Some properties of the sample median of an in-fill sequence of events with an application to high frequency financial econometrics*

Neil Shephard
Department of Economics and Department of Statistics,
Harvard University

September 4, 2022

Abstract

Using an in-fill argument, the properties of the sample median of a sequence of events are established both for the case of a fixed period of time and for a period which shrinks as the sample size grows. The results are used to study the properties of the sample median of absolute returns under stochastic volatility. This estimator is invariant, asymptotically pivotal and a 1/2 breakdown estimator. In practice it has deep robustness to jump processes even when there are jumps of α-stable type.

Keywords: High frequency financial econometrics, jumps, median, in-fill asymptotics, stochastic volatility.

1 Introduction

What does a sample median of a heterogeneous sequence of events estimate?

The classical answer, which I will not use, is based on long span asymptotics, assuming a strictly stationary process $\{Z_i\}_{i=1,2,\ldots}$ with marginal density f_{Z_1}, where the sample median, written $\widehat{\mu}$, is calculated on

$$Z_1, ..., Z_n.$$

The limit theory for $\widehat{\mu}$ is given in Sen (1968) and has the form

$$\frac{\sqrt{n}\left(\widehat{\mu}-\mu\right)}{\sqrt{V_\infty/f_{Z_1}(\mu)^2}} \xrightarrow{d} N(0,1), \quad \text{where} \quad V_\infty = \frac{1}{4} + 2\sum_{l=1}^{\infty}\left\{F_{Z_1,Z_{1+l}}(\mu,\mu) - 1/4\right\},$$

assuming $\{1_{Z_i \leq m}\}$ is m-dependent, where $\mu = Q_{Z_1}(1/2)$, is the median of Z_1. This statement deploys notation used throughout this paper: for a generic random variable A and constants a and $p \in (0,1)$, the cumulative distribution function, density function (if it exists) and p-quantile are written as $F_A(a)$, $f_A(a)$ and $Q_A(p)$, respectively.

My line of thought is based on in-fill asymptotics (this is sometimes called fixed-domain asymptotics in statistics and sometimes continuous record asymptotics in econometrics). The asymptotic thought experiment is that there are n data points per unit of time measured over the fixed interval 0 to T, and that n increases, the data is written as

$$Z_1, ..., Z_{n_T},$$

*The code for the results in this paper is in file: `median20220902.r`.

where $n_T = \lfloor nT \rfloor$. I make the assumption that

$$Z_i = \mu_{(i-1)/n} + \sigma_{(i-1)/n} v_i, \quad i = 1, 2, ..., n_T, \tag{1}$$

where v_i is i.i.d. with a zero median and independent from $\{\mu_t, \sigma_t\}_{t \geq 0}$, the μ_t is the time-t "spot median," the σ_t is the time-t "spot scale," where $t \geq 0$. Then I will show that the estimand, the "in-fill median", for the sample median of a sequence of events is

$$\frac{\frac{1}{T}\int_0^T (\mu_u/\sigma_u) f_v(\widetilde{z}_u)\mathrm{d}u}{\frac{1}{T}\int_0^T (1/\sigma_u) f_v(\widetilde{z}_u)\mathrm{d}u},$$

where f_v is the associated density of v_1 and $\widetilde{z}_t$ will be explained later. If the $\{\mu_t, \sigma_t\}_{t \geq 0}$ is a random process, then the in-fill median is random. The structure of equation (1) means that the distribution of $\left(Z_i - \mu_{(i-1)/n}\right)/\sigma_{(i-1)/n}$ is time-invariant, which will turn out to be a natural assumption for some problems in financial econometrics but is certainly a very strong assumption for general problems.

In practical applications it is often helpful to think of the length of time the data is recorded over as shrinking to zero as n increases, and then computing the sample quantile $\widehat{\mu}_t$ using data

$$Z_{\lfloor tn \rfloor - n_T + 1}, ..., Z_{\lfloor tn \rfloor},$$

available just before time t. As n increases, then the unsurprising result

$$\frac{n_T^{1/2}\left(\widehat{\mu}_t - \mu_t\right)}{\sqrt{\sigma_t^2/4f_v\left(0\right)^2}} \xrightarrow{d} N\left(0,1\right),$$

holds, where $n_T = cn^{1/2+\eta}$, with the constants $c > 0$ and the small $\eta > 0$. The statistic $\widehat{\mu}_t$ is a "median filter," in the sense introduced by Tukey (1971) for time series. An early survey of median filtering of time series is provided by Justusson (1981), while more modern work is developed and discussed by Fried et al. (2007). Related work on median filtering and smoothing on images is discussed in, for example, Koch (1996) and Arisa-Castro and Donoho (2009).

In-fill arguments over fixed intervals of time are often used in the financial econometrics literature (e.g. Barndorff-Nielsen and Shephard (2002) and Andersen et al. (2001) and the reviews by Andersen and Benzoni (2009), Bollerslev (2022)), the discretization of stochastic processes (e.g. Kloeden and Platen (1992) and Jacod and Protter (2012)), spatial statistics (e.g. Cressie (1993), Stein (1999), Gneiting et al. (2012) and Tang et al. (2021)) and regression discontinuity (e.g. Carraneo and Titijunik (2022)). In our paper the assumed non-parametric model will have a nugget or measurement error type feature, but the scale of the error changes through time — which is important in financial applications.

My motivation for thinking about this type of estimator comes from a variety of problems in financial econometrics where the robustness of sample medians is attractive due to (i) the famously heavy tailed nature of the data, (ii) the need to have real time methods which can be automatically and reliably used in the context of data feeds which occasionally have mistakes in them (e.g. a trade is delayed 30 minutes in the tape).

Here my theory of the in-fill sample median is applied to stochastic volatility (e.g. Ch. 1 of Shephard (2005) and Andersen and Benzoni (2011)) — a core model in modern financial econometrics. In the simplest Gaussian stochastic volatility model, returns are models as $Y_i = c_{(i-1)/n} v_i$, where v_i is i.i.d. standard normal. Then set $Z_i = |\sqrt{n} Y_i|^r$. Then the main properties of the sample median class of estimators

$$\widehat{c_t^r} = q^{-r} med(|\sqrt{n} Y_{\lfloor nt \rfloor - d_T + 1}|^r, ..., |\sqrt{n} Y_{\lfloor nt \rfloor}|^r), \quad r > 0, \quad q = Q_{|N(0,1)|}(1/2),$$

will be established when

$$n_T = c n^{1/2+\eta},$$

with the constants $c > 0$ and the small $\eta > 0$. The $\widehat{c_t^r}$ is invariant to r, that is $\widehat{c_t^r} = \widehat{c}_t^{\,r}$, as well as being central to an asymptotic pivot

$$\sqrt{n_T} \left(\frac{\widehat{c_t^r} - c_t^r}{c_t^r} \right) \xrightarrow{d} N\left(0, b_r^2\right), \quad \text{where} \quad b_r^2 = \frac{r^2}{16 q^2 f_{\chi_1^2}(q^2)^2},$$

and having a 1/2 (robustness) breakdown point (see section 2.2 of Hampel et al. (2005)) — I suggest in practice taking $n_T = \lceil 2n^{0.44} \rceil$ and $r = 2/3$ to produce very good finite sample properties. Further $\widehat{c_t^r}$ is robust to drift, compound Poisson jumps and a pure α-stable process under very wide and simple conditions (consistency for c_t^r always holding while the CLT holds if $\alpha < 4/3$). Hence the median based estimator is a plausible alternative to the bipower type spot-volatility estimators introduced by Barndorff-Nielsen and Shephard (2004, 2006) — which is the basis of most preliminary spot volatility estimators used in modern financial econometrics (often to set the time-varying threshold for Mancini (2001, 2004, 2009) type volatility estimators). $\widehat{c_t^r}$ is most like the innovative estimator of Andersen et al. (2012) who look at the sample average of medians of the three most recent data points — their estimator will be spelt out in Section 2.2.

The excellence of the small sample behaviour of the sample median when $r = 2/3$, suggests the practical use of the "shrunk sample median" estimators

$$\widetilde{c}_t = \widehat{c}_t / \mathrm{E}\left[\left(1 + U b_{2/3}/\sqrt{n_T}\right)^{3/2}\right], \quad \widetilde{c_t^2} = \widehat{c_t^2} / \mathrm{E}\left[\left(1 + U b_{2/3}/\sqrt{n_T}\right)^{3}\right], \quad \text{where } U \sim N(0,1),\ \widehat{c}_t = \widehat{c_t^{2/3}}^{3/2},\ \widehat{c_t^2} = \widehat{c_t^{2/3}}^{3},$$

of c_t and c_t^2, respectively, the main estimands economists care about. The $\widetilde{c}_t$ and $\widetilde{c_t^2}$ should have very good small n_T properties (obviously the shrinkage does not effect the asymptotic properties) — simulation studies reported here suggest that this is true.

The structure of the rest of this paper is as follows. In Section 2 the properties of the sample median based on data from time 0 to time T are formalized using an in-fill argument. In Section 3 the properties of the sample median are established as the length of the interval, T, gets smaller as the sample size increases. In both sections the results are applied to the special case of stochastic volatility. The important feature is that the sample median is very robust to the presence of the contribution of jump processes. Section 4 concludes, while the Appendix 5 contains the proofs of various results stated in the paper.

2 The sample median of an in-fill sequence

The following theorem drives all the results in this paper.

Theorem 1 *Assume that*

$$Z_{i,n} = \mu_{(i-1)/n} + \sigma_{(i-1)/n} v_{i,n}, \quad i = 1, 2, ..., n_T, \quad n_T = \lfloor nT \rfloor, \quad 1 \geq T > 0,$$

where $v_{1,n}, ..., v_{n_T,n}$ is, for each n, an i.i.d. sequence with the median of $v_{1,n}$ being 0, possessing a strictly positive density function $f_{v_{1,n}}$ which is bounded from above. The sequence $\{v_{1,n}, ..., v_{n_T,n}\}$ is independent of the $\{\mu_t, \sigma_t\}_{t\geq 0}$ processes and define the "in-fill median" $\mu^*_{T,n}$ which is the solution to

$$F_{T,n}(\mu^*_{T,n}) = 1/2,$$

assuming $\{\sigma_t > c\}_{t\geq 0}$, for the tiny constant $c > 0$, and defining, for all $m \in R$,

$$F_{T,n}(m) = \frac{1}{T}\int_0^T F_{v_{1,n}} \{(m - \mu_u)/\sigma_u\} \mathrm{d}u.$$

As $n \to \infty$, writing $\widehat{Q}_Z(1/2)$ as the sample median of $Z_{1,n}, ..., Z_{n_T,n}$, then unconditionally

$$\frac{\sqrt{n_T}\left\{\widehat{Q}_Z(1/2) - \mu^*_{T,n}\right\}}{\omega_{T,n}(\mu^*_{T,n})/\phi_{T,n}(\mu^*_{T,n})} \xrightarrow{d} N(0,1), \tag{2}$$

where, assuming $\phi_{T,n}(m)$ exists,

$$\begin{aligned}
\omega^2_{T,n}(\mu^*_{T,n}) &= \frac{1}{T}\int_0^T F_{v_{1,n}} \{(\mu^*_{T,n} - \mu_u)/\sigma_u\} \left[1 - F_{v_{1,n}} \{(\mu^*_{T,n} - \mu_u)/\sigma_u\}\right] \mathrm{d}u, \\
\phi_{T,n}(\mu^*_{T,n}) &= \frac{1}{T}\int_0^T \sigma_u^{-1} f_{v_{1,n}} \{(\mu^*_{T,n} - \mu_u)/\sigma_u\} \mathrm{d}u.
\end{aligned}$$

Proof. Given in Appendix 5.2.

