Tail process and its role in limit theorems

Bojan Basrak, University of Zagreb

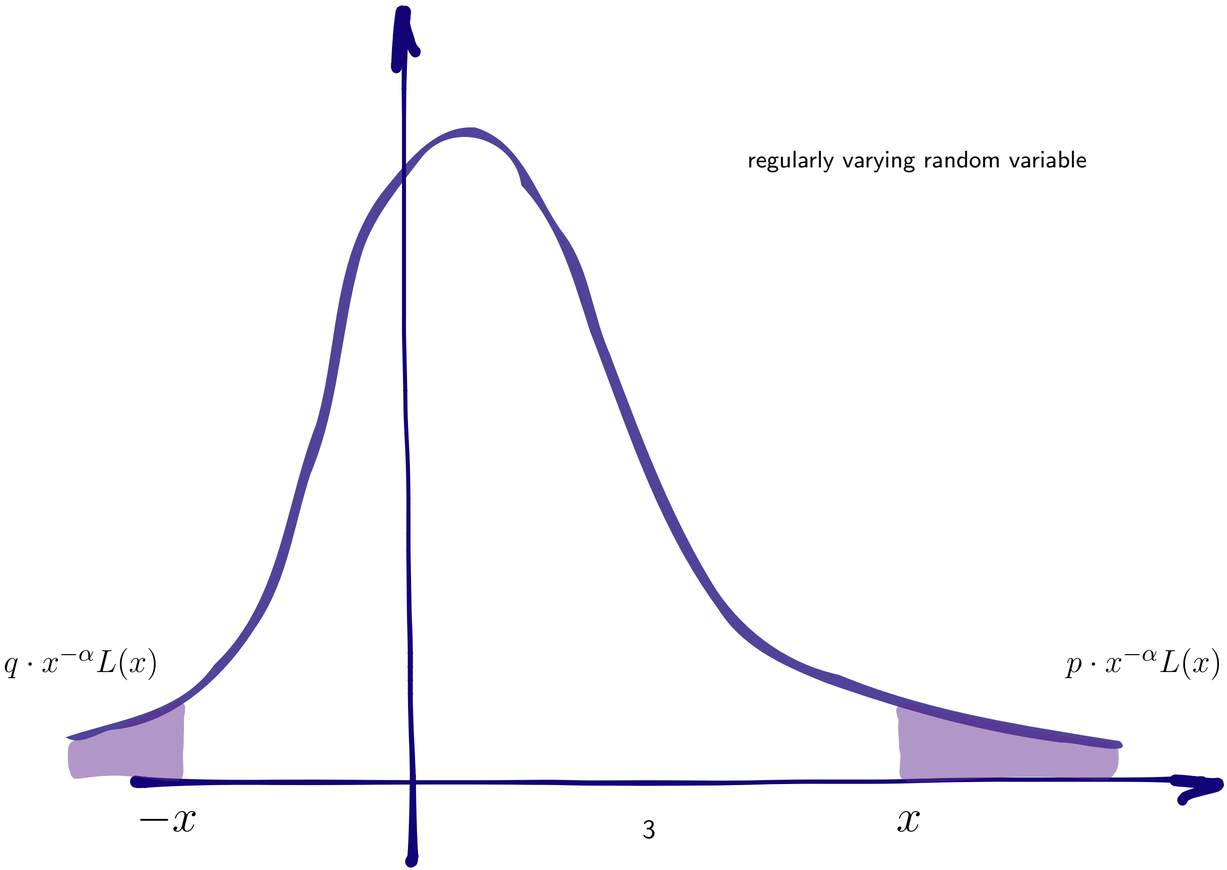
The Fields Institute
Toronto, May 2016

based on the joint work (in progress) with

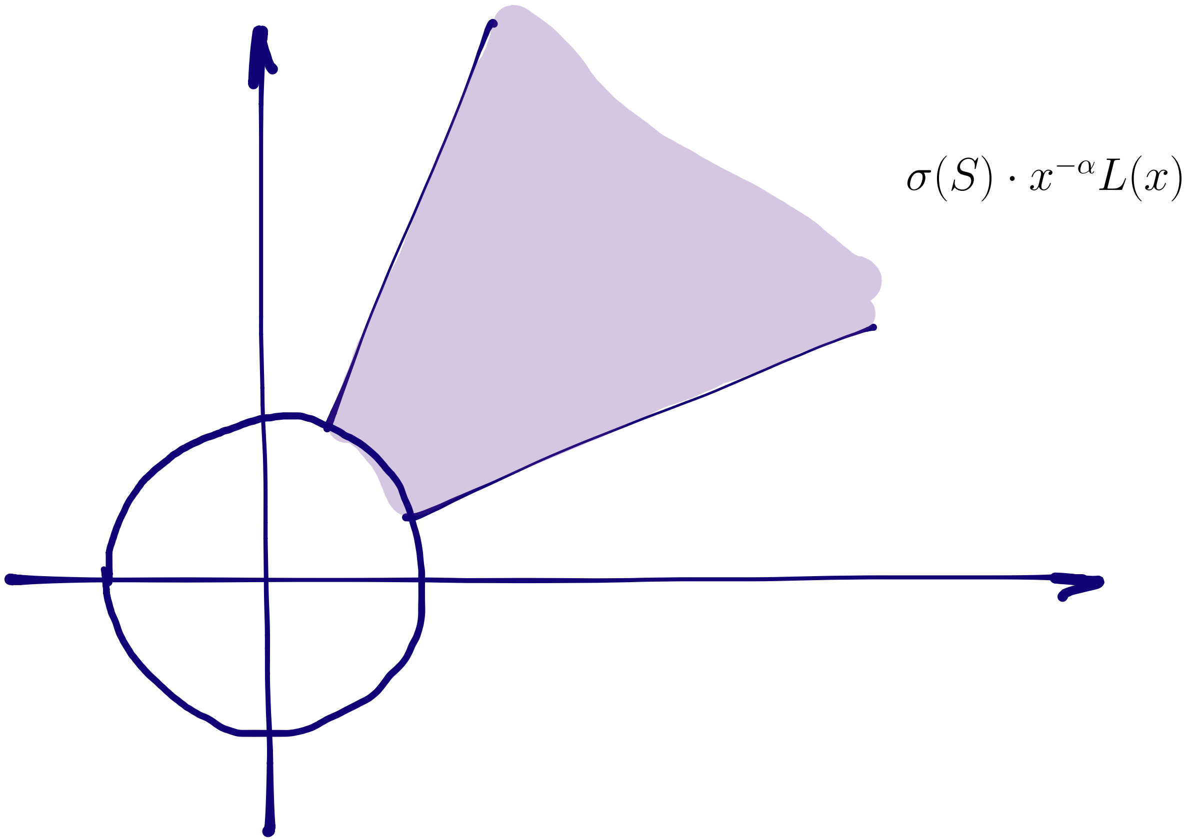
Philippe Soulier, Azra Tafro, Hrvoje Planinić

Stationary regularly varying sequences

Regular variation



Multivariate regular variation



Regularly varying process

- A stationary time series $(X_n)_n$ is said to be **regularly varying** if random vectors

$$(X_0, \ldots, X_k) \quad k \geq 0$$

are regularly varying for each k.

