Markov chains, extremal index and its estimation

Ana Ferreira*
Inst. Sup. Agronomia,
Univ. Técnica de Lisboa
e CEAUL

Statistical Modeling of Extremes,
Strasbourg, 23-26 June, 2008

*Partially supported by POCTI/MAT/58876/2004.

Outline:

- Discuss some basic notions: Markov chains, extremal index, extremal functions, max-domain of attraction, dependence characterizations.

- Markov chains: estimation of the extremal index.

- Simulations

- Generalizations

Consider:

- Stationary sequences of r.v.'s, $\{X_n\}_{n\geq 1}$, with common marginal continuous d.f. $F(x)$, $x \in (\alpha_F, \omega_F)$.

- Condition on the marginal F,

$$\forall_{\tau>0}\ \exists_{\{u_n\}_{n\geq 1}}\ :\ n(1 - F(u_n)) \to \tau \tag{1}$$

  (If X_i's are independent then (1) $\Leftrightarrow$ $F^n(u_n) = P\left(\max(X_1, \ldots, X_n) \leq u_n\right) \to e^{-\tau}$.)

- Distribution of maxima,

$$p_n(x) = P\left(\max(X_1, \ldots, X_n) \leq x\right).$$

- Markov process (in discrete time with continuous state space) of order d,

$$p_n(x) = \frac{p_{d+1}^{n-d}(x)}{p_d^{n-d-1}(x)}, \qquad n \geq d, \quad x \in (\alpha_F, \omega_F).$$

EXTREMAL INDEX $\theta \in [0, 1]$

(Leadbetter, Lindgren, and Rootzén (1983))

$\{X_n\}_{n \geq 1}$ is said to have extremal index θ if, additionally to (1),

$$P\left(\max(X_1, \ldots, X_n) \leq u_n\right) \to e^{-\theta\tau}. \qquad (2)$$

Dependence conditions related to the extremal index?

- Under $D(u_n)$ and $D'(u_n)$ for all linear normalization, if the normalized maxima converges in distribution for some non-degenerate G, then $G = G_\gamma^\theta$.

- Under $D(u_n)$, θ exists.

Dependence conditions related to the estimation of the extremal index?

- E.g. Hsing and co-authors (1988,1991,1993), Weissman & Novak (1998): establish consistency and asymptotic normality of some estimators of the extremal index (e.g. runs and blocks) under some stonger dependence conditions than $D(u_n)$, or other contitions ...

- More applied papers, e.g. Smith & Weissman (1994), Ancona-Navarrete & Tawn (2000) basically "assume" the above conditions.

Markov chains?

- All continuous state space Markov chains with non-degenerate transition probabilities satisfy D condition (O'Brien, 1987).

EXTREMAL FUNCTIONS $\{\epsilon_n(.)\}_{n\geq 1}$

Sequence of functions defined by,

$$\epsilon_n(x) = \frac{\log P\left(\max(X_1,\ldots,X_n) \leq x\right)}{\log P(X_1 \leq x)}, \qquad (3)$$

for $x \in (\alpha_F, \omega_F)$; whenever clear we omit the argument of the extremal functions and simply write ϵ_n.

Examples:

1. I.i.d. r.v.'s: $\epsilon_n(x) = n$, for all x.

2. R.v.'s totally dependent: $\epsilon_n(x) = 1$, for all x.

3. Stationary Markov chains of order d:

   $$\epsilon_n = (n-d)\epsilon_{d+1} - (n-d-1)\epsilon_d, \quad n \geq d.$$

   Note,

   $$\epsilon_n = \frac{\log p_n}{\log p_1} = (n-d)\frac{\log p_{d+1} - \log p_d}{\log p_1} + \frac{\log p_d}{\log p_1},$$

   for $n \geq d$.

Stability condition on the extremal functions

The extremal functions are called stable over the real sequence $\{u_n\}_{n\geq 1}$ if, for each integer $k \geq 1$, there exist non-negative constants Δ_k such that

$$\left|\epsilon_n(u_n) - k\epsilon_{[n/k]}(u_n)\right| \leq \Delta_k, \qquad n \to \infty, \tag{4}$$

where [ ] denotes the integer part.