The Lemma 1 is an immediate consequence of Theorem 1. It is the form that is used in this paper.

Lemma 1 *Assuming the same setup as Theorem 1, but additionally assume that as n increases, that $v_{1,n} \xrightarrow{d} v$, where v possesses a strictly positive density function f_v which is bounded from above. Then conditioning on $\{\mu_t, \sigma_t\}_{t\geq 0}$, the*

$$F_{T,n}(m) \to F_T(m) = \frac{1}{T}\int_0^T F_v \{(m - \mu_u)/\sigma_u\} \mathrm{d}u$$

and so $\mu^*_{T,n} \to \mu^*_T$, where μ^*_T solves $F_T(\mu^*_T) = 1/2$, while

$$\begin{aligned}
\omega^2_{T,n}(\mu^*_T) &\to \omega^2_T(\mu^*_T) = \frac{1}{T}\int_0^T F_v \{(\mu^*_T - \mu_u)/\sigma_u\} [1 - F_{v_1} \{(\mu^*_T - \mu_u)/\sigma_u\}] \mathrm{d}u, \\
\phi_{T,n}(\mu^*_T) &\to \phi_T(\mu^*_T) = \frac{1}{T}\int_0^T \sigma_u^{-1} f_v \{(\mu^*_T - \mu_u)/\sigma_u\} \mathrm{d}u.
\end{aligned}$$

Further, if additionally, $v_{1,n} - v = o_p(n_T^{-1/2})$, then

$$\frac{\sqrt{n_T}\left\{\widehat{Q}_Z(1/2) - \mu^*_T\right\}}{\omega_T(\mu^*_T)/\phi_T(\mu^*_T)} \xrightarrow{d} N(0,1). \tag{3}$$

As this is an asymptotic pivot, this result also holds unconditionally.

I refer to μ_T^* as the "in-fill median," over the interval 0 to T.

Lemma 2 *Assuming the same setup as Lemma 1, then conditioning on* $\{\mu_t, \sigma_t\}_{t\geq 0}$

$$\mu_T^* = \frac{\int_0^T (\mu_u/\sigma_u)\, f_v(\widetilde{z}_u)\mathrm{d}u}{\int_0^T (1/\sigma_u)\, f_v(\widetilde{z}_u)\mathrm{d}u},$$

where $\{\widetilde{z}_t\}_{t\in[0,T]}$ *is a sequence where at time* u*, the* $\widetilde{z}_u$ *is between* $(\mu_T^* - \mu_u)/\sigma_u$ *and* 0.

Proof. Given in Appendix 5.3.

In the special case where $\mu_t = \mu$ for all $u \in [0,T]$, then $\mu_T^* = \mu$ — that is it is not effected by changing volatility. In the special case where $\sigma_t = \sigma$ for all $u \in [0,T]$, then $\mu_T^* = \int_0^T \mu_u f_v(\widetilde{z}_u)\mathrm{d}u / \int_0^T f_v(\widetilde{z}_u)\mathrm{d}u$.

Lemma 2 bridges between the in-fill median μ_T^* and the spot median process $\{\mu_t\}_{t\geq 0}$. This is a type of Angrist et al. (2006) regression argument, who interpreted the traditional estimator of quantile linear regression when an exact linear quantile regression does not hold. Lemma 2 shows μ_T^* is a weighted average of the $\{\mu_t\}_{t\in[0,T]}$. The weight

$$\frac{\frac{1}{T}\,(1/\sigma_t)\, f_v(\widetilde{z}_t)}{\frac{1}{T}\int_0^T (1/\sigma_l)\, f_v(\widetilde{z}_l)\mathrm{d}l},$$

on the spot median μ_t is non-negative and integrates overtime to one.

2.1 Small $\{\widetilde{z}_t\}_{t\in[0,T]}$ approximation

In the special case where $\{\widetilde{z}_t\}_{t\in[0,T]}$ is small (e.g. σ is large compared to μ, which you would expect in some applications in financial econometrics), then

$$\mu_T^* \simeq \mu_T^+, \quad \mu_T^+ = \frac{\int_0^T (\mu_u/\sigma_u)\,\mathrm{d}u}{\int_0^T (1/\sigma_u)\,\mathrm{d}u}, \tag{4}$$

$$\omega_T^2(\mu_T^*) \simeq 1/4, \quad \phi_T(\mu_T^*) \simeq f_{v_1}(0)\frac{1}{T}\int_0^T (1/\sigma_u)\,\mathrm{d}u. \tag{5}$$

The approximation μ_T^+ is typically quite accurate if v_1 is symmetrically distributed about 0.

Under the small $\{\widetilde{z}_t\}_{t\in[0,T]}$ approach, then

$$\sqrt{n_T}\frac{\left\{\widehat{Q}_Z(1/2) - \mu_T^+\right\}}{\sqrt{H_T}} \dot{\sim} N(0,1), \quad H_T = \frac{1}{4f_{v_1}^2(0)\left\{\frac{1}{T}\int_0^T (1/\sigma_u)\,\mathrm{d}u\right\}^2}.$$

To compare to standard results it is helpful to write

$$H_T = \frac{\overline{\sigma}^2}{4f_{v_1}^2(0)\left\{\frac{1}{T}\int_0^T (\overline{\sigma}/\sigma_u)\,\mathrm{d}u\right\}^2}, \quad \overline{\sigma}^2 = \frac{1}{T}\int_0^T \sigma_u^2 \mathrm{d}u.$$

As $\left(x^2\right)^{-1/2}$ is convex in x^2, Jensen's inequality implies $\overline{\sigma}^{-1} = \left(\overline{\sigma}^2\right)^{-1/2} \leq \frac{1}{T}\int_0^T \sigma_u^{-1}\mathrm{d}u$, so

$$1 \leq \frac{1}{T}\int_0^T (\overline{\sigma}/\sigma_u)\,\mathrm{d}u, \quad \text{implying} \quad \overline{\sigma}^2 \geq \frac{\overline{\sigma}^2}{\left\{\frac{1}{T}\int_0^T (\overline{\sigma}/\sigma_u)\,\mathrm{d}u\right\}^2}.$$

Hence time-varying volatility pushes H_T down, with the sample median greatly benefitting from having periods of very low σ_u.

The properties of the sample median under the small $\{\widetilde{z}_t\}_{t\in[0,T]}$ approach, are comparable with a scaled sample mean

$$\overline{Z/\sigma} = \frac{\frac{1}{n_T}\sum_{i=1}^{n} Z_{i,n}/\sigma_{i,n}}{\frac{1}{n_T}\sum_{i=1}^{n} 1/\sigma_{i,n}},$$

which has, assuming the data is symmetric, very similar properties (conditioning on the $\{\sigma_t\}_{t\geq 0}$ process)

$$\frac{\sqrt{n_T}\left(\overline{Z/\sigma} - \mu_T^+\right)}{\sqrt{H_T^\#}} \dot{\sim} N\left(0, \frac{\overline{\sigma}^2}{\left\{\frac{1}{T}\int_0^T (\overline{\sigma}/\sigma_u)\,\mathrm{d}u\right\}^2}\right), \quad H_T^\# = \frac{\overline{\sigma}^2}{\left\{\frac{1}{T}\int_0^T (\overline{\sigma}/\sigma_u)\,\mathrm{d}u\right\}^2}.$$

From a financial perspective $\overline{Z/\sigma}$ is a risk parity sample mean (e.g. Roncalli (2014), Harvey et al. (2018) and Dachraoui (2018)) — scaling the data through time to equalize time-varying risk over time. Of course the difficulty with $\overline{Z/\sigma}$ is that it needs to know the volatilities.

2.2 Illustration: stochastic volatility

Think of $Y_{1,n}, ..., Y_{n_T,n}$ are high frequency financial returns over the interval $[0, T]$ — so each return is measured over a period of the length of time $1/n$. Then define

$$Z_{i,n} = \left|\sqrt{n}Y_{i,n}\right|, \quad i = 1, 2, ..., n_T.$$

The sample median of $Z_{1,n}^r, ..., Z_{n_T,n}^r$, written as

$$\widehat{Q}_{Z^r}(1/2) = median(Z_{1,n}^r, ..., Z_{n_T,n}^r),$$

is invariant to r, that is

$$\widehat{Q}_Z(1/2) = \left\{\widehat{Q}_{Z^{1/r}}(1/2)\right\}^r,$$

(this is similar to the maximum likelihood estimator). This result follows as the function $|x|^r$ is monotonic, so does not change the order of the data as r varies.

Looking at the median of some absolute returns is not novel. Most influentially, Andersen et al. (2012) studied sample averages of medians of three data points $n_T^{-1}\sum_{i=1}^{n_T}\{median(Z_{i,n}, Z_{i-1,n}, Z_{i-2,n})\}^2$, an inventive alternative to bipower variation.

More broadly, Barndorff-Nielsen and Shephard (2002) and Andersen et al. (2001) formalized averaging the squares of $\{Z_{i,n}\}_{i=1,2,..n_T}$ in their work on realized volatility. Andersen and Benzoni (2009) and Bollerslev (2022) review some of the large subsequent literature.

What does $\widehat{Q}_{|Z|^{1/r}}(1/2)$ estimate? Applying Theorem 1 provides an answer.

I develop the answer in two stages. First by looking at a Gaussian stochastic volatility model (e.g. Ch. 1 of Shephard (2005), Andersen and Benzoni (2011)) with no drift. After that, second, looking at the impact of adding drift, compound Poisson process jumps and a pure jump stable process. The bottom line is that the result derived under a Gaussian stochastic volatility process holds under the more general conditions plus some extremely weak conditions.

