Asymptotics and Simulation of Heavy-Tailed Processes

Henrik Hult

Department of Mathematics
KTH Royal Institute of Technology
Stockholm, Sweden

Workshop on Heavy-tailed Distributions and Extreme Value Theory
ISI Kolkata
January 14-17, 2013

Outline

Regularly Varying Sequences and Random Variables

- A sequence c_n is called regularly varying at ∞ with index $\rho \in \mathbf{R}$ if, for each $\lambda > 0$,

$$\lim_{n\to\infty} \frac{c_{[\lambda n]}}{c_n} = \lambda^\rho.$$

- A non-negative random variable Z is called regularly varying with index α if the tail $\mathbb{P}\{Z > n\}$ is regularly varying at ∞ with index $-\alpha$, $\alpha \geq 0$.

Regular Variation and Weak Convergence

- Suppose Z is regularly varying with index α.
- For any $\lambda > 0$, with $c_n = \mathbb{P}\{Z > n\}^{-1}$ it follows that

$$c_n \mathbb{P}\{Z \in n(\lambda, \infty)\} \to \lambda^{-\alpha} =: \mu_\alpha(\lambda, \infty).$$

- This convergence can be formulated as a weak convergence:

$$c_n \mathbb{P}\{n^{-1} Z \in \cdot\} \xrightarrow{w} \mu_\alpha,$$

when restricted to any subset where μ_α is finite. That is, of the form (ϵ, ∞), $\epsilon > 0$.

Regular Variation and Weak Convergence

- Suppose Z is regularly varying with index α.
- For any $\lambda > 0$, with $c_n = \mathbb{P}\{Z > n\}^{-1}$ it follows that

$$c_n \mathbb{P}\{Z \in n(\lambda, \infty)\} \to \lambda^{-\alpha} =: \mu_\alpha(\lambda, \infty).$$

- This convergence can be formulated as a weak convergence:

$$c_n \mathbb{P}\{n^{-1} Z \in \cdot\} \xrightarrow{w} \mu_\alpha,$$

when restricted to any subset where μ_α is finite. That is, of the form (ϵ, ∞), $\epsilon > 0$.

The Space $\mathbf{M}_0$[1]

- Let $(\mathbf{S}, d)$ be a complete separable metric space with its Borel σ-field.
- s_0 is the origin in $\mathbf{S}$.
- $B_{0,r} = \{s \in \mathbf{S} : d(s, s_0) < r\}$ (open ball of radius r).
- $\mathscr{C}_0$ are the real-valued bounded continuous functions on $\mathbf{S}$ vanishing on some ball $B_{0,r}$, $r > 0$.
- $\mathbf{M}_0 = \{\text{Borel measures } \mu \text{ on } \mathbf{S} \text{ with } \mu(B_{0,r}^c) < \infty \text{ for each } r > 0\}$.
- Convergence in $\mathbf{M}_0$: $\mu_n \to \mu$ in $\mathbf{M}_0$ if

$$\int f d\mu_n \to \int f d\mu, \qquad \text{for all } f \in \mathscr{C}_0.$$

[1] The details are in [4]

The Space $\mathbf{M}_0$[1]

- Let $(\mathbf{S}, d)$ be a complete separable metric space with its Borel σ-field.
- s_0 is the origin in $\mathbf{S}$.
- $B_{0,r} = \{s \in \mathbf{S} : d(s, s_0) < r\}$ (open ball of radius r).
- $\mathscr{C}_0$ are the real-valued bounded continuous functions on $\mathbf{S}$ vanishing on some ball $B_{0,r}$, $r > 0$.
- $\mathbf{M}_0 = \{$Borel measures μ on $\mathbf{S}$ with $\mu(B^c_{0,r}) < \infty$ for each $r > 0\}$.
- Convergence in $\mathbf{M}_0$: $\mu_n \to \mu$ in $\mathbf{M}_0$ if

$$\int f d\mu_n \to \int f d\mu, \qquad \text{for all } f \in \mathscr{C}_0.$$

[1] The details are in [4]

