An invariance principle for sums and record times for regularly varying stationary sequences

Hrvoje Planinić University of Zagreb, Croatia
based on the joint work with Bojan Basrak, Philippe Soulier

Introduction - Functional limit theorems

It is well known that for an i.i.d. sequence $(X_i)_{i\in\mathbb{N}}$ of regularly varying random variables with index $\alpha \in (0,2)$, the partial sum process

$$S_n(t) := \frac{1}{a_n}\sum_{i=1}^{\lfloor nt\rfloor}(X_i - b_n),\ t\in[0,1]$$

converges in distribution to an α-stable Lévy process in the space D of càdlàg functions equipped with the Skorohod J_1 topology.

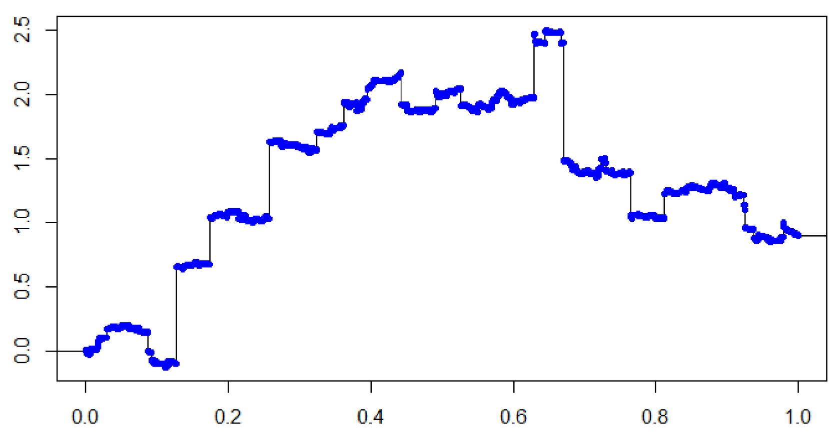

Figure 1: S_{10000} for (X_i) i.i.d. regularly varying with $\alpha = 1.1$

When there is dependence between the X_i's, even for some very simple m-dependent linear processes, the Skorohod topologies don't work.

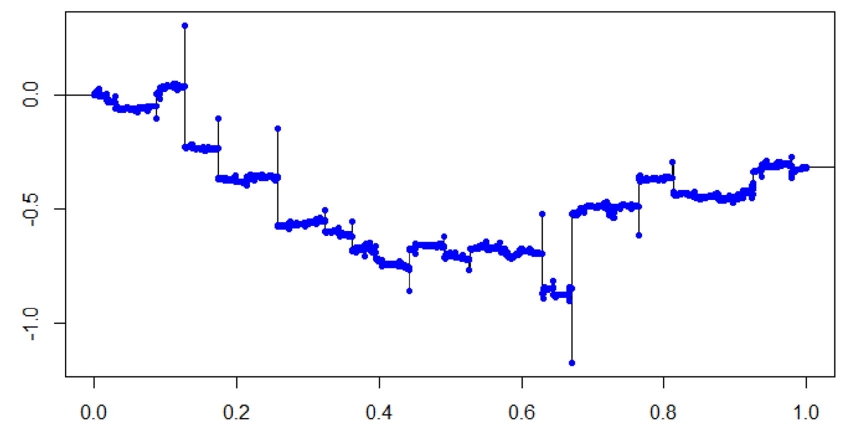

Figure 2: S_{10000} for $X_i = Z_i - 2Z_{i-1}$ for (Z_i) i.i.d. regularly varying with $\alpha = 1.1$

By going beyond the usual space D and using a new type of point process convergence theorem, we show that the limit results exist even for such processes but also for many other important time-series models for which the standard limiting theory in the space D does not apply.