Assumption 1 *(Gaussian stochastic volatility) Let $Y_{i,n} = c_{(i-1)/n}\left(B_{i/n} - B_{(i-1)/n}\right)$, where $\{B_t\}_{t\geq 0}$ is Brownian motion independent from $\{c_t\}_{t\geq 0}$, where $c_t > c > 0$ for all $t \geq 0$.*

The following Corollary follows immediately as an application of Lemma 1 by writing

$$\mu_t = q^r c_t^r, \quad \sigma_t = c_t^r, \quad v = |\varepsilon|^r - q^r, \quad q = Q_{|\varepsilon|}(1/2), \quad \varepsilon \overset{L}{\sim} N(0,1).$$

Corollary 1 *Under Assumption 1, then*

$$\begin{aligned} F_T(m) &= \frac{1}{T}\int_0^T F_v(m/c_u^r - q^r)\mathrm{d}u = \frac{1}{T}\int_0^T F_{|\varepsilon|^r}(m/c_u^r)\mathrm{d}u \\ \omega_T^2(m) &= \frac{1}{T}\int_0^T F_{|\varepsilon|^r}(m/c_u^r)\left\{1 - F_{|\varepsilon|^r}(m/c_u^r)\right\}\mathrm{d}u, \\ \phi_T(m) &= \frac{1}{T}\int_0^T c_u^{-r} f_{|\varepsilon|^r}(m/c_u^r)\,\mathrm{d}u, \quad \mu_T^* = q^r \frac{\int_0^T f_{|\varepsilon|^r}(\widetilde{z}_u)\mathrm{d}u}{\int_0^T \frac{1}{c_u^r} f_{|\varepsilon|^r}(\widetilde{z}_u)\mathrm{d}u}, \quad \mu_T^+ = \frac{q^r}{T^{-1}\int_0^T \frac{1}{c_u^r}\mathrm{d}u}. \end{aligned}$$

Then as $T/n \to 0$ and $n_T \to \infty$,

$$\widehat{Q}_{Z^r}(1/2) \overset{p}{\to} \mu_T^*, \tag{6}$$

and

$$\frac{\sqrt{n_T}\left\{\widehat{Q}_{Z^r}(1/2) - \mu_T^*\right\}}{\omega_T(m)/\phi_T(m)} \overset{d}{\to} N(0,1). \tag{7}$$

Sometimes it is convenient to express the distribution and density of $|\varepsilon|^r$ in terms of the corresponding terms for χ_1^2, then obviously

$$F_{|\varepsilon|^r}(x) = F_{\chi_1^2}(x^{2/r}), \quad f_{|\varepsilon|^r}(x) = \frac{\partial x^{2/r}}{\partial x} f_{\chi_1^2}(x^{2/r}) = \frac{2}{r}x^{2/r-1} f_{\chi_1^2}(x^{2/r}).$$

Example 1 *In the homogenous case, where $c_t = c_0$ for all $t \geq 0$, then $\mu_T^* = q^r c_0^r$, $\omega_T^2(m) = 1/4$ and $\phi_T(\mu_T^*) = c_0^{-r} f_{|\varepsilon|^r}(q^r)$, so that*

$$\sqrt{n_T}\left\{\widehat{Q}_{Z^r}(1/2) - q^r c_0^r\right\} \overset{d}{\to} N\left(0, \frac{c_0^{2r}}{4\left\{f_{|\varepsilon|^r}(q^r)\right\}^2}\right).$$

As $f_{|\varepsilon|^r}(x^r) = \frac{2}{r}x^{2-r} f_{\chi_1^2}(x^2)$, then

$$\sqrt{n_T}\left\{q^{-r}\widehat{Q}_{Z^r}(1/2) - c_0^r\right\} \overset{d}{\to} N\left(0, \frac{r^2}{16 q^2 f_{\chi_1^2}(q^2)^2} c_0^{2r}\right),$$

so the denominator is invariant to r. Notice that $q \simeq 0.6744898$ and $16q^2 f_{\chi_1^2}(q^2)^2 \simeq 0.735$. When $r = 2/3$ then $Z_{1,n}^{2/3}$ is close to being symmetrically distributed about $c_0^r(1 - 2/9)$ (e.g. Wilson and Hilferty (1931) and Terrell (2003)), implying $\widehat{Q}_{Z^{2/3}}(1/2)$ is roughly unbiased and symmetrically distributed. When $r = 2$, then $Z_{1,n}^2 = c_0^2\chi_1^2$ and the asymptotic variance of the scaled median is $5.442c_0^4$ (recall RV has an asymptotic variance of $2c_0^4$, e.g. Barndorff-Nielsen and Shephard (2002)). Finally, $\log\left\{q^{-r}\widehat{Q}_{Z^r}(1/2)\right\} - \log\sigma^r$ is an asymptotic pivot.

	no	jumps	$\alpha=$	.3	$\alpha=$	.6	$\alpha=$	.9	$\alpha=$	1.2	$\alpha=$	1.5	$\alpha=$	1.8	no jumps
n	$\mathrm{E}(\varepsilon)$	$sd(\varepsilon)$	$\mathrm{E}(\varepsilon)$	$sd(\varepsilon)$	$\mathrm{E}(\varepsilon)$	$sd(\varepsilon)$	$\mathrm{E}(\varepsilon)$	$sd(\varepsilon)$	$\mathrm{E}(\varepsilon)$	$sd(\varepsilon)$	$\mathrm{E}(\varepsilon)$	$sd(\varepsilon)$	$\mathrm{E}(\varepsilon)$	$sd(\varepsilon)$	asy $sd(\varepsilon)$
Median															
20	0.005	0.747	0.255	0.803	0.363	0.826	0.541	0.865	0.802	0.907	1.17	0.968	1.65	1.03	0.777
10^2	0.009	0.771	0.141	0.791	0.248	0.792	0.476	0.815	0.916	0.859	1.77	0.926	3.20	1.02	0.777
10^3	-0.008	0.774	0.063	0.781	0.155	0.780	0.406	0.782	1.15	0.807	3.22	0.868	8.33	0.991	0.777
10^4	-0.004	0.776	0.032	0.777	0.099	0.780	0.360	0.786	1.44	0.794	5.84	0.826	21.6	0.956	0.777
10^5	-0.002	0.781	0.018	0.769	0.064	0.775	0.320	0.777	1.81	0.782	10.4	0.802	55.8	0.925	0.777
10^6	-0.002	0.776	0.019	0.779	0.038	0.777	0.283	0.774	2.30	0.778	18.7	0.792	143	0.894	0.777

Table 1: Results from simulation experiments where $\varepsilon = \sqrt{n_T}\left\{q^{-r}\widehat{Q}_{Z^r}(1/2) - \sigma^r\right\}$ and $r = 2/3$. When $\alpha < 1$, has a Gaussian limit law with mean 0 and variance of the stated asymptotic standard deviation. Examples of this are to the left of the double line in the Table. When $\alpha \geq 1$ the estimator is consistent but the Gaussian limit law does not hold.

2.2.1 Adding drift and jumps

How do the results change when there are jumps and drift? To start, define a more general process.

Assumption 2 *Let*

$$Y_{i,n} = a_{(i-1)/n}/n + c_{(i-1)/n}\left(B_{i/n} - B_{(i-1)/n}\right) + \left(J_{i/n} - J_{(i-n)/n}\right) + \lambda_{(i-1)/n}(S_{i/n} - S_{(i-1)/n}),$$

where $\{B_t\}_{t\geq 0}$ is Brownian motion, $\{J_t\}_{t\geq 0}$ is a compound Poisson process with finite intensity $\psi < \infty$ and jumps $\{C_j\}_{j=1,2,...}$, while $\{S_t\}_{t\geq 0}$ is a α-stable process, where $\alpha \in (0,2)$. Assume $\{B_t, J_t, S_t, a, c, \lambda\}_{t\geq 0}$ are independent processes, where $c_t > c > 0$ for all $t \geq 0$.

Again take

$$\mu_t = q^r c_t^r, \quad \sigma_t = c_t^r, \quad v = |\varepsilon|^r - q^r, \quad q = Q_{|\varepsilon|}(1/2), \quad \varepsilon \overset{L}{\sim} N(0,1),$$

but now

$$v_{1,n} = \left|\varepsilon_{1/n}\right|^r - \left\{Q_{|\varepsilon_{1/n}|}(1/2)\right\}^r, \quad \varepsilon_{1/n} = (a_0/c_0)\, n^{-1/2} + \sqrt{n}B_{1/n} + \sqrt{n}(1/c_0)J_{1/n} + \sqrt{n}\,(\lambda_0/c_0)\, S_{1/n}.$$

Now apply Lemma 1, so what is left to show for the CLT to holds is give conditions that $v_{1,n} - v_1 = o_p(n_T^{-1/2})$.

Corollary 2 *(a) Under A2, $T/n \to 0$ and $n_T \to \infty$, then the consistency equation (6) still holds.*
(b) Under A2, $T/n \to 0$, $T^{1/2}n^{1-(1/\alpha)} \to 0$ and $n_T \to \infty$, the CLT equation (7) still holds.

Proof. Given in the Appendix.

The key practical result is that the CLT is always robust to finite activity jumps. For α-stable jumps and T fixed, the CLT will still hold if $\alpha < 1$ otherwise it fails and more sophisticated methods are needed.

Example 2 *Results from a simulation experiment with n and α varying, and set $r = 2/3$ are recorded in Table 1. The Table prints the sample mean and sample standard deviation of $\widehat{u} = \sqrt{n_T}\left\{q^{-r}\widehat{Q}_{Z^r}(1/2) - c_0^r\right\}$ under A2 and homogeneity, over the 25,000 replications, with, $T = 1$, $c_0 = 1$, $\lambda_0 = 1$ and no drift or compound Poisson process. The Table also states the asymptotic standard deviation. The results are in line with the theory, with the asymptotics holding rapidly when there are no jumps. When there is a stable component the* *jumps cause $\widehat{u}$ to have a positive bias which falls as n increases, while the pace of the fall becomes less sharp as α rises. When α goes beyond 1 then the bias in $\widehat{u}$ increases with n but at a slow rate in-line with the consistency of $q^{-r}\widehat{Q}_{Z^r}(1/2)$ itself. When α is less than one but close to 1 the limit theory is not a faithful guide to the finite sample behaviour.*

The $\alpha < 1$ condition also appears in Todorov and Tauchen (2011) in their analysis of the distribution of realized power variation under jumps (see also Lépingle (1976), Barndorff-Nielsen and Shephard (2004), Jacod (2007) and Jacod (2008), Todorov and Tauchen (2011))

$$\widehat{PV}_{T,n} = \{\mathrm{E}\left[|\varepsilon|^r\right]\}^{-1} \frac{1}{n_T} \sum_{j=1}^{n_T} Z_{j,n}^r, \quad r > 0,$$

but the Todorov and Tauchen (2011) CLT result also needs the additional condition that $1 > r > \alpha/(2-\alpha) > 0$, e.g. if $\alpha = 1/2$ then $r \in (1/3, 1)$ or for a specific $r \in (0,1)$ then $\alpha \in \left(\frac{2r}{r+1}, 1\right)$. Why? Because

$$\sqrt{Tn}\left|\sqrt{n}S_{1/n}\right|^r \stackrel{L}{=} \sqrt{T}n^{1/2+r(1/2-1/\alpha)}\left|S_1\right|.$$

This is frustrating in practice as it needs knowledge of α (which in practice might be time-varying and so in practice unknowable) to select an appropriate r — while the sample median version is invariant to r. On the other hand, consistency just needs $r < 2$ for any $\alpha \in (0,2)$ — which is as simple to use as the median version.

