Large deviations for point processes based on stationary sequences with heavy tails

Gennady Samorodnitsky
jointly with Henrik Hult

March 2010

Stationary processes and large deviations

Let $X_1, X_2, \ldots,$ be a stationary process. Many questions of interest about the process turn out to be large deviations questions.

Suppose, for example, that the process is ergodic, and has a finite and negative mean. Let

$$W = \sup_{n=0,1,2,\ldots} (X_1 + X_2 + \ldots + X_n) .$$

Question: what is the behaviour of the tail probability for the supremum: $P(W > x)$ for large x?

A different question: how long are the long "strange" stretches of time? For example, for some $\epsilon > 0$, what is the behaviour of the probability

$$P\Big(\frac{X_{i+1} + \ldots + X_j}{j - i} > EX + \epsilon$$

$$\text{for some } 1 \leq i < j \leq n, \;\; j - i > a_n\Big)$$

for some $a_n \uparrow \infty$ sufficiently fast (but with $a_n = o(n)$).

Note: in both cases we are talking about probabilities of events that depend not only on the sizes of $X_1, X_2, \ldots$, but also on their order.

We are specifically interested in a situation where:

- (X_n) form a stationary *heavy tailed* process;
- (X_n) are "tail dependent".

"Tail dependence" means that the extreme values of the sequence (X_n) may cluster.

Specifically, for the models we will consider, for some $k = 1, 2, \ldots$,

$$\liminf_{x \to 0} \frac{P(X_1 > x,\, X_{k+1} > x)}{P(X_1 > x)} > 0\,.$$

The tool: abstract large deviations

One of the main tools for handling large deviations questions is via *abstract large deviations*. The first level of those is the functional large deviations.

Suppose, for a moment, that $E|X| < \infty$. Define, for $n = 1, 2, \ldots$,

$$S_n(t) = \sum_{j=1}^{[nt]} (X_j - EX),\ 0 \le t \le 1\,.$$

Then each S_n is a random element of the space $D[0,1]$.

Let $\gamma_n \uparrow \infty$ so fast that

$$\frac{1}{\gamma_n}\sum_{j=1}^{n}(X_j - EX) \to 0$$

in probability. A function space large deviations result would describe the behaviour of the small probability

$$P\Big(\frac{1}{\gamma_n} S_n(\cdot) \in A\Big),\ n \to \infty\,,$$

where A is a set of functions whose closure does not contain the zero function.

A typical statement in the light tail case:

$$\frac{1}{r_n} \log P\left(\frac{1}{\gamma_n} S_n(\cdot) \in A\right) \approx - \inf_{f \in A} J(f),$$

where $r_n \uparrow \infty$, and J is a *rate function*.

A typical statement in the heavy tail case:

$$r_n P\left(\frac{1}{\gamma_n} S_n(\cdot) \in A\right) \approx \mu(A)\,,$$

where $r_n \uparrow \infty$, and μ is a *large deviations measure*.

Both give the most likely paths for the rare event to happen.

Heavy tails: multivariate regular variation

We say that a d-dimensional random vector Z has a regularly varying distribution if there exists a non-null Radon measure θ_α on $\overline{\mathbb{R}}^d \setminus \{0\}$ with $\theta_\alpha(\overline{\mathbb{R}}^d \setminus \mathbb{R}^d) = 0$ such that

$$\frac{P(u^{-1}Z \in \cdot)}{P(|Z| > u)} \xrightarrow{v} \theta_\alpha(\cdot)$$

on $\overline{\mathbb{R}}^d \setminus \{0\}$.

The limiting measure θ_α necessarily obeys a homogeneity property: there is an $\alpha > 0$ such that $\theta_\alpha(uB) = u^{-\alpha}\theta_\alpha(B)$ for all Borel sets $B \subset \overline{\mathbb{R}}^d \setminus \{0\}$.

In the one–dimensional case this is equivalent to the statement that

$$P(|Z| > u) = u^{-\alpha} L(u), \ \ u > 0$$

for $\alpha > 0$ and L a slowly varying function, and

$$\frac{P(Z > u)}{P(|Z| > u)} \to p \text{ as } u \to \infty$$

for $0 \leq p \leq 1$. If $q = 1 - p$, then

$$\theta_\alpha(dx) = \begin{cases} p\alpha x^{-(1+\alpha)} dx & x > 0, \\ q\alpha |x|^{-(1+\alpha)} dx & x < 0. \end{cases}$$

Suppose, for a moment, that (X_n) are regularly varying, nonnegative and i.i.d.

Then $r_n = \big(nP(X > \gamma_n)\big)^{-1}$, and the large deviations measure is given by

$$\mu(A) = \mathrm{Leb} \times \theta_\alpha \Big\{ (t,z) \in (0,1) \times (0,\infty) :\ z\mathbf{1}(\cdot > t) \in A \Big\}.$$

θ_α is a measure on $(0,\infty)$ with $\theta_\alpha(x,\infty) = x^{-\alpha}$, $x > 0$.

