On the extremal behavior of random variables observed at renewal times

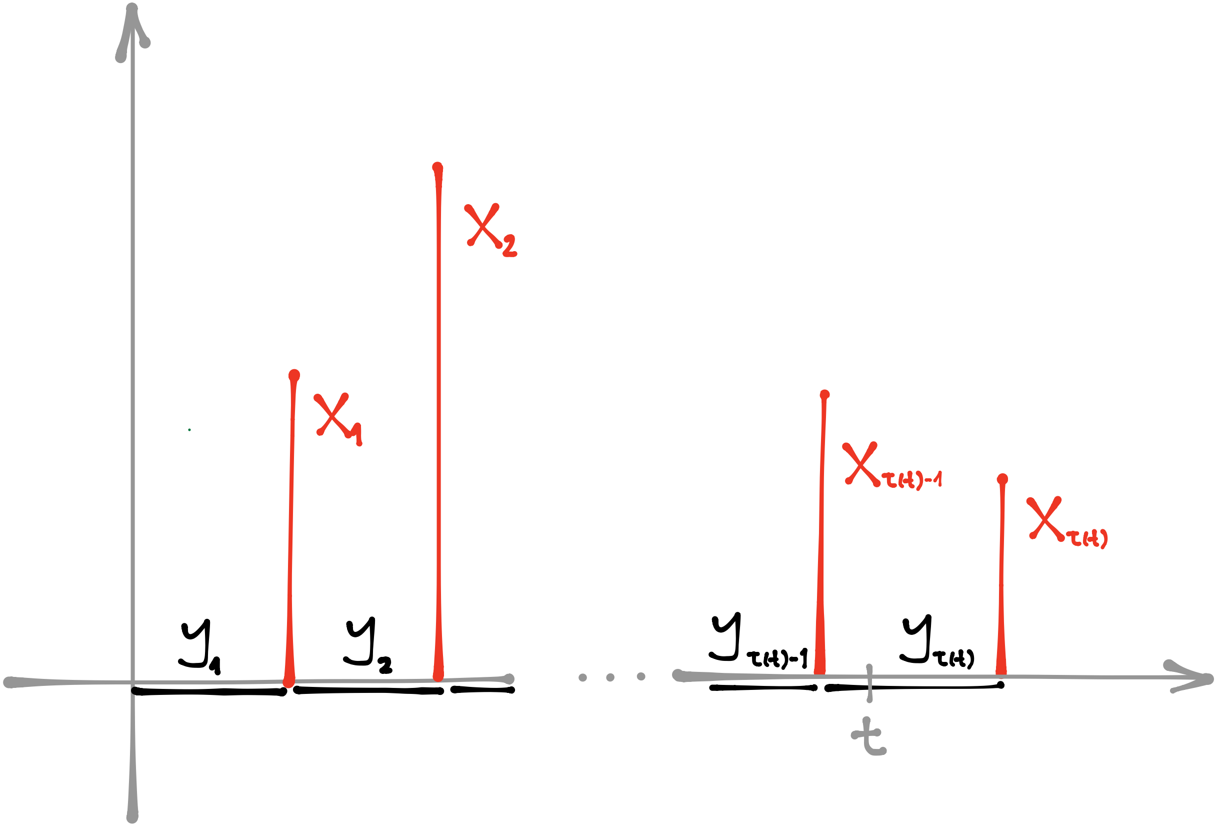
Bojan Basrak, University of Zagreb

Workshop on Extremes in Space and Time
Copenhagen, May 31, 2013

based on the joint work with

Drago Špoljarić, University of Zagreb

The problem



observations at renewal times

Main assumptions

- $((X_n, Y_n))_n$ are iid random vectors, with $Y_n \geq 0$
- $X_i \in \mathrm{MDA}(G)$ (enough to study $G = \Lambda$ or Φ_α)

For the sequence of interarrival times (Y_i) we study renewal process

$$\tau(t) = \inf\{k : Y_1 + \cdots + Y_k > t\}.$$

and partial maxima

$$M(t) = M^\tau(t) = \max_{i=1,\ldots,\tau(t)} X_i\,,$$

Berman and Barndorff-Nielsen in 1960's did the same for more general random variables $\tau(t)$.

