Statistical Inference for Extremes of Long Memory Processes

Rafał Kulik

Department of Mathematics and Statistics (University of Ottawa)

Bloomington, IN, 1/04/2017

Statistical Inference for Extremes of Long Memory Processes

Rafał Kulik

Department of Mathematics and Statistics (University of Ottawa)

Bloomington, IN, 1/04/2017

Plan

1. Introduction
   - Tail empirical process and its limit
   - Panorama of limit theorems under LRD: Partial Sums
2. Tail empirical process and LRD
3. Applications
   - Hill estimator
   - Value-at-Risk
4. Open problems

Regular variation

Problem: Assume that $\{X_j, j \geq 0\}$ is a stationary sequence with a marginal df F_X such that

$$\mathbb{P}(|X_0| > x) = \bar{F}_X(x) = x^{-\alpha}\ell(x) \; , \quad x > 0 \; , \tag{1}$$

where $\alpha > 0$ is called the *tail index* and ℓ is a function that is slowly varying at infinity.[1] Let $u_n \to \infty$ be such that $n\bar{F}_X(u_n) \to \infty$. Define

$$\widetilde{T}_n(s) = \frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} \mathbb{1}_{\{X_j > u_n s\}} \; , \quad T_n(s) = \mathbb{E}[\widetilde{T}_n(s)] \; .$$

Under the regular variation assumption (1) we have

$$\lim_{n\to\infty} T_n(s) = \lim_{n\to\infty} \frac{\bar{F}_X(u_n s)}{\bar{F}_X(u_n)} = s^{-\alpha} =: T(s) \; .$$

[1] In what follows, ℓ will be a slowly varying, different at each its occurrence

Central Limit Theorem

Result 1

Assume that $\{X_j, j \geq 0\}$ is a stationary sequence such that (1) holds. Under the appropriate weak dependence conditions and Lindeberg-type assumptions we have

$$\sqrt{n\bar{F}_X(u_n)}\left\{\widetilde{T}_n(s) - T_n(s)\right\} \Rightarrow G(s) \tag{2}$$

in $\mathbb{D}(0,\infty)$, where $G(\cdot)$ is a Gaussian process. In particular, if X_j's are i.i.d., then $G = B \circ T$, where B is a Brownian motion on $(0,\infty)$.

[2] [3] **Question:** What about long range dependent sequences (LRD)?

---

[2] See Drees (2003), Rootzen (2009), Drees and Rootzen (2010), Kulik, Soulier, Wintenberger (2016).

[3] $\Rightarrow$ denotes convergence in space $\mathbb{D}$ w.r.t. Skorohkod topology, but the space will differ in the subsequent results

Gaussian sequences

Let $\{\eta_j, j \in \mathbb{Z}\}$ be a sequence of i.i.d. standard normal random variables. Define

$$X_j = \sum_{k=1}^{\infty} a_k \eta_{j-k} \ . \tag{3}$$

If $a_k = k^{d-1}\ell_a(k)$, $d \in (0, 1/2)$, $\sum_{k=1}^{\infty} a_k^2 = 1$, then $\{X_j, j \geq 0\}$ is a stationary sequence of standard normal random variables such that $\mathrm{var}(\sum_{j=1}^{n} X_j) \sim n^{2d+1}\ell^2(n)$.

Result 2

We have

$$\frac{1}{n^{d+1/2}\ell(n)} \sum_{j=1}^{[nt]} X_j \Rightarrow C(d) B_H(t), \qquad H = 1/2 + d > 1/2,$$

where $B_H(\cdot)$ is a Fractional Brownian motion.

Functionals of Gaussian sequences

Furthermore, let ϕ be a function such that $\mathbb{E}[\phi(X_0)] = 0$, $\mathbb{E}[\phi^2(X_0)] < \infty$ and define the Hermite rank as

$$\tau = \inf\{m \in \mathbb{N} : J_m(\phi) \neq 0\} , \quad J_m(\phi) = \mathbb{E}[\phi(X_0) H_m(X_0)] ,$$

where $H_m(\cdot)$ is the mth Hermite polynomial.

