# Estimating the Integrated Variance

## 1 Setup

Let $X$ denote the stochastic process for log-prices. Let's start with the continuous case:

$$X_t = \int_0^t a_s \, ds + \int_0^t \sqrt{c_s} \, dW_s$$

As before, the first integral is the ordinary integral, the second is a stochastic integral. Throughout, $c_t$ is the local variance of the process, which is a stochastic process itself determined by the state of the economy. At high frequencies we can ignore the drift $a_t$, which is essentially a tiny constant and not detectable with high frequency data.

For a small $\tau$:

$$\begin{aligned} X_{t+\tau} &= X_t + \int_t^{t+\tau} \sqrt{c_s} \, dW_s \\ &\approx X_t + \sqrt{\int_t^{t+\tau} c_s ds} \underbrace{Z_t}_{\overset{d}{\sim} \mathcal{N}(0,1)} \end{aligned}$$

In most of what follows, we treat the $c$ process as given independently of the $W$ process. This assumption is only for simplicity and can be omitted in the advanced theory.

If $c$ is constant then:

$$X_{t+\tau} \approx X_t + \sqrt{\tau c} Z_t$$

Many authors write $\sigma \equiv \sqrt{c}$, so the above becomes:

$$X_{t+\tau} \approx X_t + \tau^{1/2} \sigma Z_t$$

Next, we replace $\tau$ with $\Delta_n$, which is the width of the sampling interval. Then, we can write:

$$X_{t+\tau} \approx X_t + \sqrt{\Delta_n c} Z_t$$

when the variance process is constant ($c_s = c$).

## 2 Sampling

Given the process for log-prices, we consider discrete and equi-spaced observations at sampling intervals given by $\Delta_n$. That is, we assume we observe $X$ at times: $X_0, X_{\Delta_n}, X_{2\Delta_n}, \ldots, X_{nT\Delta_n}$, where $n = \lfloor 1/\Delta_n \rfloor$.

The returns of this asset are given by:

$$\Delta_i^n X \equiv X_{i\Delta_n} - X_{(i-1)\Delta_n} \text{ for } i = 1, 2, \ldots, nT$$

Following Ole E. Barndorff-Nielsen and Neil Shephard, 2004 (and references therein) we can define the realized variance and bipower variance via sums of $|\Delta_i^n X|^2$ and $|\Delta_i^n X| \left|\Delta_{i-1}^n X\right|$.

# 3 The Realized Variance and Bipower Variance

The basic measure of variance in much of our work is the realized variance. Suppose for the moment $T = 1$ (a day). Then:

$$RV \equiv \sum_{i=1}^{n} |\Delta_i^n X|^2$$

Below, we see why in the continuous case:

$$RV \to IV \text{ where } IV \equiv \int_0^1 c_s \, ds$$

The above only holds if $X$ is continuous, in the case of jumps:

$$RV \to IV + \sum_p \left|\Delta_{\tau_p} X\right|^2$$

i.e., the integrated variance plus the sum of the jumps squared. (Where $(\tau_p)_{p\geq 1}$ index the jump times.) A very important jump-robust measure of integrated variance is the bipower variation

$$BV \equiv \frac{\pi}{2} \sum_{i=2}^{n} |\Delta_i^n X| \left|\Delta_{i-1}^n X\right|$$

It can be shown that even under the presence of jumps:

$$BV \to IV$$

Let's show the result for RV.

## 3.1 Law of Large Numbers (LLN) for RV in the Continuous Case

The basic and familiar law of large numbers is

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \to \mu_y \text{ where } \mu_y \equiv \mathbb{E}[Y]$$

under reasonable independence assumptions on the $Y_i$. This same law of large numbers implies that $RV \to IV$ in the continuous case.

To see why, note that

$$RV = \sum_{i=1}^{n} |\Delta_i^n X|^2$$

A reasonable question is where is the $1/n$ term to make an average? The answer is that it is already there from the sampling scheme. Suppose $c_s = c$, a constant. Then

$$\begin{aligned} RV &\approx \sum_{i=1}^{n} \left|\sqrt{\Delta_n c} Z_i\right|^2 \\ &= \Delta_n c \sum_{i=1}^{n} |Z_i|^2 \\ &= c \frac{1}{n} \sum_{i=1}^{n} |Z_i|^2 \end{aligned}$$

By the law of large numbers $\frac{1}{n}\sum_{i=1}^{n}|Z_i|^2 \to \mathbb{E}[Z^2] = 1$, so we get $RV \to c$.

