Sample Path Large Deviations for Heavy-Tailed Lévy Processes and Random Walks

Bert Zwart

Centrum Wiskunde & Informatica
Amsterdam

Joint with Chang-Han Rhee (CWI) and Jose Blanchet (Columbia U)

http://arxiv.org/abs/1606.02795

What this talk is about

Let $X(t), t \geq 0$ be a Lévy process, such that

$$\mathbf{P}(X(1) > x) = L_+(x)x^{-\alpha}, \qquad \mathbf{P}(X(1) < -x) = L_-(x)x^{-\beta}.$$

Set

$$\bar{X}_n(t) = X(nt)/n, \qquad \bar{X}_n = \{\bar{X}_n(t), t \in [0,1]\}.$$

Can we obtain sample-path large deviations for $\bar{X}_n$? i.e.

$$\mathbf{P}(\bar{X}_n \in A) \sim \quad ???, \qquad n \to \infty$$

Overview

- Motivation and introduction
- Main result: sample path large deviations principles
- Implications:
  - Random walks
  - Conditional limit theorem
  - Connections to the standard LD framework
- Examples
- Comments on proof

Large deviations

Let $U_i, i \geq 1$ be an i.i.d. sequence with $E[U_1] = 0$.

Let $S_0 = 0$ and for $n \geq 1$, let $S_n = U_1 + \ldots + U_n$.

The weak law of large numbers (WLLN) states that

$$\lim_{n\to\infty} P(|S_n/n| > \epsilon) = 0$$

for every $\epsilon > 0$.

How fast is the convergence to 0 in the WLLN?

Cramérs theorem

- Cumulant generating function of U_1: $\Lambda(s) = \log E[e^{sU_1}]$
- Convex conjugate of Λ: $\Lambda^*(a) = \sup_{s \geq 0}[as - \Lambda(s)]$.

Theorem (Cramér (1938))

For $a > 0$:

$$\lim_{n\to\infty} \frac{1}{n} \log P(S_n/n > a) = -\Lambda^*(a).$$

What about $\mathbf{P}(S_n/n \in A)$?

The large-deviations principle (Varadhan, 1966)

A sequence of r.v.'s Z_n satisfies an LDP with rate function I if I is lower semi-continuous and

$$-\inf_{\xi \in A^\circ} I(\xi) \le \liminf_{n\to\infty} \frac{\log \mathbf{P}(Z_n \in A)}{n} \le \limsup_{n\to\infty} \frac{\log \mathbf{P}(Z_n \in A)}{n} \le -\inf_{\xi \in A^-} I(\xi),$$

Example: $Z_n = S_n/n$ and $I = \Lambda^*$.

Some cornerstones:

- Contraction principle (analogue of continuous mapping principle)
- Connections with convex & variational analysis, statistical physics, ...

Often Crucial: the property that I is **good** (i.e. has **compact level sets**).

Heavy tails

Cramérs theorem does not give a precise answer (0) if

$$E[e^{\epsilon U}] = \infty$$

for all $\epsilon > 0$. In this case, we say that U has a heavy (right) tail.

Examples of heavy tails:

- Pareto: $P(U > x) \sim x^{-\alpha}$
- Lognormal: $P(U > x) \sim e^{-(\log x)^2}$
- Weibull: $P(U > x) \sim e^{-x^{\alpha}}$, $\alpha \in (0, 1)$.
- Any df with a hazard rate decreasing to 0.

The principle of a single big jump

Theorem (A. Nagaev, 1969) Let U_i be i.i.d. with mean 0 and $\mathbf{P}(U_1 > x) = L(x)x^{-\alpha}$. Let $S_n = U_1 + \ldots + U_n$ and $a > 0$. Then

$$\mathbf{P}(S_n/n > a) \sim n\mathbf{P}(U_1 > an).$$

Heavy-tailed analogue of Cramérs theorem

Most general formulation of this result in Dieker, Denisov, Shneer (2008).
Necessary (due to CLT): $\mathbf{P}(U_1 > n) \sim \mathbf{P}(U_1 > n + O(\sqrt{n}))$.

