Clustering of rare events

Arijit Chakrabarty

Joint work with Gennady Samorodnitsky

30 January 2024

The broad question

Question. If and when a system deviates from its usual behaviour, how likely it is to cause a cascade of additional deviations?

The broad question

Question. If and when a system deviates from its usual behaviour, how likely it is to cause a cascade of additional deviations?

- Let $X_0, X_1, X_2, \ldots$ be a stationary process with finite mean μ.

The broad question

Question. If and when a system deviates from its usual behaviour, how likely it is to cause a cascade of additional deviations?

- Let $X_0, X_1, X_2, \dots$ be a stationary process with finite mean μ.
- For fixed $\varepsilon > 0$, given that $\frac{1}{n}\sum_{i=0}^{n-1} X_i > \mu + \varepsilon$, how likely is it that

$$\frac{1}{n}\sum_{i=j}^{j+n-1} X_i > \mu + \varepsilon$$

for $j = 1, 2, 3, \dots$?

An example

- Consider an insurance company which earns a premium of Y_i and settles claims worth Z_i in the i-th year.

An example

- Consider an insurance company which earns a premium of Y_i and settles claims worth Z_i in the i-th year.
- If $\mathrm{E}(Y_i) > \mathrm{E}(Z_i)$, then under some assumptions,

$$\frac{1}{n}\sum_{i=1}^{n}(Y_i - Z_i) \approx \mathrm{E}(Y_1 - Z_1) > 0\,,$$

with high probability, for large n. That is, the company makes profit in the long run.

An example

- Consider an insurance company which earns a premium of Y_i and settles claims worth Z_i in the i-th year.
- If $\mathrm{E}(Y_i) > \mathrm{E}(Z_i)$, then under some assumptions,

$$\frac{1}{n}\sum_{i=1}^{n}(Y_i - Z_i) \approx \mathrm{E}(Y_1 - Z_1) > 0\,,$$

  with high probability, for large n. That is, the company makes profit in the long run.
- In case it does happen that

$$\sum_{i=1}^{n}(Y_i - Z_i) \le 0\,,$$

  the company would be interested in knowing the conditional probabilities of the following events for $j \ge 1$:

$$\sum_{i=j}^{j+n-1}(Y_i - Z_i) \le 0\,.$$

Large deviations

Theorem (Cramér (1938))

Let $X_1, X_2, \ldots$ be i.i.d. random variables with

$$\Lambda(t) = \log \mathrm{E}\left(e^{tX_1}\right) < \infty\, , t \in \mathbb{R}\, .$$

Then, for $x > \mathrm{E}(X_1)$ with $P(X_1 > x) > 0$,

$$\lim_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq x\right) = -\Lambda^*(x)\, ,$$

where

$$\Lambda^*(x) = \sup_{t\in\mathbb{R}} \left(tx - \Lambda(t)\right)\, .$$

Strong large deviations

Theorem (Bahadur and Ranga Rao (1960))

Let $X_1, X_2, \ldots$ be i.i.d. from a non-lattice distribution with

$$\Lambda(t) = \log \mathrm{E}\left(e^{tX_1}\right) < \infty\,, t \in \mathbb{R}\,.$$

Then, for $x_0 > \mathrm{E}(X_1)$ with $P(X_1 > x_0) > 0$, there exists $a > 0$ such that $\Lambda'(a) = x_0$.

Strong large deviations

Theorem (Bahadur and Ranga Rao (1960))

Let $X_1, X_2, \ldots$ be i.i.d. from a non-lattice distribution with

$$\Lambda(t) = \log \mathrm{E}\left(e^{tX_1}\right) < \infty\,, t \in \mathbb{R}\,.$$

Then, for $x_0 > \mathrm{E}(X_1)$ with $P(X_1 > x_0) > 0$, there exists $a > 0$ such that $\Lambda'(a) = x_0$. Furthermore,

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i > x_0\right) \sim \frac{1}{a\sqrt{2\pi\Lambda''(a)}} n^{-1/2} e^{-n\Lambda^*(x_0)}\,,$$

that is,

$$\lim_{n\to\infty} n^{1/2} e^{n\Lambda^*(x_0)} P\left(\frac{1}{n}\sum_{i=1}^{n} X_i > x_0\right) = \frac{1}{a\sqrt{2\pi\Lambda''(a)}}\,.$$

Non-identically distributed random variables

1. Results of Gärtner (1977) and Ellis (1984) yield large deviations of sums of independent possibly non-identically distributed random variables on the logarithmic scale.

Non-identically distributed random variables

1. Results of Gärtner (1977) and Ellis (1984) yield large deviations of sums of independent possibly non-identically distributed random variables on the logarithmic scale.
2. Chaganty and Sethuraman (1993) studied strong large deviations of sums of independent possibly non-identically distributed random variables.

Non-identically distributed random variables

1. Results of Gärtner (1977) and Ellis (1984) yield large deviations of sums of independent possibly non-identically distributed random variables on the logarithmic scale.
2. Chaganty and Sethuraman (1993) studied strong large deviations of sums of independent possibly non-identically distributed random variables.
3. Little is known about the asymptotic conditional distribution given that a large deviation event has occurred.

Folklore

If Z_n is the sum of several independent random variables,

Folklore

If Z_n is the sum of several independent random variables,

- the event $[Z_n - \mathrm{E}(Z_n) > x\mathrm{Var}(Z_n)]$ is a "large deviation" for fixed $x > 0$,

Folklore

If Z_n is the sum of several independent random variables,

- the event $[Z_n - \mathrm{E}(Z_n) > x\mathrm{Var}(Z_n)]$ is a "large deviation" for fixed $x > 0$,
- while $[Z_n - \mathrm{E}(Z_n) > \eta_n]$ is a "moderate deviation" event if

$$\sqrt{\mathrm{Var}(Z_n)} \ll \eta_n \ll \mathrm{Var}(Z_n)\,,\, n \to \infty\,.$$

Folklore

If Z_n is the sum of several independent random variables,

- the event $[Z_n - \mathrm{E}(Z_n) > x\mathrm{Var}(Z_n)]$ is a "large deviation" for fixed $x > 0$,
- while $[Z_n - \mathrm{E}(Z_n) > \eta_n]$ is a "moderate deviation" event if

$$\sqrt{\mathrm{Var}(Z_n)} \ll \eta_n \ll \mathrm{Var}(Z_n)\,, n \to \infty\,.$$

Important distinction: Large deviation probabilities depend on the distribution while moderate deviation probabilities depend only on the variance.

Clustering of rare events

- Let $(X_n : n \geq 0)$ be a stationary ergodic process with

$$\mathrm{E}\left(e^{tX_0}\right) < \infty\,, t \in \mathbb{R}\,.$$

Clustering of rare events

- Let $(X_n : n \geq 0)$ be a stationary ergodic process with
$$\mathrm{E}\left(e^{tX_0}\right) < \infty \, , t \in \mathbb{R} \, .$$
- Denote $\mu = \mathrm{E}(X_0)$, and
$$E_{j,\varepsilon}(n) = \left[\frac{1}{n} \sum_{i=j}^{j+n-1} X_i > \mu + \varepsilon \right] \, ,$$
for $\varepsilon > 0$, $j \geq 0$ and $n \geq 1$.