A disadvantage of $\widehat{PV}_{T,n}$, is that the mean only exists if $r < \alpha$, while its variance exists only if $2r < \alpha$. This suggests the asymptotic results discussed above might be poor guides to finite sample behaviour. This turns out turns out to be true.

Example 3 *Results from a simulation experiment varying n and α, and set $r = 2/3$ are recorded in Table 2. For this value of r, the CLT holds if $\alpha \in (4/5, 1)$, while consistency holds for all $\alpha < 2$. The Table prints the sample mean and sample standard deviation of $\widehat{u} = \sqrt{n_T}\left\{\widehat{PV}_{T,n} - c_0^r\right\}$, under A2 and homogeneity, over the $25,000$ replications, with $T = 1$, $c_0 = 1$, $\lambda_0 = 1$ and no drift or compound Poisson process. Table 2 also states the asymptotic standard deviation. The results are in line with the asymptotic guarantees but are terrible in terms of providing a useful guide to statistical practice except for when there is no stable process ($\alpha = 0$). In the experiment, the coverage of a nominal 95% confidence interval based on the asymptotics were also computed: they were extremely poor (when $\alpha \in (4/5, 1)$) even with n being $100,000$ — there were signs of very slow convergence to the 95% convergence but the nominal levels of coverage were still wildly off.*

3 Small T and μ_t

Recall $\widehat{Q}_{Z_1}(1/2)$ estimates μ_T^*? Suppose T is small, is $\widehat{Q}_{Z_1}(1/2)$ close to μ_0? More generally, does the median of a subsample of data around time t, which is called the median filter in the statistics literature, estimate μ_t? I state results in terms of the former question, as the answer to the latter question trivially follows from the former.

	no	jumps	$\alpha =$	.3	$\alpha =$	.6	$\alpha =$	.9	$\alpha =$	1.2	$\alpha =$	1.5	$\alpha =$	1.8	no jumps
n	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	asy $sd(\varepsilon)$
Power															
20	-0.006	0.544	6×10^5	6×10^7	150	19088	2.17	11.3	1.57	2.80	1.59	1.27	1.81	0.919	0.539
10^2	-0.001	0.538	1×10^5	2×10^7	7.17	135	2.21	17.7	1.88	1.74	2.49	1.15	3.53	0.866	0.539
10^3	-0.014	0.538	3×10^4	2×10^6	10.9	486	1.96	13.8	2.47	1.61	4.64	0.937	9.27	0.828	0.539
10^4	-0.028	0.542	1×10^5	1×10^7	15.0	1×10^4	1.65	3.43	3.12	1.06	8.47	0.797	24.1	0.785	0.539
10^5	-0.080	0.540	9×10^4	1×10^7	3.11	78.9	1.39	2.22	3.90	0.94	15.2	0.701	62.5	0.744	0.539
10^6	-0.249	0.537	6255	3×10^5	1.43	22.1	1.08	1.41	4.77	0.687	27.2	0.637	161	0.709	0.539

Table 2: Results from simulation experiments where $\varepsilon = \sqrt{n_T}\left\{\widehat{PV}_{T,n} - c_0^r\right\}$ when $r = 2/3$. When $\alpha \in (4/5, 1)$, then ε has a Gaussian limit law with mean 0 and variance of the stated asymptotic standard deviation. The Example of this is bracketed by the double line in the Table. Otherwise the estimator is consistent but limit law does not hold.

Theorem 2 *Set $T = dn^{-1/2-\eta}$, for small $\eta > 0$, then*

$$\frac{n_T^{1/2}\left\{\widehat{Q}_{Z_1}(1/2) - \mu_0\right\}}{\sqrt{\sigma_0^2/4f_v\left(0\right)^2}} \xrightarrow{d} N\left(0,1\right), \tag{8}$$

if the conditions of Lemma 1 hold and that $\left\{\mu_t, \sigma_t\right\}_{t\geq 0}$ is a bivariate Ito semimartingale. If $\left\{\mu_t, \sigma_t\right\}_{t\geq 0}$ are continuously differentiable with respect to time and set $T = dn^{-1/3-\eta}$, for small $\eta > 0$, then the form (8) again holds.

Proof. Given in Appendix 5.4.

The proof shows the need for a small $\eta > 0$ to asymptotically remove the contribution to the asymptotic distribution of the difference between μ_T^* and μ_0. Notice that now $\widehat{Q}_{Z_1}(1/2)$ directly estimates μ_0, not integrals involving the $\left\{\sigma_t\right\}_{t\geq 0}$ process. The proof of this result includes the optimal mean square error choices for T, but their form is not very instructive for empirical work.

The above states the result in terms of estimating μ_0, using data $Z_1, ..., Z_{n_T}$ from the start of the sample. The same results can also be phrased in terms of estimating μ_t, where $t \in [0, T]$, say, using n_T datapoints — in the past compared to time t, in the future, or using both the past and future.

For a sequence of events, in filtering μ_t is estimated using contemporaneous and past data. In smoothing μ_t is estimated using contemporaneous, past and future data. Typically smoothing is more precise, but often filtering is more useful as it can be used as an input into forecasting.

The smoother is likely to be less effected by systematic moves (e.g. diurnal features) in $\left\{\mu_t\right\}_{t\geq 0}$ as it balances out the impact of bias caused by the first derivative of any continuously differentiable component of $\left\{\mu_t\right\}_{t\geq 0}$ — which are important if n is only modestly large. For diffusive components the optimal choice of $T = dn^{-1/2-\eta}$ does not change moving between filtering and smoothing.

Under the strong condition that $\left\{\mu_t, \sigma_t\right\}_{t\geq 0}$ are continuously differentiable then smoothing can be carried out using $T = dn^{-1/4-\eta}$, not the $T = dn^{-1/3-\eta}$ required for filtering.

3.1 Illustration: stochastic volatility

Focus on filtering to estimate c_t^r, so define the sample median of

$$\widehat{c_t^r} = q^{-r} med(\left|\sqrt{n}Y_{\lfloor nt \rfloor - d_T + 1}\right|^r, ..., \left|\sqrt{n}Y_{\lfloor nt \rfloor}\right|^r).$$

In the SV case, where $\{\mu_t, \sigma_t\}_{t\geq 0}$ are expected to be partially driven by diffusive components which suggests taking $T = dn^{-1/2-\eta}$, to map into the above theory by writing

$$\mu_t = q^r c_t^r, \quad \sigma_t = c_t^r, \quad v = |\varepsilon|^r - q^r, \quad q = Q_{|\varepsilon|}(1/2), \quad \varepsilon \overset{L}{\sim} N(0,1).$$

The above implies immediately that when there are no jumps

$$\sqrt{n_T}\left(\frac{\widehat{c_t^r} - c_t^r}{c_t^r}\right) \xrightarrow{d} N\left(0, b_r^2\right), \quad b_r^2 = \frac{r^2}{16 q^2 f_{\chi_1^2}\left(q^2\right)^2}.$$

Further this CLT will hold when there are stable jumps and diffusion based volatility. What values of α are allowed? When $T = dn^{-1/2-\eta}$, so

$$(nT)^{1/2} n^{1/2-1/\alpha} = T^{1/2} n^{1-1/\alpha},$$

so the CLT needs $\alpha < 4/(3-2\eta)$.

To assess the practical usefulness of the asymptotic arguments some simulation experiments were carried out based on 100,000 replications. Throughout the data generating process will be governed by Assumption 3.

Assumption 3 *Take the volatility*

$$c_t = \left(\frac{1}{1.3}\right) \exp\left(W_t/2\right)\left\{0.3 + \cos(2.6t)^2\right\}, \quad t \in [0,1]. \tag{9}$$

The returns are computed using

$$Y_{j,n} = c_{(j-1)/n}\left(B_{j/n} - B_{(j-1)/n}\right) + 1_{\alpha>0}\left(S_{j/n} - S_{(j-1)/n}\right), \quad j = 1, 2, ..., n_T,$$

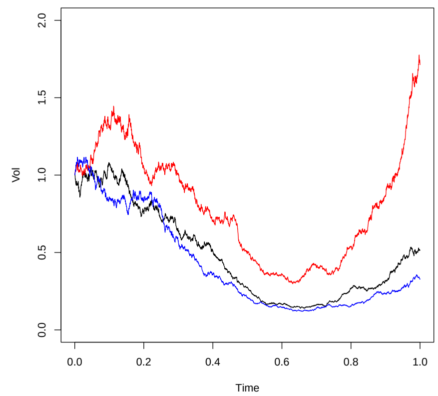
where $\{S_t\}_{t\geq 0}$ is an α-stable process, independent from the independent Brownian motions $\{B_t, W_t\}_{t\geq 0}$. Throughout $n_T = \lceil 2n^{0.44} \rceil$, that is $T \simeq 1.4n^{-0.56}$, while $r = 2/3$.

Figure 3.1 shows three independently draw simulated paths of $\{c_t\}_{t\in[0,1]}$ made under Assumption 3. The cosine in (9) mimics a strong diurnal feature often seen in high frequency datasets. This makes this Monte Carlo design challenging for filter based estimation. Focus will be on estimating c_0, corresponding to a time when volatility is sharply declining.

Throughout the sampling scheme the Monte Carlo experiments will be governed by Assumption 4.