Result 3

If $\tau(1 - 2d) > 1$, *then*

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{[nt]} \phi(X_j) \stackrel{\mathrm{d}}{\to} \sigma_0 B(t), \qquad \sigma_0^2 := \sum_{k=0}^{\infty} \mathrm{Cov}(\phi(X_0), \phi(X_k)) < \infty.$$

Functionals of Gaussian sequences

Result 4

If $\tau(1-2d)<1$, *then*

$$\frac{1}{a_{n,\tau}}\sum_{j=1}^{[nt]}\phi(X_j)\Rightarrow C(\beta)I_H^{(\tau)}(t), \qquad H=1/2+d,$$

where $a_{n,\tau}=n^{1-\tau(1/2-d)}\ell(n)$ *and* $I_H^{(\tau)}(\cdot)$ *is called Hermite-(Rosenblatt) process*

$$I_H^{(\tau)}(t):=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}Q_t^{(\tau)}(x_1,\ldots,x_\tau;H)B(dx_1)\cdots B(dx_\tau).$$

4

[4] Rosenblatt, Taqqu (1977-1979)

Stochastic volatility

Let $\{Y_j, j \geq 0\}$ be like in (3). Let $\{Z_j, j \geq 0\}$ be a sequence of i.i.d. regularly varying random variables. For $\phi \geq 0$ define

$$X_j = \phi(Y_j)Z_j \ , \quad j \geq 0 \ , \quad \mathcal{F}_j = \sigma(\{\eta_i, Z_i\}, i \leq j) \ . \tag{4}$$

Let τ_p be the Hermite rank of ϕ^p.

- Long Memory Stochastic Volatility (LMSV) model: where $\{\eta_j\}$ and $\{Z_j\}$ are independent.
- Model with leverage: where $\{(\eta_j, Z_j)\}$ is a sequence of i.i.d. random vectors. For fixed j, Z_j and Y_j are independent, but Y_j may not be independent of the past $\{Z_i, i < j\}$.

Due to Breiman's lemma, if $\mathbb{E}[\phi^{\alpha+\epsilon}(Y_0)] < \infty$, then

$$\lim_{x \to +\infty} \frac{\mathbb{P}(|X_0| > x)}{\mathbb{P}(|Z_0| > x)} = \mathbb{E}[\phi^{\alpha}(Y_0)] \ . \tag{5}$$

Stochastic volatility

If $\alpha > 1$, then

$$\mathrm{cov}(X_0, X_k) = \mathbb{E}[Z_0]\mathrm{cov}(\phi(Y_0), \phi(Y_k)) \ , \quad k \geq 1 \ .$$

If $\alpha \in (1,2)$, then $\mathrm{var}(X_0) = \infty$, yet covariances are well-defined.
If $\alpha > 2$, then $\mathrm{cov}(X_0^2, X_k^2) = \mathbb{E}[Z_0^2]\mathrm{cov}(\phi^2(Y_0), \phi^2(Y_k))$, $k \geq 1$. Define

$$S_{p,n}(t) = \sum_{j=1}^{[nt]} |X_j|^p \ .$$

Then

$$S_{p,n}(t) - \mathbb{E}[S_{p,n}(t)] =$$

$$= \underbrace{\sum_{j=1}^{[nt]} \{|X_j|^p - \mathbb{E}[|X_j|^p \mid \mathcal{F}_{j-1}]\}}_{\text{"martingale"}} + \underbrace{\mathbb{E}[|Z_0|^p] \sum_{j=1}^{[nt]} \{\phi^p(Y_j) - \mathbb{E}[\phi^p(Y_j)]\}}_{\text{"long memory part"}} \ . \tag{6}$$

Stochastic volatility

Result 5

Let $c_n = \inf\{x : \mathbb{P}(|X_0| > x) < 1/n\}$ *(so that* $c_n \approx n^{-1/\alpha}\ell(n)$*). Then*