Continuation of the Examples

1. I.i.d. r.v.'s, $\epsilon_n = n$:

$$0 \leq \epsilon_n - k\epsilon_{[n/k]} \leq \Delta_k = k, \quad \text{for all } n \text{ and } k \leq n.$$

2. Total dependence, $\epsilon_n = 1$:

$$\left|\epsilon_n - k\epsilon_{[n/k]}\right| = \Delta_k = k - 1, \quad \text{for all } n \text{ and } k \leq n.$$

3. Stationary Markov chains of order d: for $u_n \to \omega_F$, a possible choice is

$$\Delta_k = k(d+1)^2.$$

STATIONARY MARKOV CHAINS, MAX-DOMAIN OF ATTRACTION, EXTREMAL FUNCTIONS and EXTREMAL INDEX

Statement 1 The distribution of the maximum of the Markov chain conveniently normalized is in the domain of attraction of some extreme value distribution: $P\left(\max(X_1, \ldots, X_n) \leq a_n x + b_n\right) \to G_\gamma(x)$.

Corollary 1 (Ferreira (2005)). *If* $\{\epsilon_n(.)\}_{n\geq 1}$ *are stable over* $\{u_n\}_{n\geq 1}$ *and* $\lim_{n\to\infty} \epsilon_n(u_n)/n$ *exists for some* $\tau > 0$*, then* $\{X_n\}_{n\geq 1}$ *has extremal index* θ *and*

$$\theta = \lim_{n\to\infty} \frac{\epsilon_n(u_n)}{n}. \tag{5}$$

Conversely, if the process has extremal index then $\lim_{n\to\infty} \epsilon_n(u_n)/n$ *exists for some* $\tau > 0$*, and* (5) *holds.*

Note:

$$\begin{aligned}
\frac{\epsilon_n(u_n)}{n} &= \frac{\log P\left(\max(X_1, \ldots, X_n) \leq u_n\right)}{n \log P(X_1 \leq u_n)} \\
&\sim_{x\uparrow 1} \frac{\log P\left(\max(X_1, \ldots, X_n) \leq u_n\right)}{-n(1 - F(u_n))} \\
&\sim_{x\uparrow 1} \frac{-\theta\tau}{-\tau} = \theta.
\end{aligned}$$

ESTIMATION OF θ (A)

We had,

$$\theta \approx \frac{\log P\left(\max(X_1,\ldots,X_n) \le u_n\right)}{n \log P(X_1 \le u_n)},$$

which motivates,

$$\widehat{\theta}_n = \frac{\log\left(\frac{1}{k}\sum_{i=1}^{k} 1_{\{M_r^{(i)} \le v_n\}}\right)}{r \log \widehat{P(X_1 \le v_n)}}. \tag{6}$$

Rewrite a sample of size n, $(X_1,\ldots,X_n)$ by

$$\begin{aligned}&\{X_i^{(j)}\}_{i=1,\ldots,r;j=1,\ldots,k}\\ &\quad = \left(X_1^{(1)},\ldots,X_r^{(1)},X_1^{(2)},\ldots,X_r^{(2)},\ldots,X_1^{(k)},\ldots,X_r^{(k)}\right)\end{aligned}$$

that is,

n(sample size) $= r$(blocks size)$\times k$(number of blocks).

For each sub-sample of size r, i.e. for each $j = 1,\ldots,k$, denote

$$M_r^{(j)} = \max\left(X_1^{(j)},\ldots X_r^{(j)}\right),$$

and use the empirical d.f. to estimate $P\left(\max(X_1,\ldots,X_n) \le \cdot\right)$; $\{v_n\}_{n\ge1}$ is some appropriate sequence and $\log \widehat{P(X_1 \le v_n)}$ is any "consistent" estimator of the sequence of tail probabilities $\log P(X_1 \le v_n)$.

