_n} \Delta_i^W}{\Delta_{n}^{1/2}} \right) \bigg| \mathcal{F}_{(i-1)\Delta_n} \right] \right)^2
\]
\[
= \int_0^T \frac{1}{2} \left( 1 + e^{-4u^2} - e^{-2u^2} \right) ds.
\]

It is easy to see that
\[
\sum_{i=1}^{T/\Delta_n} \mathbb{E} \left[ (\zeta_i^n)^4 \bigg| \mathcal{F}_{(i-1)\Delta_n} \right] \rightarrow 0.
\]

Hence, by a CLT for MDS, we have
\[
\Delta_n^{1/2} \sum_{i=1}^{T/\Delta_n} \left( \cos \left( \frac{\sqrt{2u} \sigma_{(i-1)\Delta_n} \Delta_i^W}{\Delta_{n}^{1/2}} \right) - \mathbb{E} \left[ \cos \left( \frac{\sqrt{2u} \sigma_{(i-1)\Delta_n} \Delta_i^W}{\Delta_{n}^{1/2}} \right) \bigg| \mathcal{F}_{(i-1)\Delta_n} \right] \right)
\]
(0.10)
\[
\overset{d}{\rightarrow} \mathcal{N} \left( 0, \int_0^T \frac{1}{2} \left( 1 + e^{-4u^2} - e^{-2u^2} \right) ds \right).
\]
It remains to analyze the first term on the right-hand side of (0.8). We have (recall that $\Delta_n^n X = \lambda_i^n + \sigma_{(i-1)\Delta_n^n}^n W$)

$$
\mathbb{E} \left[ \Delta_n^n \sum_{i=1}^{T/\Delta_n} \left( \cos \left( \frac{\sqrt{2u\Delta_n^n X}}{\Delta_n^{1/2}} \right) - \cos \left( \frac{\sqrt{2u\sigma_{(i-1)\Delta_n^n}^n W}}{\Delta_n^{1/2}} \right) \right) \right] \\
\leq \Delta_n^n \sum_{i=1}^{T/\Delta_n} \mathbb{E} \left| \cos \left( \frac{\sqrt{2u\Delta_n^n X}}{\Delta_n^{1/2}} \right) - \cos \left( \frac{\sqrt{2u\sigma_{(i-1)\Delta_n^n}^n W}}{\Delta_n^{1/2}} \right) \right| \\
\leq K \Delta_n^n \sum_{i=1}^{T/\Delta_n} \mathbb{E} \left| \lambda_i^n / \Delta_n^{1/2} \right| \\
(0.11) \leq K \Delta_n^{1/2}.
$$

Hence, the first term on the right-hand side of (0.8) is $o_p(1)$. In view of (0.8), (0.9) and (0.10), we have

$$
\hat{L}_n(u) \xrightarrow{p} \int_0^T \exp \left(-u\sigma_1^2 \right) ds.
$$

Again, we see that the bound (0.11) is not “fine” enough for getting a CLT. Below, we want to show that

$$
\Delta_n^{1/2} \sum_{i=1}^{T/\Delta_n} \left( \cos \left( \frac{\sqrt{2u\Delta_n^n X}}{\Delta_n^{1/2}} \right) - \cos \left( \frac{\sqrt{2u\sigma_{(i-1)\Delta_n^n}^n W}}{\Delta_n^{1/2}} \right) \right) = o_p(1).
$$

Note that

$$
\cos \left( \frac{\sqrt{2u\Delta_n^n X}}{\Delta_n^{1/2}} \right) - \cos \left( \frac{\sqrt{2u\sigma_{(i-1)\Delta_n^n}^n W}}{\Delta_n^{1/2}} \right) \\
= -\sin \left( \frac{\sqrt{2u\sigma_{(i-1)\Delta_n^n}^n W}}{\Delta_n^{1/2}} \right) \sqrt{2u\lambda_i^n / \Delta_n^{1/2}} - \frac{1}{2} \cos \left( \xi_i^n \right) \left( \sqrt{2u\lambda_i^n / \Delta_n^{1/2}} \right)^2,
$$

where $\xi_i^n$ is a mean-value in the second-order Taylor expansion above. Note that