Related results

an incomplete list

- Shantikumar and Sumita (1983)
- Anderson (1987)
- Silvestrov and Teugels (1998, 2004)
- Meerschaert and Scheffler (2004)
- Meerschaert and Stoev (2009)
- Pancheva, Mitov, Mitov (2009)

Our goal is to move beyond (or rather below) the maximum and understand the asymptotics for all upper order statistics of observations X_i until $\tau(t)$ and relax some assumption on the dependence between (X_n) and (Y_n)

Since $X_i \in$MDA(G) there exist functions $\tilde{a}(t)$ and $\tilde{b}(t)$ such that

$$tP\left(\frac{X_1 - \tilde{b}(t)}{\tilde{a}(t)} > x\right) \to -\log G(x),$$

but for iid (X_i) this is known to be equivalent to convergence of point processes

$$N_t = \sum_{i=1}^{\infty} \delta_{\left(\frac{i}{t}, X_{t,i}\right)}$$

towards

$$N \sim \text{PRM}(Leb \times \mu_G)$$

where $X_{t,i}$ represent appropriate affine transformations of the observations and the state space depends on MDA but can be written as

$$[0, \infty) \times \mathbb{E}$$

We actually sometimes need more general time normalization, and therefore we consider

$$N_t = \sum_{i=1}^{\infty} \delta_{\left(\frac{i}{g(t)}, X_{t,i}\right)}$$

where

$$X_{t,i} = \frac{X_i - \tilde{b}(g(t))}{\tilde{a}(g(t))} .$$

for a function $g \nearrow \infty$.

Although, often

$$g(t) = t$$

For simplicity we write $a(t) = \tilde{a}(g(t))$ and $b(t) = \tilde{b}(g(t))$.

Clearly

$$\left\{\frac{M(t)-b(t)}{a(t)} \leq x\right\} = \left\{\max_{i \leq \tau(t)} \frac{X_i - b(t)}{a(t)} \leq x\right\} = \left\{\max_{\frac{i}{g(t)} \leq \frac{\tau(t)}{g(t)}} \frac{X_i - b(t)}{a(t)} \leq x\right\}$$

Therefore

$$P\left(\frac{M(t)-b(t)}{a(t)} \leq x\right) = P\left(N_t\Big|_{[0,\frac{\tau(t)}{g(t)}]\times(x,\infty]} = 0\right)$$

On the rhs above we have an object of the form

$$N_t\big|_{[0,Z_t]\times\mathbb{E}}$$

Since $N_t \stackrel{d}{\to} N$, if this happens jointly with

$$Z_t \stackrel{d}{\to} Z$$

one could expect

$$N_t\big|_{[0,Z_t]\times\mathbb{E}} \stackrel{d}{\to} N\big|_{[0,Z]\times\mathbb{E}}$$