Our setup

A stationary sequence $(X_i)_{i\in\mathbb{Z}}$ is said to be **regularly varying** with index $\alpha > 0$ if random vectors $(X_k,\dots,X_l)$ are regularly varying with index α for each $k,l\in\mathbb{Z}$. This implies the existence of the **tail process** $(Y_i)_{i\in\mathbb{Z}}$ such that for all $k,l\in\mathbb{Z}$, as $u\to\infty$,

$$\frac{(X_k,\dots,X_l)}{u}\ \Big|\ |X_0|>u \xrightarrow{d} (Y_k,\dots,Y_l).$$

We impose a technical condition on the length of the clusters of extremes, which is often called the anti-clustering (AC) condition. Together with an appropriate mixing condition, such as β-mixing, we will refer to them as the weak dependence conditions.

All of these assumptions are satisfied by many well known time-series models such as linear processes with heavy-tailed innovations, solutions to stochastic recurrence equation, ARCH and GARCH models...

Point processes

Take a sequence (a_n) such that $\mathbb{P}(|X_0|>a_n)\sim 1/n$, as $n\to\infty$ and consider a sequence of point processes

$$N_n = \sum_{i=1}^{n}\delta_{(\frac{i}{n},\frac{X_i}{a_n})}.$$

For (X_i) i.i.d., N_n converges in distribution to a suitable **Poisson point process** on $[0,1]\times\overline{\mathbb{R}}\setminus\{0\}$.

In general, when the sequence (X_i) is dependent, N_n converges to a **compound Poisson point process**. This reflects the fact that large values of the sequence tend to occur in clusters.

However, due to time scaling, there is a loss of information in the limit about the **order** at which large observations occured within the clusters.

Clustering of extreme observations

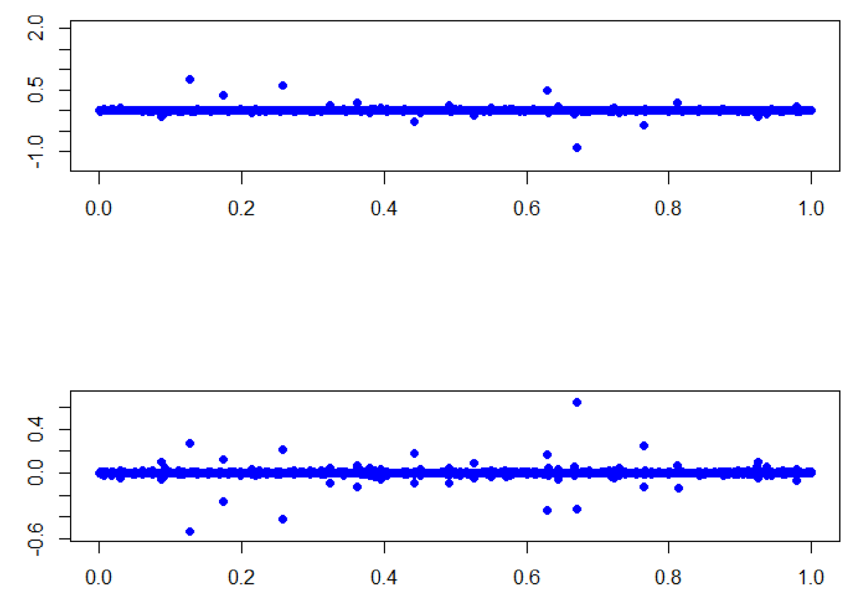

Figure 3: N_{10000} for $X_i = Z_i$ (above) and $X_i = Z_i - 2Z_{i-1}$ for the same sequence of (Z_i) i.i.d. regularly varying with index $\alpha = 1.1$. The data is the same as for the examples in the introduction.