$$
\mathbb{E} \left[ \Delta_n^{1/2} \sum_{i=1}^{T/\Delta_n} \frac{1}{2} \cos \left( \xi_i^n \right) \left( \sqrt{2u\lambda_i^n / \Delta_n^{1/2}} \right)^2 \right] \leq K \Delta_n^{1/2} \sum_{i=1}^{T/\Delta_n} \mathbb{E} \left| \lambda_i^n / \Delta_n^{1/2} \right|^2 \leq K \Delta_n^{1/2} \rightarrow 0.
$$
On the other hand,

\[
\Delta_n^{-1/2} \sum_{i=1}^{T/\Delta_n} \sin \left( \frac{\sqrt{2}u \sigma_i \Delta_i W}{\Delta_n^{1/2}} \right) \sqrt{2u} \frac{1}{\Delta_n^{1/2}} \int_{(i-1)\Delta_n}^{i\Delta_n} b_s ds
\]

\[
= \sqrt{2u} \Delta_n^{1/2} \sum_{i=1}^{T/\Delta_n} \sin \left( \frac{\sqrt{2}u \sigma_i \Delta_i W}{\Delta_n^{1/2}} \right) b_{(i-1)\Delta_n}
\]

\[
+ \sum_{i=1}^{T/\Delta_n} \sin \left( \frac{\sqrt{2}u \sigma_i \Delta_i W}{\Delta_n^{1/2}} \right) \sqrt{2u} \int_{(i-1)\Delta_n}^{i\Delta_n} (b_s - b_{(i-1)\Delta_n}) ds.
\]

Observing that the first term above is a sum of MDS, we easily see that the first term on the right-hand side is \(O_p(\Delta_n^{1/2})\). Using the smoothness of the process \(b_t\), it is easily seen that the second term on the right-hand side is \(o_p(1)\). We hence have (0.12). Therefore,

\[
\Delta_n^{-1/2} \left( \hat{I}_n(u) - \int_0^T \exp(-u\sigma_s^2) ds \right) \xrightarrow{d} N \left( 0, \int_0^T \frac{1}{2} \left( 1 + e^{-4\sigma_s^2 u} - e^{-2u\sigma_s^2} \right) ds \right).
\]

**Problem 1.** Following similar argument as in the integrated variance and the realized Laplace transform examples, derive the asymptotics for \(IQ_n = \Delta_n \sum_{i=1}^{T/\Delta_n} \left( \Delta_i^\sigma X/\Delta_n^{1/2} \right)^4 \).

**5. A toy discontinuous semimartingale model**

Consider a standard Brownian motion \(W\). Let \(\sigma > 0\) be a constant. Now, let us add some jumps into the model. Let \(\tau_k \geq 1\) be a sequence random (stopping) times, strictly increasing in time, independent of \(W\), such that \(E \left[ \sum_{k \geq 1} 1_{\{\tau_k \leq T\}} \right] < \infty\). Let \((S_k)_{k \geq 1}\) be a sequence of bounded random variables (such that \(S_k \in \mathcal{F}_{\tau_k^-}\)) also independent of \(W\). Define a pure-jump process by

\[
J_t = \sum_{k \geq 1} S_k 1_{[\tau_k, \infty)}(t).
\]

Then define

\[
X_t = \sigma W_t + J_t.
\]

**6. Realized Variance**

Let us consider the behavior of the IV estimator, but now let us rename it as the RV estimator (Realized Variance). We can assume that \(|\tau_k - \tau_{k'}| > \Delta_n\) for any \(k \neq k'\) without loss. Let \(I^n_k\) be the unique integer \(i\) such that \(\{(i - 1)\Delta_n, i\Delta_n\} \) contains \(\tau_k\); note that \(I^n_k\) is random. Let \(I^n = \{I^n_k : k \geq 1\}\). Recall that

\[
RV_n = \sum_{i=1}^{T/\Delta_n} (\Delta_i^\sigma X)^2.
\]
We decompose the estimator as follows:

\[ RV_n = \sum_{i=1}^{T/\Delta_n} (\sigma \Delta_i^n W)^2 + \sum_{i=1, i \in \mathcal{I}_n} \left( (\Delta_i^n X)^2 - (\sigma \Delta_i^n W)^2 \right) \]