In heavy-tailed world, rare events happen by **catastrophes**.
In light-tailed world, by **conspiracies**

Main motivation of this work

In many applications, we can write our object of interest as a mapping Ψ of the **sample path** of a random walk or Levy process.
Recall

$$\bar{X}_n(t) = X(nt)/n, \qquad \bar{X}_n = \{\bar{X}_n(t), t \in [0,1]\}.$$

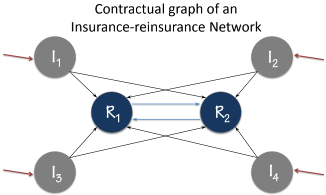
Let Ψ be a continuous mapping on $\mathbb{D}$. What can we say about

$$\mathbf{P}(\Psi(\bar{X}_n) \in B) = \mathbf{P}(\bar{X}_n \in \Psi^{-1}(B))?$$

Need

- Estimate of $\mathbf{P}(\bar{X}_n \in A)$ for sufficiently many sets A.
- A version of a contraction principle/continuous mapping theorem
- Sample-path result of Hult/Lindskog/Mikosch/Samorodnitsky (2005) applies to single big jump case only

Additional motivation: multiple big jumps



- Insurance-reinsurance networks (Blanchet & Shi 2015)
- Fluid queues with On-Off sources (Z, Borst, Mandjes 2004)
- Many-server queues (Foss & Korshunov 2006, 2012)

Our vision: a structural approach to such problems, in line with large-deviations theory for light-tailed systems and weak convergence

Spectrally positive Lévy processes

$$X(s) = sa + B(s) + \int_{0<x\le 1} x[N([0,s]\times dx) - s\nu(dx)] + \int_{x>1} xN([0,s]\times dx),$$

- $E[X(s)] = 0$,
- a is a drift parameter,
- B a Brownian motion,
- N is a Poisson random measure with mean measure Leb$\times\nu$ on $[0,1]\times(0,\infty)$;
- ν is a measure on $(0,\infty)$ satisfying $\int_0^1 x^2\nu(dx) < \infty$ and $\nu(x,\infty) = L(x)x^{-\alpha}$, $\alpha > 1$.

Spectrally positive Lévy processes (2)

Some notation:

- $\mathbb{D}_j$: subspaces of the Skorokhod space $\mathbb{D}$ consisting of nondecreasing step functions, vanishing at the origin, with exactly j jumps
- $\mathbb{D}_{\leqslant j} \triangleq \bigcup_{0 \leq i \leq j} \mathbb{D}_i$
- $\mathcal{D}_+(\xi)$: the number of upward jumps of an element ξ in $\mathbb{D}$.
- Finally, set

$$\mathcal{J}(A) \triangleq \inf_{\xi \in \mathbb{D}_{<\infty} \cap A} \mathcal{D}_+(\xi).$$

Main result for one-sided case

Theorem 1

Suppose that A is a measurable set. If $\mathcal{J}(A) < \infty$, and if A is bounded away from $\mathbb{D}_{\leqslant \mathcal{J}(A)-1}$, then

$$C_{\mathcal{J}(A)}(A^\circ) \le \liminf_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in A)}{(n\nu[n,\infty))^{\mathcal{J}(A)}} \le \limsup_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in A)}{(n\nu[n,\infty))^{\mathcal{J}(A)}} \le C_{\mathcal{J}(A)}(A^-).$$

Recall

$$\mathcal{J}(A) \triangleq \inf_{\xi \in \mathbb{D}_{<\infty} \cap A} \mathcal{D}_+(\xi).$$

$\mathcal{J}(A)$ is the number of jumps needed to achieve the event $\bar{X}_n \in A$.

The pre-factor and a conditional limit theorem

- ν_α^j: restriction to $\mathbb{R}_+^{j\downarrow}$ of the j-fold product measure of ν_α, where $\nu_\alpha(x,\infty) \triangleq x^{-\alpha}$.
- For $j \geq 1$, $C_j(\cdot) \triangleq \mathbf{E}\Big[\nu_\alpha^j\{y \in (0,\infty)^j : \sum_{i=1}^j y_i 1_{[U_i,1]} \in \cdot\}\Big]$, with $U_i, i \geq 1$ are i.i.d. uniform on $[0,1]$.

If, in addition, $C_{\mathcal{J}(A)}(A^\circ) = C_{\mathcal{J}(A)}(A^-)$ and if A is bounded away from $\mathbb{D}_{\leqslant \mathcal{J}(A)-1}$, Theorem 1 implies

$$\mathbf{P}(\bar{X}_n \in \cdot \mid \bar{X}_n \in A) \xrightarrow{d} \frac{C_{\mathcal{J}(A)}(\cdot \cap A)}{C_{\mathcal{J}(A)}(A)}.$$

Two-sided Lévy processes

$$X(s) = sa + B(s) + \int_{0<|x|\leq 1} x[N([0,s]\times dx) - s\nu(dx)] + \int_{|x|>1} xN([0,s]\times dx),$$

- $E[X(s)] = 0$,
- a is a drift parameter,
- B a Brownian motion,
- N is a Poisson random measure with mean measure Leb$\times\nu$ on $[0,1]\times(-\infty,\infty)$;
- ν is a measure on $(-\infty,\infty)$ satisfying $\int_{-1}^{1} x^2\nu(dx) < \infty$,

$$\nu(x,\infty) = L_+(x)x^{-\alpha}, \qquad \nu(-\infty,-x) = L_-(x)x^{-\beta},$$

$\alpha, \beta > 1$.

