Joseph *and* Noah:

Long Memory Heavy-Tailed Time Series

T. S. McElroy

1 Joseph and Noah in Time Series

The time series plots that accompany this paper exhibit both the Joseph and Noah effects, otherwise known as long range dependence and heavy tails, respectively. A rough definition of **long range dependency** is the presence of a slowly decaying autocorrelation/autocovariance function; indeed, the decay may be so slow (e.g. a polynomial rate) that the covariances are not summable. We don't strictly require the word "auto" above, which is appropriate for stationary data - more general heteroskedastic data may also be thought of as having long range dependence. It is common to measure dependence via strong mixing coefficients; roughly speaking, exponential decay corresponds to short-range dependence (e.g. Markov dependence), while polynomial decay corresponds to long-range dependence. This persistence over long gaps or lags will also be seen in the sample correlation and covariance functions. This has been called the **Joseph Effect** due to the persistency of phenomenon over time, viz. "Behold, there shall come seven years of great plenty throughout all the land of Egypt: And there shall arise after them seven years of famine..." (Genesis 41:29-30).

A non-rigorous definition of **heavy tails** is any random variable whose pdf has tails of polynomial decay (Mandelbrot coined the term "hyperbolic" random variables — viz. Mandelbrot (1983), p.204 — due to the asymptotic shape of the tail); in particular they are not exponential, and hence lack certain moments. An example is the Cauchy distribution, which has no mean. Nearly all the heavy-tailed random variables lack a second moment, and thus are used in models which exhibit "infinite variance." There are ways to estimate the tail exponent, e.g. the Hill estimator (viz. Resnick (1997)). This has been called the **Noah Effect** due to the infinite variance, which shows high variability, commonly encountered in catastrophic events, viz. "...all the fountains of the great deep [were] broken up, and the windows of heaven were opened. And the rain was upon the earth forty days and forty nights." (Genesis 7:11-12).

1.1 Examples of Long Range Dependent Data

Many example of long range dependence can be found in Jan Beran's book (Beran 1994), where in Chaper One we see that the sample autocorrelations have a persistency over a large number of lags. This phenomena arises in the diverse fields of hydrology, video conferencing , Ethernet networks, governmental standardization, and climatology. It crops up in a number of other areas, e.g. economics (viz. Mandelbrot (1969) and Taqqu and Levy (1985)), turbulence (viz. Mandelbrot (1974)), weather (viz. Lovejoy (1982)), and communications (viz. Mandelbrot (1965)). In finance, "The existence of long term dependence in common stock price series is not surprising since stock prices are related directly or indirectly to climatological variables, such as rainfall, in which the existence of long term dependence is well established." (Greene and Fielitz (1979)) See Mandelbrot and Wallis (1968,1969) for examples in hydrology.

1.2 Examples of Heavy-Tailed Data

Some instances of heavy-tailed data can be found in Willinger et al. (1997), who examine Ethernet traffic data, and in Embrechts et al. (1997). In the S and P data there are values significantly larger than others, so that the marginal distribution is not light-tailed. In the arena of finance, much debate is continuing over infinite versus finite variance models. "It is a fact...that most financial data are heavy-tailed!" (viz. Embrechts et al. (1997) p. 406) In some cases second moments may exist, but third or fourth moments do not.

2 A Long Memory Heavy-Tailed Process

Here are some mathematical definitions for the above concepts. For a stricly stationary discrete time stochastic process $\{X_t\}$ with second moments, we can write down its autocovariance function γ and autocorrelation function ρ, which depend only on the time lags between variables:

$$\gamma(h) = \mathbb{E}[X_{t+h}X_t], \quad \rho(h) = \gamma(h)/\gamma(0)$$

For simplicity, we have assumed that the process has mean zero.

Definition 1 *A* **long memory process** *satisfies*

$$\rho(h) \sim Kh^{\beta-1} \tag{1}$$

as $h \to \infty$ and $K \neq 0$. Here $\beta \in [0,1)$; if $\beta < 0$, this would imply that the autocovariance function is absolutely summable – these are sometimes called intermediate memory processes.

Remark 1 I have been vague about *long range dependence*, but quite specific about *long memory*. The viewpoint here is that the latter is an instance of the former concept.

Remark 2 It is common to use $2d$ instead of β, and use d to parametrize memory – see Brockwell and Davis (1991).

A random variable has heavy tails when the tail decreases at a slow polynomial rate, which depends on a *characteristic exponent* α (which is the reciprocal of the well-studied *extreme value index*).