\[ = \sum_{i=1}^{T/\Delta_n} (\sigma \Delta_i^n W)^2 + \sum_{k \geq 1, \tau_k \leq T} \left( (S_k + \sigma \Delta_{T_k}^n W)^2 - (\sigma \Delta_{T_k}^n W)^2 \right) \]

\[ = \sum_{i=1}^{T/\Delta_n} (\sigma \Delta_i^n W)^2 + \sum_{k \geq 1, \tau_k \leq T} S_k^2 + 2S_k \sigma \Delta_{T_k}^n W \]

\[ = \sum_{i=1}^{T/\Delta_n} (\sigma \Delta_i^n W)^2 + \sum_{k \geq 1, \tau_k \leq T} S_k^2 + \sum_{k \geq 1, \tau_k \leq T} 2S_k \sigma \Delta_{T_k}^n W. \]

Note that \( \sum_{i=1}^{T/\Delta_n} (\sigma \Delta_i^n W)^2 \overset{p}{\to} \sigma^2 T, \sum_{k \geq 1, \tau_k \leq T} S_k^2 = \sum_{s \leq T} (\Delta X_s)^2 \). Moreover,

\[ \mathbb{E} \left| \sum_{k \geq 1, \tau_k \leq T} 2S_k \sigma \Delta_{T_k}^n W \right| \leq \sum_{k \geq 1, \tau_k \leq T} \mathbb{E} \left| \Delta_{T_k}^n W \right| \leq K \Delta_n^{1/2} \to 0. \]

Hence,

\[ RV_n \overset{p}{\to} \sigma^2 T + \sum_{s \leq T} (\Delta X_s)^2 = \left( \int_0^T \sigma_s^2 ds + \sum_{s \leq T} (\Delta X_s)^2 \right) \quad \text{quadratic variation of } X \]

Now, let us consider the CLT. Note that we have made a (quite strong) simplifying assumption that \( S_k \) and \( \tau_k \) are independent of \( W \). Under this assumption, if we condition on \((S_k, \tau_k), W\) is still a Brownian motion, and we can treat \((S_k, \tau_k)\) as constants. Note that

\[ \Delta_n^{-1/2} \left( RV_n - \left( \int_0^T \sigma_s^2 ds + \sum_{s \leq T} (\Delta X_s)^2 \right) \right) \]

\[ = \Delta_n^{-1/2} \left( \sum_{i=1}^{T/\Delta_n} \sigma_i^2 \left( (\Delta_i^n W)^2 - \Delta_n \right) + \sum_{k \geq 1, \tau_k \leq T} 2S_k \sigma \Delta_{T_k}^n W \right) \]

\[ = \Delta_n^{-1/2} \left( \sum_{i=1, i \in \mathcal{I}_n} \sigma_i^2 \left( (\Delta_i^n W)^2 - \Delta_n \right) + \sum_{i=1, i \in \mathcal{I}_n} \sigma_i^2 \left( (\Delta_i^n W)^2 - \Delta_n \right) + \sum_{k \geq 1, \tau_k \leq T} 2S_k \sigma \Delta_{T_k}^n W \right). \]

Note that

\[ \mathbb{E} \left| \Delta_n^{-1/2} \sum_{i=1, i \in \mathcal{I}_n} \sigma_i^2 \left( (\Delta_i^n W)^2 - \Delta_n \right) \right| \leq K \Delta_n^{1/2} \to 0, \]

\[ \Delta_n^{-1/2} \sum_{i=1, i \in \mathcal{I}_n} \sigma_i^2 \left( (\Delta_i^n W)^2 - \Delta_n \right) = o_p(1). \]
Also note that
\[
\Delta_n^{-1/2} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} \sigma^2 \left( (\Delta_i^n W)^2 - \Delta_n \right) \xrightarrow{d} \mathcal{N} \left( 0, 2 \int_0^T \sigma_s^4 ds \right),
\]
\[
\Delta_n^{-1/2} \sum_{k \geq 1, \tau_k \leq T} 2S_k \sigma \Delta_t^n W \sim \mathcal{N} \left( 0, 4 \sum_{k \geq 1, \tau_k \leq T} S_k^2 \sigma^2 \right) = \mathcal{N} \left( 0, 4 \sum_{s \leq T} \sigma_s^2 \Delta X_s^2 \right).
\]
Furthermore, note that the Brownian increments involved in the two lines above do not overlap. Hence, the convergence above hold jointly.