For the general case, the required law of large numbers is as follows. For random variables $Y_{n,i}$ that are independent across $i$ for each $n$, then

$$\frac{1}{n}\sum_{i=1}^{n} Y_{n,i} \to \mu \text{ where } \mu = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[Y_{n,i}]$$

If $c_t$ varies over the interval $[0,1]$, then $RV$ can be expressed as

$$RV = \sum_{i=1}^{n} \bar{c}_{n,i} Z_{n,i}^2$$

where

$$\bar{c}_{n,i} = \int_{t_{i-1}}^{t_i} c_s\, ds, \qquad t_i = i\Delta_n.$$

The $i^{th}$ increment in $X$ is

$$\Delta_i^n X = \sqrt{\bar{c}_{n,i}} Z_{n,i}$$

Locally, $c_s$ is approximately a constant $c_{n,i}$, on $[(i-1)\Delta_n, i\Delta_n]$, so $\bar{c}_{n,i} = \int_{t_{i-1}}^{t_i} c_s\, ds = \Delta_n c_{n,i}$

$$RV = \sum_{i=1}^{n} \Delta_n c_{n,i} Z_{n,i}^2$$

or

$$RV = \frac{1}{n}\sum_{i=1}^{n} c_{n,i} Z_{n,i}^2$$

Setting $Y_{n,i} = c_{n,i}Z_{n,i}^2$, then

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[Y_{n,i}] = \Delta_n \sum_{1}^{n} c_{n,i} \to \int_0^1 c_s\, ds.$$

by ordinary (calculus) integration.

## 3.2 The Central Limit Theorem

We only now know that $RV \to IV$, but the result is only useful only if the convergence is fast enough and we can make inferences, i.e., form confidence intervals. Across statistics the best rate we generally achieve $\sqrt{n}$ on $n$ observations. We will see that the $RV$ actually achieves this rate.

The central limit theorem (CLT) problem is to determine the rate of convergence and the limiting distribution, if available. Suppose $t \in [0,1]$. The claim is that the rate is $\Delta_n^{-\frac{1}{2}}$, i.e., $n^{\frac{1}{2}}$ and the asymptotic distribution as given below.

To see this, consider

$$\Delta_n^{-\frac{1}{2}}\left(\sum_{i=1}^{n}|\Delta_i^n X|^2 - \int_0^1 c_s\ ds\right)$$

Remember $\bar{c}_{n,i} = \int_{t_{i-1}}^{t_i} c_s\ ds$, where $t_i = i\Delta_n$, and $\Delta_i^n X = \sqrt{\bar{c}_{n,i}}Z_{n,i}$. Thus we can write the above as

$$\Delta_n^{-\frac{1}{2}}\sum_{i=1}^{n}\bar{c}_{n,i}\left(Z_{n,i}^2 - 1\right)$$

### 3.2.1 The Case of Constant Local Variance

Suppose for the moment that $c_s = c$, a randomly selected (by the economy) constant but we condition on the realization. Then $c_{n,i} = c$ and $\bar{c}_{n,i} = c\Delta_n$, and the previous equation becomes:

$$\Delta_n^{-\frac{1}{2}} \Delta_n \left( \sum_{i=1}^{n} c_{n,i} \left( Z_{n,i}^2 - 1 \right) \right)$$

or

$$c\Delta_n^{\frac{1}{2}} \sum_{i=1}^{n} \left( Z_{n,i}^2 - 1 \right)$$

Equivalently it is

$$\frac{c}{\sqrt{n}} \sum_{i=1}^{n} \left( Z_{n,i}^2 - 1 \right).$$

By the ordinary central limit theorem we have that the above $\xrightarrow{d} \mathrm{N}(0, 2c^2)$. The 2 comes from $\mathrm{Var}\left(Z^2\right) = 2$.

### 3.2.2 The Case of Non-Constant Local Variance

If $c_t$ is not a constant then things are a little more delicate:

$$\Delta_n^{-\frac{1}{2}} \sum_{i=1}^{n} \bar{c}_{n,i} \left( Z_{n,i}^2 - 1 \right) \quad = \quad \sum_{i=1}^{n} \Delta_n^{-\frac{1}{2}} \bar{c}_{n,i} \left( Z_{n,i}^2 - 1 \right)$$