Definition 2 *A random variable X with distribution function F has* **heavy tails** *if*

$$\mathbb{P}[X > x] \sim pCx^{\alpha}$$

$$\mathbb{P}[X < -x] \sim qCx^{\alpha}$$

as $x \to \infty$. The constant $C > 0$ is the dispersion, and has to do with the scale of X, whereas the numbers p and q describe the asymmetry of F; they are both in $[0,1]$ and sum to unity. The characteristic exponent α is positive, and less than two.

Remark 3 All moments strictly less than α exist, and all moments greater than or equal to α do not exist. For example, the Cauchy distribution is heavy-tailed with $\alpha = 1$. More generally, the class of stable variables and the Pareto family furnish two groups of heavy-tailed variables.

2.1 Construction

With these definitions, let us now construct a process with both properties. Note that long memory is defined for second moment processes, which is not the case for leptokurtic (= heavy-tailed) distributions. One solution is to stipulate that after a certain number of lags, the autocovariance function γ is always defined, and it satisfies (1).

It is a fact that a symmetric α-stable random variable X (write X is sαs) can be represented by the product of a totally right skewed $\frac{\alpha}{2}$-stable r.v. and a centered Gaussian (if the Gaussian is standard, so is X) :

$$X = A^{\frac{1}{2}} \cdot Z$$

where

$$\begin{aligned} X &\sim S_\alpha(\sigma^{\frac{1}{\alpha}}, 0, 0) \\ A &\sim S_{\frac{\alpha}{2}}(1, 1, 0) \\ Z &\sim \mathcal{N}(0, \sigma^2). \end{aligned}$$

Here, the r.v.'s A and Z must be independent. Thus we construct the following series:

$$\begin{aligned} \varepsilon_t &\sim iid\, S_{\frac{\alpha}{2}}(1, 1, 0) \\ Z_t &\sim id\, \mathcal{N}(0, \sigma^2) \end{aligned}$$

and the two sequences are chosen independently of one another. Now the Gaussians are not drawn *iid*, but have the following dependence stucture: they are weakly stationary with autocovariance function

$$\gamma_Z(h) := h^{\beta-1} \quad \forall h \geqslant 1 \qquad \gamma_Z(0) = \sigma^2,$$

where $\beta \in [0, 1)$ parametrizes the long range dependence. Then define

$$X_t := \sqrt{\varepsilon_t} Z_t.$$

Lemma 1 *If we let $\alpha \in (1, 2)$, then the series X_t is a sαs stationary time series with autocovariance function*

$$\gamma_X(0) = \infty \qquad \gamma_X(h) = \mu^2 \gamma_Z(h) \, \forall h \geqslant 1$$

for some where $\mu = \mathbb{E}\sqrt{\varepsilon_t}$.

Proof First note that $\mu < \infty$, since

$$\mathbb{E}|\sqrt{\varepsilon_t}| = \|\varepsilon_t\|_{\frac{1}{2}}^{\frac{1}{2}} < \infty$$

since $\frac{1}{2} < \frac{\alpha}{2}$. However, the mean of ε is not finite; it is definitely infinite (due to its heavy tails). Now X has mean zero due to independence of the centered Gaussian element. Finally, if $h \geqslant 1$ (since $h = 0$ gives the variance, which we already know is infinite) ,

$$\begin{aligned}
\gamma_X(h) &= Cov\left(X_t, X_{t+h}\right) = \mathbb{E}\left(X_t \cdot X_{t+h}\right) \\
&= \mathbb{E}\left(\sqrt{\varepsilon_t}\sqrt{\varepsilon_{t+h}} Z_t Z_{t+h}\right) \\
&= \mathbb{E}\left(\sqrt{\varepsilon_t}\right)\mathbb{E}\left(\sqrt{\varepsilon_{t+h}}\right)\mathbb{E}\left(Z_t Z_{t+h}\right) \\
&= \mu^2 \gamma_Z(h)
\end{aligned}$$

Thus X has the required long memory properties. And since the $\sqrt{\varepsilon_t}$ sequence and the Gaussians are independent, we see that X is sαs. †

Remark 4 We can easily construct other such series with infinite autocovariance for all lags less than k by, running the following $MA(k)$ filter: generate long memory Gaussian sequences independently of one another, multiply them by one stream of $\sqrt{\varepsilon}$, and then put the result through a standard $MA(k)$ filter with these dependent sαs random processes as the inputs. The following is actually a sum of independent sαs r.v.'s:

$$Y_t := \psi_0 X_t^0 + \psi_1 X_{t-1}^1 + \cdots + \psi_{k-1} X_{t-k+1}^{k-1}.$$

Hence Y_t is sαs with scale

$$\sigma^{\frac{1}{\alpha}} \left(\sum_{j=0}^{k-1} |\psi_j|^\alpha \right)^{\frac{1}{\alpha}}$$

and autocovariance

$$\gamma_X(h) \sim C h^{\beta - 1} \qquad \forall h \geqslant k$$

and infinity otherwise, where $C > 0$ is a constant. This is assuming that all the Gaussian sequences asymptotically have the same decay behavior. As usual in time series, the ψ's are constant coefficients, real or complex.