Clustering of rare events

- Let $(X_n : n \geq 0)$ be a stationary ergodic process with

$$\mathrm{E}\left(e^{tX_0}\right) < \infty\,, t \in \mathbb{R}\,.$$

- Denote $\mu = \mathrm{E}(X_0)$, and

$$E_{j,\varepsilon}(n) = \left[\frac{1}{n}\sum_{i=j}^{j+n-1} X_i > \mu + \varepsilon\right]\,,$$

  for $\varepsilon > 0$, $j \geq 0$ and $n \geq 1$. That is, $E_{j,\varepsilon}(n)$ is the event that the mean of the sample $X_j, \ldots, X_{j+n-1}$ deviates from the population mean by at least ε.

The question

- For a fixed $\varepsilon > 0$ and large n, if $E_{0,\varepsilon}(n)$ occurs, how many of the subsequent $E_{j,\varepsilon}(n)$'s are made likely by it?

The question

- For a fixed $\varepsilon > 0$ and large n, if $E_{0,\varepsilon}(n)$ occurs, how many of the subsequent $E_{j,\varepsilon}(n)$'s are made likely by it?
- Regardless of whether $E_{0,\varepsilon}(n)$ occurs or not, ergodicity implies if $P(E_{0,\varepsilon}(n)) > 0$, then

$$\sum_{j=1}^{\infty} 1_{E_{j,\varepsilon}(n)} = \infty \text{ a.s.}$$

The question

- For a fixed $\varepsilon > 0$ and large n, if $E_{0,\varepsilon}(n)$ occurs, how many of the subsequent $E_{j,\varepsilon}(n)$'s are made likely by it?
- Regardless of whether $E_{0,\varepsilon}(n)$ occurs or not, ergodicity implies if $P(E_{0,\varepsilon}(n)) > 0$, then

$$\sum_{j=1}^{\infty} 1_{E_{j,\varepsilon}(n)} = \infty \text{ a.s.}$$

Occurrences of the "nearby" $E_{j,\varepsilon}(n)$'s are to be considered, therefore.

The question

- For a fixed $\varepsilon > 0$ and large n, if $E_{0,\varepsilon}(n)$ occurs, how many of the subsequent $E_{j,\varepsilon}(n)$'s are made likely by it?
- Regardless of whether $E_{0,\varepsilon}(n)$ occurs or not, ergodicity implies if $P(E_{0,\varepsilon}(n)) > 0$, then

$$\sum_{j=1}^{\infty} 1_{E_{j,\varepsilon}(n)} = \infty \text{ a.s.}$$

  Occurrences of the "nearby" $E_{j,\varepsilon}(n)$'s are to be considered, therefore.
- The answer depends on the "memory" of $(X_n : n \geq 0)$.

Memory of a Gaussian process

A stationary Gaussian process $(X_n : n \in \mathbb{Z})$ has

- "short memory" if

$$\sum_{n=1}^{\infty} |\mathrm{Cov}(X_0, X_n)| < \infty \,,$$

  and

$$\sum_{n=-\infty}^{\infty} \mathrm{Cov}(X_0, X_n) \neq 0 \,,$$

Memory of a Gaussian process

A stationary Gaussian process $(X_n : n \in \mathbb{Z})$ has

- "short memory" if

$$\sum_{n=1}^{\infty} |\mathrm{Cov}(X_0, X_n)| < \infty\,,$$

  and

$$\sum_{n=-\infty}^{\infty} \mathrm{Cov}(X_0, X_n) \neq 0\,,$$

- and "long memory" if

$$\mathrm{Cov}(X_0, X_n) \sim cn^{-\alpha}\,, n \to \infty\,,$$

  for some $\alpha \in (0, 1)$ and $0 < c < \infty$.

Moving average process

- Let $(Z_n : n \in \mathbb{Z})$ be a collection of i.i.d. zero mean random variables with finite exponential moments.

Moving average process

- Let $(Z_n : n \in \mathbb{Z})$ be a collection of i.i.d. zero mean random variables with finite exponential moments.
- Let $a_0, a_1, \ldots \in \mathbb{R}$ be such that

$$\sum_{j=0}^{\infty} a_j^2 < \infty\,.$$

Moving average process

- Let $(Z_n : n \in \mathbb{Z})$ be a collection of i.i.d. zero mean random variables with finite exponential moments.
- Let $a_0, a_1, \ldots \in \mathbb{R}$ be such that

$$\sum_{j=0}^{\infty} a_j^2 < \infty\,.$$

- Set

$$X_n = \mu + \sum_{j=0}^{\infty} a_j Z_{n-j}\,, n \geq 0\,.$$

Moving average process

- Let $(Z_n : n \in \mathbb{Z})$ be a collection of i.i.d. zero mean random variables with finite exponential moments.
- Let $a_0, a_1, \ldots \in \mathbb{R}$ be such that

$$\sum_{j=0}^{\infty} a_j^2 < \infty \,.$$

- Set

$$X_n = \mu + \sum_{j=0}^{\infty} a_j Z_{n-j} \,, n \geq 0 \,.$$

- Then, $(X_n : n \geq 0)$ is a moving average (M.A.) process with inputs (Z_n) and coefficients (a_j). In particular, it is ergodic and the marginal has mean μ.

Moving average process (contd.)

- If

$$\sum_{j=0}^{\infty} |a_j| < \infty \,,$$

and

$$\sum_{j=0}^{\infty} a_j \neq 0 \,,$$

then (X_n) is a “short memory” process.

Moving average process (contd.)

- If
$$\sum_{j=0}^{\infty} |a_j| < \infty \, ,$$
and
$$\sum_{j=0}^{\infty} a_j \neq 0 \, ,$$
then (X_n) is a "short memory" process.
- On the contrary, if
$$a_j \sim j^{-\alpha} \, , j \to \infty \, ,$$
for some $\frac{1}{2} < \alpha < 1$, then (X_n) has a "long memory".

The short memory regime

In this regime, the clustering is studied in 3 steps.

The short memory regime

In this regime, the clustering is studied in 3 steps.

1. Given $E_{0,\varepsilon}(n)$, for a fixed $\varepsilon > 0$,

$$\left(1_{E_{j,\varepsilon}(n)} : j = 1, 2, 3, \ldots\right)$$

   is shown to have an asymptotic non-degenerate weak limit as $n \to \infty$.

The short memory regime

In this regime, the clustering is studied in 3 steps.

