Central Limit Theorem for the Realized Volatility based on Tick Time Sampling

Masaaki Fukasawa

University of Tokyo

An outline of this talk is as follows.

- What is the Realized Volatility (RV) ?

First, we recall the definition of the realized volatility,

An outline of this talk is as follows.

- What is the Realized Volatility (RV) ?
- Known facts on the RV.

and then, I explain several known facts on the RV.

An outline of this talk is as follows.

- What is the Realized Volatility (RV) ?
- Known facts on the RV.
- What is the Tick Time Sampling (TTS) ?

Next, I introduce a specific sampling scheme, named TTS.

An outline of this talk is as follows.

- What is the Realized Volatility (RV) ?
- Known facts on the RV.
- What is the Tick Time Sampling (TTS) ?
- Robustness to market microstructure noise.

The RV based on the TTS is robust to microstructure noise.

An outline of this talk is as follows.

- What is the Realized Volatility (RV) ?
- Known facts on the RV.
- What is the Tick Time Sampling (TTS) ?
- Robustness to market microstructure noise.
- Main result - asymptotic mixed normality.

Finally, I state the main result of this talk.

What is the RV?

Let $\{\tau_j\}$ be a partition of $[0, T]$ with

$$0 = \tau_0 < \tau_1 < \cdots < \tau_n < \tau_{n+1} = T.$$

Define the RV as the sum of squared log-returns, i.e.,

$$\text{RV} = \sum_{j=1}^{n-1} \left| \log(S_{\tau_{j+1}}) - \log(S_{\tau_j}) \right|^2 .$$

It is well known that if $X_t \colon = \log(S_t)$ satisfies

$$dX_t = \mu(t, X)dt + \sigma(t, X)dW_t$$

and $\sup_j |\tau_{j+1} - \tau_j| \to 0$, it holds that

$$\text{RV} \to \text{IV} \colon = \int_0^T \sigma(t, X)^2 dt.$$

What is the RV?

- Since the IV relates the risk of the corresponding stock, we want to estimate it by historical data.
- Recently, the RV has attracted researchers because of the availability of high frequency data .
- Theoretically, the higher the sampling frequency, the more accurate the RV as an estimator of the IV.
- The consistency holds for not only deterministic but random sampling schemes (partitions).

An example of high frequency data: FX rates

EUR/JPY 10 Sep 2007

DATE	TIME	VOLUME	OPEN	CLOSE	MIN	MAX
09/09/2007	23:59:51	8	155.50	155.55	155.50	155.56
09/10/2007	00:00:01	15	155.53	155.55	155.49	155.58
09/10/2007	00:00:11	6	155.55	155.56	155.53	155.57
09/10/2007	00:00:21	2	155.51	155.50	155.50	155.52
09/10/2007	00:00:31	13	155.50	155.52	155.50	155.56
09/10/2007	00:00:41	21	155.53	155.52	155.49	155.54
09/10/2007	00:00:51	15	155.52	155.53	155.51	155.57
09/10/2007	00:01:01	10	155.57	155.57	155.54	155.58

Here prices are recorded every 10 seconds. Higher frequency data are also available. We bear this type of data in mind.

Known facts on the RV.