Lemma

Assume

$$(N_t, Z_t) \xrightarrow{d} (N, Z)$$

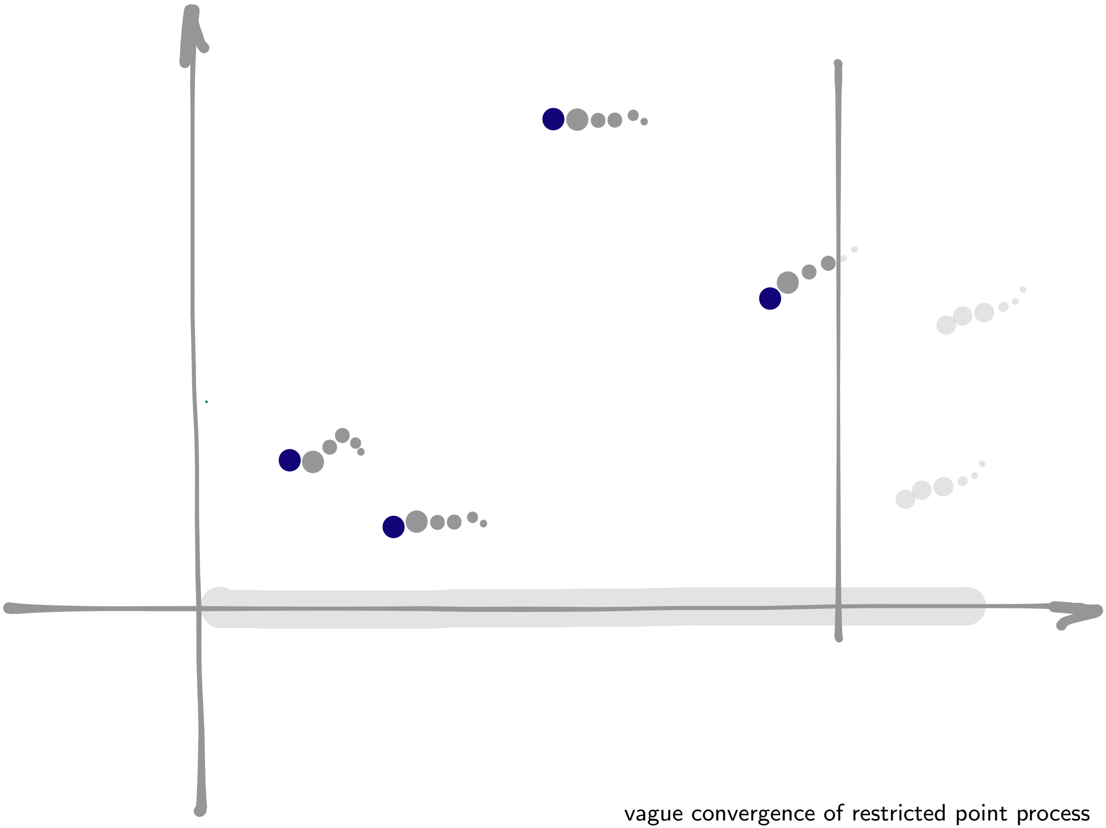
as $t \to \infty$ and

$$P(N(\{Z\} \times \mathbb{E}) > 0) = 0$$

then

$$N_t\big|_{[0,Z_t]\times\mathbb{E}} \xrightarrow{d} N\big|_{[0,Z]\times\mathbb{E}}$$

as $t \to \infty$.



vague convergence of restricted point process

The finite mean case

Assume $\mu = EY_1 < \infty$, the by SLLN

$$\frac{\tau(t)}{t} \stackrel{\text{a.s.}}{\rightarrow} \frac{1}{\mu},$$

and therefore for $g(t) = t$

$$(N_t, \frac{\tau(t)}{t}) \stackrel{d}{\rightarrow} (N, \frac{1}{\mu})$$

Remark i) dependence between X_i's and Y_i's is irrelevant ii) all upper order statistics are covered iii) Y_i's do not have to be iid actually