For a stationary regularly varying sequence there exists a **tail process**

$$(Y_t)_{t\in\mathbb{Z}}$$

such that

$$\left(\frac{X_t}{x}\right)_{t\in\mathbb{Z}}\Bigg|\,|X_0|>x \xrightarrow{d} (Y_t)_{t\in\mathbb{Z}}$$

and a **spectral tail process**

$$(\theta_t)_{t\in\mathbb{Z}}$$

independent of $|Y_0|$ such that

$$(Y_t)_t \stackrel{d}{=} |Y_0|(\theta_t)_t\,.$$

Moreover

$$|\theta_0|=1 \quad \text{and} \quad |Y_0|\sim \text{Pareto}(\alpha)\,.$$

There exists a sequence (a_n) such that

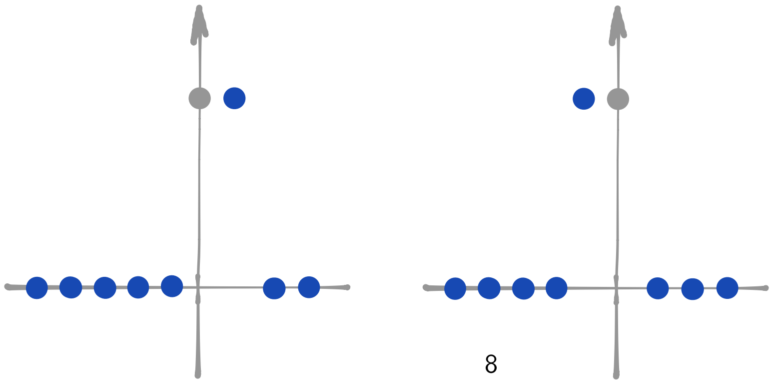
$$\left(\frac{X_t}{a_n}\right)_{t\in\mathbb{Z}} \Bigg| \, |X_0| > a_n \xrightarrow{d} (Y_t)_{t\in\mathbb{Z}}.$$

Examples (for simplicity, assume $\theta_0 = 1$)

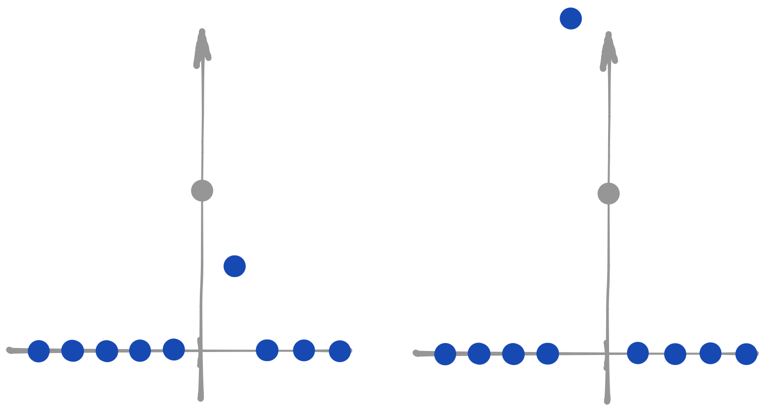
a) X_t iid RV(α), $\quad \theta_t = 0$, for $t \neq 0$.

b) $X_t = Z_t \vee Z_{t-1}$, Z_t iid RV(α),

$$\ldots, \theta_{-1}, \theta_0, \theta_1, \ldots \sim \begin{cases} \ldots, 0, 0, 1, 1, 0, \ldots & \text{w.p. } 1/2 \\ \ldots, 0, 1, 1, 0, 0, \ldots & \text{w.p. } 1/2 \end{cases}$$



c) $X_t = Z_t + \frac{1}{2} Z_{t-1}$, Z_t iid RV(α).



Independent observations

Regular variation assumption determines limiting behavior of

- point processes
- sums and random walks $S_n = X_1 + \cdots + X_n$
- maxima and other extremes $M_n = \max\{X_1, \ldots, X_n\}$
- records and record times

Complete convergence theorem 1

simple but powerful – cf. Leadbetter-Rootzén, Resnick

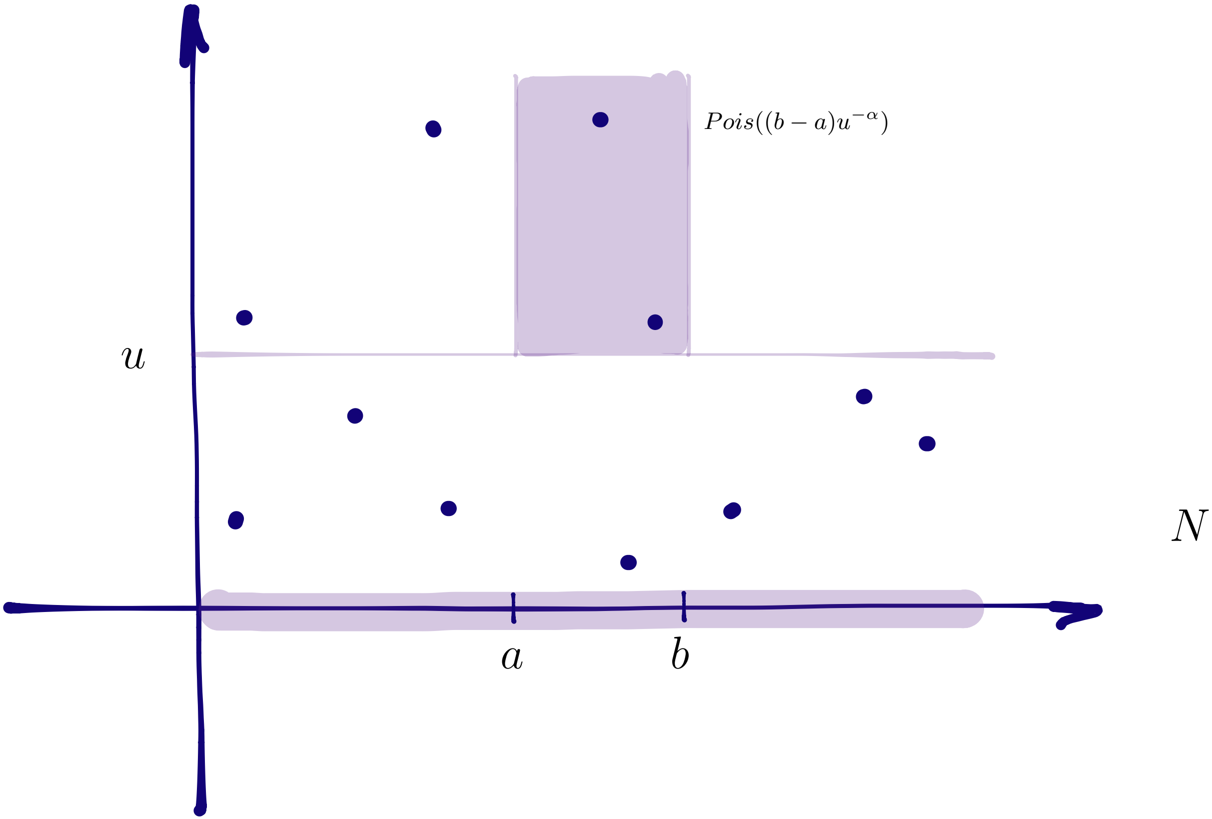
Theorem For iid $X_t \geq 0$, X_0 is reg. varying is if and only if