Assumption 4 *Set $n_T = \lceil 2n^{0.44} \rceil$, that is $T \simeq 1.4n^{-0.56}$, while $r = 2/3$.*

Assumption 4 is interesting as statistical theory and our previous Monte Carlo results suggest selecting $r = 2/3$ yields good small sample performance. As $T \simeq 1.4n^{-0.56}$, under Assumption 3, the CLT to work needs $\alpha \preceq 1.39$, while the estimator is always consistent.



		no	jumps	$\alpha =$	.3	$\alpha =$	.6	$\alpha =$	.9	$\alpha =$	1.2	$\alpha =$	1.5	$\alpha =$	1.8	no jumps
n	n_T	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	asy $sd(\varepsilon)$
Median																
200	21	-0.041	0.804	-0.004	0.812	0.032	0.818	0.102	0.833	0.284	0.862	0.675	0.924	1.38	1.01	0.777
10^3	42	-0.005	0.802	0.009	0.803	0.031	0.806	0.087	0.813	0.234	0.829	0.665	0.882	1.72	0.986	0.777
10^4	116	0.001	0.800	0.006	0.804	0.011	0.803	0.043	0.805	0.165	0.815	0.634	0.843	2.33	0.958	0.777
10^5	317	0.002	0.799	0.007	0.800	0.005	0.800	0.019	0.799	0.109	0.802	0.590	0.826	3.15	0.930	0.777
10^6	874	0.000	0.793	0.005	0.796	0.002	0.792	0.010	0.793	0.069	0.793	0.558	0.811	4.24	0.904	0.777

Table 3: Results from the spot volatility simulation experiments where $\varepsilon = \sqrt{n_T}\left\{q^{-r}\widehat{Q}_{Z^r}(1/2) - c_0^r\right\}$ when $r = 2/3$. When $\alpha \preceq 1.39$, then ε has a Gaussian limit law with mean 0 and variance of the stated asymptotic standard deviation. When $1.39 \preceq \alpha < 2$ the estimator is consistent but limit law does not hold.

Example 4 *The Monte Carlo results using Assumptions 3 and 4, are given in Table 3. They are in-line with the asymptotic theory, with the theory kicking in quite fast for $\alpha < 1$. For high values of α the CLT is not a useful guide, but of course the estimator is still consistent.*

3.1.1 Comparison to power variation

In the literature $\widehat{PV}_{T,n}$ for small T (or its bipower variation version) is often used to estimate c_0^r. Again, the mean of $\widehat{PV}_{T,n}$ will only exist if $r < \alpha$. Now $\sqrt{n_T}\left|\sqrt{n}S_{1/n}\right|^r \stackrel{L}{=} n^{-1/4-\eta/2}n^{1/2+r(1/2-1/\alpha)}\left|S_1\right|^r$, so for the CLT to be valid it needs that

$$0 > -1/4 - \eta/2 + 1/2 + r(1/2 - 1/\alpha) = r\left(\frac{1+2r-2\eta}{4r} - 1/\alpha\right).$$

So for fixed r, the need is for $\alpha < 4r/(1+2r-2\eta)$, e.g. if $r = 2/3$ and $\eta = 0.06$, then α must be between around 2/3 and 1.2. The same type of asymptotics holds for the corresponding bipower statistic (Barndorff-Nielsen and Shephard (2004)), which is what is typically used in empirical practice, but now $\sqrt{n_T}\left|\sqrt{n}S_{1/n}\right|^{r/2}\left|\sqrt{n}\left(S_{2/n} - S_{1/n}\right)\right|^{r/2}$ equals in law $n^{-1/4-\eta/2}n^{1/2+r/2(1/2-1/\alpha)}\left|S_1\right|^{r/2}\left|S_2 - S_1\right|^{r/2}$ which means that the mean will exist if $r/2 < \alpha$, so if $r = 2/3$ and $\eta = 0.06$, then α must be between around 1/3 and 1.2.

Example 5 *Table 4 has the same structure as that from the experiment in Example 4 but uses the realized power variation estimator, using the same choice of n_T. The results for $\widehat{PV}_{T,n}$ are better than for $q^{-r}\widehat{Q}_{Z^r}(1/2)$ in the no jump case, with around a 50% smaller standard deviation. But when there are jumps the realized power*

		no	jumps	$\alpha =$	.3	$\alpha =$	.6	$\alpha =$	.9	$\alpha =$	1.2	$\alpha =$	1.5	$\alpha =$	1.8	no jumps
n	n_T	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	asy $sd(\varepsilon)$
Power																
200	21	-0.047	0.601	1×10^5	2×10^6	4.08	658	0.695	23.6	0.638	1.72	0.956	1.20	1.51	0.872	0.538
10^3	42	-0.010	0.594	1×10^4	2×10^6	18.3	4860	0.373	6.34	0.497	1.31	0.947	0.918	1.90	0.840	0.538
10^4	116	-0.003	0.580	7904	2×10^6	0.294	12.8	0.169	1.51	0.341	0.951	0.912	0.819	2.59	0.787	0.538
10^5	317	-0.003	0.574	32.3	7325	0.086	6.54	0.071	1.01	0.226	0.892	0.857	0.713	3.52	0.753	0.538
10^6	874	-0.008	0.563	1.75	334	0.023	2.07	0.031	0.732	0.141	0.782	0.811	0.659	4.76	0.719	0.538

Table 4: Results from the spot volatility simulation experiments where $\varepsilon = \sqrt{n_T}\left\{\widehat{PV}_{T,n} - c_0^r\right\}$ when $r = 2/3$. When α is between around 2/3 and 1.2, then ε has a Gaussian limit law with mean 0 and variance of the stated asymptotic standard deviation. Otherwise the estimator is consistent but limit law does not hold.

		no	jumps	$\alpha =$	.3	$\alpha =$	.6	$\alpha =$	.9	$\alpha =$	1.2	$\alpha =$	1.5	$\alpha =$	1.8	no jumps
n	n_T	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	E(ε)	$sd(\varepsilon)$	asy $sd(\varepsilon)$
Bipower																
200	21	-0.044	0.674	1.76	129	0.248	2.24	0.271	0.968	0.458	0.857	0.831	0.859	1.46	0.883	0.612
10^3	42	-0.008	0.664	0.886	90.2	0.146	4.15	0.183	0.779	0.365	0.762	0.828	0.793	1.83	0.846	0.612
10^4	116	0.000	0.649	0.301	50.3	0.044	0.728	0.089	0.681	0.253	0.692	0.796	0.726	2.50	0.804	0.612
10^5	317	0.004	0.643	0.039	2.73	0.017	0.655	0.041	0.647	0.169	0.658	0.750	0.685	3.38	0.771	0.612
10^6	874	0.004	0.633	0.015	0.82	0.011	0.635	0.025	0.636	0.112	0.637	0.713	0.661	4.58	0.741	0.612

Table 5: Results from the spot volatility simulation experiments where $\varepsilon = \sqrt{n_T}\left\{\widehat{BPV}_{T,n} - c_0^r\right\}$ when $r = 2/3$. When α is between around 1/3 and 1.2, then ε has a Gaussian limit law with mean 0 and variance of the stated asymptotic standard deviation. Otherwise the estimator is consistent but limit law does not hold.

variation statistic is extremely fragile, as the above theory indicates. Table 5 repeats the above but now reports results from the corresponding realized bipower statistic

$$\widehat{BPV}_n = \left\{\mathrm{E}\left[|\varepsilon|^{r/2}\right]\right\}^{-2} \frac{1}{n_T - 1}\sum_{j=2}^{n_T} Z_{j,n}^{r/2} Z_{j-1,n}^{r/2}, \quad r > 0$$

The results are better than for power variation when $\alpha > 0$, particularly for small α, and very slightly less accurate when there is no stable component. Overall the results are less widely reliable than the median case, but the difference is much less stark. In practice the bipower variation statistic is usually used in empirical work with $r = 2$ (in which case power variation is realized variance and so has a different estimand) in which case the limit theory will be valid if $1 < \alpha < 4/3.88$ when $\eta = 0.06$, which is a very narrow range.

3.1.2 Spot variance and the sample median

Appendix 5.5 contains Table 6 which has the corresponding results for the sample median and for the bipower case when $r = 2$, that is the focus is on estimating the spot variance c_t^2. Economists often have this as their preferred estimand.

The results for bipower variation when their are α-stable jumps are terrible, while the sample median are quite good but display some material biases when n_T is small. Here we introduce a simple shrunk sample median which has excellent finite sample properties.

Recall when $r = 2/3$, then when there are jumps $|N(0,1)|^{2/3}$ is roughly Gaussian using the ideas from Wilson and Hilferty (1931) and Terrell (2003), so

$$\widehat{c_t^{2/3}}/c_t^{2/3} \overset{\cdot}{\sim} N(1, b_{2/3}^2/n_T),$$

should perform reasonably well for small n_T. Simulation experiments in Example 4 strongly support the accuracy of this approximation. By invariance

$$\begin{aligned}\mathrm{E}\left[\widehat{c}_t^{\,2}/c_t^2\right] &= \mathrm{E}\left[\left(\widehat{c_t^{2/3}}/c_t^{2/3}\right)^3\right] \\ &\simeq \mathrm{E}\left[\left(1+Ub_{2/3}/\sqrt{n_T}\right)^3\right] = 1+3b_{2/3}^2/n_T, \quad U\sim N(0,1),\end{aligned}$$

so defining

$$\widetilde{c_t^2} = \widehat{c_t^{2/3}}^{\,3}/(1+3b_{2/3}^2/n_T),$$

then $\mathrm{E}\left[\widetilde{c_t^2}|c_t^2\right] \simeq c_t^2$, is roughly conditionally unbiased for c_t^2. I call $\widetilde{c_t^2}$ a "shrunk sample median estimator". Obviously the shrinkage makes no difference to the asymptotic distribution, it is just a small sample correction. Table 6 shows the shrunk sample median performs excellently for small samples both when there are, and are not, jumps.

The same style of argument applies more generally to produce roughly unbiased estimators of c_t^s for any power $s > 0$. The second leading case is when $s = 1$, where the spot volatility c_t is the estimand, then I suggest using

$$\widetilde{c}_t = \widehat{c_t^{2/3}}^{\,3/2}/\mathrm{E}\left[\left(1+Ub_{2/3}/\sqrt{n_T}\right)^{3/2}\right].$$

The expectation in the denominator will have to be computed numerically, but it is time invariant, so this can be carried out trivially to any degree of accuracy by simulation.

There has been an active econometric literature on producing inference methods for realized volatility and other high frequency statistics for small n_T. Most of this important work has focused on bootstrapping methods, e.g. Goncalves and Meddahi (2009) and Dovonon et al. (2019), but there is also an interesting strand of work on fixed n_T methods, e.g. Bollerslev et al. (2021). That work is complementary to the innovations presented in this paper.

