Proof of Proposition 5.1 from the paper 'Computer-intensive rate estimation, diverging statistics, and scanning' by McElroy and Politis (2005).

Tucker McElroy
U.S. Bureau of the Census

Dimitris N. Politis
University of California, San Diego

Set-up: Consider a time series defined as $X_t = \sigma_t G_t$ for $t \in \mathbf{Z}$ where: (A) The series $\{\sigma_t\}$ and $\{G_t\}$ are independent; (B) $\sigma_t = \sqrt{\epsilon_t}$, where ϵ_t are positive and i.i.d. with cdf in $D(\alpha/2)$ for some $\alpha \in (0,2)$; and (C) $\{G_t, t \in \mathbf{Z}\}$ is a stationary, mean zero, purely non-deterministic Gaussian time series.

Also consider the condition $LM(\zeta)$ defined for some $\zeta \in [0,1)$ as

$$\sum_{|h|<n} R(h) \sim C n^{\zeta} \quad \textit{and} \quad \sum_{|h|<n} |R(h)| = O(n^{\zeta}) \quad \text{as} \quad n \to \infty$$

where $C > 0$ is a constant, and R is the autocovariance function of $\{G_t\}$. The series $\{G_t\}$ and $\{X_t\}$ are said to have *long memory* if $LM(\zeta)$ holds with $\zeta \in (0,1)$; $LM(0)$ will denote intermediate memory, i.e., weak dependence.

One easy consequence of $LM(\zeta)$ for $\zeta > 0$ is that $R(n) \sim \frac{\zeta C}{2} n^{\zeta - 1}$ since $\sum_{|h|<n} R(h) = 2\sum_{h=1}^{n-1} R(h)$ and L'Hopital's Rule gives

$$\frac{C}{2} = \lim_{n\to\infty} \frac{\sum_{h=1}^{n} R(h)}{n^{\zeta}} = \lim_{n\to\infty} \frac{R(n)}{\zeta n^{\zeta-1}}.$$

Interestingly, when appropriately normalized, the sample second moment converges in distribution in this general setting as the following Proposition demonstrates.

Proposition 5.1 *Assume assumptions (A) through (C), as well as* $LM(\zeta)$. *Also suppose that* ϵ_t *is absolutely continuous with a bounded pdf* f_ϵ *that is ultimately monotone, i.e.,* f_ϵ *is monotone on* (z,∞) *for some* $z > 0$*, and is monotone on* $(-\infty, u)$ *for some* $u < 0$*. Then, we have*

$$a_n^{-2} \sum_{t=1}^{n} X_t^2 \stackrel{\mathcal{L}}{\Longrightarrow} W \text{ as } n \to \infty,$$

where $a_n = n^{1/\alpha} K(n)$ *for some slowly varying function* $K(n)$*. In the above,* W *is* $\alpha/2$*-stable with scale* $C_{\alpha/2}^{-2/\alpha}(E|G_t|^\alpha)^{2/\alpha}$*, skewness* 1*, and location zero, and the constants* C_p^{-1} *are defined by*

$$C_p^{-1} = \frac{\Gamma(2-p)cos(\pi p/2)}{1-p}.$$

The limit theorem of Proposition 5.1 is interesting as it shows that the convergence of the sample second moment does *not* depend on the long memory parameter ζ.

Proof of Proposition 5.1 Since the cdf of ϵ_t is $F_\epsilon \in D(\alpha/2)$, we can find a slowly-varying function L and constant $C > 0$ such that

$$1 - F_\epsilon(x) = \frac{C + o(1)}{x^{\alpha/2}} L(x)$$

for $x > 0$, by Theorem 2.2.8 of Embrechts, Klüppelberg, and Mikosch (1997). Define a new slowly-varying function H by $H(x) = (1 - F_\epsilon(x))\frac{x^{\alpha/2}}{C}$. Then we can write $1 - F_\epsilon(x) = x^{-\alpha/2} H(x) C$ for $x > 0$. Applying Theorem $A.3.7$ of