- CLT for deterministic sampling; if e.x.,

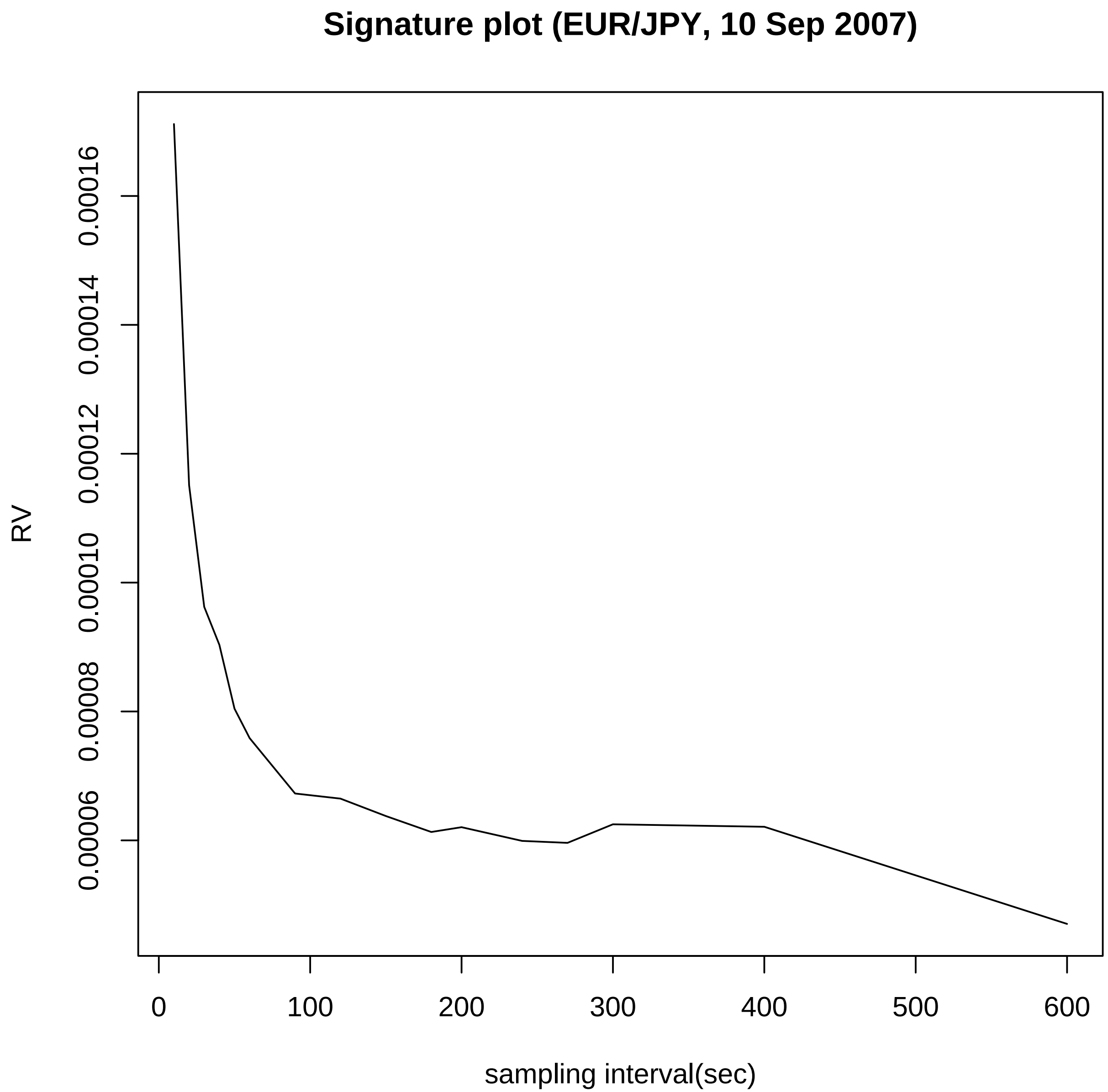
$$\tau_j = \tau_j^n = Tj/(n+1),$$

  that is, if the partition is equidistant in time, then

$$\sqrt{n}(\mathrm{RV}(n) - \mathrm{IV}) \Rightarrow \mathcal{MN}(0, 2T\mathrm{IQ}), \quad \mathrm{IQ} = \int_0^T \sigma(t,X)^4 dt$$

  as $n \to \infty$. This type of result is useful for, say, constructing confidence intervals.

- In practice, however, the RV based on $\{\tau_j^n\}$ seems not consistent. It's because the observed values of $S_{\tau_j^n}$ are contaminated by market microstructure noise.



Signature plot (EUR/JPY, 10 Sep 2007)
RV
0.00006
0.00008
0.00010
0.00012
0.00014
0.00016
0
100
200
300
400
500
600
sampling interval(sec)

1. Central Limit Theorems:
   (a) Jacod and Protter (1998)
   (b) Barndorff-Nielsen and Shephard (2002)
   (c) Mykland and Zhang (2002)
2. Microstructure noise: e.x.
   (a) Zhou (1996)
   (b) Bandi and Russell (2003)
   (c) Barndorff-Nielsen, Hansen, Lunde and Shephard (2004)
   (d) Zhang, Mykland and Aït-Sahalia (2005)
   (e) Hansen and Lunde (2006)
3. Rounding error: Delattre and Jacod (1997)

The easiest way to deal with the problem is to throw away a large fraction of the available data (sampling less frequently). However, this approach is not sound in statistical point of view.