Regularly Varying Measures

- A sequence of measures ν_n in $\mathbf{M}_0$ is regularly varying with index $-\alpha$ if there exists a sequence $\{c_n\}$ of positive numbers, which is regularly varying with index $\alpha \geq 0$, and a nonzero $\mu \in \mathbf{M}_0$ such that

$$c_n \nu_n \to \mu, \text{ in } \mathbf{M}_0.$$

- A measure $\nu \in \mathbf{M}_0$ is called regularly varying if the sequence $\{\nu(n\,\cdot)\}$ is regularly varying with index $-\alpha$. In this case the limiting measure μ satisfies the scaling property: for any $\lambda > 0$ and Borel set $B \subset \mathbf{S} \setminus \{s_0\}$

$$\mu(\lambda B) = \lambda^{-\alpha} \mu(B).$$

Regularly Varying Measures

- A sequence of measures ν_n in $\mathbf{M}_0$ is regularly varying with index $-\alpha$ if there exists a sequence $\{c_n\}$ of positive numbers, which is regularly varying with index $\alpha \geq 0$, and a nonzero $\mu \in \mathbf{M}_0$ such that
$$c_n \nu_n \to \mu, \text{ in } \mathbf{M}_0.$$

- A measure $\nu \in \mathbf{M}_0$ is called regularly varying if the sequence $\{\nu(n\,\cdot)\}$ is regularly varying with index $-\alpha$. In this case the limiting measure μ satisfies the scaling property: for any $\lambda > 0$ and Borel set $B \subset \mathbf{S} \setminus \{s_0\}$
$$\mu(\lambda B) = \lambda^{-\alpha} \mu(B).$$

Outline

Multivariate Regular Variation

- Let X be a random vector in $\mathbf{R}^d$.
- The distribution of X is called multivariate regularly varying if $\mathbb{P}\{n^{-1}X \in \cdot\}$ is a regularly varying measure: there exists a nonzero $\mu \in \mathbf{M}_0(\mathbf{R}^d)$ and a regularly varying sequence c_n with index $\alpha \geq 0$ such that

$$c_n \mathbb{P}\{n^{-1}X \in \cdot\} \to \mu, \qquad \text{in } \mathbf{M}_0(\mathbf{R}^d).$$

Multivariate Regular Variation

Independent Components[2]

- Let $Z = (Z_1, \dots, Z_d)'$ be a random vector in $\mathbf{R}^d$ with iid regularly varying components.
- Take $c_n = \mathbb{P}\{Z_1 > n\}^{-1}$.
- For any set of the form $A_i = \{x : x_i > a\}$ it follows that

$$c_n \, \mathbb{P}\{n^{-1} Z \in A_i\} = \frac{\mathbb{P}\{Z_1 > an\}}{\mathbb{P}\{Z_1 > n\}} \to a^{-\alpha} = \mu_\alpha(a, \infty),$$

whereas for any set which is a subset of some $A_{i,j} = \{x : x_i > \epsilon_1, x_j > \epsilon_2, i \neq j\}$ it follows that

$$c_n \, \mathbb{P}\{n^{-1} Z \in A_{i,j}\} \leq \frac{\mathbb{P}\{Z_1 > n\epsilon_1\}\mathbb{P}\{Z_1 > n\epsilon_2\}}{\mathbb{P}\{Z_1 > n\}} \to 0.$$

[2] See e.g.[1]

Multivariate Regular Variation

Independent Components

- One can show that $\mathbb{P}\{n^{-1}Z \in \cdot\}$ is regularly varying with limiting measure μ which is concentrated on the union of the coordinate axis.
- More precisely,

$$c_n \mathbb{P}\{n^{-1}Z \in \cdot\} \to \mu, \qquad \text{in } \mathbf{M}_0(\mathbf{R}^d),$$

with

$$\mu(B) = \sum_{k=1}^{d} \int_0^\infty I\{ze_k \in B\}\mu_\alpha(dz).$$

Outline

Regular Variation on $\mathbf{D}[0,1]$

A Heavy-Tailed Lévy Process[3]

- Let $\mathbf{D}[0,1]$ be the space of càdlàg functions $[0,1] \to \mathbf{R}$ equipped with the Skorohod J_1-metric.
- Consider a Lévy process X with regularly varying Lévy measure ν:

$$c_n \nu(n,\infty) \to 1,$$

for a regularly varying sequence c_n.