Embrechts et al. (1997), we find that $f_\epsilon(x) \sim C\frac{\alpha}{2}x^{-(\alpha/2+1)}H(x)$ as $x \to \infty$. Hence it follows that

$$\frac{x^{1+\alpha/2}f_\epsilon(x)}{H(x)} \to \frac{\alpha}{2}C \tag{1}$$

as $x \to \infty$. Now by Proposition 2.2.13 of Embrechts et al. (1997), we can choose the normalizing sequence a_n for the partial sums by solving $\frac{1}{n} = P[\epsilon > a_n^2] = a_n^{-\alpha}H(a_n^2)C$ so that $C = \frac{a_n^\alpha}{nH(a_n^2)}$. We also see that $a_n \sim n^{1/\alpha}C^{1/\alpha}K(n)$ for a slowly-varying function K; hence, the scale C appears in the rate a_n, so that the limit of the normalized sample moments will not depend on the scale of the original data. The characteristic function of the sample second moments, normalized by the sequence a_n^2 and conditional on $\mathcal{G} = \{G_t, t \in \mathbf{Z}\}$, is

$$\phi(\nu) = E\left[\exp\{i\,\nu a_n^{-2}\sum_{t=1}^n X_t^2\}|\mathcal{G}\right] = \Pi_{t=1}^n E\left[\exp\{i\,\nu a_n^{-2}\epsilon_t G_t^2\}|\mathcal{G}\right].$$

Taking the complex logarithm yields

$$\begin{aligned}\log\phi(\nu) &= \sum_{t=1}^n \log\left(E\left[\exp\{i\,\nu a_n^{-2}\epsilon_t G_t^2\}|\mathcal{G}\right]\right)\\ &= \sum_{t=1}^n \log\left(1 + \int_0^\infty (\exp\{i\,\nu a_n^{-2}xG_t^2\} - 1)\,f_\epsilon(x)dx\right),\end{aligned}$$

which uses the fact that f_ϵ integrates to one over the positive real line. After a change of variable, the integral becomes

$$\begin{aligned}&\int_0^\infty (\exp\{i\nu y\} - 1)\,f_\epsilon(a_n^2 y G_t^{-2})\,a_n^2 G_t^{-2}dy \qquad (2)\\ &= \frac{nH(a_n^2G_t^{-2})}{a_n^\alpha}|G_t|^\alpha n^{-1}\int_0^\infty (\exp\{i\,\nu y\} - 1)\,y^{-(1+\alpha/2)}\\ &\frac{H(a_n^2yG_t^{-2})}{H(a_n^2G_t^{-2})}\,\frac{(a_n^2yG_t^{-2})^{1+\alpha/2}f_\epsilon(a_n^2yG_t^{-2})}{H(a_n^2yG_t^{-2})}dy.\end{aligned}$$

Now for each fixed $y > 0$ and any G_t,

$$\frac{H(a_n^2 y G_t^{-2})}{H(a_n^2 G_t^{-2})} \frac{(a_n^2 y G_t^{-2})^{1+\alpha/2} f_\epsilon(a_n^2 y G_t^{-2})}{H(a_n^2 y G_t^{-2})} \to \frac{\alpha}{2} C$$

as $n \to \infty$, due to (1) and the fact that H is slowly-varying. This convergence is uniformly bounded in y and G_t. In order to apply the Dominated Convergence Theorem and take the limit inside the integral, we must bound the integrand by an integrable function. For y in a neighborhood of zero, use the bound $|\exp\{i\,\nu y\} - 1| \leq \tilde{C} y$ for some constant $\tilde{C} > 0$. Since $y y^{-(1+\alpha/2)} = y^{\alpha/2}$ is integrable about zero if $\alpha < 2$, we can use $y^{\alpha/2}$ times an appropriate constant as our dominating function near the origin. Outside a compact set, use the bound $|\exp\{i\,\nu y\} - 1| \leq 2$. Since $y^{-(1+\alpha/2)}$ is integrable outside a compact set if $\alpha > 0$, we can use $2y^{-(1+\alpha/2)}$ as our dominating function.