Example 6 *Figure 1 plots $\widetilde{c}_t$ through time, together with the corresponding c_t and the square root of the corresponding realized bipower variation statistic and the realized volatility statistic. The top line of pictures highlights c_t and $\widetilde{c}_t$, the bottom line has all four quantities against time. Throughout $n_T = \lceil 2n^{0.44} \rceil$. The day has a strong diurnal in the volatility — which is one of the reasons estimating the spot volatility is in practice difficult. The returns include α-stable jumps where $\alpha = 0.8$. Most of the time the median based statistic is very close to the realized bipower estimator, but sometimes when there are large jumps the bipower statistic is disrupted. Throughout the sample median type estimator provides sensible answers, even in this very challenging environment.*

4 Conclusion

This paper formalizes the sample median as an in-fill estimator under some strong conditions. These conditions seem well setup to solve problems in financial econometrics. For fixed interval T the in-fill median μ_T^* is quite

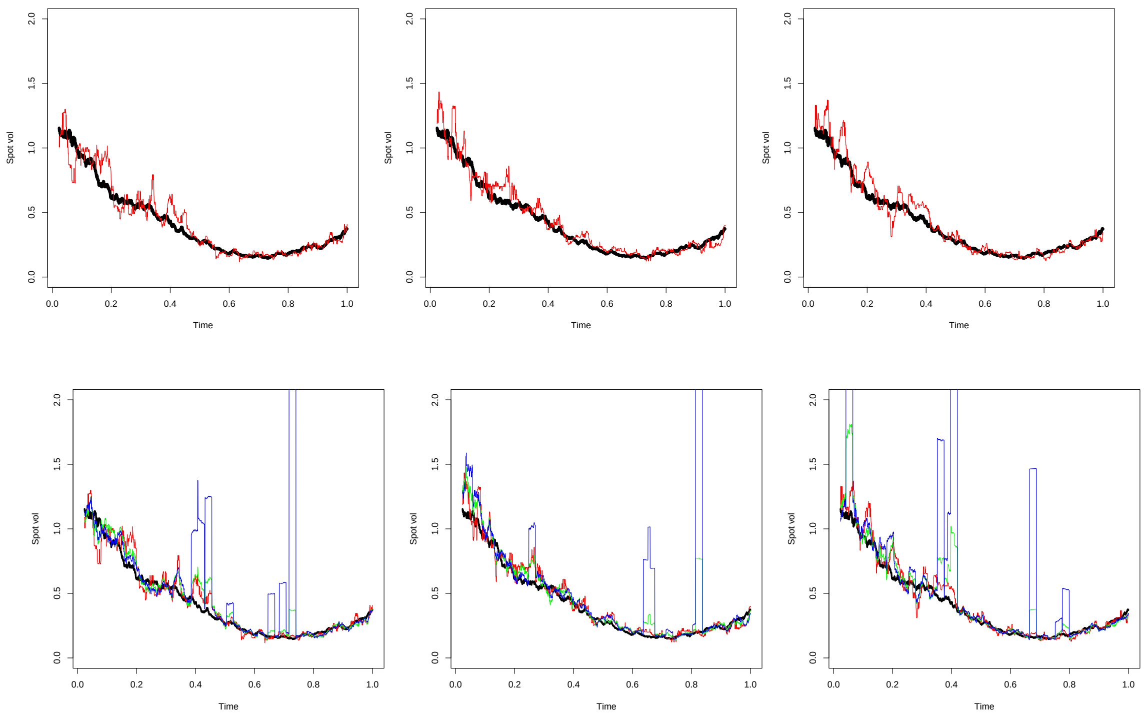


Figure 1: Three replications of estimating the same intra-day spot volatility, using different simulated returns which include the presence of a α-stable component, where $\alpha = 0.8$. Top row: thick black line is the path of c_t, thinner red line is the shrunk sample median estimator $\widetilde{c}_t$. Bottom row: true c_t (black line) together with spot realized volatility (blue), spot square root of realized bipower variation (green) and the sample median estimator (red).

complicated, but can be estimated robustly to jumps. For T getting small as n increases the in-fill median is the spot median μ_t and so has many practical applications. Throughout this paper all the concepts were illustrated in the context of a stochastic volatility model. Some results were compared to various power and bipower variation estimators.

5 Appendix

5.1 Proof of Lemma 2

By invariance, only study $r = 1$ case (note we could use the $r = 2$ case where the derivations are even easier) – then the other results hold by the delta method. The reverse triangular inequality says that

$$\begin{aligned} \left| \left| \sqrt{n} Y_{1,n} \right| - \left| a_0/n^{1/2} + c_0 \sqrt{n} B_{1/n} \right| \right| &\leq \left| \sqrt{n} J_{1/n} + \lambda_0 \sqrt{n} S_{1/n} \right| \\ &\leq \left| \sqrt{n} J_{1/n} \right| + \left| \lambda_0 \right| \left| \sqrt{n} S_{1/n} \right|. \end{aligned}$$

But jumps in the compound Poisson process only happen with probability of $O(n^{-1})$ (the size of the jumps is scaled up by $n^{r/2}$ but it does not change the number of jumps) while the stable increments $S_{1/n} \overset{L}{=} (1/n)^{1/\alpha} S_1$, so

$$\left|\sqrt{n}S_{1/n}\right| \overset{L}{=} n^{\{1/2-(1/\alpha)\}}\left|S_1\right|,$$

while $\sqrt{n}B_{1/n} \overset{L}{=} B_1$. As $\alpha < 2$ and $r > 0$ the scaled stable increments become irrelevant for consistency.

For the CLT we have two tasks: (i) show the drift does not matter and (ii) give conditions for the scaled stable increments do not matter. The former will be dealt with in a moment, focus first on the stable increments.

For the CLT to be not effected by the scaled increments we need that

$$\sqrt{n_T}n^{\{1/2-(1/\alpha)\}}S_1 = \sqrt{T}n^{\{1-(1/\alpha)\}}S_1,$$

not to matter, which needs

$$\sqrt{T}n^{1-(1/\alpha)} \to 0$$

as stated in the Lemma.

Now focus on the drift. Now

$$\left|a_0/n^{1/2} + c_0\sqrt{n}B_{1/n}\right| \overset{L}{=} c_0\left|\left(a_0/c_0\right)/n^{1/2} + B_1\right| = c_0 q + c_0 v_{1,n},$$

where

$$\begin{aligned} v_{1,n} &= \left|\left(a_0/c_0\right)/n^{1/2} + B_1\right| - q \\ v &= \left|B_1\right| - q \end{aligned}$$

What remains to show for the CLT is that

$$\sqrt{Tn}\left\{F_{v_{1,n}}(x) - F_v(x)\right\} \to 0.$$

Now

$$\begin{aligned} F_{v_{1,n}}(x) &= P\left(\left|\left(a_0/c_0\right)/n^{1/2} + B_1\right| \leq x + q\right) \\ &= P(-x - q \leq \left(a_0/c_0\right)/n^{1/2} + B_1 \leq x + q) \\ &= P(-a + b \leq B_1 \leq a + b), \quad a = x + q, \quad b = -\left(a_0/c_0\right)/n^{1/2}. \end{aligned}$$

So

$$\begin{aligned} F_{v_{1,n}}(x) &= P(B_1 \leq a + b) - P(B_1 \leq -a + b) \\ &\simeq P(B_1 \leq a) - P(B_1 \leq -a) \\ &\quad + b\left\{f_{N(0,1)}(a) - f_{N(0,1)}(-a)\right\} + \frac{b^2}{2}\left\{f'_{N(0,1)}(a) - f'_{N(0,1)}(-a)\right\} \\ &= \left\{P(B_1 \leq a) - P(B_1 \leq -a)\right\} + b^2 f'_{N(0,1)}(a) \\ &= F_v(x) + \frac{\left(a_0/c_0\right)^2}{n} f'_{N(0,1)}(x+q). \end{aligned}$$

So the result needs that $\sqrt{T/n} \to 0$.

5.2 Proof of Theorem 1

Following, for example, p. 71 Koenker (2005) define

$$G_{n_T}(m) = \frac{1}{n_T}\sum_{i=1}^{n_T}\left(1_{Z_{i,n}\leq m} - 1/2\right)$$

which is monotonically non-decreasing in $m \in R$, so by monotonicity

$$\widehat{Q}_Z(1/2) - m > 0 \Leftrightarrow G_{n_T}(m) < 0.$$

Thus

$$P\left(\sqrt{n_T}\left(\widehat{Q}_Z(1/2) - m\right) > \delta\right) = P\left(\widehat{Q}_Z(1/2) > m + \delta/\sqrt{n_T}\right) = P\left(G_{n_T}(m + \delta/\sqrt{n_T}) < 0\right).$$

But what is the law of $G_{n_T}(m + \delta/\sqrt{n_T})$?

Now, conditioning on $\{\mu_u, \sigma_u\}_{u\in[0,1]}$, the

$$\begin{aligned}
\mathrm{E}[G_{n_T}(\mu_T^* + \delta/\sqrt{n_T})] &= \frac{1}{n_T}\sum_{t=1}^{n_T} F_{Z_{i,n}}(\mu_T^* + \delta/\sqrt{n_T}) - 1/2 \\
&= \frac{1}{n_T}\sum_{t=1}^{n_T} F_{v_{1,n}}((\mu_T^* + \delta/\sqrt{n_T} - \mu_{i/n})/\sigma_{i/n}) - 1/2 \\
&= \frac{1}{n_T}\sum_{i=1}^{n_T} F_{v_{1,n}}((\mu_T^* - \mu_{i/n})/\sigma_{i/n}) - 1/2 \\
&\quad + (\delta/\sqrt{n_T})\frac{1}{n_T}\sum_{t=1}^{n_T}\frac{1}{\sigma_{i/n}} f_{v_{1,n}}((\mu_T^* - \mu_{i/n})/\sigma_{i/n}) + O(n_T^{-1}).
\end{aligned}$$

Think abstractly in terms of Riemann sums

$$\left|\frac{1}{\lfloor nT\rfloor}\sum_{i=1}^{\lfloor nT\rfloor} g(i/n) - \frac{1}{T}\int_0^T g(u)\mathrm{d}u\right| = \left|\frac{1}{T}\sum_{i=1}^{\lfloor nT\rfloor}\left\{\frac{1}{n}g(i/n) - \int_{(i-1)/n}^{i/n} g(u)\mathrm{d}u\right\}\right| + O(n_T^{-1})$$

$$\left|\frac{1}{T}\sum_{i=1}^{\lfloor nT\rfloor}\left\{\frac{1}{n}g(i/n) - \int_{(i-1)/n}^{i/n} g(u)\mathrm{d}u\right\}\right| \leq \frac{1}{T}\sum_{i=1}^{\lfloor nT\rfloor}\left|\frac{1}{n}g(i/n) - \int_{(i-1)/n}^{i/n} g(u)\mathrm{d}u\right|.$$

If g is continuous, let $\varepsilon > 0$. Then there exists a $\delta > 0$ such that $|g(x) - g(y)| < \varepsilon/(b-a)$ for all $x, y \in [b, a]$ such that $|x - y| < \delta$.

$$\left|\frac{1}{n}g(i/n) - \int_{(i-1)/n}^{i/n} g(u)\mathrm{d}u\right| = \left|\int_{(i-1)/n}^{i/n}\{g(i/n) - g(u)\}\,\mathrm{d}u\right| \leq \frac{\varepsilon}{n}\int_{(i-1)/n}^{i/n} 1du = \frac{\varepsilon}{n^2}$$