$$N_n = \sum_1^n \delta_{\frac{i}{n}, \frac{X_i}{a_n}} \xrightarrow{d} N = \sum_i \delta_{T_i, P_i} \, ,$$

where N is PRM(Leb$\times d(-x^{-\alpha})$).

So

$$P(M_n/a_n \leq u) = P(N_n([0,1] \times (u, \infty)) = 0) \to P(N([0,1] \times (u, \infty)) = 0) = e^{-u^{-\alpha}}$$



$Pois((b-a)u^{-\alpha})$
u
N
a
b

Extremes of dependent sequences cluster

Strongly mixing observations

Mori (1977) showed if

$$N_n \xrightarrow{d} N$$

then

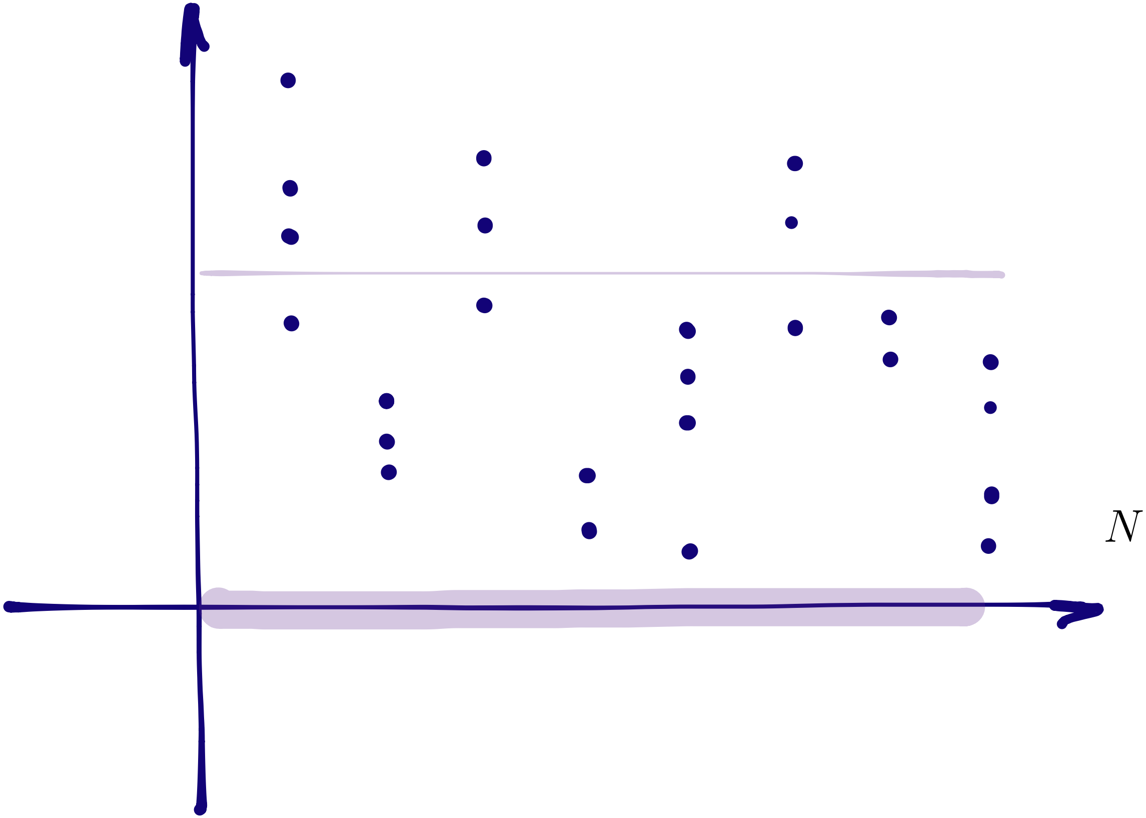
$$N = \sum_i \sum_j \delta_{T_i, P_i Q_{i,j}}$$

where

- ▷ $\sum_i \delta_{T_i,P_i}$ is a PRM($\vartheta \cdot \text{Leb} \times d(-x^{-\alpha})$) with $\vartheta \in (0,1]$
- ▷ $\sum_j \delta_{Q_{ij}}$ is an iid sequence of point processes in $[-1,1]$, independent of PRM above.

It was not really clear

- what is ϑ
- what is the distribution of $\sum_j \delta_{Q_{ij}}$
- what would be a sufficient condition for such a convergence



N

Anti-clustering condition

or finite mean cluster size condition

High level exceedances are not clustering for "too long", i.e for some $r_n \to \infty$ and $r_n/n \to 0$:

$$\lim_{m\to\infty} \limsup_{n\to\infty} P\left(\bigvee_{m \le |i| \le r_n} |X_i| > a_n u \,\middle|\, |X_0| > a_n u \right) = 0\,, \quad u > 0\,. \quad (1)$$

It implies

$$Y_m \xrightarrow{P} 0, \qquad \text{as } |m| \to \infty.$$

Main technical lemma

Denote $L_Y = \sup_{i\in\mathbb{Z}} |Y_i|$ and $M_n = \max |X_1|, \ldots, |X_n|$, then under the assumptions above

$$\left(\sum_{i=1}^{r_n} \delta_{X_i/M_{r_n}}, \frac{M_{r_n}}{a_n} \,\middle|\, M_{r_n} > a_n \right) \Rightarrow \left(\sum_i \delta_{Y_i/L_Y}, L_Y \,\middle|\, \sup_{j<0} |Y_j| \le 1 \right)$$

with two components on the right hand side being independent – B. & Tafro (2016).