1. Given $E_{0,\varepsilon}(n)$, for a fixed $\varepsilon > 0$,

$$\left(1_{E_{j,\varepsilon}(n)} : j = 1, 2, 3, \ldots\right)$$

is shown to have an asymptotic non-degenerate weak limit as $n \to \infty$.

Consequently, for fixed $K \in \mathbb{N}$,

$$P\left(\sum_{j=1}^{K} 1_{E_{j,\varepsilon}(n)} \in \cdot \middle| E_{0,\varepsilon}(n)\right) \Rightarrow \nu_{K,\varepsilon}(\cdot)\,,$$

for some probability measure $\nu_{K,\varepsilon}$ on $\mathbb{R}$.

The short memory regime (contd.)

2. For fixed ε, the "total cluster size" is finite, that is,

The short memory regime (contd.)

2. For fixed ε, the "total cluster size" is finite, that is,

$$\nu_{K,\varepsilon} \Rightarrow \nu_\varepsilon \, , K \to \infty \, .$$

The short memory regime (contd.)

2. For fixed ε, the "total cluster size" is finite, that is,

$$\nu_{K,\varepsilon} \Rightarrow \nu_\varepsilon \,, K \to \infty \,.$$

Letting $n \to \infty$ and $K \to \infty$ in this order makes precise that ν_ε is the law of the total number of events whose occurrence has been caused by that of $E_{0,\varepsilon}(n)$, for large n.

The short memory regime (contd.)

2. For fixed ε, the "total cluster size" is finite, that is,

$$\nu_{K,\varepsilon} \Rightarrow \nu_\varepsilon \, , K \to \infty \, .$$

   Letting $n \to \infty$ and $K \to \infty$ in this order makes precise that ν_ε is the law of the total number of events whose occurrence has been caused by that of $E_{0,\varepsilon}(n)$, for large n.

3. The behaviour of the total cluster size ν_ε is studied, after appropriate scaling, as $\varepsilon \downarrow 0$.

The short memory setup

- Let $(X_n : n \geq 0)$ be an M.A. process with i.i.d. zero mean inputs $(Z_n : n \in \mathbb{Z})$ having all exponential moments finite, and coefficients $(a_j : j \geq 0)$ satisfying

$$\sum_{j=0}^{\infty} |a_j| < \infty \,, \text{ and } A := \sum_{j=0}^{\infty} a_j \neq 0 \,.$$

The short memory setup

- Let $(X_n : n \geq 0)$ be an M.A. process with i.i.d. zero mean inputs $(Z_n : n \in \mathbb{Z})$ having all exponential moments finite, and coefficients $(a_j : j \geq 0)$ satisfying

$$\sum_{j=0}^{\infty} |a_j| < \infty\,, \text{ and } A := \sum_{j=0}^{\infty} a_j \neq 0\,.$$

- Assume $\mu = 0$ without loss of generality.

The short memory setup

- Let $(X_n : n \geq 0)$ be an M.A. process with i.i.d. zero mean inputs $(Z_n : n \in \mathbb{Z})$ having all exponential moments finite, and coefficients $(a_j : j \geq 0)$ satisfying

$$\sum_{j=0}^{\infty} |a_j| < \infty \,, \text{ and } A := \sum_{j=0}^{\infty} a_j \neq 0 \,.$$

- Assume $\mu = 0$ without loss of generality.
- Assume Z_0 is not supported on a lattice, that is,

$$\left| \mathrm{E}\left(e^{\iota t Z_0} \right) \right| < 1 \,, t \in \mathbb{R} \setminus \{0\} \,, \iota = \sqrt{-1} \,.$$

The short memory setup

- Let $(X_n : n \geq 0)$ be an M.A. process with i.i.d. zero mean inputs $(Z_n : n \in \mathbb{Z})$ having all exponential moments finite, and coefficients $(a_j : j \geq 0)$ satisfying

$$\sum_{j=0}^{\infty} |a_j| < \infty \,, \text{ and } A := \sum_{j=0}^{\infty} a_j \neq 0 \,.$$

- Assume $\mu = 0$ without loss of generality.
- Assume Z_0 is not supported on a lattice, that is,

$$\left| \mathrm{E}\left(e^{\iota t Z_0} \right) \right| < 1 \,, t \in \mathbb{R} \setminus \{0\} \,, \iota = \sqrt{-1} \,.$$

- Without loss of generality, assume $A > 0$.

- For all $\theta \in \mathbb{R}$, let G_θ be the probability measure on $\mathbb{R}$ obtained by "exponentially tilting" the distribution of Z_0 by θ, that is,

$$G_\theta(dx) = \left[\mathrm{E}\left(e^{\theta Z_0}\right)\right]^{-1} e^{\theta x} P(Z_0 \in dx)\,.$$

- For all $\theta \in \mathbb{R}$, let G_θ be the probability measure on $\mathbb{R}$ obtained by "exponentially tilting" the distribution of Z_0 by θ, that is,

$$G_\theta(dx) = \left[\mathrm{E}\left(e^{\theta Z_0}\right)\right]^{-1} e^{\theta x} P(Z_0 \in dx)\,.$$

- Denote

$$s_0 = \sup\{z \in \mathbb{R} : P(Z_0 \le z) < 1\} \in (0, \infty]\,.$$

- For all $\theta \in \mathbb{R}$, let G_θ be the probability measure on $\mathbb{R}$ obtained by "exponentially tilting" the distribution of Z_0 by θ, that is,

$$G_\theta(dx) = \left[\mathrm{E}\left(e^{\theta Z_0}\right)\right]^{-1} e^{\theta x} P(Z_0 \in dx)\,.$$

- Denote

$$s_0 = \sup\left\{z \in \mathbb{R} : P(Z_0 \le z) < 1\right\} \in (0, \infty]\,.$$

- Fix ε such that

$$0 < \frac{\varepsilon}{A} < s_0\,.$$

- There exists unique $\tau(\varepsilon) > 0$ such that

$$\int_{-\infty}^{\infty} G_{\tau(\varepsilon)}(dx)\, x = \frac{\varepsilon}{A}\,.$$

- There exists unique $\tau(\varepsilon) > 0$ such that

$$\int_{-\infty}^{\infty} G_{\tau(\varepsilon)}(dx)\, x = \frac{\varepsilon}{A}\,.$$

- Let $\{Z_j^u : j \in \mathbb{Z}, u = + \text{ or } -\}$ be a collection of independent random variables with distributions given as follows:

$$\begin{aligned}
Z_{-j}^{-} &\sim G_{(1-A^{-1}A_{j-1})\tau(\varepsilon)}\,, && j \geq 1\,,\\
Z_{j}^{-} &\sim G_{\tau(\varepsilon)}\,, && j \geq 0\,,\\
Z_{-j}^{+} &\sim G_{A^{-1}A_{j-1}\tau(\varepsilon)}\,, && j \geq 1\,,\\
Z_{j}^{+} &\sim G_{0}\,, && j \geq 0\,,
\end{aligned}$$

where

$$A_j = \sum_{i=0}^{j} a_i\,,\; j \geq 0\,.$$

- Let T^* follow exponential with parameter $\tau(\varepsilon)/A$ independently of the above family.