[3] See [2, 4]

Regular Variation on $\mathbf{D}[0, 1]$

A Heavy-Tailed Lévy Process[3]

- Let $\mathbf{D}[0, 1]$ be the space of càdlàg functions $[0, 1] \to \mathbf{R}$ equipped with the Skorohod J_1-metric.
- Consider a Lévy process X with regularly varying Lévy measure ν:

$$c_n \nu(n, \infty) \to 1,$$

for a regularly varying sequence c_n.

[3] See [2, 4]

Regular Variation on $\mathbf{D}[0,1]$

A Heavy-Tailed Lévy Process

- Then, the heavy tailed heuristics **(one big jump + the heaviest tail wins)** can be made precise by showing that $\mathbb{P}\{n^{-1}X \in \cdot\}$ is regularly varying:

$$c_n \mathbb{P}\{n^{-1}X \in \cdot\} \to m, \qquad \text{in } \mathbf{M}_0(\mathbf{D}[0,1]),$$

where m is supported on step functions with one step.

$$m(B) = \int_0^1 \int_0^\infty I\{zI_{[\tau,1]}(\cdot) \in B\} \mu_\alpha(dz) d\tau,$$

where B is any Borel subset of $\mathbf{D}[0,1] \setminus \{0\}$.

Stochastic Integrals (c.f. [5])

- Consider a Lévy process X with regularly varying Lévy measure ν (index α).
- Let Y be an adapted process with lighter tails than ν:

$$E[\sup_{t\in[0,1]} Y_t^{\alpha+\epsilon}] < \infty, \qquad \text{some } \epsilon > 0.$$

- The stochastic integral process $(Y\cdot X)_t = \int_0^t Y_s dX_s$ is regularly varying with index α. In particular

$$c_n \mathbb{P}\{n^{-1}(Y\cdot X) \in \cdot\} \to m, \qquad \text{in } \mathbf{M}_0(\mathbf{D}[0,1]),$$

where m is supported on step functions with one step.

$$m(B) = E\Big[\int\int I\{Y_\tau z I_{[\tau,1]}(\cdot) \in B\}\mu_\alpha(dz)d\tau\Big],$$

where B is any Borel subset of $\mathbf{D}[0,1] \subset \{0\}$.

Outline

Large Deviations for the Empirical Measure (c.f. [6])

- Let $\{Z_k\}$ be iid with a regularly varying distribution on $\mathbf{R}^d$ with limiting measure μ and $\alpha > 1$.
- The empirical measure is

$$N_n = \sum_{k=1}^{n} \delta_{n^{-1}Z_k},$$

where δ_z is a unit point mass at z.

- Consider N_n as a random element taking values in the space of $\mathbf{N}_p$ of point measure on $\mathbf{R}^d \setminus \{0\}$ equipped with the vague topology.
- Then, the sequence $\mathbb{P}\{N_n \in \cdot\}$ is regularly varying:

$$c_n \mathbb{P}\{N_n \in \cdot\} \to m, \qquad \text{in } \mathbf{M}_0(\mathbf{N}_p),$$

with $m(B) = \int I\{\delta_z \in B\}\mu(dz)$.