So

$$\left|\frac{1}{\lfloor nT\rfloor}\sum_{i=1}^{\lfloor nT\rfloor} g(i/n) - \frac{1}{T}\int_0^T g(u)du\right| \leq \frac{\varepsilon}{n} + O(1/(Tn)).$$

Now

$$F_{v_1}((\mu_T^* - \mu_u)/\sigma_u)$$

is continuous in u, so using the above properties of Riemann sums

$$\begin{aligned}\frac{1}{n_T}\sum_{i=1}^{n_T} F_{v_{1,n}}((\mu_T^* - \mu_{i/n})/\sigma_{i/n}) - 1/2 &= \frac{1}{T}\int_0^T F_{v_{1,n}}\left\{(\mu_T^* - \mu_u)/\sigma_u\right\} \mathrm{d}u - 1/2 + O(n^{-1}) + O(n_T^{-1}) \\ &= F_{T,n}(\mu_{T,n}^*) - 1/2 + O(n^{-1}) + O(n_T^{-1}) \\ &= O(n^{-1}) + O(n_T^{-1}),\end{aligned}$$

where the first line is due to the properties of Riemann integrals. As $\sigma_u > 0$ and $f_{v_1}((\mu_T^* - \mu_u)/\sigma_u)$ is again continuous in u, then

$$\frac{1}{n_T}\sum_{t=1}^{n_T} \frac{1}{\sigma_{\Delta i}} f_{v_{1,n}}((\mu_T^* - \mu_{i/n})/\sigma_{i/n}) = \frac{1}{T}\int_0^T \frac{1}{\sigma_u} f_{v_{1,n}}((\mu_T^* - \mu_u)/\sigma_u)\mathrm{d}u + O(n^{-1}) + O(n_T^{-1}),$$

so

$$\mathrm{E}[G_n(\mu_T^* + \delta/\sqrt{n_T})] = (\delta/\sqrt{n_T})\,\phi_{T,n}(\mu_T^*) + O(n^{-1}) + O(n_T^{-1}).$$

Now

$$\begin{aligned}n_T \times \mathrm{Var}[G_n(\mu_T^* + \delta/\sqrt{n_T})] &= \frac{1}{n_T}\sum_{i=1}^{n_T} F_{Z_{i,n}}(\mu_T^* + \delta/\sqrt{n_T})\left\{1 - F_{Z_{i,n}}(\mu_T^* + \delta/\sqrt{n_T})\right\} \\ &= \frac{1}{n_T}\sum_{i=1}^{n_T} F_{v_{1,n}}((\mu_T^* + \delta/\sqrt{n_T} - \mu_{i/n})/\sigma_{i/n})\left\{1 - F_{v_{1,n}}((\mu_T^* + \delta/\sqrt{n_T} - \mu_{i/n})/\sigma_{i/n})\right\} \\ &= \omega_{T,n}^2(\mu_{T,n}^*) + O(1/\sqrt{n_T})\end{aligned}$$

Then the triangular array CLT applies as each term in the average is bounded. Thus

$$\frac{\sqrt{n_T}\left\{G_n(\mu_T^* + \delta/\sqrt{n_T}) - \mathrm{E}[G_n(\mu_T^* + \delta/\sqrt{n_T})]\right\}}{\sqrt{n \times \mathrm{Var}[G_n(\mu_T^* + \delta/\sqrt{n_T})]}} \xrightarrow{d} N(0,1).$$

So, conditioning on $\{\mu_u, \sigma_u\}_{u\in[0,1]}$,

$$\begin{aligned}&P\left(\sqrt{n}\left(\widehat{Q}_{Z_n}(1/2) - \mu_T^*\right) > \delta\right) \\ =\ &P\left(G_n(\mu_T^* + \delta/\sqrt{n_T}) < 0\right) \\ =\ &P\left(\frac{G_n(\mu_T^* + \delta/\sqrt{n_T}) - \mathrm{E}[G_n(\mu_T^* + \delta/\sqrt{n_T})]}{\sqrt{n_T \times \mathrm{Var}[G_n(\mu_T^* + \delta/\sqrt{n_T})]}} < \frac{-\mathrm{E}[G_n(\mu_T^* + \delta/\sqrt{n_T})]}{\sqrt{n_T \times \mathrm{Var}[G_n(\mu_T^* + \delta/\sqrt{n_T})]}}\right) \\ =\ &F_{N(0,1)}\left(\frac{-\mathrm{E}[G_n(\mu_T^* + \delta/\sqrt{n_T})]}{\sqrt{n_T \times \mathrm{Var}[G_n(\mu_T^* + \delta/\sqrt{n_T})]}}\right) + O(1/\sqrt{n_T}) \\ =\ &F_{N(0,1)}\left(-\delta\frac{\phi_{T,n}(\mu_{T,n}^*)}{\sqrt{\omega_{T,n}^2(\mu_{T,n}^*)}}\right) + O(1/\sqrt{n_T}).\end{aligned}$$

Thus, conditionally

$$\frac{\sqrt{n_T}\left(\widehat{Q}_Z(1/2) - \mu_{T,n}^*\right)}{\omega_{T,n}(\mu_{T,n}^*)/\phi_{T,n}(\mu_T^*)} \xrightarrow{d} N(0,1).$$

As the right hand side is a pivot, this result also holds unconditionally.

5.3 Proof of Lemma 2

This follows as $F(\mu) = 1/2$, then expanding out by mean value theorem

$$\begin{aligned} F(\mu) &= \frac{1}{T}\int_0^T F_{v_1}\left\{(\mu-\mu_u)/\sigma_u\right\}\mathrm{d}u, \\ &= \frac{1}{T}\int_0^T F_{v_1}(0)\,\mathrm{d}u + \frac{1}{T}\int_0^T \left\{(\mu-\mu_u)/\sigma_u\right\} f_{v_1}(\widetilde{z}_u)\,\mathrm{d}u \\ &= 1/2 + \frac{1}{T}\int_0^T (\mu/\sigma_u) f_{v_1}(\widetilde{z}_u)\,\mathrm{d}u - \frac{1}{T}\int_0^T (\mu_u/\sigma_u) f_{v_1}(\widetilde{z}_u)\,\mathrm{d}u, \end{aligned}$$

so solving gives the stated representation.

5.4 Proof of Theorem 2

Recall μ_T^* solves $F_T(\mu_T^*) = 1/2$, where

$$F_T(m) = \frac{1}{T}\int_0^T F_v(m, x_u)\mathrm{d}u, \quad F_v(m,x) = F_v((m-x_1)/x_2),$$

where $x_u = (\mu_u, \sigma_u)^{\mathrm{T}}$. Let

$$\begin{aligned} F^{(x)}(m,x) &= \frac{\partial F_v(m,x)}{\partial x} = \begin{pmatrix} \frac{\partial (m-x_1)/x_2}{\partial x_1} \\ \frac{\partial (m-x_1)/x_2}{\partial x_2} \end{pmatrix} f_{v_1}(m,x) = -c(x)\frac{1}{x_2} f_{v_1}(m,x), \\ F^{(m)}(m,x) &= \frac{\partial F_v(m,x)}{\partial m} = \frac{1}{x_2} f_{v_1}(m,x), \quad c(x) = \begin{pmatrix} 1 \\ \frac{x_1}{x_2} \end{pmatrix}. \end{aligned}$$

Notice that for all vectors γ,

$$\frac{F^{(x)}(m,x)^{\mathrm{T}}\gamma}{F^{(m)}(m,x)} = -c(x)^{\mathrm{T}}\gamma,$$

which is invariance to F_v and m.

Now

$$\begin{aligned} G_t(m) &= F(m, X_0) - 1/2 + \frac{1}{t}\int_0^t \left\{F(m,X_u) - F(m,X_0)\right\}\mathrm{d}u \\ &= F(m, X_0) - F(m_0, X_0) + \frac{1}{t}\int_0^t \left\{F(m,X_u) - F(m,X_0)\right\}\mathrm{d}u, \end{aligned}$$

as $F(m_0, X_0) = 1/2$. Now, by Ito's lemma,

$$\begin{aligned} \mathrm{d}F(m, X_t) &= F^{(x)}(m,X_t)^{\mathrm{T}}\mathrm{d}X_t + \frac{1}{2}\iota^{\mathrm{T}} F^{(x,x)}(m,X_t)\iota \mathrm{d}t \\ &= F^{(x)}(m,X_t)^{\mathrm{T}}\Sigma_{X,0}^{1/2}\mathrm{d}W_{X,t} + \mu_{F,t}\mathrm{d}t, \quad \mu_{F,t} = F^{(x)}(m,X_t)^{\mathrm{T}}\mu_{X,t} + \frac{1}{2}\iota^{\mathrm{T}} F^{(x,x)}(m,X_t)\iota. \end{aligned}$$

So for small t,

$$\begin{aligned} \frac{1}{t}\int_0^t \left\{F(m,X_u) - F(m,X_0)\right\}\mathrm{d}u &\simeq F^{(x)}(m,X_0)\Sigma_{X,0}^{1/2}\frac{1}{t}\int_0^t \mathrm{d}W_{X,u} + \mu_{F,0}\frac{1}{t}\int_0^t u\mathrm{d}u \\ &\sim N(t\mu_{F,0}/2, tF^{(x)}(m,X_t)^{\mathrm{T}}\Sigma_{X,0}F^{(x)}(m,X_t)/3), \end{aligned}$$

as $\int_0^t \mathrm{d}W_{X,u} \sim N(0, t^3/3)$. Likewise

$$F(m_t, X_0) = F(m_0, X_0) + (m_t - m_0)\, F^{(m)}(m_0, X_0),$$

so, for small t,

$$\{\mu_T^* - \mu_0\}|X_0, \mu_0 \dot{\sim} N\left(-\frac{t\left\{F^{(x)}(\mu_0, X_0)^{\mathrm{T}}\mu_{X,0} + \frac{1}{2}\iota^{\mathrm{T}}F^{(x,x)}(\mu_0, X_0)\iota\right\}/2}{F^{(m)}(\mu_0, X_0)}, t\frac{F^{(x)}(\mu_0, X_0)^{\mathrm{T}}\Sigma_{X,0}F^{(x)}(\mu_0, X_0)}{3F^{(m)}(\mu_0, X_0)^2}\right).$$

In the diffusion case, the mean square error (MSE) of $\widehat{Q}_{Z_1}(1/2) - \mu_0$ is roughly

$$c_2 T + \frac{1}{nT}\frac{1}{4\left\{\sigma_0^{-1} f_v(0)\right\}^2},$$

which is minimized at

$$T = d_2 n^{-1/2}, \quad d_2 = \frac{1}{\sqrt{c_2} 2\sigma_0^{-1} f_v(0)}.$$

In the differentiable case, the MSE roughly

$$c_3 T^2 + \frac{1}{nT}\frac{1}{4\left\{\sigma_0^{-1} f_v(0)\right\}^2},$$

which is minimized at

$$T = d_3 n^{-1/3}, \quad d_3 = \frac{1}{\left[8c_3\left\{\sigma_0^{-1} f_v(0)\right\}^2\right]^{1/3}}.$$

In the differentiable case where smoothing is carried out, then MSE is

$$c_4 T^4 + \frac{1}{nT}\frac{1}{4\left\{\sigma_0^{-1} f_v(0)\right\}^2},$$

which is minimized at $T = d_4 n^{-1/5}$. In the diffusion case then the filter and the smoother have the same bandwidth, although in finite samples the smoother case will likely perform better as it removes drift type terms.