- Let T^* follow exponential with parameter $\tau(\varepsilon)/A$ independently of the above family.
- Define

$$U_n^- = \sum_{i=0}^{\infty} a_i Z_{n-i}^-\,, n \geq 0\,,$$

$$U_n^+ = \sum_{i=0}^{\infty} a_i Z_{n-i}^+\,, n \geq 0\,.$$

Theorem (C. and Samorodnitsky (2022+))

For fixed ε *such that* $0 < A^{-1}\varepsilon < s_0$*, as* $n \to \infty$*,*

$$P\left((1(E_{j,\varepsilon}(n)) : j \in \mathbb{N}) \in \cdot \middle| E_{0,\varepsilon}(n)\right)$$

$$\Rightarrow P\big((V_1(\varepsilon), V_2(\varepsilon), \ldots) \in \cdot\big)\,,$$

where

$$V_j(\varepsilon) = 1\left(T^* > \sum_{i=0}^{j-1}(U_i^- - U_i^+)\right)\,, j \ge 1\,.$$

The special case of i.i.d.

When $X_0, X_1, X_2, \ldots$ are i.i.d., $V_j(\varepsilon)$ has the following simple form:

$$V_j(\varepsilon) = 1\left(T^* > \sum_{i=0}^{j-1}(Y_i^- - Y_i^+)\right), j \geq 1\,.$$

The special case of i.i.d.

When $X_0, X_1, X_2, \ldots$ are i.i.d., $V_j(\varepsilon)$ has the following simple form:

$$V_j(\varepsilon) = 1\left(T^* > \sum_{i=0}^{j-1}(Y_i^- - Y_i^+)\right), j \geq 1 .$$

Here $Y_0^-, Y_1^-, \ldots$ are i.i.d. copies of X_0, and $Y_0^+, Y_1^+, \ldots$ are i.i.d. from the tilted distribution of X_0 so that the mean is ε, and the 2 families are independent of each other, and of T^*.

Intuition

- For i.i.d. zero mean $X_1, X_2, \ldots$, conditionally on the event $[X_1 + \ldots + X_n > n\varepsilon]$,

$$(X_1, \ldots, X_k) \Rightarrow (X_1^*, \ldots, X_k^*), n \to \infty, \tag{1}$$

Intuition

- For i.i.d. zero mean $X_1, X_2, \ldots$, conditionally on the event $[X_1 + \ldots + X_n > n\varepsilon]$,

$$(X_1, \ldots, X_k) \Rightarrow (X_1^*, \ldots, X_k^*)\,, n \to \infty\,, \tag{1}$$

where $X_1^*, \ldots, X_k^*$ are i.i.d. from a distribution obtained by exponentially tilting that of X_1 by an amount such that the mean is ε.

Intuition

- For i.i.d. zero mean $X_1, X_2, \ldots$, conditionally on the event $[X_1 + \ldots + X_n > n\varepsilon]$,

$$(X_1, \ldots, X_k) \Rightarrow (X_1^*, \ldots, X_k^*)\,, n \to \infty\,, \qquad (1)$$

  where $X_1^*, \ldots, X_k^*$ are i.i.d. from a distribution obtained by exponentially tilting that of X_1 by an amount such that the mean is ε.

- Furthermore, conditionally on the above event,

$$\sum_{i=1}^{n} X_i - n\varepsilon \Rightarrow T^*\,,$$

  together with (1), where T^* follows exponential with some parameter, independently of $(X_1^*, \ldots, X_k^*)$.

The limiting cluster size

- For fixed ε, the "limiting cluster size" as $n \to \infty$ is

$$D_\varepsilon = \sum_{j=1}^{\infty} V_j(\varepsilon)\,.$$

The limiting cluster size

- For fixed ε, the "limiting cluster size" as $n \to \infty$ is

$$D_\varepsilon = \sum_{j=1}^{\infty} V_j(\varepsilon)\,.$$

- The right hand side is finite a.s.

The limiting cluster size

- For fixed ε, the "limiting cluster size" as $n \to \infty$ is

$$D_\varepsilon = \sum_{j=1}^{\infty} V_j(\varepsilon) \, .$$

- The right hand side is finite a.s.
- **Next question.** How does D_ε behave as $\varepsilon \downarrow 0$?

Theorem (C. and Samorodnitsky (2022+))

As $\varepsilon \to 0$,

$$\varepsilon^2 D_\varepsilon \Rightarrow A^2 \sigma_Z^2 \int_0^\infty dt\, 1\left(T_0 \geq (\sqrt{2} B_t + t) \right) ,$$

Theorem (C. and Samorodnitsky (2022+))

As $\varepsilon \to 0$,

$$\varepsilon^2 D_\varepsilon \Rightarrow A^2 \sigma_Z^2 \int_0^\infty dt \, 1\left(T_0 \ge (\sqrt{2} B_t + t) \right) ,$$

where

$$\sigma_Z^2 = \mathrm{Var}(Z_0) \, ,$$

Theorem (C. and Samorodnitsky (2022+))

As $\varepsilon \to 0$,

$$\varepsilon^2 D_\varepsilon \Rightarrow A^2 \sigma_Z^2 \int_0^\infty dt\, 1\left(T_0 \geq (\sqrt{2} B_t + t)\right) ,$$

where

$$\sigma_Z^2 = \mathrm{Var}(Z_0) ,$$

T_0 *follows standard exponential*

Theorem (C. and Samorodnitsky (2022+))

As $\varepsilon \to 0$,

$$\varepsilon^2 D_\varepsilon \Rightarrow A^2 \sigma_Z^2 \int_0^\infty dt \, 1\left(T_0 \geq (\sqrt{2} B_t + t)\right) ,$$

where

$$\sigma_Z^2 = \mathrm{Var}(Z_0) \, ,$$

T_0 *follows standard exponential and* $(B_t : t \geq 0)$ *is a standard Brownian motion independent of* T_0.

Understanding the invariance

Question. For ε fixed, the limiting distribution as $n \to \infty$ depend on the law of Z_0. However, the subsequent limit as $\varepsilon \to 0$ does not. Why ?

Understanding the invariance

Question. For ε fixed, the limiting distribution as $n \to \infty$ depend on the law of Z_0. However, the subsequent limit as $\varepsilon \to 0$ does not. Why ?

Answer. For fixed ε, $[X_1 + \ldots + X_n > n\varepsilon]$ is a large deviation event, and hence the limiting law as $n \to \infty$ depends on the distribution of Z_0.

Understanding the invariance

Question. For ε fixed, the limiting distribution as $n \to \infty$ depend on the law of Z_0. However, the subsequent limit as $\varepsilon \to 0$ does not. Why ?
Answer. For fixed ε, $[X_1 + \ldots + X_n > n\varepsilon]$ is a large deviation event, and hence the limiting law as $n \to \infty$ depends on the distribution of Z_0.
As $\varepsilon \to 0$, the above becomes a moderate deviation, and hence limit depends on the distribution of Z_0 only through its variance.