Example

Erdös and Rényi 1970

Note

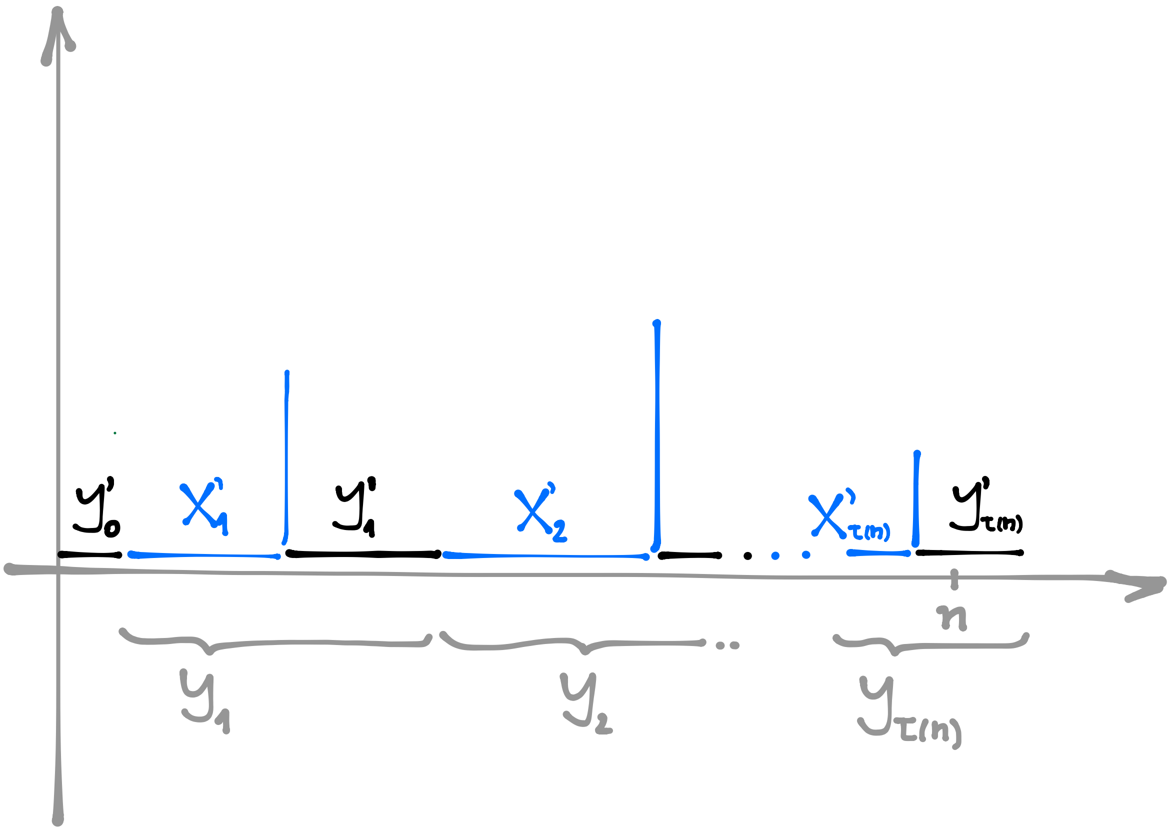
- ▷ runs of heads $X_i' \sim \text{Geom}(p)$
- ▷ runs of tails $Y_i' \sim \text{Geom}(q)$
- ▷ initial $Y_0' \sim \text{Geom}(q)$ but on $0, 1, \ldots$

We study

$$L_n = M_{\tau(n)} = \max\{X_i' : 1 \leq i \leq \tau(n)\}$$

where

$$\tau(n) = \inf\{k : Y_0' + \sum_{i=1}^{k}(X_i' + Y_i') > n\}$$



Y'_0
X'_1
Y'_1
X'_2
$X'_{\tau(n)}$
$Y'_{\tau(n)}$
n
Y_1
Y_2
$Y_{\tau(n)}$

Although geometric rv's do not belong to any MDA, still they are not far since we can always write

$$X_i' = \lfloor X_i \rfloor + 1$$

for an iid sequence

$$X_i \sim \mathsf{Exp}(-\ln p)$$

which satisfies $X_i \in \mathsf{MDA}(\Lambda)$

Set $Y_0 = Y_0'$ and

$$Y_i = X_i' + Y_i'$$

clearly corresponding renewal process satisfies

$$\frac{\tau(t)}{t} \stackrel{\text{a.s.}}{\to} \frac{1}{\mu}$$

with

$$\mu = EY_1 = \frac{1}{pq}$$

Finally

$$M^{\tau}(t) - \log_{1/p}(npq) \leq L_n - \log_{1/p}(npq) \leq M^{\tau}(t) + 1 - \log_{1/p}(npq)$$

and since the lhs and the rhs converge to G and $G + 1$ by our results, it follows that

$$L_n - \log_{1/p}(npq)$$

is tight in distribution and not far from G as well cf. EKM, but much more can be said.