Large Deviations for the Empirical Measure (c.f. [6])

- Let $\{Z_k\}$ be iid with a regularly varying distribution on $\mathbf{R}^d$ with limiting measure μ and $\alpha > 1$.
- The empirical measure is

$$N_n = \sum_{k=1}^{n} \delta_{n^{-1}Z_k},$$

where δ_z is a unit point mass at z.
- Consider N_n as a random element taking values in the space of $\mathbf{N}_p$ of point measure on $\mathbf{R}^d \setminus \{0\}$ equipped with the vague topology.
- Then, the sequence $\mathbb{P}\{N_n \in \cdot\}$ is regularly varying:

$$c_n \mathbb{P}\{N_n \in \cdot\} \to m, \qquad \text{in } \mathbf{M}_0(\mathbf{N}_p),$$

with $m(B) = \int I\{\delta_z \in B\}\mu(dz)$.

Large Deviations for the Empirical Measure (c.f. [6])

- Let $\{Z_k\}$ be iid with a regularly varying distribution on $\mathbf{R}^d$ with limiting measure μ and $\alpha > 1$.
- The empirical measure is

$$N_n = \sum_{k=1}^{n} \delta_{n^{-1}Z_k},$$

where δ_z is a unit point mass at z.
- Consider N_n as a random element taking values in the space of $\mathbf{N}_p$ of point measure on $\mathbf{R}^d \setminus \{0\}$ equipped with the vague topology.
- Then, the sequence $\mathbb{P}\{N_n \in \cdot\}$ is regularly varying:

$$c_n \mathbb{P}\{N_n \in \cdot\} \to m, \qquad \text{in } \mathbf{M}_0(\mathbf{N}_p),$$

with $m(B) = \int I\{\delta_z \in B\}\mu(dz)$.

Large Deviations for the Empirical Measure

Keeping Track of Time

- We may keep track of time in the sense that

$$N_n = \sum_{k=1}^{n} \delta_{(\frac{k}{n}, n^{-1}Z_k)}.$$

- Then, the sequence $\mathbb{P}\{N_n \in \cdot\}$ is regularly varying:

$$c_n \mathbb{P}\{N_n \in \cdot\} \to m, \qquad \text{in } \mathbf{M}_0(\mathbf{N}_p),$$

with

$$m(B) = \int_0^1 \int_{\mathbf{R}^d} I\{\delta_{(t,z)} \in B\} \mu(dz)dt.$$

Moving Averages

- MA(2) process: Let $\{Z_k\}$ be iid regularly varying ($\alpha > 1$, μ) and $A_0 > 0, A_1 < 0$ constants. Put

$$X_k = A_0 Z_k + A_1 Z_{k-1}, \qquad k \geq 1, S_n = X_1 + \cdots + X_n.$$

- Tempted to consider $S^{(n)}(t) = n^{-1} S_{[nt]}$ as an element in $\mathbf{D}[0,1]$ and study the convergence in $\mathbf{M}_0(\mathbf{D}[0,1])$ of

$$c_n \mathbb{P}\{S^{(n)} \in \cdot\}.$$

- WARNING: loosing tightness. Why?

Moving Averages

- MA(2) process: Let $\{Z_k\}$ be iid regularly varying ($\alpha > 1$, μ) and $A_0 > 0, A_1 < 0$ constants. Put

$$X_k = A_0 Z_k + A_1 Z_{k-1}, \qquad k \geq 1, S_n = X_1 + \cdots + X_n.$$

- Tempted to consider $S^{(n)}(t) = n^{-1} S_{[nt]}$ as an element in $\mathbf{D}[0,1]$ and study the convergence in $\mathbf{M}_0(\mathbf{D}[0,1])$ of

$$c_n \mathbb{P}\{S^{(n)} \in \cdot\}.$$

- WARNING: loosing tightness. Why?

Moving Averages

- MA(2) process: Let $\{Z_k\}$ be iid regularly varying ($\alpha > 1$, μ) and $A_0 > 0, A_1 < 0$ constants. Put

$$X_k = A_0 Z_k + A_1 Z_{k-1}, \qquad k \geq 1, S_n = X_1 + \cdots + X_n.$$

- Tempted to consider $S^{(n)}(t) = n^{-1} S_{[nt]}$ as an element in $\mathbf{D}[0,1]$ and study the convergence in $\mathbf{M}_0(\mathbf{D}[0,1])$ of

$$c_n \mathbb{P}\{S^{(n)} \in \cdot\}.$$

- WARNING: loosing tightness. Why?