Gaussian short-memory processes

For a Gaussian process $(X_n : n \in \mathbb{Z})$, the stated results hold whenever

$$\sum_{n=0}^{\infty} |\mathrm{Cov}(X_0, X_n)| < \infty\,,$$

and

$$\sum_{n=-\infty}^{\infty} \mathrm{Cov}(X_0, X_n) \neq 0\,.$$

Gaussian short-memory processes

For a Gaussian process $(X_n : n \in \mathbb{Z})$, the stated results hold whenever

$$\sum_{n=0}^{\infty} |\mathrm{Cov}(X_0, X_n)| < \infty \,,$$

and

$$\sum_{n=-\infty}^{\infty} \mathrm{Cov}(X_0, X_n) \neq 0 \,.$$

The above is a weaker assumption than that in terms of coefficients of the M.A. process, in the Gaussian case.

Conclusions in the short memory regime

1. Conditionally on $E_{0,\varepsilon}(n)$, as $n \to \infty$,

$$\left(1(E_{1,\varepsilon}(n)), 1(E_{2,\varepsilon}(n)), \ldots\right) \Rightarrow \left(V_1(\varepsilon), V_2(\varepsilon), \ldots\right),$$

   where $V_1(\varepsilon), V_2(\varepsilon), \ldots$ are 0-1 valued random variables whose distribution depends on that of Z_0.

Conclusions in the short memory regime

1. Conditionally on $E_{0,\varepsilon}(n)$, as $n \to \infty$,

$$\left(1(E_{1,\varepsilon}(n)), 1(E_{2,\varepsilon}(n)), \ldots\right) \Rightarrow \left(V_1(\varepsilon), V_2(\varepsilon), \ldots\right),$$

   where $V_1(\varepsilon), V_2(\varepsilon), \ldots$ are 0-1 valued random variables whose distribution depends on that of Z_0.

2. As $\varepsilon \to 0$,

$$\varepsilon^2 \sum_{j=1}^{\infty} V_j(\varepsilon) \Rightarrow A^2 \sigma_Z^2 \int_0^{\infty} dt\, 1\left(T_0 \geq (\sqrt{2} B_t + t)\right).$$

Conclusions in the short memory regime

1. Conditionally on $E_{0,\varepsilon}(n)$, as $n \to \infty$,

$$\left(1(E_{1,\varepsilon}(n)), 1(E_{2,\varepsilon}(n)), \ldots\right) \Rightarrow \left(V_1(\varepsilon), V_2(\varepsilon), \ldots\right),$$

   where $V_1(\varepsilon), V_2(\varepsilon), \ldots$ are 0-1 valued random variables whose distribution depends on that of Z_0.

2. As $\varepsilon \to 0$,

$$\varepsilon^2 \sum_{j=1}^{\infty} V_j(\varepsilon) \Rightarrow A^2 \sigma_Z^2 \int_0^\infty dt\, 1\left(T_0 \geq (\sqrt{2} B_t + t)\right).$$

   The RHS depends on the distribution of Z_0 only through its variance.

Conclusions in the short memory regime

1. Conditionally on $E_{0,\varepsilon}(n)$, as $n \to \infty$,

$$\left(1(E_{1,\varepsilon}(n)), 1(E_{2,\varepsilon}(n)), \dots\right) \Rightarrow \left(V_1(\varepsilon), V_2(\varepsilon), \dots\right),$$

   where $V_1(\varepsilon), V_2(\varepsilon), \dots$ are 0-1 valued random variables whose distribution depends on that of Z_0.

2. As $\varepsilon \to 0$,

$$\varepsilon^2 \sum_{j=1}^{\infty} V_j(\varepsilon) \Rightarrow A^2 \sigma_Z^2 \int_0^{\infty} dt\, 1\left(T_0 \geq (\sqrt{2} B_t + t)\right).$$

   The RHS depends on the distribution of Z_0 only through its variance.

3. For Gaussian processes, the results can be stated in terms of the correlations.

Long memory

- Let $(Z_n : n \in \mathbb{Z})$ be as before, $a_0, a_1, \ldots \in \mathbb{R}$ be such that
$$a_j \sim j^{-\alpha}\,, j \to \infty\,,$$
for some $\frac{1}{2} < \alpha < 1$, and (X_n) be constructed from the above as before.

Long memory

- Let $(Z_n : n \in \mathbb{Z})$ be as before, $a_0, a_1, \ldots \in \mathbb{R}$ be such that
$$a_j \sim j^{-\alpha}\,, j \to \infty\,,$$
for some $\frac{1}{2} < \alpha < 1$, and (X_n) be constructed from the above as before.
- It turns out in this regime that for all fixed $j \in \mathbb{N}$,
$$\lim_{n\to\infty} P\left(E_{j,\varepsilon}(n) \middle| E_{0,\varepsilon}(n)\right) = 1\,.$$

Long memory

- Let $(Z_n : n \in \mathbb{Z})$ be as before, $a_0, a_1, \ldots \in \mathbb{R}$ be such that
$$a_j \sim j^{-\alpha}\,, j \to \infty\,,$$
for some $\frac{1}{2} < \alpha < 1$, and (X_n) be constructed from the above as before.
- It turns out in this regime that for all fixed $j \in \mathbb{N}$,
$$\lim_{n\to\infty} P\left(E_{j,\varepsilon}(n)\middle|E_{0,\varepsilon}(n)\right) = 1\,.$$
That is, infinitely many $E_{j,\varepsilon}(n)$'s occur due to the occurrence of $E_{0,\varepsilon}(n)$.

Long memory

- Let $(Z_n : n \in \mathbb{Z})$ be as before, $a_0, a_1, \ldots \in \mathbb{R}$ be such that
$$a_j \sim j^{-\alpha}, j \to \infty,$$
for some $\frac{1}{2} < \alpha < 1$, and (X_n) be constructed from the above as before.
- It turns out in this regime that for all fixed $j \in \mathbb{N}$,
$$\lim_{n\to\infty} P\left(E_{j,\varepsilon}(n) \middle| E_{0,\varepsilon}(n)\right) = 1\,.$$
That is, infinitely many $E_{j,\varepsilon}(n)$'s occur due to the occurrence of $E_{0,\varepsilon}(n)$.
- Therefore, the cluster analysis done in the short memory regime does not make sense any more.

Why the difference?

- In the long memory regime,

$$\left[\frac{1}{n}\sum_{i=1}^{n} X_i > \mu + \varepsilon\right]$$

is actually a moderate deviation event.

Why the difference?

- In the long memory regime,

$$\left[\frac{1}{n}\sum_{i=1}^{n} X_i > \mu + \varepsilon\right]$$

is actually a moderate deviation event.