Complete convergence theorem 2

building on Davis & Resnick, Davis & Hsing, Davis & Mikosch

tafro,b.(2016) /krizmanić, segers, b. (2012)

Theorem Under strong mixing and a.c., as $n \to \infty$,

$$N_n \xrightarrow{d} N = \sum_{i,j} \delta_{(T_i, P_i Q_{ij})} ,$$

where

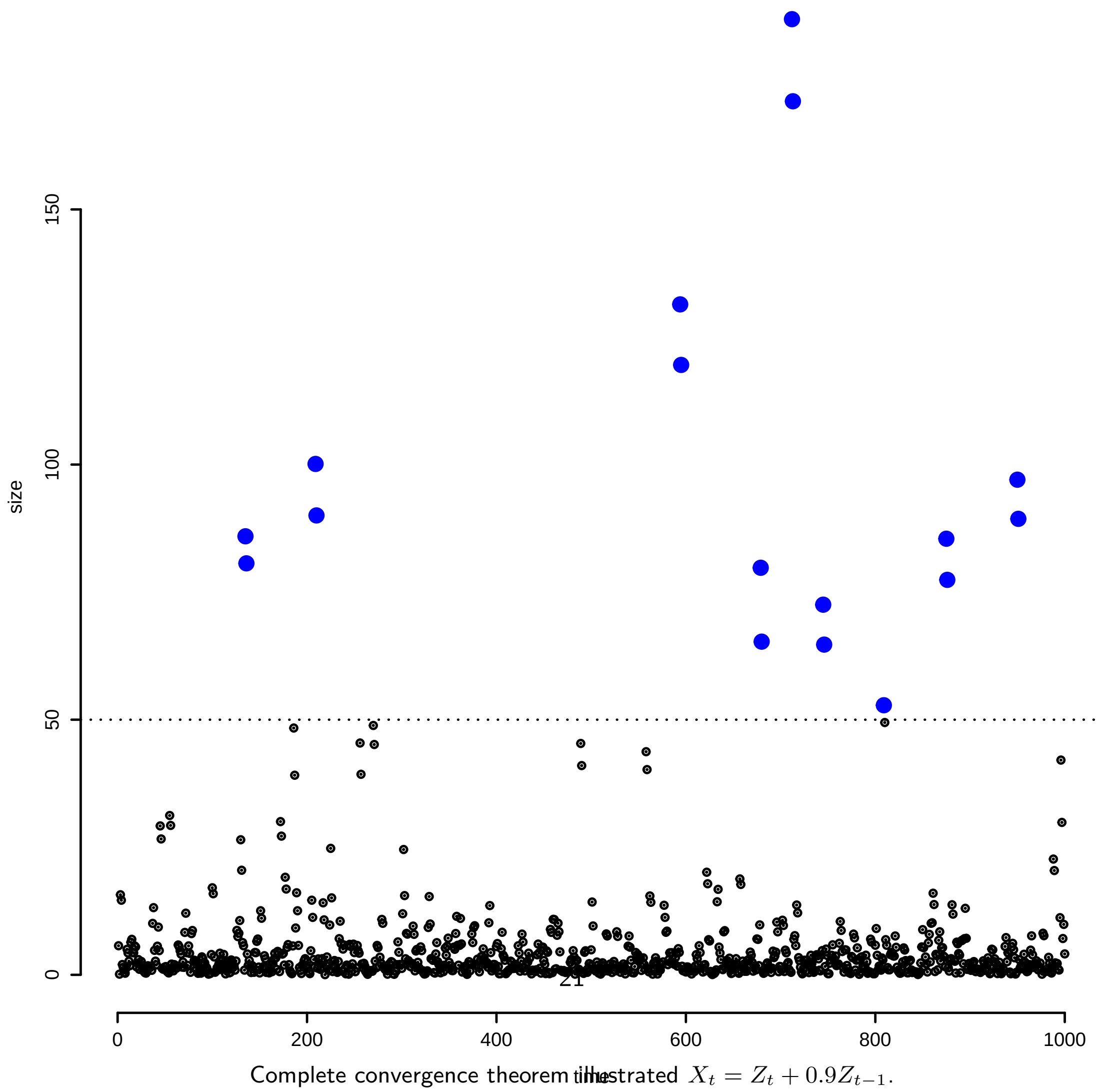
- $\sum_i \delta_{T_i, P_i}$ is a Poisson process on $[0,1] \times (0,\infty]$ with intensity ϑLeb $\times$ $d(-x^{-\alpha})$
- $(\sum_j \delta_{Q_{ij}})_i$ is an iid sequence of point processes independent of the process above

Here ϑ is **the extremal index** of the sequence $|X_t|$ with representation

$$\vartheta = P(\bigvee_{i \geq 1} |Y_i| \leq 1) = P(\bigvee_{i \leq -1} |Y_i| \leq 1) > 0.$$

While cluster shapes satisfy

$$\sum_j \delta_{Q_j} \stackrel{d}{=} \sum_i \delta_{Y_i/L_Y} \Big| \sup_{j<0} |Y_j| \leq 1$$



Complete convergence theorem illustrated $X_t = Z_t + 0.9Z_{t-1}$.

As in the iid case one can prove (functional) limit theorems for

- ▷ partial maxima $M_{\lfloor nt \rfloor}$
- ▷ partial sums $S_{\lfloor nt \rfloor}$ under additional conditions and unusual topologies (Avram & Taqqu, B. Krizmanić, Segers, or Jakubowski...)

However...

- in the limit there is a loss of information about the order
- one cannot find the limit of $S_{\lfloor nt \rfloor}$ even for some very simple models
- it is difficult to say much about records or record times

Space for ordered clusters

We introduce a new space

$$\tilde{l}_0 = l_0/\sim \qquad \text{where } l_0 = \{\boldsymbol{x} = (x_i)_{i\in\mathbb{Z}} : \lim_{|i|\to\infty} x_i = 0\}$$

and $\boldsymbol{x} \sim \boldsymbol{y}$ if $d(\boldsymbol{x}, \boldsymbol{y}) = 0$ with

$$d(\boldsymbol{x}, \boldsymbol{y}) = \inf_k \sup_j |x_j - y_{j+k}|.$$

Adapting d to $\tilde{l}_0$ produces a separable and complete metric space.