The large deviations measure is concentrated on step functions with a single step.

This corresponds to exactly one of the "steps" $X_1, \ldots, X_n$ in $S_n(t) = \sum_{j=1}^{[nt]} (X_j - EX)$, $0 \le t \le 1$, being unusually large, and causing the rare event $\gamma_n^{-1} S_n(\cdot) \in A$.

The usefulness of the method is reduced when the "steps" are dependent.

Suppose the "steps" (X_n) follow the MA(1) model:

$$X_n = Z_n + \phi Z_{n-1}, \quad (Z_n) \text{ i.i.d. with regularly varying tails.}$$

Then the rare event will be caused by exactly one of $Z_0, Z_1, \ldots, Z_n$ (say, Z_k) being unusually large. This will mean that

$$\begin{aligned} X_1 + \ldots + X_j &\approx 0, & 0 < j < k \\ X_1 + \ldots + X_k &\approx Z_k, & \\ X_1 + \ldots + X_j &\approx (1+\phi)Z_k, & k+1 \le j \le n\,. \end{aligned}$$

The problem: the large deviations measure has to be of the form

$$\mu(A) = \mathrm{Leb} \times \theta_\alpha \Big\{ (t, z) \in (0,1) \times (0,\infty) : \, (1+\phi) z \mathbf{1}(\cdot > t) \in A \Big\}.$$

In the limit the two jumps coalesce into one, and the order of arrival of the large "steps" $X_1, \ldots, X_n$ is lost.

To solve this problem one needs a different abstract large deviations setup, where the order of arrivals of the large "steps" is preserved in the limit.

We replace function space large deviations with measure space large deviations.

Large deviations for point processes

Given a stationary process with regularly varying tails (X_n), with values in $\mathbb{R}^d$, we construct, for an integer $q \geq 1$, a sequence of point processes

$$N_n^q = \sum_{k=1}^{n} \delta_{(k/n, \gamma_n^{-1} X_k, \gamma_n^{-1} X_{k-1}, \ldots, \gamma_n^{-1} X_{k-q})}, \ n = 1, 2, \ldots,$$

on the space $\mathbf{E}_q = [0,1] \times (\mathbf{R}^{d(q+1)} \setminus \{0\})$. The sequence (γ_n) grows fast enough to assure that $\gamma_n^{-1} \max(X_1, \ldots, X_n) \to 0$ in probability.

The idea: to preserve the information on the order of at least q largest "steps" in the limit.

Technical framework

Let $\mathbf{N}_p$ be the space of Radon point measures on $\mathbf{E}_q = [0,1] \times (\mathbf{R}^{d(q+1)} \setminus \{0\})$. We equip $\mathbf{N}_p$ with the (metrizable) topology of vague convergence.

Let ξ_0 be the null measure in $\mathbf{N}_p$. By the choice of the sequence (γ_n), the sequence of point processes N_n^q converges in probability to ξ_0.

Let $\mathbf{M}_0 = \mathbf{M}_0(\mathbf{N}_p \setminus \{\xi_0\})$ be the space of Radon measures on $\mathbf{N}_p \setminus \{\xi_0\}$ whose restriction to $\mathbf{N}_p \setminus B_{\xi_0,r}$ is finite for each $r > 0$.

We define convergence in $\mathbf{M}_0$ ($m_n \to m$) by requiring the convergence $m_n(f) \to m(f)$ for all $f \in C_0(\mathbf{N}_p \setminus \{\xi_0\})$, the space of bounded continuous functions on $\mathbf{N}_p \setminus \{\xi_0\}$ that vanish in a neighborhood of "the origin" ξ_0.

We are looking for a sequence $r_n \uparrow \infty$ and a measure $m \in \mathbf{M}_0$ such that

$$r_n P\left(N_n^q \in \cdot\right) \to m \ \text{ in } \mathbf{M}_0.$$

This will be the point process level large deviation principle for the stationary heavy tailed process (X_n), and it will provide answers to many interesting large deviation questions in queuing and other areas of applications.

Setup for the main result

We consider a stationary d-dimensional stochastic process $(X_k)_{k\in\mathbf{Z}}$ with the stochastic representation

$$X_k = \sum_{j\in\mathbf{Z}} A_{k,j} Z_{k-j},\ k \in \mathbf{Z},\ \text{ where:}$$

- the sequence $(Z_j)_{j\in\mathbf{Z}}$ consists of independent and identically distributed random vectors with values in $\mathbf{R}^p$, that are multivariate regularly varying with exponent $\alpha > 0$;
- the sequence $(\mathbb{A}_k)_{k\in\mathbf{Z}}$ is stationary and each $\mathbb{A}_k$ is itself a sequence of random $(d \times p)$ matrices, $\mathbb{A}_k = (A_{k,j})_{j\in\mathbf{Z}}$;
- the sequence $(\mathbb{A}_k)_{k\in\mathbf{Z}}$ is independent of the sequence $(Z_k)_{k\in\mathbf{Z}}$.