The infinite mean case

As before assume

- $((X_n, Y_n))_n$ are iid random vectors, with $Y_n \geq 0$
- $X_i \in \text{MDA}(G)$ (enough to study $G = \Lambda$ or Φ_α)
- $Y_i \sim \text{RV}(\alpha)$, $\alpha \in (0, 1)$

Therefore for (d_n) such that

$$nP(Y_1 > d_n) \to 1$$

we have

$$d_n^{-1}(Y_1 + \cdots + Y_n) \xrightarrow{d} S_\alpha$$

Moreover

$$S_n(s) = \frac{1}{d_n} \sum_{i=1}^{\lfloor ns \rfloor} Y_i \xrightarrow{d} S_\alpha(s), \quad s \geq 0$$

in $D[0,\infty)$ with J_1 metric, for a positive α-stable process $S_\alpha(\cdot)$

It is known that

$$S_\alpha(s) = \sum_{T_i \le s} P_i$$

where

$$\sum_i \delta_{T_i, P_i} \sim \mathsf{PRM}(Leb \times d(-u^{-\alpha}))$$

It will be useful to study S_t indexed over all $t \in [0, \infty)$, with normalization

$$d_t = d_{\lfloor t \rfloor} \,.$$

One can also find an asymptotic inverse $\tilde{d}$ of d such that

$$d(\widetilde{d}(t)) \sim \widetilde{d}(d(t)) \sim t$$

see Seneta (1976) for instance.

Recall

$$\tau(t) = \inf\{k : Y_1 + \cdots + Y_k > t\}.$$

thus

$$\frac{\tau(t)}{\tilde{d}(t)} = \frac{1}{\tilde{d}(t)} \inf\{k : Y_1 + \cdots + Y_k > t\}$$

$$= \inf\{s : \frac{1}{t} \sum_{i=1}^{\lfloor \tilde{d}(t)s \rfloor} Y_i > 1\} \approx S^{\leftarrow}_{\tilde{d}(t)}(1)$$

since $d(\widetilde{d}(t)) \sim t$.

It is known that

$$S_t^{\leftarrow}(1) \xrightarrow{d} W_\alpha := S_\alpha^{\leftarrow}(1)$$

where W_α has Mittag-Leffler distribution. This is even true on the level of stoch. processes. Since $\tilde{d}(t) \to \infty$ also

$$\frac{\tau(t)}{\tilde{d}(t)} \xrightarrow{d} W_\alpha \, .$$

Suppose X_i and Y_i are independent. Then jointly

$$N_t \stackrel{d}{\to} N \sim \mathsf{PRM}(Leb \times \mu)$$

and

$$\frac{\tau(t)}{\tilde{d}(t)} \stackrel{d}{\to} W_\alpha \,,$$

with N independent of W_α.

In particular

$$P\left(\frac{M(t)-b(t)}{a(t)} \leq x\right) = P\left(N_t\big|_{[0,\frac{\tau(t)}{\tilde{d}(t)}]\times(x,\infty]} = 0\right)$$

as $t\to\infty$ converges to

$$P\left(N\big|_{[0,W_\alpha]\times(x,\infty]} = 0\right) = \cdots = E\left(G(x)^{W_\alpha}\right)$$

cf. Berman (1962).

Dependence between X_i and Y_i

Define

$$U_X = \frac{1}{1 - F_X} \text{ and } U_Y = \frac{1}{1 - F_Y}$$

Note that $\widetilde{d}(t) \sim U_Y(t)$. One can describe the limit quite precisely even under some sorts of dependence between observations and interarrival times.