1. *If* $p < \alpha < 2p$ *and* $1 - \tau_p(1/2 - d) < p/\alpha$*, then*

$$c_n^{-p}(S_{p,n} - n\mathbb{E}[|X_0|^p]) \Rightarrow L_{\alpha/p} \; ,$$

where $L_{\alpha/p}$ *is a totally skewed to the right* α/p*-stable Lévy process.*

2. *If* $p < \alpha < 2p$ *and* $1 - \tau_p(1/2 - d) > p/\alpha$*, then*

$$a_{n,\tau}^{-1}(S_{p,n} - n\mathbb{E}[|X_0|^p]) \Rightarrow \frac{J_{\tau_p}(\phi^p)\mathbb{E}[|Z_0|^p]}{\tau_p!} I_H^{(\tau_p)} \; .$$

3. *If* $p > \alpha$*, then* $c_n^{-p} S_{p,n} \Rightarrow L_{\alpha/p}$*, where* $L_{\alpha/p}$ *is a positive* α/p*-stable Lévy process.*

Stochastic volatility

- Let us focus on $p = 1$. Then the theorem is applicable to $\alpha \in (0,1)$ and $\alpha \in (1,2)$. If the long memory is not strong enough, that is $1 - \tau_p(1/2 - d) < 1/\alpha$, then we have the classical limit as in i.i.d. case. If the long memory is strong enough then the limiting process has finite variance! (the long memory phase)
- To illustrate this, take $\phi(x) = \exp(x)$ so that $\tau_p = 1$ for all p. Take $p = 1$. Then the condition $1 - \tau_p(1/2 - d) < p/\alpha$ becomes $d + 1/2 < 1/\alpha$. Hence, keeping in mind that $d \in (0, 1/2)$, if α is close to 2, we can have the long memory phase only.
- Furthermore, if we consider $\alpha > 2p$ (*finite variance case*) then only the long memory phase occurs.
- Other interesting phenomena occur when sample covariances are considered. [5]

[5] See Kulik and Soulier (2012).

Infinite variance functionals of Gaussian sequences

Let's go back to the sequence given in (3), that is the long memory Gaussian sequence. Let ϕ be such that

$$\mathbb{P}(\phi(X_0) > x) = x^{-\alpha}\ell(x)\ .$$

For example, $\phi(x) = \exp(\alpha x^2)$ or $\phi(x) = 1/|x|^r$, $r > 0$. Then the similar result as Result 5 also holds. The proof is based on point processes and hypercontractivity, but in principle it is based on the decomposition (6). [6]

[6] See Davis (1983), Sly and Heyde (2008)

TEP for Stochastic Volatility

Theorem 6

Consider the stochastic volatility model given in (4). Assume that $\mathbb{E}[\phi^{\alpha+\epsilon}(Y_0)] < \infty$. *Let* $\tau = \tau_\alpha$ *be the Hermite rank of* ϕ^α.

- *If* $a_{n,\tau}\sqrt{\frac{\bar{F}_X(u_n)}{n}} \to 0$, *then (cf. Result 2)*

$$\sqrt{n\bar{F}_X(u_n)}\left\{\widetilde{T}_n(s) - T_n(s)\right\} \Rightarrow B \circ T(s)\,.$$

- *If* $a_{n,\tau}\sqrt{\frac{\bar{F}_X(u_n)}{n}} \to \infty$, *then*

$$\frac{n}{a_{n,\tau}}\left\{\widetilde{T}_n(s) - T_n(s)\right\} \Rightarrow \frac{J_\tau(\phi^\alpha)}{\tau!\mathbb{E}[\phi^\alpha(Y_0)]} s^{-\alpha} I_H^{(\tau)}(1)\,.$$

TEP for Stochastic Volatility - random levels

The proof of Theorem 6 is based on (more involved than before) martingale-LRD decomposition.