5.5 Simulation results for stochastic volatility when $r = 2$

Economists typically focus on the $r = 2$, that is σ_t^2 is the estimand. Table 6 contains the results from the experiment results reported in Examples 1 and 5 but now when $r = 2$, not $r = 2/3$. Only results for the sample median and realized bipower statistics are reported as the realized power variation statistic has a different estimand when $r = 2$.

The results for bipower variation are terrific when there are no jumps, but the CLT result is simply not a helpful guide when there are stable jumps for any value of α.

The results for the sample median based estimator, $\widehat{c_t}^2$, are in line with the theory again, but their finite sample behaviour is quite weak. The reason of course is that the sample median is based on scaled χ_1^2 variables and so the sample median is quite biased for small sample sizes.

The shrunk sample median $\widetilde{c_t^2} = \widehat{c_t}^2/(1 + 3b_{2/3}^2/n_T)$ has much better finite properties — of course the shrinkage factor makes no difference to the limit theory. The results in Table 6 show quite good results for $\alpha < 1$, but $\sqrt{n_T}$ times the bias increasing as α increases.

n	n_T	no jumps E(ε)	no jumps $sd(\varepsilon)$	$\alpha=0.3$ E(ε)	$\alpha=0.3$ $sd(\varepsilon)$	$\alpha=0.6$ E(ε)	$\alpha=0.6$ $sd(\varepsilon)$	$\alpha=0.9$ E(ε)	$\alpha=0.9$ $sd(\varepsilon)$	$\alpha=1.2$ E(ε)	$\alpha=1.2$ $sd(\varepsilon)$	$\alpha=1.5$ E(ε)	$\alpha=1.5$ $sd(\varepsilon)$	$\alpha=1.8$ E(ε)	$\alpha=1.8$ $sd(\varepsilon)$	no jumps asy $sd(\varepsilon)$
Median																
200	21	0.300	2.61	0.425	2.68	0.545	2.75	0.786	2.88	1.43	3.21	2.98	4.01	6.41	5.61	2.33
10^3	42	0.296	2.52	0.342	2.53	0.410	2.55	0.592	2.62	1.07	2.80	2.62	3.37	7.27	4.94	2.33
10^4	116	0.187	2.43	0.203	2.45	0.216	2.45	0.316	2.47	0.696	2.56	2.23	2.88	8.95	4.31	2.33
10^5	317	0.116	2.41	0.130	2.41	0.123	2.41	0.165	2.41	0.439	2.44	1.95	2.66	11.3	3.89	2.33
10^6	874	0.065	2.38	0.079	2.39	0.072	2.38	0.095	2.38	0.273	2.39	1.77	2.53	14.7	3.55	2.33
Bipower																
200	21	-0.061	1.85	1×10^9	5×10^{11}	5299	1×10^6	31.4	4207	4.69	78.8	5.00	82.1	6.79	8.90	1.61
10^3	42	0.020	1.81	9×10^8	2×10^{11}	1×10^5	3×10^7	13.3	2771	3.22	43.4	4.26	11.1	8.21	8.16	1.61
10^4	116	0.014	1.75	6×10^8	2×10^{11}	31.0	3×10^3	2.27	83.8	2.05	23.7	4.15	54.2	10.6	5.81	1.61
10^5	317	0.017	1.72	2×10^5	7×10^7	5.91	644	1.08	62.7	1.53	66.6	3.52	11.0	13.8	6.60	1.61
10^6	874	0.007	1.69	4098	9×10^5	3.20	439	0.51	27.6	0.962	36.7	3.31	6.90	18.1	5.74	1.61
Shrunk Median																
200	21	-0.088	2.40	0.026	2.46	0.136	2.53	0.358	2.65	0.951	2.96	2.38	3.69	5.5	5.16	2.33
10^3	42	0.003	2.41	0.047	2.42	0.112	2.45	0.286	2.51	0.749	2.68	2.23	3.22	6.6	4.73	2.33
10^4	116	0.015	2.40	0.031	2.41	0.044	2.41	0.142	2.43	0.517	2.52	2.02	2.83	8.6	4.24	2.33
10^5	317	0.014	2.40	0.028	2.40	0.021	2.40	0.062	2.40	0.335	2.43	1.83	2.64	11.2	3.87	2.33
10^6	874	0.003	2.38	0.018	2.39	0.010	2.37	0.033	2.38	0.211	2.39	1.70	2.52	14.6	3.54	2.33

Table 6: Results from the spot volatility simulation experiments where $\varepsilon = \sqrt{n_T}\left(\widehat{c^2_{0}}_{T,n} - c_0^2\right)$, for three estimators: the sample median, realized bipower variation and the shrunk sample median. When α is less than around 1.2 then ε has a Gaussian limit law with mean 0 and variance of the stated asymptotic standard deviation in the median and the shrunk median case. Otherwise these estimators are consistent but limit law does not hold.

References

Andersen, T. and L. Benzoni (2009). Realized volatility. In T. G. Andersen, R. A. Davis, J. P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp. 555–575. Springer-Verlag.

Andersen, T. and L. Benzoni (2011). Stochastic volatility. In R. A. Meyers (Ed.), *Complex Systems in Finance and Econometrics*, pp. 694–726. Springer.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of exchange rate volatility. *Journal of the American Statistical Association 96*, 42–55.

Andersen, T. G., D. Dobrev, and E. Schaumburg (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics 169*, 75–93.

Angrist, J., V. Chernozhukov, and I. Fernandez-Val (2006). Quantile regression under misspecification, with an application to the U.S. wage structure. *Econometrica 74*, 539–563.

Arisa-Castro, E. and D. L. Donoho (2009). Does median filtering theory truly preserve edges better than linear filtering. *Annals of Statistics 37*, 1172–1206.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B 64*, 253–280.

Barndorff-Nielsen, O. E. and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics 2*, 1–48.

Barndorff-Nielsen, O. E. and N. Shephard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics 4*, 1–30.

Bollerslev, T. (2022). Realized semi(co)variation: Signs that all volatilities are not created equal. *Journal of Financial Econometrics 20*, 219–252.

Bollerslev, T., J. Li, and Z. Liao (2021). Fixed-k inference for volatility. *12*, 1053–1084.

Carraneo, M. and R. Titijunik (2022). Regression discontinuity designs. *Annal Review of Economics 14*, 821–851.

Cressie, N. (1993). *Statistics for Spatial Data*. New York: Wiley.

Dachraoui, K. (2018). On the optimality of target volatility strategies. *Journal of Portfolio Management 44*, 58–67.

Dovonon, P., S. Goncalves, U. Hounyo, and N. Meddahi (2019). Bootstrapping high-frequency jump tests. *Journal of the American Statistical Association 114*, 793–803.

Fried, R., J. Einbeck, and U. Gather (2007). Weighted repeated median smoothing and filtering. *Journal of the American Statistical Association 102*, 1300–1308.

Gneiting, T., H. Sevcikova, and D. B. Percival (2012). Estimators of fractal dimension: Assessing the roughness of time series and spatial data. *Statistical Science 27*, 247–277.

Goncalves, S. and N. Meddahi (2009). Bootstrapping realized volatility. *Econometrica 77*, 283–306.

Hampel, F., E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel (2005). *Robust Statistics: The Approach based on Influence Functions*. New York: Wiley.

Harvey, C. R., E. Hoyle, R. Korgaonkar, S. Rattray, M. Sargaison, and O. Van Hemert (2018). The impact of volatility targeting. *Journal of Portfolio Management 45*, 14–33.

Jacod, J. (2007). Asymptotic properties of power variations of Lévy processes. *ESAIM — Probability and Statistics 11*, 173–196.

Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Processes and Their Applications 118*, 517–559.

Jacod, J. and P. Protter (2012). *Discretization of Processes*. Heidelberg: Springer.

Justusson, B. (1981). Median filtering: Statistical properties. In *Two-Dimensional Digital Signal Prcessing II: Topics in Applied Physics*, Volume 43. Springer.

Kloeden, P. E. and E. Platen (1992). *Numerical Solutions to Stochastic Differential Equations*. New York: Springer.

Koch, I. (1996). On the asymptotic performance of median smoothers in image analysis and nonparametric regression. *Annals of Statistics 24*, 1648–1666.

Koenker, R. (2005). *Quantile Regression*. Cambridge: Cambridge University Press.

Lépingle, D. (1976). La variation d'ordre p des semi-martingales. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete 36*, 295–316.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric Brownian motion. *Giornale dell'Istituto Italiano degi Attuari LXIV*, 19–47.

Mancini, C. (2004). Estimation of the characteristics of jump of a general Poisson-diffusion process. *Scandinavian Actuarial Journal 1*, 42–52.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics 36*, 270–296.

Roncalli, T. (2014). *Introduction to Risk Parity and Budgeting.* Boca Raton: Chapman & Hall.

Sen, P. K. (1968). Asymptotic normality of sample quantiles for m-dependent processes. *Annals of Mathematical Statistics 39*, 1724–1730.

Shephard, N. (2005). *Stochastic Volatility: Selected Readings.* Oxford: Oxford University Press.

Stein, M. L. (1999). *Interpolation of Spatial Data.* Springer-Verlag.

Tang, W., L. Zhang, and S. Banerjee (2021). On identifiability and consistency of the nugget in Gaussian spatial process models. *Journal of the Royal Statistical Society, Series B 83*, 1044–1070.

Terrell, G. R. (2003). The Wilson-Hilferty transformation is locally saddlepoint. *Biometrika 90*, 445–453.

Todorov, V. and G. Tauchen (2011). Limit theorems for power variations of pure-jump processes with application to activity estimation. *The Annals of Applied Probability 21*, 546–588.

Tukey, J. W. (1971). *Exploratory Data Analysis.* Pearson.

Wilson, E. and M. Hilferty (1931). The distribution of chi-square. *Proc. Nat. Acad. 17*, 684–688.