▷ **Asymptotic tail independence**

$$\lim_{x \to \infty} P(X_1 > U_X^{\leftarrow}(x) | Y_1 > U_Y^{\leftarrow}(x)) = 0 \ .$$

▷ **Asymptotic full tail dependence**

$$\lim_{x \to \infty} P(X_1 > U_X^{\leftarrow}(x) | Y_1 > U_Y^{\leftarrow}(x)) = 1 \ .$$

Now we need to include interarrival times in the point processes and define

$$N_t = \sum_{i=1}^{\infty} \delta_{\left(\frac{i}{\widetilde{d}(t)}, X_{t,i}, Y_{t,i}\right)}$$

with

$$Y_{t,i} = \frac{Y_i}{t}$$

Tail independence

In this case (see Thm 6.2.3 in de Haan & Ferreira)

$$\widetilde{d}(t)P\left((X_{t,i},Y_{t,i})\in\cdot\right)\xrightarrow{v}\mu_0(\cdot)$$

for a measure μ_0 **concentrated on the axes** and given as

$$\mu_0(([-\infty,x]\times[0,y])^c)=e^{-x}+y^{-\alpha}\ .$$

if $X_1\in$ MDA(Λ) or as

$$\mu_0(([0,x]\times[0,y])^c)=x^{-\beta}+y^{-\alpha}\ .$$

if $X_1\in \mathrm{MDA}(\Phi_\beta)$

Recall that

$$\widetilde{d}(t)P\left((X_{t,i}, Y_{t,i}) \in \cdot\right) \xrightarrow{v} \mu_0(\cdot)$$

is necessary and sufficient for

$$N_t \xrightarrow{d} N\,,$$

where N is $\mathrm{PRM}(\lambda \times \mu_0)$

Similarly as before we obtain the joint convergence

$$(N_t, \frac{\tau(t)}{\tilde{d}(t)}) \xrightarrow{d} (N, W_\alpha)$$

with $W_\alpha =$ the first hitting time of the level 1 by a positive α-stable process S_α (note $\tau(t)/\tilde{d}(t)$ is just a transformation of N_t)

Moreover, if

$$N = \sum_i \delta_{T_i, P_i^X, P_i^Y}$$

then

$$N' = \sum_i \delta_{T_i, P_i^X} \text{ and } W_\alpha$$

are independent.

Therefore

$$P\left(\frac{M(t)-b(t)}{a(t)} \leq x\right) = P\left(N'_t\Big|_{[0,\frac{\tau(t)}{\tilde{d}(t)}]\times(x,\infty]} = 0\right)$$

as $t \to \infty$ converges to

$$E\left(G(x)^{W_\alpha}\right)$$

as before.

Full dependence

de Haan & Resnick 1977

Again

$$N_t \stackrel{d}{\to} N \,,$$

where N is $\mathrm{PRM}(\lambda \times \mu_0)$ with μ_0 concentrated on the set

$$C = \{(u,v) \in (-\infty,\infty) \times (0,\infty) : e^{-u} = v^{-\alpha}\}, \quad \text{if } G_1 = \Lambda$$

$$C = \{(u,v) \in (0,\infty) \times (0,\infty) : u^{-\beta} = v^{-\alpha}\}, \quad \text{if } G_1 = \Phi_\beta \,.$$

and

$$\mu_0(\{(u,v) : (u,v) \in C,\ v > y\}) = y^{-\alpha} \,,$$

for $y > 0$.

Again

$$(N_t, \frac{\tau(t)}{\tilde{d}(t)}) \xrightarrow{d} (N, W_\alpha)$$

However, N' and W_α are not independent, just the opposite since if $X_i, Y_i \in$ MDA(α) for instance, then

$$N = \delta_{T_i, P_i, P_i}$$

Conclusion

- ▷ point processes approach turns out to be simple and very efficient
- ▷ all upper order statistics are covered
- ▷ extremal process can be understood as well