Combining the above, we have
\[
\Delta_n^{-1/2} \left( RV_n - \left( \int_0^T \sigma_s^2 ds + \sum_{s \leq T} (\Delta X_s)^2 \right) \right) \xrightarrow{d} \mathcal{N} \left( 0, 2 \int_0^T \sigma_s^4 ds + 4 \sum_{s \leq T} \sigma_s^2 \Delta X_s^2 \right).
\]

### 7. Bipower Variation

We now consider the following estimator, the so-call bipower estimator:

\[
BV_n = \frac{1}{m_1^2} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor-1} |\Delta_i^n X| |\Delta_{i+1}^n X|,
\]

where \( m_1 = \mathbb{E}[U], \ U \sim \mathcal{N}(0, 1). \) We shall show that \( BV_n \xrightarrow{p} \sigma^2 T = \int_0^T \sigma_s^2 ds \) even though there are jumps. Moreover, we derive the associated rate of convergence.

Let \( \beta_i^n = \sigma \Delta_t^n W \). We define

\[
\xi_n = \frac{1}{m_1^2} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor-1} |\beta_i^n| |\beta_{i+1}^n|.
\]

Note that \( |\beta_i^n| |\beta_{i+1}^n| \) is 1-dependent with

\[
\mathbb{E} |\beta_i^n| |\beta_{i+1}^n| = \sigma^2 m_1^2,
\]

\[
\text{Var} |\beta_i^n| |\beta_{i+1}^n| = \mathbb{E} |\beta_i^n|^2 |\beta_{i+1}^n|^2 - \sigma^4 m_1^4 = \sigma^4 (1 - m_1^4)
\]

\[
\text{Cov} |\beta_i^n| \beta_{i+1}^n, |\beta_{i+1}^n| \beta_{i+2}^n| = \mathbb{E} |\beta_i^n|^2 |\beta_{i+1}^n|^2 - m_1^4 = \sigma^4 (m_1^2 - m_1^4).
\]

Hence,
\[
\Delta_n^{-1/2} \left( \frac{1}{m_1^2} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor-1} |\beta_i^n| |\beta_{i+1}^n| - \sigma^2 T \right) \xrightarrow{d} \mathcal{N} \left( 0, \frac{\sigma^4 T \cdot (1 + 2m_1^2 - 3m_1^4)}{m_1^4} \right).
\]

The above gives the CLT if there are no jump.

We now show that

\[
\Delta_n^{-1/2} (BV_n - \xi_n) = O_p(1).
\]

\(^1\)Let \( m_r = \mathbb{E}|U|^r, \ U \sim \mathcal{N}(0, 1). \) We have \( m_r = \pi^{-1/2} 2^{r/2} \Gamma((r + 1)/2). \) In particular, \( m_1 = \sqrt{2}/\pi. \)
Note that under our simple model, \( \Delta_i^n X \neq \beta_i^n \) only if \( \Delta_i^n X \) contains a jump. Hence,

\[
\mathbb{E} \left| \Delta_n^{-1/2} \sum_{i=1}^{[T/\Delta_n]-1} (|\Delta_i^n X| |\Delta_{i+1}^n X| - |\beta_i^n| |\beta_{i+1}^n|) \right| \\
\leq \mathbb{E} \left| \Delta_n^{-1/2} \sum_{i \in \mathcal{I}_n} (|\Delta_i^n X| |\beta_{i+1}^n| - |\beta_i^n| |\beta_{i+1}^n|) \right| + \mathbb{E} \left| \Delta_n^{-1/2} \sum_{i+1 \in \mathcal{I}_n} (|\beta_i^n| |\Delta_{i+1}^n X| - |\beta_i^n| |\beta_{i+1}^n|) \right| \\
\leq K \mathbb{E} \sum_{k \geq 1: \tau_k \leq T} |S_k| \leq K.
\]

We are done. The CLT in the jump case does not exist, see Jacod and Protter (2012) page 313.