Two issues:

- We want to replace $T_n(s)$ with its limit $T(s) = s^{-\alpha}$.
- We want to get rid of u_n and $\bar{F}_X(u_n)$.

Choose $k = k_n \to \infty$ such that $k/n \to 0$. Choose u_n such that $k = n\bar{F}_X(u_n)$. Let $X_{n:n} \ge X_{n:n-k} \ge \cdots \ge X_{n:1}$ be the order statistics. Then Theorem 6 implies

$$X_{n:n-k}/u_n \xrightarrow{\text{p}} 1 \ .$$

Hence, define

$$\widehat{T}_n(s) = \frac{1}{k} \sum_{j=1}^{n} \mathbb{1}_{\{X_j > X_{n:n-k} s\}} \ .$$

TEP for Stochastic Volatility - random levels

Theorem 7

Assume that the conditions of Theorem 6 are satisfied. Assume moreover that

$$\lim_{n\to\infty}\left(\frac{n}{a_{n,\tau}}+\sqrt{n\bar{F}(u_n)}\right)\sup_{t\ \epsilon>0}\left|\frac{\bar{F}_X(u_n s)}{\bar{F}_X(u_n)}-s^{-\alpha}\right|=0\ . \tag{7}$$

Then

$$\sqrt{k}\left\{\widehat{T}_n(s)-T(s)\right\}\Rightarrow B_0\circ T(s)\ ,$$

where B_0 is a Brownian bridge.

[7] [8]

[7] Kulik and Soulier (2009), Bilayi and Kulik (2017)
[8] Condition (7) is linked to *second order regular variation*

Hill estimator

Note that $\int_1^\infty (T(s)/s)\mathrm{d}s = \alpha^{-1}$. Hence, we can consider

$$\widehat{\alpha}^{-1} = \int_1^\infty (\widehat{T}_n(s)/s)\mathrm{d}s = \frac{1}{k}\sum_{j=1}^{n} \log\left(\frac{X_{n:n-j+1}}{X_{n:n-k}}\right) .$$

It is called the Hill estimator.

Corollary 8

Under the conditions of Theorem 7 we have

$$\sqrt{k}\left\{\widehat{\alpha}^{-1} - \alpha^{-1}\right\} \stackrel{\mathrm{d}}{\to} \int_1^\infty (B_0 \circ T(s)/s)\mathrm{d}s \stackrel{\mathrm{d}}{=} \mathcal{N}(0, \alpha^{-2}) .$$

Conclusion: Long memory (in the stochastic volatility) does not influence the behaviour of the Hill estimator!!!

Value-at-Risk

Let $\{X_j, j \geq 0\}$ be the Stochastic Volatility model Let $p \in (0,1)$ be *small*. Then we are interested in estimation of $Q(p) = F_Y^{\leftarrow}(1-p)$. The extreme order statistics cannot be really used directly. Hence, the idea is to *go to the intermediate statistics and scale*. Let $p = p_n \to 0$ and $k/n \to 0$. The approximation

$$\frac{Q(p)}{Q(k/n)} \approx \left(\frac{k}{np}\right)^{1/\alpha}$$

suggests the following estimator for $Q(p)$:

$$\widehat{Q}_n(p) = X_{n:n-k}\left(\frac{k}{np}\right)^{1/\widehat{\alpha}} .$$

Unfortunately, the limiting distribution of $X_{n:n-k}$ has to be deduced from Theorem 6, hence the long memory influences estimation of the Value-at-Risk.

Open problems

1. Tail empirical process for infinite variance functionals of Gaussian sequences;
2. Tail empirical process for infinite variance moving averages:

$$X_j = \sum_{k=1}^{\infty} a_k \eta_{j-k} \ ,$$

where $\{\eta_j, j \in \mathbb{Z}\}$ are i.i.d. regularly varying random variables with tail index $\alpha > 0$.

It is not clear to me if we have dichotomous behaviour as in Theorem 6!!!

Thank you!!!