Two-sided Lévy processes (2)

Some notation:

- $\mathbb{D}_{j,k}$: step functions vanishing at the origin with exactly j upward jumps and k downward jumps.
- $\mathbb{D}_{<j,k} = \bigcup_{(l,m)\in I_{<j,k}} \mathbb{D}_{l,m}$ and $I_{<j,k} = \{(l,m) \in \mathbb{Z}_+^2 \setminus (j,k) : (\alpha-1)l+(\beta-1)m \leq (\alpha-1)j+(\beta-1)k\}$.
- Let $\mathcal{I}(j,k) \triangleq (\alpha-1)j + (\beta-1)k$, and consider

$$(\mathcal{J}(A), \mathcal{K}(A)) \in \underset{(j,k)\in\mathbb{Z}_+^2;\mathbb{D}_{j,k}\cap A\neq\emptyset}{\arg\min} \mathcal{I}(j,k).$$

Main result for two-sided case

Theorem 2 Suppose that A is a measurable set. If the argument minimum $(\mathcal{J}(A), \mathcal{K}(A))$ is unique and A is bounded away from $\mathbb{D}_{<\mathcal{J}(A),\mathcal{K}(A)}$, then

$$\liminf_{n\to\infty} \frac{P(\bar{X}_n \in A)}{(n\nu[n,\infty))^{\mathcal{J}(A)}(n\nu(-\infty,-n])^{\mathcal{K}(A)}} \geq C_{\mathcal{J}(A),\mathcal{K}(A)}(A^\circ)$$

$$\limsup_{n\to\infty} \frac{P(\bar{X}_n \in A)}{(n\nu[n,\infty))^{\mathcal{J}(A)}(n\nu(-\infty,-n])^{\mathcal{K}(A)}} \leq C_{\mathcal{J}(A),\mathcal{K}(A)}(A^-).$$

$$C_{j,k}(\cdot) \triangleq \mathbf{E}\Big[\nu_\alpha^j \times \nu_\beta^k\{(x,y) \in (0,\infty)^{j+k} : \sum_{i=1}^{j} x_i 1_{[U_i,1]} - \sum_{i=1}^{k} y_i 1_{[V_i,1]} \in \cdot\}\Big]$$

The rate function

- The minimization problem

$$\underset{(j,k)\in\mathbb{Z}_+^2;\mathbb{D}_{j,k}\cap A\neq\emptyset}{\arg\min} \quad (\alpha-1)j+(\beta-1)k$$

  can be cast as a **deterministic impulse control** problem.

- Related framework: Barles (1985).
  [Stochastic IC: Bensoussan & Lions (1984), Dai & Yao (2014)]

- Optimality of such problems can be described by quasi-variational inequalities.

- In examples we solve the minimization problem directly.

Random walks

- $S(k), k \geq 0$: centered random walk
- $\bar{S}_n(t) = S([nt])/n, t \geq 0$, and $\bar{S}_n = \{\bar{S}_n(t), t \in [0,1]\}$.
- $N(t), t \geq 0$: unit rate Poisson process.
- $X(t) \triangleq S_{N(t)}, t \geq 0$, $\bar{X}_n(t)$ and $\bar{X}_n$ are as defined before

$\bar{S}_n$ satisfies the same limit theorem as $\bar{X}_n$.