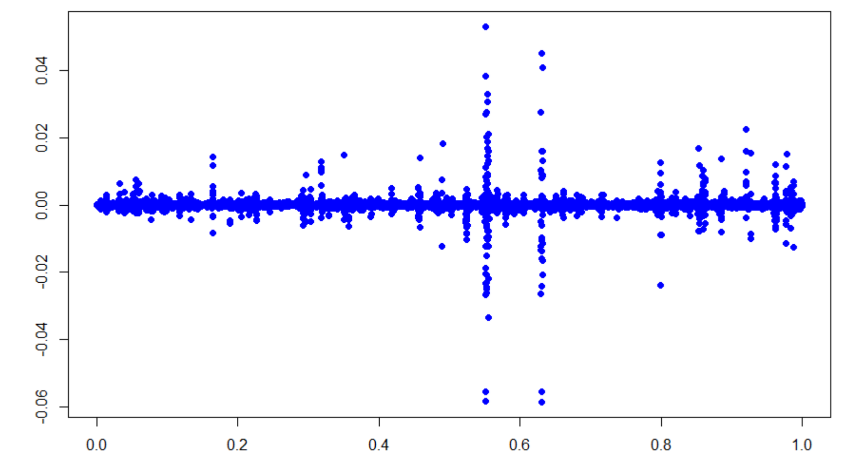

Figure 4: N_{10000} for ARCH(1) process with index of regular variation $\alpha \approx 1$

A new space for clusters - Preserving the order

Divide the sample in $k_n = \lfloor n/r_n\rfloor$ blocks of size r_n with $\lim_{n\to\infty} r_n = \lim_{n\to\infty} n/r_n = \infty$ and consider them as elements of the space

$$l_0 = \{\boldsymbol{x} = (x_j)_{j\in\mathbb{Z}} : \lim_{|j|\to\infty} x_j = 0\}.$$

Set $\boldsymbol{x}\sim\boldsymbol{y}$ if $\boldsymbol{y} = \theta^k\boldsymbol{x}$ for some $k\in\mathbb{Z}$, where θ is the shift operator and denote by $\tilde{l}_0$ the quotient space $l_0/\sim$. Let $M_n = \max_{1\le i\le n} X_i$.

Proposition 1 (Basrak, Soulier, P. (2016)) *Under regular variation and (AC) as $n\to\infty$,*

$$\frac{(X_1,\dots,X_{r_n})}{a_n}\ \Big|\ M_{r_n}>a_n \xrightarrow{d} Y^*(Q_j)_{j\in\mathbb{Z}}$$

*in the space $\tilde{l}_0$ where $(Q_j)_{j\in\mathbb{Z}}$ is characterized by the **tail process** and **independent** of the Pareto random variable Y^* with tail index α.*

Theorem 1 (B. , S. , P. (2016)) *Under regular variation and weak dependence conditions as $n\to\infty$, the **point processes of clusters***

$$N_n' = \sum_{i=1}^{k_n}\delta_{(\frac{i}{k_n},\frac{X_{(i-i)r_n+1},\dots,X_{ir_n}}{a_n})}$$

*converge in distribution to a **Poisson point process** $N' = \sum_i \delta_{(T_i,P_i(Q_{ij})_{j\in\mathbb{Z}})}$ on $[0,1]\times\tilde{l}_0\setminus\{\tilde{\boldsymbol{0}}\}$ where*

- *$\sum_i\delta_{(T_i,P_i)}$ is a Poisson point process on $[0,1]\times(0,\infty]$ with intensity measure $\mu(dx,dy) = dx\times\theta\alpha y^{-\alpha-1}dy$ with $\theta\in(0,1]$ being the extremal index of the sequence $(|X_i|)$*
- *$((Q_{ij})_{j\in\mathbb{Z}})_i$ is an i.i.d. sequence of **elements in $\tilde{l}_0$ independent** of the above point process and distributed as $(Q_j)_{j\in\mathbb{Z}}$.*

Partial sums converge - the space E (Whitt)

Elements of the space $E[0,1]$ are triples

$$x' = (x,\{t_i\},\{x(t_i-)+[m_i,M_i]\}),$$

where x is a càdlàg function and $\{t_i\}$ a countable subset of $[0,1]$ with $Disc(x)\subseteq\{t_i\}$ (see Figures 2 and 5).

Theorem 2 (B. , S. , P. (2016)) *For $\alpha\in(0,2)$, under regular variation and weak dependence conditions, and with some additional conditions if $\alpha\ge 1$, as $n\to\infty$,*

$$S_n \xrightarrow{d} S_\alpha'$$

in the space $E[0,1]$ with the M_2 topology where the limit S_α' is characterized by the limiting point process $N' = \sum_i\delta_{(T_i,P_i(Q_{ij})_{j\in\mathbb{Z}})}$ and the first coordinate being an α-stable Levy process.