Remark 5 The algorithm for generating n data from this series is fairly straightforward. Since the Splus function "rstab" doesn't work for skewed stables, I use another algorithm [1]:

- (1) Generate the ε series by generating n uniforms U on $(0, 1)$ and n unit exponentials E, and calling

$$eps \longleftarrow rstable\,(U, E, \frac{\alpha}{2})$$

- (2) Generate n standard iid Gaussians and specify your Gaussian autocovariances by writing them into a huge Toeplitz covariance matrix. Then do a singular value decomposition of this matrix into

$$\Sigma = R \Lambda R^T$$

  and come up with new Gaussians via

$$Z := R \Lambda^{\frac{1}{2}} W.$$

  You can use

$$gau \longleftarrow gauss\,(n, \beta).$$

- (3) Write

$$xdata \longleftarrow (sqrt\,eps) * gau$$

[1] The Splus file is "depstabseries.s", and this code can be found by following the link.

Remark 6 This construction is reminiscent of an ARCH model. Indeed, one can view the above decomposition of a sαs r.v. as

$$X \sim \mathcal{N}(0, \sigma^2 \varepsilon)$$

informally. Thus we generate dependent Gaussians which are then "fattened up" by a positive heavy-tailed variable. In the literature, if one takes $\{\varepsilon_t = \varepsilon\}$ for all times, we get a **sub-Gaussian** process. Perhaps this could be a useful way of modelling long memory heavy-tailed data.

3 Conclusion

This paper has discussed the Joseph and Noah effects in Time Series data, and proposed a class of stationary models which exhibit both phenomena. The simulation of such time series is quite easy, and the accompanying plots give an idea of the structure. These plots have positive dependence, and are fairly short; also the characteristic exponent was taken fairly large.

Below is a list of relevant references for long range dependence and heavy tails in statistics and science; included are some papers that have not been referenced in the above text.

References

[1] Beran, J. (1994) Statistics for Long-Memory Processes. Chapman and Hall.

[2] Beran, J., Sherman R., Taqqu, M., and Willinger W. (1995) Long-range dependence in Variable-bit rate video traffic. *IEEE Trans. Comm.* **43** 1566–1579.

[3] Billingsley, P. (1995) Probability and Measure. John Wiley and Sons.

[4] Brockwell, P. and Davis, R. (1991) Time Series: Theory and Methods. Springer.

[5] Castillo, E. (1988) Extreme Value Theory in Engineering. Academic Press, San Diego.

[6] Davis, R. and Hsing, T. (1995) Point Process and Partial Sum Convergence for Weakly Dependent Random Variables with Infinite Variance. *Annals of Probability* **23**, No. 2, 879–917.

[7] Davis, R., and Mikosch, T. (1998) The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Annals of Statistics* **26**, No. 5, 2049–2080.

[8] Davis, R., and Resnick, S. (1985) Limit theory for moving averages of random variables with regularly varying tail probabilities. *Annals of Probability* **13**, No. 1, 179–195.

[9] Davis, R., and Resnick, S. (1986) Limit theory for the sample covariance and correlation functions of moving averages. *Annals of Statistics* **14**, No. 2, 533–558.

[10] Duffy, D., McIntosh, A., Rosenstein, M., and Willinger, W. (1993) Analyzing telecommunications traffic data from working common channel signaling networks. In *Proceedings of the 25th Interface, San Diego*. Interface Foundation of North America.

[11] Duffy, D., McIntosh, A., Rosenstein, M., and Willinger, W. (1994) Statistical analysis of CCSN/SS7 traffic data from working CCS subnetworks. *IEEE Journal on Selected Areas in Communications* **12** 544–551.

[12] Durrett, R. (1996) Probability: Theory and Examples. Duxbury Press.

[13] Embrechts, Mikosh, and Kluppelberg (1997) Modeling Extrema Events for Insurance and Finance. Springer-Verlag, Berlin.