In order to compute this integral, note that $\int_0^\infty \sin(x)\, x^{-p} dx = C_p^{-1}$, so that

$$\begin{aligned}
\int_0^\infty (\cos(\nu y) - 1) y^{-(1+\alpha/2)} dy &= \int_0^\infty (\cos(|\nu| y) - 1) y^{-(1+\alpha/2)} dy \\
&= |\nu|^{\alpha/2} \int_0^\infty (\cos(v) - 1) v^{-(1+\alpha/2)} dv \\
&= |\nu|^{\alpha/2} \left(-\frac{2}{\alpha}\right) \int_0^\infty \sin(v)\, v^{-\alpha/2} dv \\
&= |\nu|^{\alpha/2} \left(-\frac{2}{\alpha}\right) C_{\alpha/2}^{-1}
\end{aligned}$$

and

$$\begin{aligned}
\int_0^\infty \sin(\nu y)\, y^{-(1+\alpha)} dy &= sign(\nu) \int_0^\infty \sin(|\nu| y)\, y^{-(1+\alpha/2)} dy \\
&= sign(\nu)\, |\nu|^{\alpha/2} \int_0^\infty \sin(v)\, v^{-(1+\alpha/2)} dv \\
&= sign(\nu)\, |\nu|^{\alpha/2} C_{1+\alpha/2}^{-1}
\end{aligned}$$

where $sign(x)$ denotes the sign function. At this point it is useful to observe, using the definition of C_p^{-1} and the Gamma function, that

$$C_{p-1}^{-1} = \frac{\Gamma(3-p)\cos(\pi(p-1)/2)}{2-p} = \Gamma(2-p)\sin(\pi p/2) = (1-p)C_p^{-1}\tan(\pi p/2).$$

Therefore

$$\int_0^\infty (\exp\{i\,\nu y\}-1)y^{-(1+\alpha/2)}\,dy = -|\nu|^{\alpha/2}\frac{2}{\alpha}C_{\alpha/2}^{-1}\left(1-i\,sign(\nu)\,\tan(\pi\alpha/4)\right).$$

Now it is easy to show that $\frac{nH(a_n^2G_t^{-2})}{a_n^\alpha} \to C^{-1}$ as $n\to\infty$ for each fixed t, since H is slowly-varying and by our previous expression for C. Now (2) is $O(1/n)$, so that

$$\begin{aligned}\log\phi(\nu) &= o(1) + \frac{1}{n}\sum_{t=1}^n |G_t|^\alpha C^{-1}\left(-|\nu|^{\alpha/2}\,\frac{2}{\alpha}\,C_{\alpha/2}^{-1}\left(1-isign(\nu)\,\tan(\pi\alpha/4)\right)\frac{\alpha}{2}C\right)\\ &= o(1) + \frac{1}{n}\sum_{t=1}^n |G_t|^\alpha\left(-|\nu|^{\alpha/2}C_{\alpha/2}^{-1}\left(1-isign(\nu)\,\tan(\pi\alpha/4)\right)\right).\end{aligned}$$

It is convenient to exponentiate $\log\phi(\nu)$ at this point. Next, we note that G_t obeys the weak law of large numbers. To see this, note that by Theorem 3.1 of Taqqu (1975) for $H > 1/2$, the variance of $\sum_{t=1}^n |G_t|^\alpha - E|G|^\alpha$ is $O(n^{2H}K(n))$ for a slowly varying function K and H satisfying $m(\zeta-1) = 2H-2$, where m is the Hermite rank of $h(x) = |x|^\alpha - E|Z|^\alpha$ (Z denotes a standard normal random variable). We claim that $m=2$. Recall that the Hermite rank of a function h is the index of the first nonzero coefficient in the Hermite polynomials expansion of h. Following Taqqu (1975), let

$$J(q) = E[h(Z)H_q(Z)] \qquad H_q(x) = (-1)^q e^{x^2/2}\frac{d^q}{dx^q}e^{-x^2/2}$$

where H_q is the qth Hermite polynomial. In particular,

$$H_0(x) = 1 \qquad H_1(x) = x \qquad H_2(x) = x^2-1$$

are the first several Hermite polynomials, which are orthonormal with respect to the standard normal distribution. Then we have

$$
\begin{aligned}
J(0) &= 0 \\
J(1) &= E[Z\,|Z|^{\alpha}] = 0 \\
J(2) &= E[(|Z|^{\alpha} - E|Z|^{\alpha})(Z^2 - 1)] \\
&= E[|Z|^{2+\alpha}] - E[|Z|^{\alpha}] \\
&= 2^{\frac{\alpha+2}{2}}\pi^{-1/2}\Gamma((\alpha+3)/2) - 2^{\alpha/2}\pi^{-1/2}\Gamma((\alpha+1)/2) \\
&= \pi^{-1/2}2^{\alpha/2}\Gamma((\alpha+1)/2)\alpha \neq 0
\end{aligned}
$$