The stationary process $(\mathbb{A}_k)_{k\in\mathbf{Z}}$ is assumed to have lighter tails than $(Z_k)_{k\in\mathbf{Z}}$.

Some examples

Example 1 Linear processes Let (A_j) be a sequence of deterministic real-valued $d \times p$-matrices. Let (Z_j) be a sequence of i.i.d. p-dimensional random vectors that are multivariate regularly varying with exponent $\alpha > 0$. Then

$$X_k = \sum_{j=-\infty}^{\infty} A_j Z_{k-j}, \; k \geq 0$$

is a (stationary) d-dimensional linear process.

Example 2 Stochastic recursions Let $(Y_k, Z_k)_{k\in\mathbb{Z}}$ be a sequence of independent and identically distributed pairs of $d \times d$-matrices and d-dimensional random vectors, with Z being multivariate regularly varying with exponent $\alpha > 0$. The stochastic recursion

$$X_k = Y_k X_{k-1} + Z_k, \quad k \in \mathbb{Z}$$

has, under some conditions, a stationary solution, and we consider the corresponding stationary process.

Theorem Assume that there is $0 < \varepsilon < \alpha$ such that

$$\sum E\|A_j\|^{\alpha-\varepsilon} < \infty \quad \text{and} \quad \sum E\|A_j\|^{\alpha+\varepsilon} < \infty, \quad \alpha \in (0,1) \cup (1,2),$$

$$E\Big(\sum \|A_j\|^{\alpha-\varepsilon}\Big)^{\frac{\alpha+\varepsilon}{\alpha-\varepsilon}} < \infty, \quad \alpha \in \{1,2\},$$

$$E\Big(\sum \|A_j\|^{2}\Big)^{\frac{\alpha+\varepsilon}{2}} < \infty, \quad \alpha \in (2,\infty).$$

Then the series defining (X_k) converges a.s. and

$$\frac{P(u^{-1}X \in \cdot\,)}{P(|Z| > u)} \to E\Big[\sum \theta_\alpha \circ A_j^{-1}(\cdot)\Big],$$

in $\mathbf{M}_0(\mathbf{R}^d)$.

Main result

Assume that sequence (γ_n) grows so fast that

$$\left.\begin{array}{rl} (Z_1 + \cdots + Z_n)/\gamma_n \to 0, & \text{in probability and} \\ \gamma_n/\sqrt{n^{1+\varepsilon}} \to \infty, & \text{for some } \varepsilon > 0 \text{ if } \alpha = 2, \\ \gamma_n/\sqrt{n \log n} \to \infty, & \text{if } \alpha > 2. \end{array}\right\}$$

Let

$$r_n = \frac{1}{nP(|Z| > \gamma_n)},\ n = 1, 2, \ldots .$$

For $q \geq 0$ define a random map from $[0,1] \times \mathbf{R}^p \setminus \{0\})$ into the space $\mathbf{N}_p$ by

$$T_{\mathbb{A},q}(t,z) = \sum_{j \in \mathbf{Z}} \delta_{(t, A_{j,j}z, A_{j-1,j-1}z, \ldots, A_{j-q,j-q}z)} \,.$$

Main Theorem The following large deviation principle holds:

$$m_n^q(\cdot) = r_n P(N_n^q \in \cdot\,) \to E[(\mathrm{Leb} \times \theta_\alpha) \circ T_{\mathbb{A},q}^{-1}(\cdot)] =: m^q(\cdot)$$

in $\mathbf{M}_0$, where θ_α is the tail measure of the noise variables $(Z_k)_{k \in \mathbf{Z}}$.

Applications

Large deviations for partial sums

Consider the large deviations of the partial sums $S_n = X_1 + \cdots + X_n$, $n = 1, 2 \ldots$. These are obtain from the large deviations for point processes by summing up the points and applying the continuous mapping argument.

Theorem Under certain additional technical assumptions,

$$r_n P(\gamma_n^{-1} S_n \in \cdot\,) \to E\Big[\theta_\alpha\Big(z : \sum_{j \in \mathbf{Z}} A_{j,j} z \in \cdot\,\Big)\Big]$$

in $\mathbf{M}_0(\mathbf{R}^d)$.

Ruin probability, or stationary workload

Consider the one-dimensional case $p = d = 1$. Let $\alpha > 1$.

For and $c > 0$ and $u > 0$ denote

$$\psi(u) = P(\sup_n (X_1 + \ldots + X_n - cn) > u) .$$

Under certain additonal technical assumptions,

$$\lim_{u\to\infty} \frac{\psi(u)}{uP(|Z| > u)} =$$

$$E\Big[w\Big(\sup_{j\in\mathbf{Z}} \sum_{k=-\infty}^{j} A_{k,k} \Big)^{\alpha} + (1-w)\Big(\sup_{j\in\mathbf{Z}} \sum_{k=-\infty}^{j} -A_{k,k} \Big)^{\alpha} \Big] \frac{1}{c(\alpha-1)} .$$