Proof idea: Skorokhod distance between $\bar{S}_n$ and $\bar{X}_n$ is bounded by

$$\sup_{t \in [0,1]} [N(nt)/n - t].$$

Connection with large deviations framework

Let $\bar{X}_n$ be a centered and scaled two sided Lévy process with Lévy measure ν satisfying

$$\nu(x,\infty) = L_+(x)x^{-\alpha}, \qquad \nu(-\infty,-x) = L_-(x)x^{-\beta},$$

Define

$$I(\xi) \triangleq \begin{cases} (\alpha-1)\mathcal{D}_+(\xi) + (\beta-1)\mathcal{D}_-(\xi), & \text{if } \xi \text{ is a step function, } \xi(0)=0, \\ \infty, & \text{otherwise.} \end{cases}$$

I is lower semi-continuous on $\mathbb{D}$ but does not have compact level sets
$\Rightarrow$ I is a rate function, but not a good rate function

Weak large deviations principle

Theorem 3
The scaled process $\bar{X}_n$ satisfies the weak large deviations principle with rate function I and speed $\log n$, i.e.,

$$\liminf_{n\to\infty} \frac{\log \mathbf{P}(\bar{X}_n \in G)}{\log n} \geq - \inf_{x\in G} I(x)$$

for every open set G, and

$$\limsup_{n\to\infty} \frac{\log \mathbf{P}(\bar{X}_n \in K)}{\log n} \leq - \inf_{x\in K} I(x)$$

for every compact set K.

Proof idea: reduce everything to balls, and use Theorem 2.

Nonexistence of strong large deviations principle

The upper bound in Theorem 3 can not be extended to all closed sets.

- Take $\alpha = \beta = 2$ for ease of exposition
- Set $\pi(\xi) \triangleq \Big(\sup_{t\in(0,1]} \big(\xi(t) - \xi(t-)\big), \sup_{t\in(0,1]} \big(\xi(t-) - \xi(t)\big) \Big)$.
- If $\bar{X}_n$ satisfies a strong LDP, the contraction principle implies that $\pi(\bar{X}_n)$ satisfies an LDP with rate function I' given by

$$I'(y_1, y_2) = \mathbb{I}(y_1 > 0) + \mathbb{I}(y_2 > 0).$$

Nonexistence of strong large deviations principle (ctd)

- Consider the closed set $A \triangleq \bigcup_{k=2}^{\infty}[\log k, \infty) \times [k^{-1/2}, \infty)$.
- $\mathbf{P}(\pi(\bar{X}_n) \in A) \geq \mathbf{P}(\pi(\bar{X}_n) \in [\log n, \infty) \times [n^{-1/2}, \infty))$,
- The log of the RHS behaves like $-1 * \log n$ (one big jump needed),
- This contradicts with $-\inf_{(y_1,y_2)\in A} I'(y_1, y_2) = -2$ (suggesting two big jumps are needed),

We need the condition that A is bounded away from the set of step functions with at most one big jump

Examples

- We consider several functionals of $\bar{X}_n$, leading to specific A.
- Need to check that A is bounded away from $\mathbb{D}_{\leqslant \mathcal{J}(A)-1}$
- It suffices to check that $A^\delta \cap \mathbb{D}_{<\infty}$ is bounded away from $\mathbb{D}_{\leqslant \mathcal{J}(A)-1}$.

Moderate jumps

Motivated by a reinsurance problem we consider

$$Q(n) \triangleq \mathbf{P}\left(\sup_{t\in[0,1]} [\bar{X}_n(t) - ct] \geq a; \sup_{t\in[0,1]} [\bar{X}_n(t) - \bar{X}_n(t-)] \leq b \right),$$

$$A \triangleq \{\xi \in \mathbb{D} : \sup_{t\in[0,1]} [\xi(t) - ct] \geq a; \sup_{t\in[0,1]} [\xi(t) - \xi(t-)] \leq b\},$$

$$\mathcal{J}(A) = \lceil a/b \rceil.$$

Condition on A holds iff a/b is not an integer, in which case

$$Q(n) \sim C_{\lceil a/b \rceil}(A)(n\nu[n,\infty))^{\lceil a/b \rceil}.$$

A barrier digital option

Consider a Lévy-driven Ornstein-Uhlenbeck process of the form

$$d\bar{Y}_n(t) = -\kappa d\bar{Y}_n(t) + d\bar{X}_n(t), \qquad \bar{Y}_n(0) = 0.$$

We apply our results to estimate

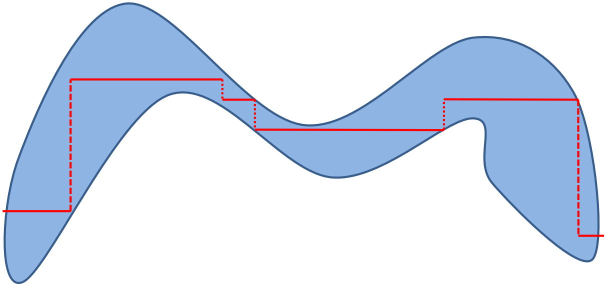
$$b(n) = \mathbf{P}(\inf_{0\le t\le 1} \bar{Y}_n(t) \le -a_-, \bar{Y}_n(1) \ge a_+).$$

Theorem 2 applies, and we obtain

$$b(n) \sim C_{1,1}(A)\, n\nu[n,\infty)n\nu(-\infty,-n]$$

as $n \to \infty$.