- That is,

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) \gg n\,, n \to \infty\,.$$

Persistence time

- Define

$$I_n(\varepsilon) = \inf \{ j \geq 1 : E_{j,\varepsilon}(n) \text{ does not occur } \} \, .$$

Persistence time

- Define

$$I_n(\varepsilon) = \inf\left\{j \geq 1 : E_{j,\varepsilon}(n) \text{ does not occur }\right\} .$$

- For ε fixed, conditional on $E_{0,\varepsilon}(n)$,

$$I_n(\varepsilon) \to \infty \text{ a.s.}$$

Persistence time

- Define
$$I_n(\varepsilon) = \inf\{j \geq 1 : E_{j,\varepsilon}(n) \text{ does not occur }\} .$$
- For ε fixed, conditional on $E_{0,\varepsilon}(n)$,
$$I_n(\varepsilon) \to \infty \text{ a.s.}$$
- **Question.** At what rate ?

Assumptions

1. The sequence (a_n) is eventually non-increasing.

Assumptions

1. The sequence (a_n) is eventually non-increasing.
2. The family $(Z_n : n \in \mathbb{Z})$ is i.i.d. from a distribution with all exponential moments finite

Assumptions

1. The sequence (a_n) is eventually non-increasing.
2. The family $(Z_n : n \in \mathbb{Z})$ is i.i.d. from a distribution with all exponential moments finite satisfying

$$\sup_{|\theta|\le\theta_0} \int_{-\infty}^{\infty} dt\, t^2 \left| \int_{-\infty}^{\infty} P(Z_0 \in dz)\, e^{(\iota t+\theta)z} \right| < \infty\,,$$

for some $\theta_0 > 0$, where $\iota = \sqrt{-1}$.

Assumptions

1. The sequence (a_n) is eventually non-increasing.
2. The family $(Z_n : n \in \mathbb{Z})$ is i.i.d. from a distribution with all exponential moments finite satisfying

$$\sup_{|\theta| \le \theta_0} \int_{-\infty}^{\infty} dt\, t^2 \left| \int_{-\infty}^{\infty} P(Z_0 \in dz)\, e^{(\iota t + \theta) z} \right| < \infty \,,$$

for some $\theta_0 > 0$, where $\iota = \sqrt{-1}$.
3. The first κ moments of Z_0 match with those of $N(0, \sigma_Z^2)$, where

$$\kappa = \left[\frac{1 + 2\alpha}{2 - 2\alpha} \right] \,,$$

and

$$\sigma_Z^2 = \mathrm{Var}(Z_0) \,.$$

Let

- $$\beta = \frac{4-4\alpha}{3-2\alpha}\,,$$
- $$H = \frac{3}{2} - \alpha\,,$$

Let

- $$\beta = \frac{4-4\alpha}{3-2\alpha},$$
- $$H = \frac{3}{2} - \alpha,$$
- T_0 be a standard exponential random variable,

Let

- $$\beta = \frac{4-4\alpha}{3-2\alpha}\,,$$
- $$H = \frac{3}{2} - \alpha\,,$$
- T_0 be a standard exponential random variable,
- and $(B_H(t) : t \geq 0)$ be a fractional Brownian motion, independent of T_0, with Hurst index H,

Let

- $$\beta = \frac{4-4\alpha}{3-2\alpha}\,,$$
- $$H = \frac{3}{2} - \alpha\,,$$
- T_0 be a standard exponential random variable,
- and $(B_H(t) : t \geq 0)$ be a fractional Brownian motion, independent of T_0, with Hurst index H, that is, it is a zero mean Gaussian process with continuous paths and

$$\mathrm{E}\left(B_H(s)B_H(t)\right) = \frac{1}{2}\left(s^{2H} + t^{2H} - |s-t|^{2H}\right)\,, s,t \geq 0\,.$$

Furthermore,

- set

$$C = \frac{\sigma_Z^2}{(1-\alpha)(3-2\alpha)} \mathrm{B}(1-\alpha, 2\alpha-1)\,,$$

Furthermore,

- set

$$C = \frac{\sigma_Z^2}{(1-\alpha)(3-2\alpha)} \mathrm{B}(1-\alpha, 2\alpha-1)\,,$$

- where $\mathrm{B}(\cdot,\cdot)$ is Euler's Beta-function,

Furthermore,

- set

$$C = \frac{\sigma_Z^2}{(1-\alpha)(3-2\alpha)} \mathrm{B}(1-\alpha, 2\alpha-1)\,,$$

- where $\mathrm{B}(\cdot,\cdot)$ is Euler's Beta-function,
- $\sigma_Z^2 = \mathrm{Var}(Z_0)$,

Furthermore,

- set
$$C = \frac{\sigma_Z^2}{(1-\alpha)(3-2\alpha)}\mathrm{B}(1-\alpha, 2\alpha-1)\,,$$
- where $\mathrm{B}(\cdot,\cdot)$ is Euler's Beta-function,
- $\sigma_Z^2 = \mathrm{Var}(Z_0)$,
- and
$$\tau_\varepsilon = \inf\left\{t \geq 0 : B_H(t) \leq (2C)^{-1/2}\varepsilon t^{2H} - \varepsilon^{-1}C^{1/2}2^{-1/2}T_0\right\}\,, \varepsilon > 0\,.$$

Theorem (C. and Samorodnitsky (2022+))

For $\varepsilon > 0$ fixed,

$$P\left(n^{-\beta} I_n(\varepsilon) \in \cdot \middle| E_{0,\varepsilon}(n)\right) \Rightarrow P\left(\tau_\varepsilon \in \cdot\right) ,$$

as $n \to \infty$.

Unsurprisingly, the limiting law depends on Z_0 only through its variance.

- Let

$$S_n(j) = \sum_{i=j}^{j+n-1} X_i \, , n \geq 1 \, , j \geq 0 \, .$$

- Let

$$S_n(j) = \sum_{i=j}^{j+n-1} X_i \,, n \geq 1 \,, j \geq 0 \,.$$

- Conditionally on $E_{0,\varepsilon}(n)$, as $n \to \infty$,

$$\left(n^{-(2-2\alpha)} \left(S_n([n^\beta t]) - S_n(0) \right) : t \geq 0 \right)$$

$$\Rightarrow \left((2C)^{1/2} B_H(t) - \varepsilon t^{3-2\alpha} : t \geq 0 \right) \,.$$

- Let

$$S_n(j) = \sum_{i=j}^{j+n-1} X_i \,, n \geq 1 \,, j \geq 0 \,.$$

- Conditionally on $E_{0,\varepsilon}(n)$, as $n \to \infty$,

$$\left(n^{-(2-2\alpha)}\left(S_n([n^\beta t]) - S_n(0)\right) : t \geq 0\right)$$

$$\Rightarrow \left((2C)^{1/2} B_H(t) - \varepsilon t^{3-2\alpha} : t \geq 0\right) \,.$$

- Conditional on the same event,

$$n^{-(2-2\alpha)} \left(S_n(0) - n\varepsilon\right) \Rightarrow \frac{C}{\varepsilon} T_0 \,,$$

jointly with the above.