Moving Averages

Loosing tightness

- By the **one big jump heuristic** you expect $S^{(n)}$ to be large because

$$X_k \approx A_0 Z_k, \quad \text{and} \quad X_{k+1} \approx A_1 Z_k$$

- For the partial sum process S_n you expect

$$S_k \approx A_0 Z_k, \quad \text{and} \quad S_{k+1} \approx (A_0 + A_1) Z_k,$$

so it takes two big jumps of opposite sign within a short period of time... loosing tightness in $\mathbf{D}[0,1]$.

Moving Averages

The Empirical Measure Level

- The problem with tightness can be resolved on the empirical measure level.
- We may consider

$$N_n = \sum_{k=1}^{n} \delta_{(\frac{k}{n}, n^{-1}X_k, n^{-1}X_{k-1})}.$$

- Then, the sequence $\mathbb{P}\{N_n \in \cdot\}$ is regularly varying:

$$c_n \mathbb{P}\{N_n \in \cdot\} \to m, \qquad \text{in } \mathbf{M}_0(\mathbf{N}_p),$$

with

$$m(B) = \int_0^1 \int_{\mathbf{R}^d} I\{\delta_{(t, A_0 z, 0)} + \delta_{(t, 0, A_1 z)} \in B\} \mu(dz) dt.$$

The Quality of Asymptotic Approximations

- The heavy-tailed asymptotics presented here are based on the heavy-tailed heuristics.
- One can anticipate that the approximations are good far out in the tail. How far?
- We will provide a small numerical study to illustrate the quality of the asymptotic approximations.

Asymptotic Approximations

$\alpha = 2$

Let $\{Z_k\}$ be iid $\mathbb{P}\{Z > z\} = (1+z)^{-\alpha}$, $z > 0$. Put $S_n = Z_1 + \cdots + Z_n$. Approximate $\mathbb{P}\{S_n > b\}$ by $n\mathbb{P}\{Z > b\}$.

$\alpha = 2, n = 5$			
	$n\mathbb{P}\{Z > b\}$	$\mathbb{P}\{S_n > b\}$	RE
$b = 25$	0.74e-2	1.05e-2	30%
$b = 100$	4.90e-4	5.34e-4	8%
$b = 5000$	1.999e-7	2.002e-7	0.16%
$\alpha = 2, n = 20$			
	$n\mathbb{P}\{Z > b\}$	$\mathbb{P}\{S_n > b\}$	RE
$b = 400$	1.24e-4	1.38e-4	10%
$b = 4000$	1.249e-6	1.261e-6	1%

Asymptotic Approximations

$\alpha = 4$, $\alpha = 6$

Let $\{Z_k\}$ be iid $\mathbb{P}\{Z > z\} = (1+z)^{-\alpha}$, $z > 0$. Put $S_n = Z_1 + \cdots + Z_n$. Approximate $\mathbb{P}\{S_n > b\}$ by $n\mathbb{P}\{Z > b\}$.

$\alpha = 4$, $n = 20$			
	$n\mathbb{P}\{Z > b\}$	$\mathbb{P}\{S_n > b\}$	RE
$b = 50$	2.95e-6	4.90e-6	28%
$b = 200$	1.23e-8	1.36e-8	10%
$b = 1000$	1.99e-11	2.05e-11	3%
$\alpha = 6$, $n = 20$			
	$n\mathbb{P}\{Z > b\}$	$\mathbb{P}\{S_n > b\}$	RE
$b = 25$	6.47e-8	1.18e-7	45%
$b = 40$	4.21e-9	6.15e-9	31%
$b = 100$	1.88e-11	2.26e-11	16%