$A = \{\xi : l \leq \xi \leq u\}$: only jump when you must



1

Proof: M-convergence (Lindskog, Resnick, Roy 2014)

- Let $(\mathbb{S}, d)$ be a complete separable metric space, and $\mathscr{S}$ be the Borel σ-algebra on $\mathbb{S}$.
- Given a closed subset $\mathbb{C}$ of $\mathbb{S}$, define $\mathbb{C}^r \triangleq \{x \in \mathbb{S} : d(x, \mathbb{C}) < r\}$ for $r \geq 0$, and let $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ be the class of measures defined on $\mathscr{S}_{\mathbb{S}\setminus\mathbb{C}}$ whose restrictions to $\mathbb{S} \setminus \mathbb{C}^r$ are finite for all $r > 0$.
- $\mathcal{C}_{\mathbb{S}\setminus\mathbb{C}}$ is the set of real-valued, non-negative, bounded, continuous functions whose support is bounded away from $\mathbb{C}$ (i.e., $f(\mathbb{C}^r) = \{0\}$ for some $r > 0$).
- A sequence of measures $\mu_n \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ converges to $\mu \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ if $\mu_n(f) \to \mu(f)$ for each $f \in \mathcal{C}_{\mathbb{S}\setminus\mathbb{C}}$.
- For Theorem 1, we take $\mathbb{S} = \mathbb{D}$ and $\mathbb{C} = \mathbb{D}_{\leq j-1}$.

Characterization of M convergence (LRR2014)

Let $\mu, \mu_n \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})$. Then $\mu_n \to \mu$ in $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ as $n \to \infty$ if and only if

$$\limsup_{n\to\infty} \mu_n(F) \leq \mu(F) \tag{1}$$

for all closed $F \in \mathscr{S}_{\mathbb{S}\setminus\mathbb{C}}$ bounded away from $\mathbb{C}$ and

$$\liminf_{n\to\infty} \mu_n(G) \geq \mu(G) \tag{2}$$

for all open $G \in \mathscr{S}_{\mathbb{S}\setminus\mathbb{C}}$ bounded away from $\mathbb{C}$.

For Theorem 1, we take $\mathbb{S} = \mathbb{D}$ and $\mathbb{C} = \mathbb{D}_{\leq j-1}$.

Asymptotic equivalence (RBZ2016)

Suppose that X_n and Y_n are random elements taking values in a complete separable metric space $(\mathbb{S}, d)$. Y_n is said to be asymptotically equivalent to X_n with respect to ϵ_n and $\mathbb{C}$, if, for each $\delta > 0$ and $\gamma > 0$,

$$\limsup_{n\to\infty} \epsilon_n^{-1} \mathbf{P}(X_n \in (\mathbb{S} \setminus \mathbb{C})^{-\gamma}, d(X_n, Y_n) \geq \delta) = 0$$

$$\limsup_{n\to\infty} \epsilon_n^{-1} \mathbf{P}(Y_n \in (\mathbb{S} \setminus \mathbb{C})^{-\gamma}, d(X_n, Y_n) \geq \delta) = 0.$$

For Theorem 1 it suffices to take $\mathbb{C} = \emptyset$.

Proof: M-convergence: the one-sided case

Theorem 1 follows from

Theorem 1'
For each $j \geq 1$,

$$(n\nu[n,\infty))^{-j}\mathbf{P}(\bar{X}_n \in \cdot) \to C_j(\cdot), \tag{3}$$

in $\mathbb{M}(\mathbb{D} \setminus \mathbb{D}_{\leqslant j-1})$, as $n \to \infty$.
Proof:

- $\bar{X}_n$ is asymptotically equivalent to J_n^j which is the process obtained from $\bar{X}_n$ keeping its j biggest jumps.
- Show both are asymptotically equivalent
- Use representation for J_n^j and many detailed technical estimates

Final Comments

- M convergence does not seem to deal easily with continuous maps of superpositions of processes
- Consequently, proof much more technical in two-sided case
- Some current/future topics:
  - application to rare-event simulation
  - subexponential (Weibull) tails
  - more exotic examples requiring infinitely many jumps, e.g.

$$\mathbf{P}(t \leq \bar{X}_n(t) \leq 2t, t \in [0,1]). \tag{4}$$