### 8. Multipower variation

Let us now consider the following estimator:

\[
MV_n = \frac{1}{m_{2/3}^{3/2}} \sum_{i=1}^{[T/\Delta_n]-2} |\Delta_i^n X|^{2/3} |\Delta_{i+1}^n X|^{2/3} |\Delta_{i+2}^n X|^{2/3} ;
\]

recall that \( m_{2/3} = \mathbb{E}|U|^{2/3} \) for \( U \sim N(0, 1) \). Similarly as in the bipower case, it is not hard to show that

\[ MV_n \overset{p}{\to} \int_0^T \sigma_2^2 ds. \]

It is also not hard to derive a CLT for the variables

\[
\Delta_n^{-1/2} \left( \frac{1}{m_{2/3}^{3/2}} \sum_{i=1}^{[T/\Delta_n]-2} |\beta_i^{2/3}| |\beta_{i+1}^{2/3}| |\beta_{i+2}^{2/3}| - \int_0^T \sigma_2^2 ds \right).
\]

The focus here is whether \( MV_n \) also has a CLT. Unlike the bipower case, the answer is yes.

It is clear that it suffices to show the following:

\[
(0.13) \quad \Delta_n^{-1/2} \sum_{i=1}^{[T/\Delta_n]-2} \left( |\Delta_i^n X|^{2/3} |\Delta_{i+1}^n X|^{2/3} |\Delta_{i+2}^n X|^{2/3} - |\beta_i^{2/3}| |\beta_{i+1}^{2/3}| |\beta_{i+2}^{2/3}| \right) = o_p(1).
\]

Consider the event \( A_n = \{ \omega \in \Omega : \text{for any } s \in [0, T], \text{there is at most one jump in the interval } [s, (s + 3\Delta_n) \wedge T] \} \). Under our simple model, there are only finitely many jumps on \([0, T] \). Since \( \Delta_n \to 0 \), it is easy to see that \( A_n \) is w.p.a.1. Recall a lemma encountered earlier in this course, in order to show \((0.13)\), we can restrict our calculation on \( A_n \).

Before moving forward, consider a simple deterministic inequality:

**Lemma 1.** Let \( r \in (0, 1] \). Then for any \( x, y \in \mathbb{R}, |x + y|^r - |y|^r \leq |x|^r \).

**Proof.** By the Cr-inequality,

\[
|x + y|^r \leq |x|^r + |y|^r \quad |y|^r = |(x + y) - x|^r \leq |x + y|^r + |x|^r.
\]
So, \(-|x|^r \leq |x+y|^r - |y|^r \leq |x|^r\). 

Now, consider

\[
\left| \Delta_n^{-1/2} \sum_{i \in I_n} \left( (\Delta_i^n X)^{2/3} |\beta_{i+1}^n|^{2/3} |\beta_{i+2}^n|^{2/3} - |\beta_i^n|^{2/3} |\beta_{i+1}^n|^{2/3} |\beta_{i+2}^n|^{2/3} \right) \right|
\]

\[
\leq \Delta_n^{-1/2} \sum_{k \geq 1: \tau_k \leq T} \left( (\Delta_k^n X)^{2/3} - |\beta_k^n|^{2/3} \right) |\beta_{i+1}^n|^{2/3} |\beta_{i+2}^n|^{2/3}
\]

\[
= O_p \left( \Delta_n^{2/3-1/2} \right) = o_p(1).
\]

Similarly,

\[
\Delta_n^{-1/2} \sum_{i+1 \in I_n} \left( |\beta_i^n|^{2/3} |\Delta_{i+1} X|^{2/3} |\beta_{i+2}^n|^{2/3} - |\beta_i^n|^{2/3} |\beta_{i+1}^n|^{2/3} |\beta_{i+2}^n|^{2/3} \right) = o_p(1)
\]

\[
\Delta_n^{-1/2} \sum_{i+1 \in I_n} \left( |\beta_i^n|^{2/3} |\beta_{i+1}^n|^{2/3} |\Delta_{i+2} X|^{2/3} - |\beta_i^n|^{2/3} |\beta_{i+1}^n|^{2/3} |\beta_{i+2}^n|^{2/3} \right) = o_p(1)
\]

Combining the estimates in the above two displays, we have (0.13) as wanted.

9. Truncation

We keep working with the jump model. Consider the following “truncation-based” estimator:

\[
\sum_{i=1}^{[T/\Delta_n]} (\Delta_i^n X)^2 1\{|\Delta_i^n X| \leq \alpha \Delta_i^n\},
\]

where

\(\alpha > 0\) and \(\varpi \in (0,1/2)\) are constants.