The above will converge in distribution to a normal with asymptotic variance

$$2 \times \lim_{n\to\infty} \sum_{i=1}^{n} \left( \Delta_n^{-\frac{1}{2}} \bar{c}_{n,i} \right)^2$$

so long as the limit is positive and finite. Suppose $c_t$ is smooth enough that it acts like a constant on $[t_{n,i-1}, t_{n,i}]$: $c_s \approx c_{n,i} I\left[t_{n,i-1} \leq s \leq t_{n,i}\right]$. Then $\bar{c}_{n,i} \approx c_{n,i}\Delta_n$, and we get:

$$2 \times \lim_{n\to\infty} \sum_{i=1}^{n} \left( \Delta_n^{-\frac{1}{2}} c_{n,i} \Delta_n \right)^2$$

Finally:

$$2 \times \lim_{n\to\infty} \sum_{i=1}^{n} c_{n,i}^2 \Delta_n \to 2 \times \int_0^1 c_s^2 \, ds$$

by the regular definition of the integral. To summarize, we have:

$$\Delta_n^{-\frac{1}{2}} \sum_{i=1}^{n} \bar{c}_{n,i} \left( Z_{n,i}^2 - 1 \right) \xrightarrow{d} \mathrm{N}\left( 0, 2 \int_0^1 c_s^2 \, ds \right)$$

Common practice is to set $\sigma_s = \sqrt{c_s}$ so the limit is written as

$$\mathrm{N}\left( 0, 2 \int_0^1 \sigma_s^4 \, ds \right)$$

the now classic result. The result was initially developed and extended for econometrics by O. Barndorff-Nielsen and N. Shephard, 2002a; O. Barndorff-Nielsen and N. Shephard, 2002b; Ole E. Barndorff-Nielsen and Neil Shephard, 2002; Ole E. Barndorff-Nielsen and Neil Shephard, 2004; O. Barndorff-Nielsen, N. Shephard, and M. Winkel, 2006; O. Barndorff-Nielsen and N. Shephard, 2006; Ole E Barndorff-Nielsen, Neil Shephard, and Matthias Winkel, 2006; O. Barndorff-Nielsen, Graversen, et al., 2005. It can be derived using different methods based on Jacod and Protter, 1998 and presented in general form in J. Jacod, 2008, Jean Jacod and Philip Protter, 2012, and Ait-Sahalia and Jean Jacod, 2014 .

# References

Ait-Sahalia, Yacine and Jean Jacod (2014). *High Frequency Financial Econometrics*. Princeton.

Barndorff-Nielsen, O.E., S. Graversen, et al. (2005). "A Central Limit Theorem for Realised Power and Bipower Variations of Continuous Semimartingales". In: *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev*. Ed. by Y. Kabanov and R. Lipster. Springer.

Barndorff-Nielsen, O.E. and N. Shephard (2002a). "Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models". In: *Journal of the Royal Statistical Society* Series B, 64, pp. 253–280.

— (2002b). "Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models". In: *Journal of the Royal Statistical Society* Series B, 64, pp. 253–280.

— (2006). "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation". In: *Journal of Financial Econometrics* 4, pp. 1–30.

Barndorff-Nielsen, O.E., N. Shephard, and M. Winkel (2006). "Limit Theorems for Multipower Variation in the Presence of Jumps in Financial Econometrics". In: *Stochastic Processes and Their Applications* 116, pp. 796–806.

Barndorff-Nielsen, Ole E. and Neil Shephard (2002). "Estimating Quadratic Variation Using Realized Variance." In: *Journal of Applied Econometrics* 17.5, pp. 457–477. ISSN: 08837252. URL: `http://search.ebscohost.com.proxy.lib.duke.edu/login.aspx?direct=true&db=eoh&AN=0628412&site=ehost-live&scope=site`.

— (2004). "Power and Bipower Variation with Stochastic Volatility and Jumps." In: *Journal of Financial Econometrics* 2.1, pp. 1–37. ISSN: 14798409. URL: `http://search.ebscohost.com.proxy.lib.duke.edu/login.aspx?direct=true&db=eoh&AN=0783511&site=ehost-live&scope=site`.

Barndorff-Nielsen, Ole E, Neil Shephard, and Matthias Winkel (2006). "Limit theorems for multipower variation in the presence of jumps". In: *Stochastic processes and their applications* 116.5, pp. 796–806.

Jacod, J. (2008). "Asymptotic Properties of Power Variations and Associated Functionals of Semimartingales". In: *Stochastic Processes and their Applications* 118, pp. 517–559.

Jacod, J and P Protter (1998). "Asymptotic error distributions for the Euler method for stochastic differential equations". In: *Annals of Probability* 26.1, 267–307. ISSN: 0091-1798.

Jacod, Jean and Philip Protter (2012). *Discretization of Processes*. Springer-Verlag. URL: `http://dx.doi.org/10.1007/978-3-642-24127-7`.