Illustrations of the One Big Jump

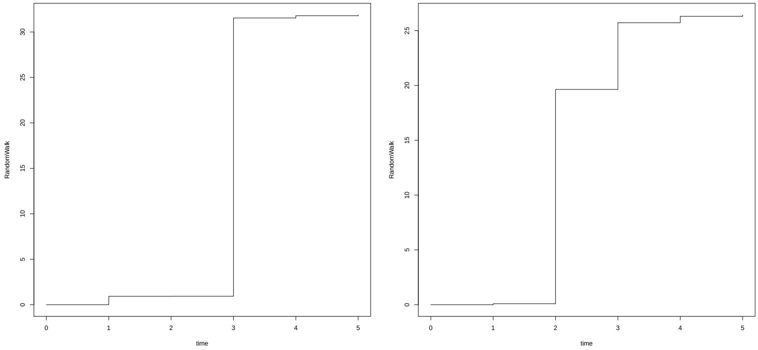

Figure: Trajectories of a random walk exceeding the level. $n = 5, b = 25, \alpha = 2$.

Illustrations of the One Big Jump

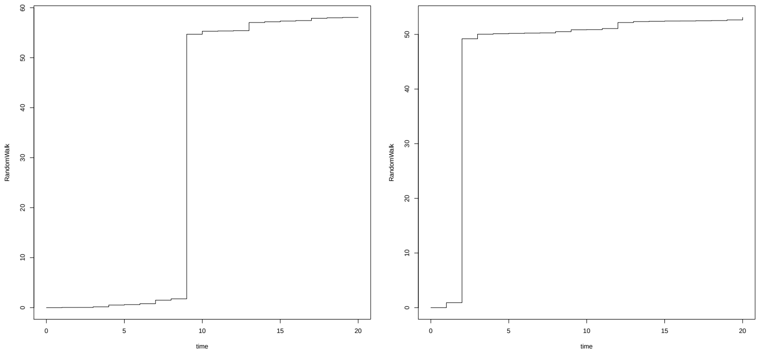

Figure: Trajectories of a random walk exceeding the level. $n = 20, b = 50, \alpha = 4$.

Illustrations of the One Big Jump

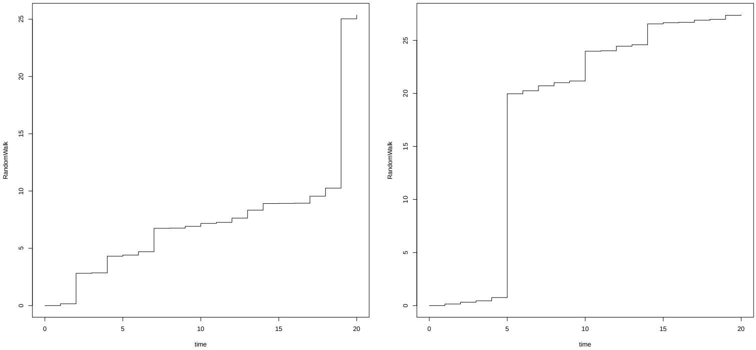

Figure: Trajectories of a random walk exceeding the level.
$n = 20, b = 25, \alpha = 4$. $\mathbb{P}\{S_n > b\} = 1.65\text{e-}4$

For Further Reading I

S. Resnick
Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.
Springer, New York, 2006.

H. Hult and F. Lindskog
Extremal behavior for regularly varying stochastic processes.
Stochastic Process. Appl. 115, 249-274, 2005.

H. Hult, F. Lindskog, T. Mikosch, and G. Samorodnitsky
Functional large deviations for multivariate regularly varying random walks
Ann. Appl. Probab. 15(4), 2651-2680.

For Further Reading II

- H. Hult and F. Lindskog
  Regular variation for measures on metric spaces.
  Publ. Inst. Math. 80, 121-140, 2006.
- H. Hult and F. Lindskog
  Extremal behavior of stochastic integrals driven by regularly varying Lévy processes.
  Ann. Probab., 35, 309-339, 2007.
- H. Hult and G Samorodnitsky
  Large deviations for point processes based on stationary sequences with heavy tails.
  J. Appl. Prob. 47, 1-40, 2010.