Furthermore, in the familar space D with the M_1 topology

$$(\sup_{s\le t} S_n(s))_{t\in[0,1]} \xrightarrow{d} (\sup_{s\le t} S_\alpha'(s))_{t\in[0,1]}.$$

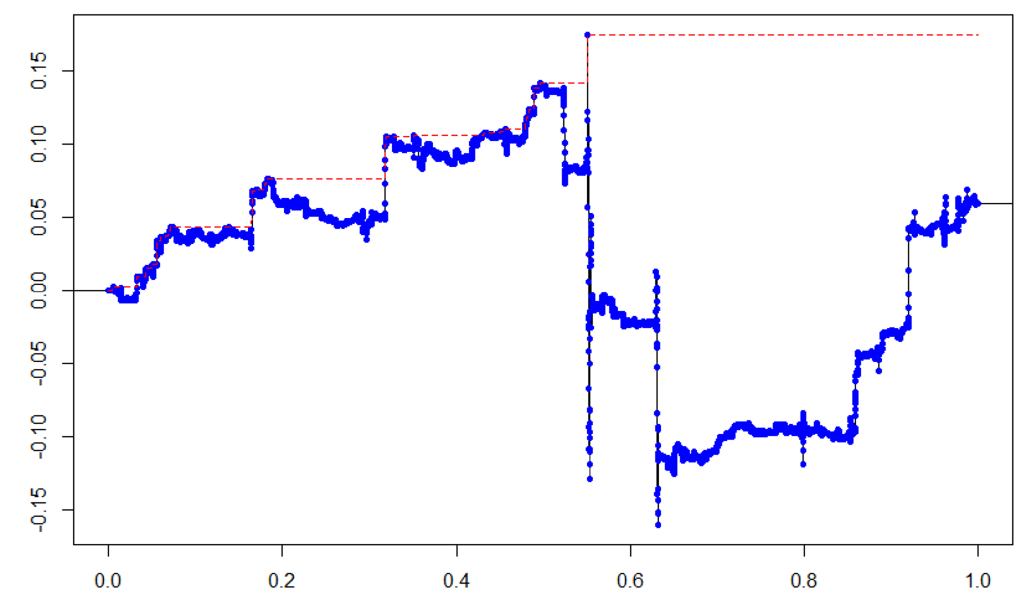

Figure 5: S_{10000} for ARCH(1) process from Figure 4.

Record times

It is well known that the record times of any i.i.d. sequence (X_i) from a continuous distribution are asymptotically Poisson in the following sense

$$R_n = \sum_{i=1}^{\infty}\delta_{\frac{i}{n}}1_{\{X_i\text{ is a record}\}} \xrightarrow{d} R = \sum_i\delta_{\tau_i},$$

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

Theorem 3 (B. , S. , P. (2016)) *Under regular variation and weak dependence conditions, and additionaly assuming that almost surely, all nonzero values of the tail process are mutually different, as $n\to\infty$,*

$$R_n \xrightarrow{d} R' = \sum_i\delta_{\tau_i}\kappa_i\ ,$$

- *$\sum_i\delta_{\tau_i}$ is the scale invariant Poisson point process on $(0,\infty)$ from the i.i.d. case.*
- *$(\kappa_i)_i$ is an **i.i.d.** sequence of random variables on $\{1,2,3,\dots\}$ **independent** of the point process above and with distribution characterized by the **tail process**.*

Summary

A stationary regularly varying sequence (X_i)

- has a tail process (Y_i)
- point processes have a limit characterized by (Y_i) with order preserved in the space $[0,1]\times\tilde{l}_0\setminus\{\tilde{\boldsymbol{0}}\}$
- partial sums converge, but in the space E with the M_2 topology
- record times have a very simple compound Poisson structure in the limit

Preprint, Acknowledgement, Contact

- Basrak, B., Planinic, H., & Soulier, P. (2016). *An invariance principle for sums and record times of regularly varying stationary sequences.* arXiv preprint.
- This work has been supported in part by Croatian Science Foundation under the project 3526.

Hrvoje Planinić
planinic@math.hr