Technical lemma in $\tilde{l}_0$

large deviations result

Under a.c. assumption as $n \to \infty$

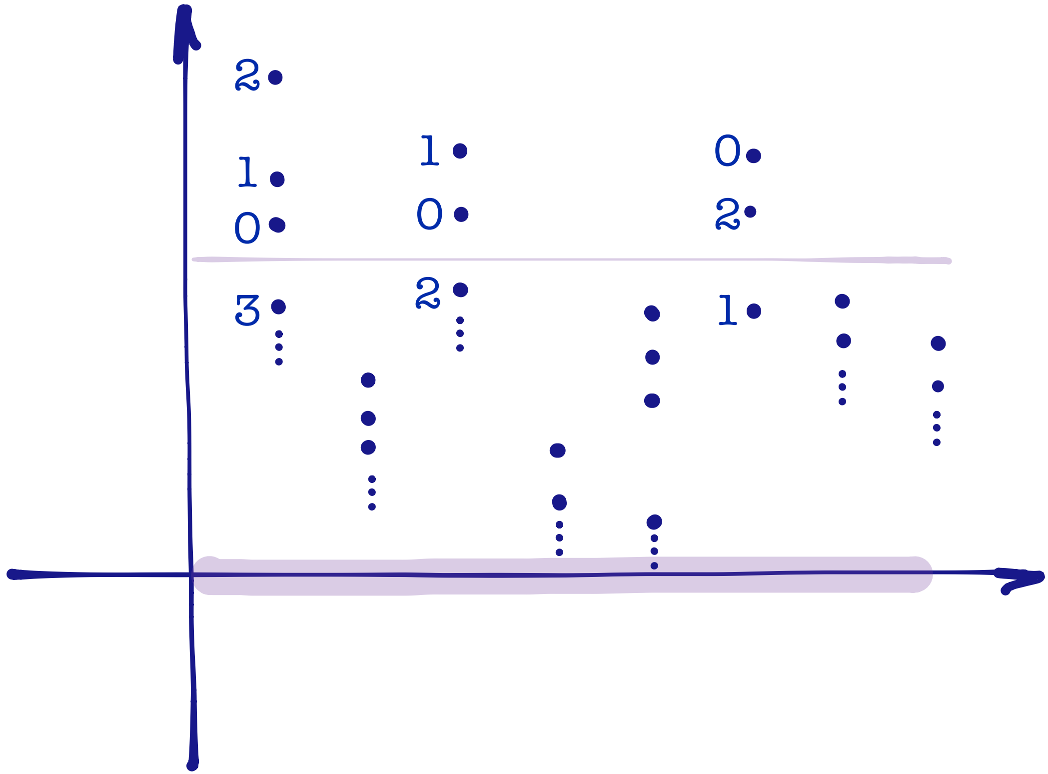
$$\left(\frac{X_1, \ldots, X_{r_n}}{a_n} \,\Big|\, M_{r_n} > a_n \right) \Rightarrow \left(Y_i,\ i \in \mathbb{Z} \,\Big|\, \sup_{j<0} |Y_j| \le 1 \right)$$

in $\tilde{l}_0 \setminus \{\mathbf{0}\}$. The space is not locally compact, so we need to use $w^{\#}$ topology (cf. Daley–Vere Jones). Related to Hult & Samorodnitsky (2010) and Mikosch & Wintenberger (2016) large deviations results.

As before, conditionally on $\sup_{j<0} |Y_j| \le 1$, random variable $L_Y = \sup_{i \in \mathbb{Z}} |Y_i|$ and random cluster $(Y_i / L_Y)_i$ are independent.

Complete convergence theorem 3

very technical but order is preserved



Partial sums converge again

Under assumptions above (for $\alpha \in (0,1)$ plus an additional one for $\alpha \in [1,2)$) we prove functional limit theorem for

$$\frac{S_{\lfloor nt \rfloor}}{a_n}, \quad t \geq 0,$$

but the limit is a "process" in a new space – $E[0,1]$, M_2. Assume for convenience that X_i's are symmetric if $\alpha \geq 1$.

Similarly one can prove the limiting theorem in a more standard space $(D[0,1], M_1)$ for

$$\sup_{s \leq t} \frac{S_{\lfloor ns \rfloor}}{a_n}, \quad t \geq 0.$$

Space $E[0, 1]$

Whitt 2002

Elements are triples

$$(x, T, \{I(t) : t \in T\}),$$

where $x \in D[0, 1]$, T is a countable subset of $[0, 1]$ with

$$Disc(x) \subseteq T,$$

and, for each $t \in T$, $I(t)$ is a closed bounded interval such that

$$x(t), x(t-) \in I(t) .$$

Moreover, for each $\varepsilon > 0$, there are at most finitely many times t with diameter of $I(t)$ greater than ε.

Functional limit theorem

in the space E

planinić, soulllier, b. (2016)

Theorem Under assumptions above for $\alpha \in (0, 2)$ as $n \to \infty$,

$$\frac{S_{\lfloor nt \rfloor}}{a_n} \xrightarrow{d} S^E \,,$$

for some $S^E = (S, T_S, \{I_S(t) : t \in T_S\})$ with S being an α–stable Lévy process.