[14] Greene, Myron T., and Fielitz, Bruce D. (1977) Long-Term Dependence in Common Stock Returns. *Journal of Financial Economics* **4**, 339-349.

[15] Greene, Myron T., and Fielitz, Bruce D. (1979) The Effect of Long Term Dependence on Risk-Return Models of Common Stocks. *Operations Research* **25**, No. 5, 944–951.

[16] Grigoriu, M. (1995) Applied Non-Gaussian Processes. Prentice Hall, Englewood Cliffs, NJ.

[17] Heath, D., Resnick, S., and Samorodnitsky, G. (1997) Heavy Tails and Long Range Dependence in On/Off Processes and Associated Fluid Models. website: http://www.orie.cornell.edu/trlist/

[18] Koedijk, K., Schafgans, M., and De Vries, C. (1990) The tail index of exchange rate returns. *Journal of International Economics* **29** 93–108.

[19] Lovejoy, S. (1982) Area-Perimeter relation for rain and cloud areas. *Science* **216**, 185–187.

[20] Mandelbrot, Benoit B. (1965) Self-similar error clusters in communications systems and the concept of conditional systems and the concept of conditional stationarity. *I.E.E.E. Transcations on Communications Technology* COM-13, 71–90.

[21] Mandelbrot, Benoit B. (1969) Long-run linearity, locally Gaussian processes, H-spectra and infinite variances. *International Economic Rev* **10**, 82-113.

[22] Mandelbrot, Benoit B. (1974) Intermittent turbulence in self similar cascades; divergence of high moments and dimension of the carrier. *Journal of Fluid Mechanics* **62**, 331–358.

[23] Mandelbrot, Benoit B. (1983) The Fractal Geometry of Nature. W.H.Freeman and Company.

[24] Mandelbrot, Benoit B., and Wallis, J.R. (1968) Noah, Joseph, and Operational Hydrology. *Water Resources Res.* **4**, 909-917.

[25] Mandelbrot, Benoit B., and Wallis, J.R. (1969) Some Long-Run Properties of Geophysical Records. *Water Resources Res* **5**, 321-340.

[26] Meier-Hellstern, K., Wirth, P., Yan, Y., and Hoeflin, D. (1991) Traffic models for ISDN data users: office automation application. In *Teletraffic and Datatraffic in a Period of Change. Proceedings of the 13th ITC* (A. Jensen and V.B. Iversen, eds.) 167–192. North Holland, Amsterdam.

[27] Politis, D. and Romano, J. (1994) Large Sample Confidence Regions Based on Subsamples Under Minimal Assumptions. *Annals of Statistics* **22**, No. 4, 2031–2050.

[28] Politis, D., Romano, J., and Wolf, M. (2000+) Subsampling. Unpublished manuscript.

[29] Resnick, S. (1986) Point processes, regular variation, and weak convergence. *Advances in Applied Probability* **18**, 66–138.

[30] Resnick, S. (1987) Extreme Values, Regular Variation, and Point Processes. Springer–Verlag.

[31] Resnick, S. (1997) Special Invited Paper: Heavy Tail Modeling and Teletraffic Data. *Annals of Statistics* **25**, No. 5, 1805–1849.

[32] Resnick, S., and Samorodnitsky, G. (1997) Performance decay in a single server exponential queuing model with long-range dependence. website: http://www.orie.cornell.edu/trlist/

[33] Resnick, S., and Stairca, C. (1998) Tail index estimation for dependent data. *Annals of Applied Probability* **8**, No. 4, 1156–1183.

[34] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences* **42**, 43–47.

[35] Rosenblatt, M. (1984). Asymptotic normality, strong mixing and spectral density estimates, *Annals of Probability* **12**, 1167–1180.

[36] Rosenblatt, M. (1985). *Stationary Sequences and Random Fields.* Birkhäuser, Boston.

[37] Samorodnitsky, G. and Taqqu, M. (1994) Stable Non-Gaussian Random Processes. Chapman and Hall.

[38] Taqqu, M., and Levy, J. (1985) Using renewal processes to generate long-range dependence and high variability. Dependence in Probability and Statistics. Birkhauser.

[39] Willinger, W., Taqqu, M., Leland, W., and Wilson, D. (1995) Self-similarity in high-speed packet traffic: analysis and modeling of ethernet traffic measurements. *Statist. Sci.* **10** 67–85.

[40] Willinger, W., Taqqu, M., Sherman, R., and Wilson, D. (1997) Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. *IEEE/ACM Transactions on Networking* **5** 71–96.