This shows that $m = 2$. Hence $H = \zeta$, and the variance of the centered sum is $O(n^{2\beta}K(n))$; actually, examining the proof, we see that K is unnecessary in this case. However, this argument only holds for $\zeta > 1/2$ (since H is greater than 1/2 in Taqqu (1975)). When $\zeta \leq 1/2$, a similar bound for the variance also holds. Following the proof of Theorem 3.1 in Taqqu (1975), we see that

$$Var[\sum_{t=1}^{n} h(G_t)] = \sum_{s,t}^{n} \sum_{q \geq m} \frac{J(q)^2}{q!} R_G^q(t-s) = \sum_{q \geq m} \frac{J(q)^2}{q!} \sum_{s,t}^{n} R_G^q(t-s).$$

Now, if $q \geq 2$, $R_G^q(h)$ is summable if $\zeta < 1/2$, but grows at a logarithmic rate if $\zeta = 1/2$ and $q = 2$. Thus a conservative bound is $\sum_{t,s}^{n} R_G^q(t-s) = O(n \log n)$ (if $\beta < 1/2$, we may dispense with the log) for all $q \geq 2$. Also by Bessel's inequality, we have

$$\sum_{q \geq m} \frac{J(q)^2}{q!} \leq \frac{1}{m!} \sum_{q} (E[h(Z)H_q(Z)])^2 \leq \frac{1}{m!}\|h\|_2^2 < \infty.$$

So the variance of the sum is of order $n^{2\beta}$, $n \log n$, or n depending on whether $\beta > 1/2$, $\beta = 1/2$, or $\beta < 1/2$ respectively. In any event, the Weak Law of Large Numbers holds for all these cases. Therefore $\frac{1}{n}\sum_{t=1}^{n} |G_t|^{\alpha} \xrightarrow{P} E|G|^{\alpha}$

and hence it also converges in distribution. Now using the boundedness of the function $\exp\{-|\nu|^{\alpha/2} C_{\alpha/2}^{-1}(1 - i\, sign(\nu)\, \tan(\pi\alpha/4))\cdot\}$, it follows from weak convergence that

$$E\left[\exp\{-i\,\nu a_n^{-2}\sum_{t=1}^{n} Y_t^2\}\right] \to E\left[\exp\{-|\nu|^{\alpha/2} C_{\alpha/2}^{-1} E|G|^{\alpha}(1 - i\, sign(\nu)\, \tan(\pi\alpha/4))\}\right]$$
$$= \exp\{-|\nu|^{\alpha/2} C_{\alpha/2}^{-1} E|G|^{\alpha}(1 - i\, sign(\nu)\, \tan(\pi\alpha/4))\},$$

which is the characteristic function of an $\alpha/2$ stable variable with skewness one, location zero, and scale $C_{\alpha/2}^{-2/\alpha}(E|G|^{\alpha})^{2/\alpha}$. Finally, from Samorodnitsky and Taqqu (1994, p. 142), we have $E|G|^{\alpha} = (R_G(0))^{\alpha/2} 2^{\alpha/2} \pi^{-1/2} \Gamma(1 + \alpha/2)$ so that the limiting scale is $C_{\alpha/2}^{-2/\alpha} R_G(0) 2\pi^{-1/\alpha}(\Gamma(1 + \alpha/2))^{2/\alpha}$. If α is known, the above constants can be computed, except for $R_G(0)$, which depends on the model. □

References

[1] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events.* Springer, Berlin.

[2] McElroy, T. and Politis, D. (2005). Computer intensive rate estimation, diverging statistics, and scanning. Technical Report available from: www.math.ucsd.edu/∼politis/PAPER/scans.pdf

[3] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes.* Chapman and Hall, New York.

[4] Taqqu, M.S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahr. Verw. Gebiete* **31**, 287 – 302.