Verification of certain conditions from the paper ‘Computer-intensive rate estimation, diverging statistics, and scanning’ by

Tucker McElroy
U.S. Bureau of the Census

Dimitris N. Politis
University of California, San Diego

Set-up: Assume that the data $X_1, ..., X_n$ are i.i.d. from some distribution $F_X \in ND(\alpha)$ for $\alpha \in (0, 2]$, where $ND(\alpha)$ is the so-called “normal” domain of attraction of an α-stable law; see e.g. Embrechts, Klüppelberg, and Mikosch (1997, Chap. 2). In this context, it is well known that there exist sequences a_n and b_n such that $a_n^{-1}\left(\sum_{t=1}^n X_t - b_n\right) \overset{\mathcal{L}}{\Longrightarrow} S_\alpha$, where S_α denotes a generic α-stable law with unspecified scale, location and skewness. Since F_X is in the normal domain of attraction, $a_n = n^{1/\alpha}$ and the centering sequence b_n can be taken to be zero if either $\alpha < 1$ or $\alpha > 1$ and X_t has mean zero. When $\alpha = 1$, we can only let $b_n = 0$ if X_t is symmetric about zero.

If $\alpha = 2$ and and the data $X_1, ..., X_n$ have finite variance, assumption (8) of McElroy and Politis (2005) holds with $\gamma_1 = 0$ and $\gamma_2 = 1$; the verification is immediate by δ–method.

For the case of a general $\alpha \in (0, 2]$, some additional structure is needed. To this end, suppose that the cdf F_X has no point masses, and that the corresponding pdf f_X is bounded. We also assume that $f_X(x)/x$ is bounded in a neighborhood of zero, which ensures that the pdf of X^2 is bounded. A condition of this sort is needed to ensure the convergence of log moments. With these assumptions, we now proceed to verify assumptions (6) and (7) of McElroy and Politis (2005). In terms of the notation of that paper, let $\lambda = \alpha$, $g(y) = 2/y$, and $T_k = S_k^2$ in eq. (5). Then we see from eq. (1) that $U_n \overset{\mathcal{L}}{\Longrightarrow} U$, where U has the distribution of a logged $\alpha/2$ positively skewed stable random variable. Now let $V_n = \exp U_n$ and $V = \exp U$, and consider EU_n. We write

$$E[U_n] = \int_0^\infty \mathbb{P}[U_n > t]\, dt - \int_{-\infty}^0 \mathbb{P}[U_n < t]\, dt, \tag{1}$$

and it will be seen that this integral exists by what follows. For the first integral we have

$$\int_0^\infty \mathbb{P}[U_n > t]\, dt = \int_0^\infty \mathbb{P}[V_n > e^t]\, dt \to \int_0^\infty \mathbb{P}[V > e^t]\, dt$$

as $n \to \infty$ by dominating $\mathbb{P}[V_n > x]$ by a constant times $x^{\delta - \alpha/2}$, which is valid for any $\delta > 0$ and all $x > 0$ and $n \geq 1$ by Lemma 2 of Meerschaert and Scheffler (1998). This same bound shows that this portion of the expectation of U_n is finite for all n. For the other integral, first note that

$$\int_{-\infty}^{0} \left(\mathbb{P}[U_n < t] - \mathbb{P}[U < t]\right) dt = \int_0^\infty \left(\mathbb{P}[V_n < e^{-t}] - \mathbb{P}[V < e^{-t}]\right) dt = \int_0^\infty \int_0^{e^{-t}} \left(f_{V_n}(u) - f_V(u)\right) du\, dt, \tag{2}$$

which incidentally shows the integrability of this other portion of U_n, since f_{V_n} is bounded by assumption. Now by a result of Gnedenko (1953/54) – see also Macht and Wolf (1989) for an English translation – the convergence of pdfs, given our setup, is uniform given that there exists some n such that f_{V_n} is bounded. Now f_{V_1} is proportional to the pdf of X^2, which is bounded as long as $f_X(x)/x$ is bounded in a neighborhood of zero, since $f_{X^2}(x) = (f_X(\sqrt{x}) + f_X(-\sqrt{x}))/\sqrt{x}$. Note that it is possible to that the condition of Gnedenko (1953/54) is satisfied under less stringent assumptions on f_X, but it is clear that some condition on the probability content at zero is necessary, since we apply the logarithm. Note that Gnedenko (1953/54) assumes independent random variables, but may possibly be true for moving average models as well. Thus, a bound on (2) is

$$\sup_x |f_{V_n}(x) - f_V(x)|$$

which tends to zero as $n \to \infty$. This shows that $EU_n \to EU$.