Remark

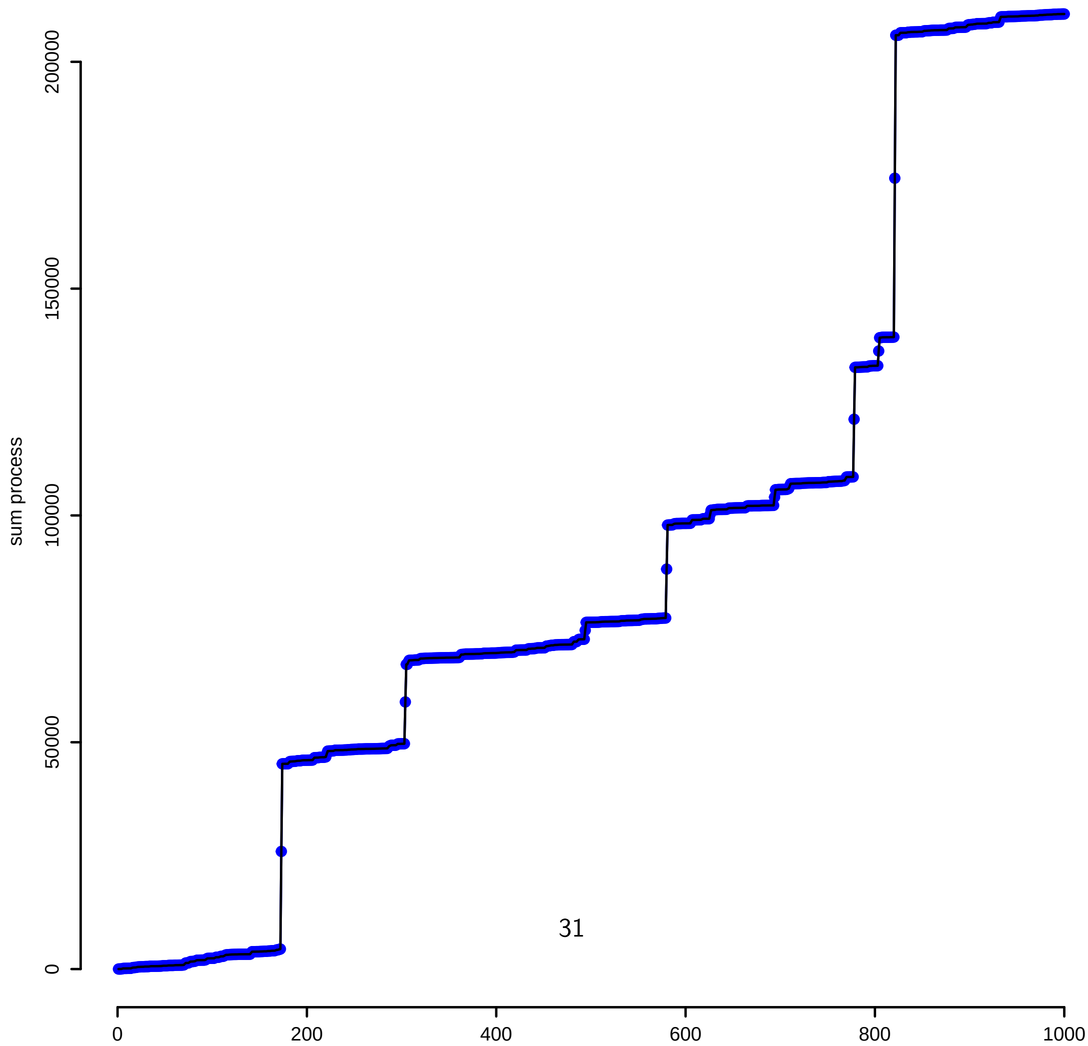
In the limit $S^E = (S, T_S, \{I_S(t) : t \in T_S\})$, countable set T_S includes all the discontinuities of the stable process S, and S can be trivial.

All three components of the process S^E can be expressed in the terms of the tail process.

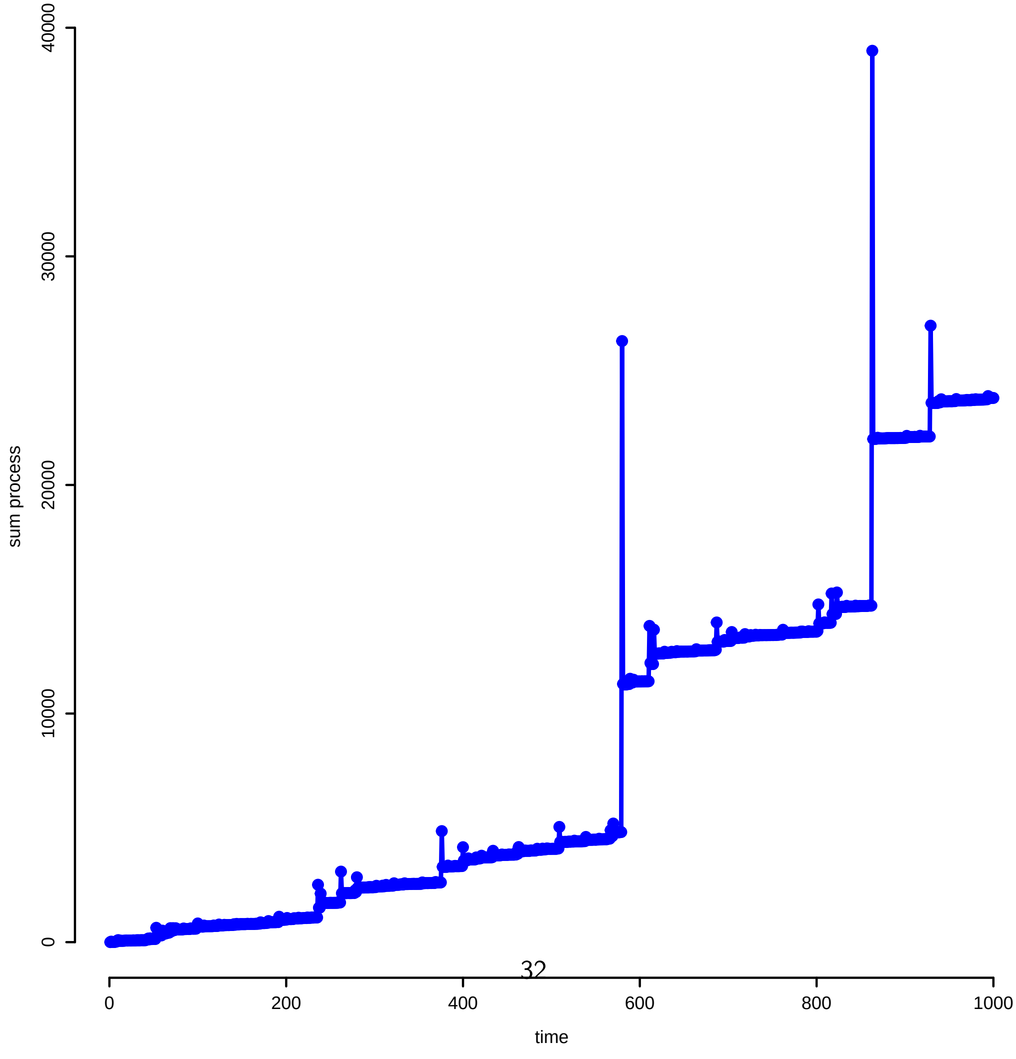
One can prove that in $D[0,1]$ with M_1 topology

$$\sup_{s \le t} \frac{S_{\lfloor ns \rfloor}}{a_n} \xrightarrow{d} \sup_{s \le t} S^E(s)\,.$$