- Combining the two convergences,

$$\left(n^{-(2-2\alpha)}\left(S_n([n^\beta t])-n\varepsilon\right):t\geq 0\right)$$

$$\Rightarrow\left((2C)^{1/2}B_H(t)-\varepsilon t^{3-2\alpha}+\frac{C}{\varepsilon}T_0:t\geq 0\right),$$

conditionally on $E_{0,\varepsilon}(n)$, as $n\to\infty$.

- Combining the two convergences,

$$\left(n^{-(2-2\alpha)}\left(S_n([n^\beta t]) - n\varepsilon\right) : t \geq 0\right)$$

$$\Rightarrow \left((2C)^{1/2} B_H(t) - \varepsilon t^{3-2\alpha} + \frac{C}{\varepsilon} T_0 : t \geq 0\right),$$

conditionally on $E_{0,\varepsilon}(n)$, as $n \to \infty$.

- Continuous mapping theorem:

$$n^{-\beta} I_n(\varepsilon) \Rightarrow \tau_\varepsilon .$$

Remarks

- Self-similarity of fractional Brownian motion implies

$$\tau_\varepsilon \stackrel{d}{=} \varepsilon^{-1/H}\tau_1 \, .$$

Remarks

- Self-similarity of fractional Brownian motion implies

$$\tau_\varepsilon \stackrel{d}{=} \varepsilon^{-1/H}\tau_1 \,.$$

  Recall that in the short memory regime, the cluster size $\approx \varepsilon^{-2}$ for small ε.

Remarks

- Self-similarity of fractional Brownian motion implies

$$\tau_\varepsilon \stackrel{d}{=} \varepsilon^{-1/H}\tau_1 \,.$$

Recall that in the short memory regime, the cluster size $\approx \varepsilon^{-2}$ for small ε.

- If $\alpha \uparrow 1$, then

$$\beta \to 0 \text{ and } H \to \frac{1}{2}\,.$$

Remarks

- Self-similarity of fractional Brownian motion implies

$$\tau_\varepsilon \stackrel{d}{=} \varepsilon^{-1/H} \tau_1 \,.$$

Recall that in the short memory regime, the cluster size $\approx \varepsilon^{-2}$ for small ε.

- If $\alpha \uparrow 1$, then

$$\beta \to 0 \text{ and } H \to \frac{1}{2} \,.$$

Growth rate of $I_n(\varepsilon)$ becomes slower and the fBM approaches a B.M.

Why additional assumptions?

- The behaviour in the moderate deviations regime is in line with the central limit theorem, that is, like a normal distribution.

Why additional assumptions?

- The behaviour in the moderate deviations regime is in line with the central limit theorem, that is, like a normal distribution.
- For technical reasons, the approximation by normal was needed to be stronger than that provided by the standard Berry-Eseen theorem.

Why additional assumptions?

- The behaviour in the moderate deviations regime is in line with the central limit theorem, that is, like a normal distribution.
- For technical reasons, the approximation by normal was needed to be stronger than that provided by the standard Berry-Eseen theorem.
- Matching of the first few moments with those of normal enables an argument as in the Edgeworth expansions.

Conclusions in the long memory regime

- The deviations are moderate, and hence the behaviour depends only on the variance.

Conclusions in the long memory regime

- The deviations are moderate, and hence the behaviour depends only on the variance.
- The total cluster size is infinite.

Conclusions in the long memory regime

- The deviations are moderate, and hence the behaviour depends only on the variance.
- The total cluster size is infinite.
- Fluctuations of the persistence time are governed by a fBM.

Conclusions in the long memory regime

- The deviations are moderate, and hence the behaviour depends only on the variance.
- The total cluster size is infinite.
- Fluctuations of the persistence time are governed by a fBM.
- Additional assumptions are required for technical reasons.

A main ingredient of all the proofs

- Consider the simplest case: $X_1, X_2, \ldots$ are i.i.d. zero mean with all exponential moments finite.

A main ingredient of all the proofs

- Consider the simplest case: $X_1, X_2, \ldots$ are i.i.d. zero mean with all exponential moments finite.
- Fix $a_n \gg \sqrt{n}$. We want to study

$$P\left(S_n > a_n\right) ,$$

where $S_n = X_1 + \ldots + X_n$.

A main ingredient of all the proofs

- Consider the simplest case: $X_1, X_2, \ldots$ are i.i.d. zero mean with all exponential moments finite.
- Fix $a_n \gg \sqrt{n}$. We want to study

$$P\left(S_n > a_n\right),$$

where $S_n = X_1 + \ldots + X_n$.

- Let θ_n be such that

$$\left[\mathrm{E}\left(e^{\theta_n X_1}\right)\right]^{-1}\mathrm{E}\left(X_1 e^{\theta_n X_1}\right) = \frac{a_n}{n}.$$

A main ingredient of all the proofs

- Consider the simplest case: $X_1, X_2, \ldots$ are i.i.d. zero mean with all exponential moments finite.
- Fix $a_n \gg \sqrt{n}$. We want to study

$$P\left(S_n > a_n\right) \,,$$

where $S_n = X_1 + \ldots + X_n$.

- Let θ_n be such that

$$\left[\mathrm{E}\left(e^{\theta_n X_1}\right)\right]^{-1} \mathrm{E}\left(X_1 e^{\theta_n X_1}\right) = \frac{a_n}{n} \,.$$

Such a θ_n exists if $P(S_n > a_n) > 0$.

- Let F_n be the measure on $\mathbb{R}$ defined by

$$F_n(dx) = \left[\mathrm{E}\left(e^{\theta_n X_1}\right)\right]^{-1} e^{\theta_n x} P(X_1 \in dx)\,.$$

- Let F_n be the measure on $\mathbb{R}$ defined by

$$F_n(dx) = \left[\mathrm{E}\left(e^{\theta_n X_1}\right)\right]^{-1} e^{\theta_n x} P(X_1 \in dx) .$$

- For $n \geq 1$, let $X_{n1}^*, X_{n2}^*, \ldots$ be i.i.d. from F_n.