In spite of a lot of studies, there is no decisive solution.

Here we propose another treatment of the microstructure noise to tackle this problem.

A similar approach to Griffin and Oomen (2006), Large (2005).

One of major source of this inconsistency is price discreteness.

Let us assume the observed bid and ask prices B_t and A_t to be given as

$$B_t = \delta\lfloor \beta S_t/\delta \rfloor, \quad A_t = \delta\lceil \alpha S_t/\delta \rceil$$

for $\delta > 0$, representing tick size, $\beta \in (0, 1]$ and $\alpha \geq 1$, representing costs of quote exposure. It holds $B_t \leq \beta S_t \leq S_t \leq \alpha S_t \leq A_t$.

The efficient price S_t is latent in this model. Since

$$\log(B_{\tau^n_{j+1}}) - \log(B_{\tau^n_j}) = \log(S_{\tau^n_{j+1}}) - \log(S_{\tau^n_j}) + O(\delta),$$

while

$$\log(S_{\tau^n_{j+1}}) - \log(S_{\tau^n_j}) = O(\frac{1}{\sqrt{n}}),$$

the noise term becomes dominant as $n \to \infty$ with δ fixed. The same situation occurs also in the use of A_t or $(A_t + B_t)/2$.

[What is the TTS?]

Let us solve the problem by changing the sampling scheme.

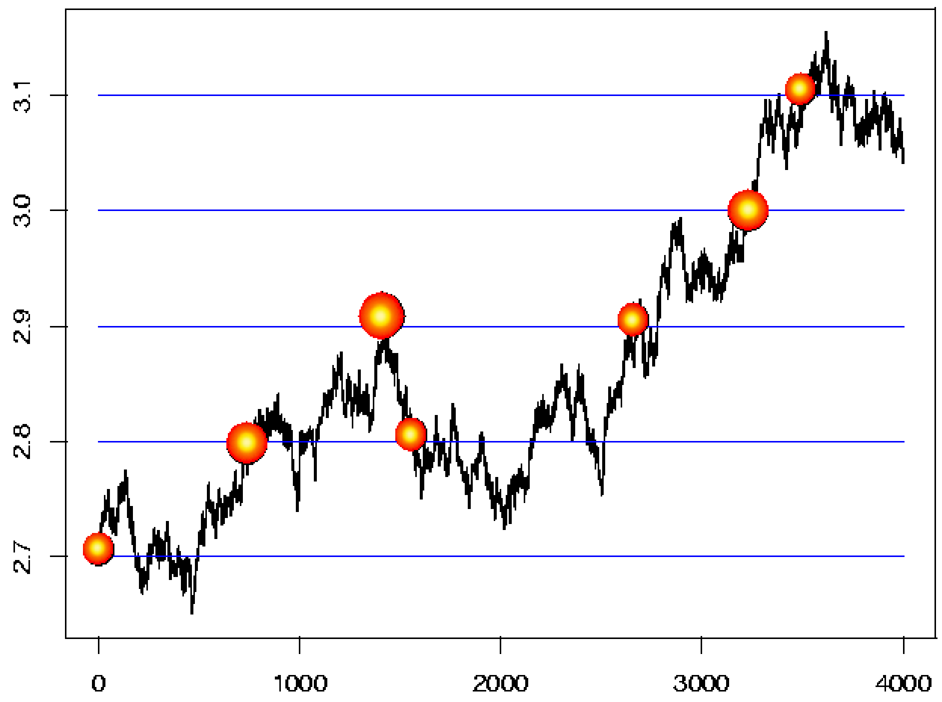
1. Calender Time Sampling (CTS): $\tau_j^n = Tj/(n+1)$, which we have considered so far, a partition equidistant in time.

2. Tick Time Sampling (TTS): based on the price (quote) changes, a partition equidistant in space.

In the following, we consider a version of the TTS and show that the RV based on the TTS is free from the noise problem.