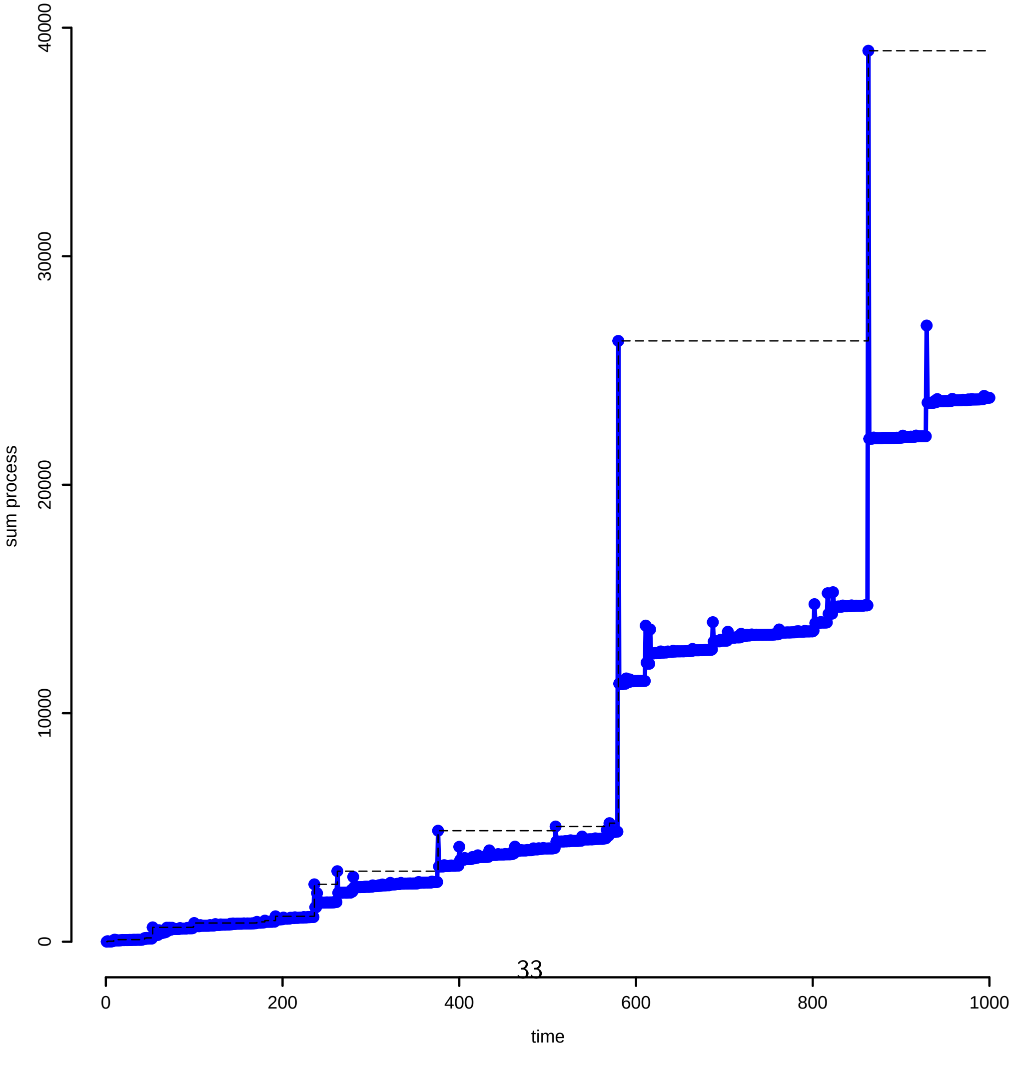
Partial sums for $X_t = Z_t + 0.9Z_{t-1}$.



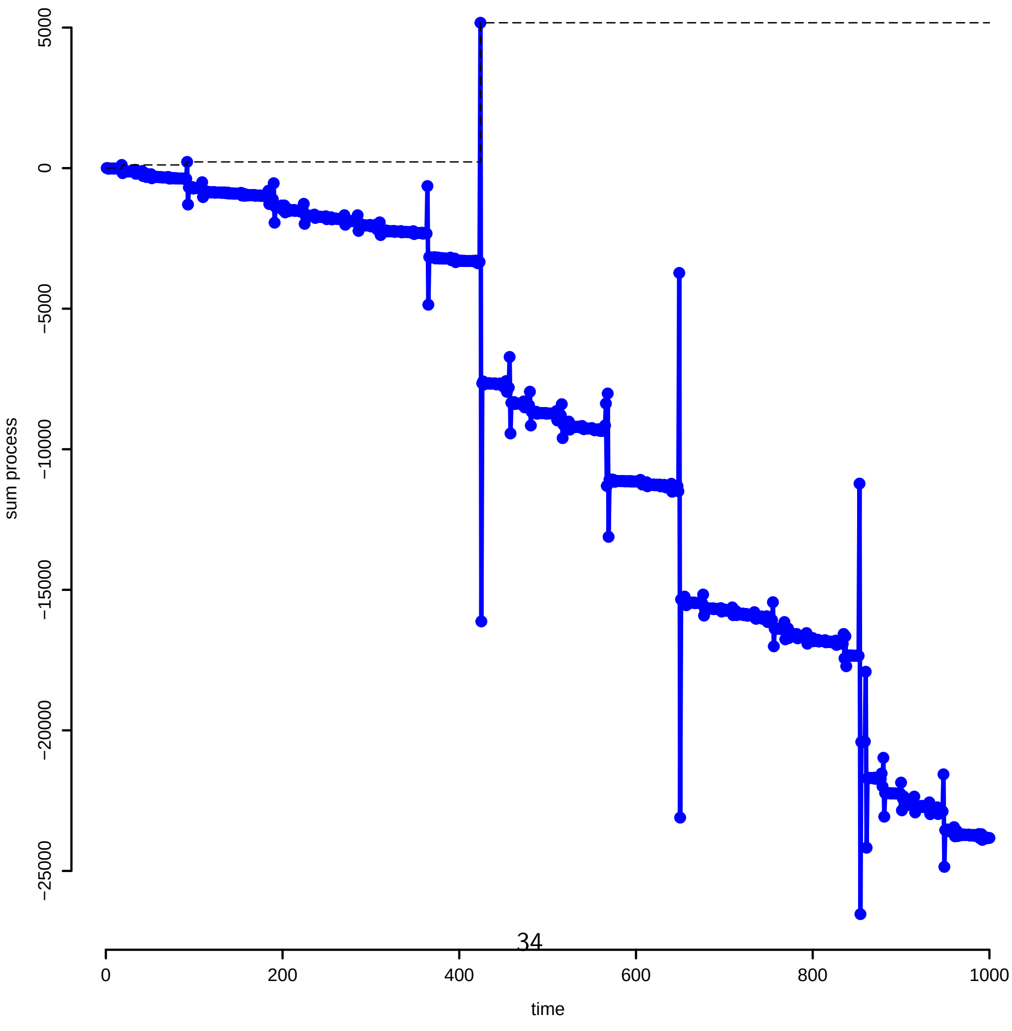
Partial sums for $X_t = Z_t - 0.7Z_{t-1}$.



