8

High frequency data

How to measure the volatility of high frequency data?

Range volatility: diff. b/w highest & lowest log of variable
↳ simple & easy to calculate but not very insightful

Realised volatility:

M = Sample size for day
↳ 1-sec freq. = 23400 $(6.5 \times 60 \times 60)$
↳ 5-sec freq. = 4680 $(23400/5)$
↳ 1-min freq = 390 $(23400/60)$
↳ 5-min freq. = 78 $(23400/(60 \times 5))$

$$\boxed{RV(M) = M \times Var(r) \approx \sum_{i=1}^{M} r_i^2}$$

↳ RV is the sum of the squared log returns during that day

$$\boxed{Var(r) = \frac{1}{M}\sum_{i=1}^{M}(r - \bar{r})^2 \approx \frac{1}{M}\sum_{i=1}^{M} r^2}$$

↗ mechanism for biasing volatility estimates

Microstructure noise

→ Unintuitively, as M increases, RV estimates become more volatile as $RV(M)$ converges towards Integrated volatility.

→ Additional movements in price at higher frequencies caused by trading, not market movements are microstructure noise
↳ makes RV very unstable, P_t = fundamental price$_t$ + ε_t
↳ 5-10 min frequencies generally considered unaffected by micro. noise

Bipower volatility estimators

→ Jumps (policy, new CEO, new data) can also bias volatility estimates as one-off movements in price caused by ext. shock.
↳ If jumps occur, RV estimate represents sum of two terms due to mkt & jumps respectively.

→ A jump-robust estimator of RV must detect persistent dynamics of daily volatility of asset returns

→ The bipower volatility estimator does this, based off AR of absolute log returns

↳ Rationale: jumps are infrequent $\therefore$ if $|r|$ is large it's likely offset by $|r_{i-1}|$ (which is expected to be small)

$$\boxed{BV(M) = \frac{\pi}{2} \sum_{i=2}^{M} |r| \times |r_{i-1}|}$$