For the second moments, we use a similar argument, but now the expression is

$$E[U_n^2] = \int_0^\infty \mathbb{P}[U_n^2 > t]\, dt = \int_0^\infty \mathbb{P}[V_n > e^{\sqrt{t}}]\, dt + \int_0^\infty \mathbb{P}[V_n < e^{-\sqrt{t}}]\, dt \tag{3}$$

and each integral is handled by arguments similar to those used above. So $EU_n^2 \to EU^2$ as well, and condition (6) of McElroy and Politis (2005) is verified.

Regarding condition (7), an additional assumption is required: suppose that the cdf F_{X^2} and F_U satisfy

$$\int_0^\infty |F_{X^2}(x) - F_U(x)|\, dx < \infty. \tag{4}$$

By Satybaldina (1972), it follows that

$$\sup_x |F_{U_n}(x) - F_U(x)| = O(n^{1-2/\alpha}).$$

Next we write

$$\mathbb{E}\epsilon_n - \mathbb{E}\epsilon = \int_0^\infty \mathbb{P}\left([\epsilon_n > t] - \mathbb{P}[\epsilon > t]\right) dt - \int_{-\infty}^0 \left(\mathbb{P}[\epsilon_n < t] - \mathbb{P}[\epsilon < t]\right) dt. \tag{5}$$

Focus on the first term. Let b_n be a positive sequence tending to infinity as $n \to \infty$. Then the first term can be broken up as

$$\int_0^{b_n} \mathbb{P}[\epsilon_n > t] - \mathbb{P}[\epsilon > t]\, dt + \int_{b_n}^\infty \mathbb{P}[\epsilon_n > t]\, dt - \int_{b_n}^\infty \mathbb{P}[\epsilon > t]\, dt,$$

the first term of which is $\int_0^{b_n} \mathbb{P}[U_n < e^t] - \mathbb{P}[U < e^t]\, dt = O(b_n\, n^{1-2/\alpha})$. Using Lemma 2 of Meerschaert and Scheffler (1998), we have

$$\int_{b_n}^{\infty} \mathbb{P}[\epsilon_n > t]\, dt = \int_{e^{b_n}}^{\infty} \mathbb{P}[U_n > x]/x\, dx \leq C \int_{e^{b_n}}^{\infty} x^{-(1+\alpha/2)+\delta}\, dx$$

for any $\delta > 0$ and some constant $C > 0$, for all $n \geq 1$. Naturally, we choose $\delta < \alpha/2$ to ensure integrability of the bounding function. Thus the bound is $C \exp\{(\delta - \alpha/2) b_n\}/(\alpha/2 - \delta)$. The same argument provides a similar bound for $\int_{b_n}^{\infty} \mathbb{P}[\epsilon > t]\, dt$. If we let $b_n = \log(n)$, for example, we obtain a $O(\log(n)\, n^{1-2/\alpha})$ term and a $O(n^{\delta-\alpha/2})$ term. For the second term of (5), we decompose as follows:

$$\int_{-b_n}^{0} \mathbb{P}[\epsilon_n < t] - \mathbb{P}[\epsilon < t]\, dt + \int_{-\infty}^{-b_n} \mathbb{P}[\epsilon_n > t]\, dt - \int_{-\infty}^{-b_n} \mathbb{P}[\epsilon > t]\, dt.$$

The first term here is handled analogously to the above argument. Using the boundedness of the pdf for all n, we have

$$\int_{-\infty}^{-b_n} \mathbb{P}[\epsilon_n < t]\, dt = \int_{0}^{e^{-b_n}} \mathbb{P}[U_n < x]/x\, dx = \int_{0}^{e^{-b_n}} \frac{1}{x} \int_{0}^{x} f_{U_n}(u)\, du\, dx \leq \sup_u f_{U_n}(u)\, e^{-b_n}$$

with a similar bound for $\int_{-\infty}^{-b_n} \mathbb{P}[\epsilon < t]\, dt$. Letting $b_n = \log(n)$ gives polynomial order bounds for all the terms. This completes the verification of assumption (7) of McElroy and Politis (2005), and we see that the power p can be taken (approximately) to be the minimum of $2/\alpha - 1$ and $\alpha/2$, which are equal at $\alpha = \sqrt{5} - 1$.

References

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

[2] Gnedenko, B. (1953/1954). On the Local Limit Theorem for Independent and Identically Distributed Summands (in Russian). *Wiss. Zt. Humboldt-Univ. Berlin, Math.-Nat. Reihe* **3**, 287–293.

[3] Macht, W. and Wolf, W. (1989). The Local Limit Problem and Hölder-Continuity. *Probability Theory and Related Fields* **82**, 295–305.

[4] 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

[5] Meerschaert, M. and Scheffler, H. (1998). A simple robust estimator for the thickness of heavy tails. *J. Statist. Plan. Inf.* **71**, 19–34.

[6] Satybaldina, K. (1972). Absolute Estimates of the Rate of Convergence to Stable Laws. *Theory of Probability and its Applications* **17**, 726 – 728.