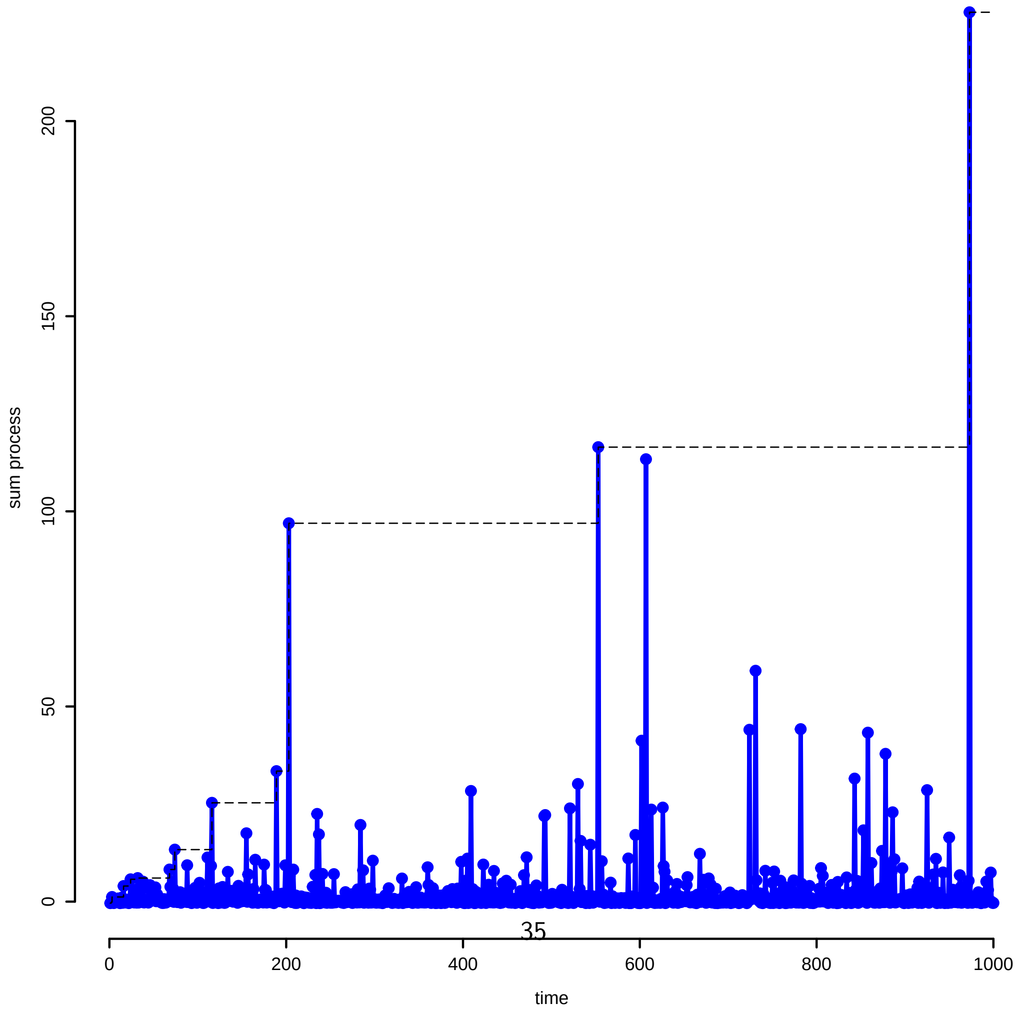
Running max of sums for $X_t = Z_t - 0.7Z_{t-1}$.



Sums of $X_t = Z_t - 2.5Z_{t-1} + Z_{t-2}$.



Sums of $X_t = Z_t - Z_{t-1}$.



Record times

It is well known (Resnick, 1987) that for any iid sequence from a continuous distribution, point process of record times converges, ie

$$R_n = \sum_{i=1}^{\infty} \delta_{i/n} \mathbb{I}_{\{X_i \text{ is a record}\}}$$

satisfies

$$R_n \xrightarrow{d} R = \sum_{i \in \mathbb{Z}} \delta_{R_i}$$

where R is a so called **scale invariant Poisson process** on $(0, \infty)$ with intensity dx/x .

Record times are much more difficult to handle when dependence is present, therefore we assume

▷ sequence (X_n), maybe after monotone transformation, is regularly varying, strong mixing and a.c. holds.

▷ with probability one all nonzero values of the tail process Y_i are mutually different .

The limit of record times

planinić, soulllier, b. (2016)

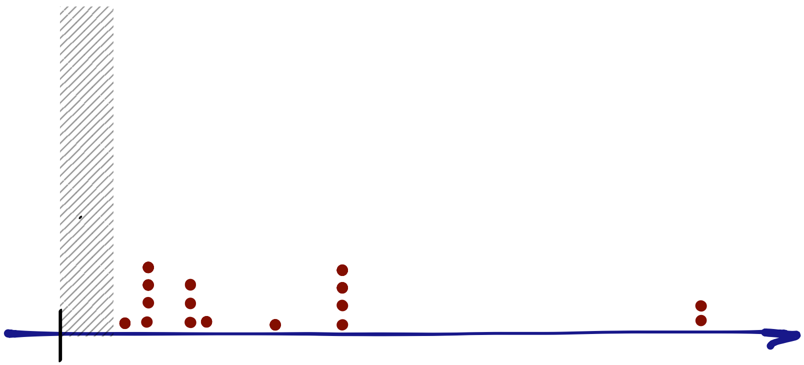
Theorem Under assumptions above as $n \to \infty$,

$$R_n \xrightarrow{d} R' = \sum_{i \in \mathbb{Z}} \delta_{R_i \kappa_i} ,$$

where

- ▷ $\sum_{i \in \mathbb{Z}} \delta_{R_i}$ is Poisson process on $(0, \infty)$ with intensity dx/x and
- ▷ $(\kappa_i)_{i \in \mathbb{Z}}$ is an iid sequence independent of it.

Records are broken in clusters.



Remark

The limit R' does not depend on ϑ directly.

Moreover, κ_i have the same distribution as

$$\sum_{j=-\infty}^{\infty} \mathbb{I}_{\{Q_j > \sup_{i<j} Q_i \vee e^{-W/\alpha}\}} \,,$$

where W is standard exponential and

$$(Q_j)_j \stackrel{d}{=} \left(\left(\frac{Y_i}{L_Y} \right)_i \,\middle|\, \sup_{j<0} |Y_j| \leq 1 \right)$$

Summary

A stationary regularly varying sequence (X_t)

- has a tail process (Y_t)
- the clusters of extremes can be described by (Y_t)
- point processes N_n have a limit characterized by (Y_t) with order preserved in the space $[0,1] \times \tilde{l}_0$.
- random walks with steps (X_t) have an "α–stable" limit for $\alpha \in (0,2)$ but in M_2 on $E[0,1]$ (càdlàg functions are ok only for some special cases).
- record times have a surprisingly simple compound Poisson structure in the limit.

Thanks