Idea; no rounding error occurs if we sample the price when it lies on the price grid.

[What is the TTS?] Definition; $\tau^{\epsilon}_{j+1} = \inf\{t > \tau^{\epsilon}_{j}; |S_t - S_{\tau^{\epsilon}_{j}}| \geq \epsilon\}$



[Robustness to market microstructure noise]

The key fact is that $B_{\tau_j^\epsilon} = \beta S_{\tau_j^\epsilon}$ (no rounding error) for $\epsilon = \delta/\beta$. Moreover,

$$\begin{aligned}\left|\log(B_{\tau_{j+1}^\epsilon}) - \log(B_{\tau_j^\epsilon})\right|^2 &= \left|\log(\beta S_{\tau_{j+1}^\epsilon}) - \log(\beta S_{\tau_j^\epsilon})\right|^2 \\ &= \left|\log(S_{\tau_{j+1}^\epsilon}) - \log(S_{\tau_j^\epsilon})\right|^2.\end{aligned}$$

The log-returns based on the bid price coincides with that based on the efficient price.

This means that the TTS yields an appropriate RV even if we use the bid price to calculate the RV.

Moreover, since

$$\begin{aligned}\tau_{j+1}^{\epsilon} &= \mathrm{inf}\{t > \tau_j^{\epsilon}; |S_t - S_{\tau_j^{\epsilon}}| \geq \epsilon\} \\ &= \mathrm{inf}\{t > \tau_j^{\epsilon}; B_t \geq B_{\tau_j^{\epsilon}} + \delta\} \wedge \mathrm{inf}\{t > \tau_j^{\epsilon}; B_t \leq B_{\tau_j^{\epsilon}} - 2\delta\},\end{aligned}$$

the sampling can be done by observing only the bid quote.

We can use also the ask quotes instead of the bid.

We have seen so far that the RV based on the TTS is robust to microstructure noise induced by the price discreteness and bid-ask spreads.

It is then natural to ask if the corresponding CLT holds.

Main result - Asymptotic mixed normality.

Recall

$$dX_t = \mu(t,X)dt + \sigma(t,X)dW_t$$

and let $X_t = \varphi(S_t)$ (e.x. $\varphi = \log$). If

$$\tau_{j+1}^{\epsilon} = \inf\{t > \tau_j^{\epsilon}; |S_t - S_{\tau_j^{\epsilon}}| \geq \epsilon\}$$

and Girsanov's theorem holds for X, then it holds

$$\epsilon^{-1}(\mathrm{RV}(\epsilon) - \mathrm{IV}) \Rightarrow \mathcal{MN}(0,\Sigma) \qquad (1)$$

as $\epsilon \to 0$, where

$$\Sigma = \frac{2}{3}\int_0^T \varphi'(S_t)^2 d\langle\varphi(S)\rangle_t.$$

- It is noteworthy that putting $M_T = \max\{j; \tau_{j+1} \le T\}$,

$$\frac{\sum_{j=1}^{M_T} |X_{\tau_{j+1}} - X_{\tau_j}|^2 - \mathrm{IV}}{\sqrt{\sum_{j=1}^{M_T} |X_{\tau_{j+1}} - X_{\tau_j}|^4}} \Rightarrow \mathcal{N}(0, 2/3)$$

for both the CTS and the TTS. The result for the CTS is already known.

- A key lemma for the proof is

$$E\left[\left||X_{\tau_{j+1}} - X_{\tau_j}|^2 - \left(\tau_{j+1} - \tau_j\right)\right|^2\right] = \frac{2}{3} E\left[\left|X_{\tau_{j+1}} - X_{\tau_j}\right|^4\right]$$

for both the CTS and the TTS, if X is a Brownian motion. This identity is trivial for the CTS where $\tau_{j+1} - \tau_j = T/n$.

Concluding remarks:

1. The general case reduces to the Brownian motion case by the Girsanov transformation and the time-change method.

2. It converges $\sigma(X)$-stably in $\mathcal{D}[0, T]$.

3. The result can be extended to the stochastic volatility model.

4. It remains for further research to investigate the performance in real data.