- Let F_n be the measure on $\mathbb{R}$ defined by

$$F_n(dx) = \left[\mathrm{E}\left(e^{\theta_n X_1}\right)\right]^{-1} e^{\theta_n x} P(X_1 \in dx)\,.$$

- For $n \geq 1$, let $X_{n1}^*, X_{n2}^*, \ldots$ be i.i.d. from F_n.
- Set

$$S_n^* = \sum_{i=1}^{n} X_{ni}^* - a_n\,.$$

Write

$$
\begin{aligned}
& P(S_n > a_n) \\
= \quad & \int_{-\infty}^{\infty} P(X_1 \in dx_1) \dots \int_{-\infty}^{\infty} P(X_n \in dx_n) 1 \left(\sum_{i=1}^{n} x_i > a_n \right)
\end{aligned}
$$

Write

$$
\begin{aligned}
& P(S_n > a_n) \\
= \quad & \int_{-\infty}^{\infty} P(X_1 \in dx_1) \dots \int_{-\infty}^{\infty} P(X_n \in dx_n) 1 \left(\sum_{i=1}^{n} x_i > a_n \right) \\
= \quad & \int_{-\infty}^{\infty} P(X_1 \in dx_1) e^{\theta_n x_1} \dots \int_{-\infty}^{\infty} P(X_n \in dx_n) e^{\theta_n x_n} 1 \left(\sum_{i=1}^{n} x_i > a_n \right) \\
& \exp\left(-\theta_n \sum_{i=1}^{n} x_i \right)
\end{aligned}
$$

Write

$$
\begin{aligned}
& P(S_n > a_n) \\
= \ & \int_{-\infty}^{\infty} P(X_1 \in dx_1) \dots \int_{-\infty}^{\infty} P(X_n \in dx_n) 1\left(\sum_{i=1}^{n} x_i > a_n\right) \\
= \ & \int_{-\infty}^{\infty} P(X_1 \in dx_1) e^{\theta_n x_1} \dots \int_{-\infty}^{\infty} P(X_n \in dx_n) e^{\theta_n x_n} 1\left(\sum_{i=1}^{n} x_i > a_n\right) \\
& \exp\left(-\theta_n \sum_{i=1}^{n} x_i\right) \\
= \ & \mathrm{E}\left(e^{\theta_n S_n}\right) \int_{-\infty}^{\infty} F_n(dx_1) \dots \int_{-\infty}^{\infty} F_n(dx_n) 1\left(\sum_{i=1}^{n} x_i > a_n\right) \\
& \exp\left(-\theta_n \sum_{i=1}^{n} x_i\right)
\end{aligned}
$$

$$= \mathrm{E}\left(e^{\theta_n S_n}\right)\mathrm{E}\left[\exp\left(-\theta_n \sum_{i=1}^{n} X_{ni}^*\right) 1\left(\sum_{i=1}^{n} X_{ni}^* > a_n\right)\right]$$

$$
\begin{aligned}
&= \mathrm{E}\left(e^{\theta_n S_n}\right) \mathrm{E}\left[\exp\left(-\theta_n \sum_{i=1}^{n} X_{ni}^*\right) 1\left(\sum_{i=1}^{n} X_{ni}^* > a_n\right)\right] \\
&= e^{-\theta_n a_n} \mathrm{E}\left(e^{\theta_n S_n}\right) \mathrm{E}\left[e^{-\theta_n S_n^*} 1(S_n^* > 0)\right] .
\end{aligned}
$$

$$
\begin{aligned}
&= \mathrm{E}\left(e^{\theta_n S_n}\right) \mathrm{E}\left[\exp\left(-\theta_n \sum_{i=1}^{n} X_{ni}^*\right) 1\left(\sum_{i=1}^{n} X_{ni}^* > a_n\right)\right] \\
&= e^{-\theta_n a_n} \mathrm{E}\left(e^{\theta_n S_n}\right) \mathrm{E}\left[e^{-\theta_n S_n^*} 1(S_n^* > 0)\right] .
\end{aligned}
$$

- By definition, S_n^* is the sum of i.i.d. zero mean random variables.

$$
\begin{aligned}
&= \mathrm{E}\left(e^{\theta_n S_n}\right) \mathrm{E}\left[\exp\left(-\theta_n \sum_{i=1}^n X_{ni}^*\right) 1\left(\sum_{i=1}^n X_{ni}^* > a_n\right)\right] \\
&= e^{-\theta_n a_n} \mathrm{E}\left(e^{\theta_n S_n}\right) \mathrm{E}\left[e^{-\theta_n S_n^*} 1(S_n^* > 0)\right].
\end{aligned}
$$

- By definition, S_n^* is the sum of i.i.d. zero mean random variables.
- Use Berry-Eseen type bounds to estimate the error incurred in replacing S_n^* by a zero mean normal random variable with the same variance in

$$
\mathrm{E}\left[e^{-\theta_n S_n^*} 1(S_n^* > 0)\right].
$$

Future work

- For fixed $N \geq 1$, in the long memory regime,

$$\sum_{j=1}^{N} 1\left(E_{j,\varepsilon}(n)\right) \xrightarrow{P} N\,,$$

conditionally on $E_{0,\varepsilon}(n)$, as $n \to \infty$.

Future work

- For fixed $N \geq 1$, in the long memory regime,

$$\sum_{j=1}^{N} 1\left(E_{j,\varepsilon}(n)\right) \xrightarrow{P} N\,,$$

  conditionally on $E_{0,\varepsilon}(n)$, as $n \to \infty$.
- **Question.** Is it possible to scale

$$\sum_{j=1}^{N} 1\left(E_{j,\varepsilon}(n)\right) - N$$

  in a way such that there is a limiting distribution which is not degenerate at 0 ?

Future work (contd.)

- **Question.** What happens in the "negative memory" regime, that is, when

$$a_j \sim j^{-\alpha}\,,$$

for some $\alpha > 1$ and

$$\sum_{j=0}^{\infty} a_j = 0\,?$$

Future work (contd.)

- **Question.** What happens in the "negative memory" regime, that is, when

$$a_j \sim j^{-\alpha},$$

for some $\alpha > 1$ and

$$\sum_{j=0}^{\infty} a_j = 0\,?$$

- In this regime,

$$\lim_{n\to\infty} P\left(E_{1,\varepsilon}(n) \middle| E_{0,\varepsilon}(n)\right) = 0\,.$$

Future work (contd.)

- **Question.** What happens in the "negative memory" regime, that is, when

$$a_j \sim j^{-\alpha},$$

for some $\alpha > 1$ and

$$\sum_{j=0}^{\infty} a_j = 0\,?$$

- In this regime,

$$\lim_{n\to\infty} P\left(E_{1,\varepsilon}(n)\middle|E_{0,\varepsilon}(n)\right) = 0\,.$$

- The precise question to ask is not clear a priori.

Future work (contd.)

- In the negative memory regime, $E_{0,\varepsilon}(n)$ is a "huge deviation" event because the deviation therein is much larger than the variance.

Future work (contd.)

- In the negative memory regime, $E_{0,\varepsilon}(n)$ is a "huge deviation" event because the deviation therein is much larger than the variance.
- Huge deviations is a largely uncharted territory.

References

- Chakrabarty, A. and Samorodnitsky, G. Clustering of large deviations in moving average processes: the short memory regime. To appear in the Annals of Applied Probability.
- Chakrabarty, A. and Samorodnitsky, G. Clustering of large deviations in moving average processes: the long memory regime. Stochastic Processes and their Applications (2023) 163: 387-423.