The asymptotics for \(\sum_{i=1}^{[T/\Delta_n]} (\beta_i^n)^2\) is simple. Our interest is the “elimination error”:

\[
\sum_{i=1}^{[T/\Delta_n]} \left( (\Delta_i^n X)^2 1\{|\Delta_i^n X| \leq \alpha \Delta_i^n\} - (\beta_i^n)^2 \right)
\]

\[
= \sum_{i \in I_n} \left( (\Delta_i^n X)^2 1\{|\Delta_i^n X| \leq \alpha \Delta_i^n\} - (\beta_i^n)^2 \right) + \sum_{i \in I_n} \left( (\beta_i^n)^2 1\{|\beta_i^n| \leq \alpha \Delta_i^n\} - (\beta_i^n)^2 \right).
\]

11
Let $G$ be the information set generated by the jumps (time and size). Under our simple assumption that jumps are independent of the Brownian motion,

$$
E \left( \sum_{i \in \mathcal{I}_n} \left( (\Delta^n_i X)^2 1_{\{|\Delta^n_i X| \leq \alpha \Delta^n_i \}} - (\beta^n_i)^2 \right) \right) \\
\leq \sum_{i \in \mathcal{I}_n} E \left[ (\Delta^n_i X)^2 1_{\{|\Delta^n_i X| \leq \alpha \Delta^n_i \}} - (\beta^n_i)^2 \mid G \right] \\
\leq K \Delta_n^{2\omega}.
$$

So,

$$
\sum_{i \in \mathcal{I}_n} \left( (\Delta^n_i X)^2 1_{\{|\Delta^n_i X| \leq \alpha \Delta^n_i \}} - (\beta^n_i)^2 \right) = O_p \left( \Delta_n^{2\omega} \right).
$$

On the other hand, for any $\delta > 0$,

$$
E \left[ (\beta^n_i)^2 1_{\{|\beta^n_i| \leq \alpha \Delta^n_i \}} - (\beta^n_i)^2 \right] \\
= E \left[ (\beta^n_i)^2 1_{\{|\beta^n_i| > \alpha \Delta^n_i \}} \right] \\
\leq K \frac{E[|\beta^n_i|^{2+\delta}]}{\Delta^n_i^{\delta}} \leq K \Delta_n^{1+\delta/2-\omega \delta} = K \Delta_n^{1+(1/2-\omega)\delta}.
$$

Hence, for any $\delta > 0$,

$$
\sum_{i \in \mathcal{I}_n} \left( (\beta^n_i)^2 1_{\{|\beta^n_i| \leq \alpha \Delta^n_i \}} - (\beta^n_i)^2 \right) = O_p \left( \Delta_n^{1/2-\omega}) \right).
$$

Hence,

$$
\sum_{i=1}^{[T/\Delta_n]} \left( (\Delta^n_i X)^2 1_{\{|\Delta^n_i X| \leq \alpha \Delta^n_i \}} - (\beta^n_i)^2 \right) = O_p \left( \Delta_n^{2\omega} + \Delta_n^{(1/2-\omega)\delta} \right).
$$

It is then clear that

$$
\sum_{i=1}^{[T/\Delta_n]} \left( (\Delta^n_i X)^2 1_{\{|\Delta^n_i X| \leq \alpha \Delta^n_i \}} - (\beta^n_i)^2 \right) = o_p \left( 1 \right).
$$

Hence, the jumps are “negligible” under truncation for LLN.

Also,

$$
\Delta_n^{-1/2} \sum_{i=1}^{[T/\Delta_n]} \left( (\Delta^n_i X)^2 1_{\{|\Delta^n_i X| \leq \alpha \Delta^n_i \}} - (\beta^n_i)^2 \right) = O_p \left( \Delta_n^{2(\omega-1/4)} + \Delta_n^{(1/2-\omega)\delta-1/2} \right).
$$

Taking $\omega \in (1/4, 1/2)$ and $\delta$ large enough, we have

$$
\Delta_n^{-1/2} \sum_{i=1}^{[T/\Delta_n]} \left( (\Delta^n_i X)^2 1_{\{|\Delta^n_i X| \leq \alpha \Delta^n_i \}} - (\beta^n_i)^2 \right) = o_p \left( 1 \right).
$$

Hence, the jumps are “negligible” under truncation also for the CLT.