A concrete example of $\widehat{\theta}_n$ is, with $v_n = X_{n-\tau+1,n}$, $\tau/n \to 0$ (in particular τ can be fixed),

$$\widehat{\theta}_n^A = \frac{\sum_{i=1}^k 1_{\left\{M_r^{(i)} > X_{n-\tau+1,n}\right\}}}{\tau}.$$

It gives a kind of proportion of block maxima which lie above the threshold. Note that the total of block maxima above $X_{n-\tau+1,n}$ can not exceed τ.

Other example is,

$$\widehat{\theta}_{SW} = \frac{\log\left(k^{-1}\sum_{i=1}^k 1_{\left\{M_r^{(i)} \le X_{n-\tau+1,n}\right\}}\right)}{r\log\left(\frac{n-\tau+1}{n}\right)}$$

(e.g. Smith and Weissman, 1994).

Theorem 1 (Consistency of $\widehat{\theta}_n$, F. & F. (2007))**.** *Assume that for some sequences $\{v_n\}_{n\geq 1}$ and $r = \{r_n\}_{n\geq 1}$, with $r \geq d$,*

$$\frac{\epsilon_r(v_n)}{r} \to \theta \in [0,1], \qquad n \to \infty.$$

1. *For $p_r \to p_0 \in (0,1)$, take $k = \{k_n\}_{n\geq 1} \to \infty$ and $\log p_1 \to 0$ or,*

2. *for $p_r \to 1$, take $k(1-p_r) \to \infty$ and $(1-p_r)^{-2} \log p_1 \to 0$,*

and let $\log \widehat{P(X_1 \leq v_n)}$ be any consistent estimator of $\log p_1$ (i.e. $\log \widehat{P(X_1 \leq v_n)} / \log p_1 \to^P 1$). Then

$$\widehat{\theta}_n \to^P \theta, \quad n \to \infty.$$

ESTIMATION OF θ (B)

Denote by F_{d+1} the joint d.f. of $d+1$ consecutive r.v.'s from a stationary Markov chain of order d;

$F_{d+1} \in \mathcal{D}(\mathbf{G}_\gamma)$ if, there exists $(d+1)-$dimensional sequences $\{\mathbf{a}_n\}_{n\geq 1} > 0$ and $\{\mathbf{b}_n\}_{n\geq 1}$, and a $(d+1)$-dimensional d.f. $\mathbf{G}_\gamma$ with non-degenerate marginals, such that,

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{\max(\mathbf{X}_1, \ldots, \mathbf{X}_n) - \mathbf{b}_n}{\mathbf{a}_n} \leq \mathbf{x}\right) = \mathbf{G}_\gamma(\mathbf{x}) \qquad (7)$$

(consider all operations taken componentwise) for all continuity points $\mathbf{x} = (x_1, \ldots, x_{d+1})$ of $\mathbf{G}_\gamma$; the components of $\gamma = (\gamma_1, \ldots, \gamma_{d+1})$ are the marginal extreme value indices.

$F_{d+1} \in \mathcal{D}(\mathbf{G}_\gamma)$ in terms of "Fréchet marginals",

$$\lim_{n\to\infty} P\left(\max_{1\leq i\leq n}\left(\frac{1}{1-F(X_1^{(i)})}, \ldots, \frac{1}{1-F(X_d^{(i)})}\right) \leq n\mathbf{x}\right)$$
$$= \mathbf{G}\left(\frac{\mathbf{x}^\gamma - \mathbf{1}}{\gamma}\right), \quad \mathbf{x} \in (0,\infty)^d.$$

The dependence function $\mathbf{L}$ introduced by Huang and co-authors (cf. Huang (1992)) is defined by

$$\mathbf{L}(\mathbf{x}) = -\log \mathbf{G}\left(\frac{\mathbf{x}^{-\gamma} - \mathbf{1}}{\gamma}\right), \quad \mathbf{x} \in (0,\infty)^d. \qquad (8)$$

When the components of the vectors $\mathbf{x}$ and γ are all equal, x and γ say, respectively, we shall simply denote $L_d(x)$ in the left-hand side of (8).

Proposition 1. *If $F_{d+1} \in \mathcal{D}(\mathbf{G}_\gamma)$,*

$$\lim_{n\to\infty} \frac{\epsilon_n(u_n)}{n} = \lim_{n\to\infty} \epsilon_{d+1}(u_n) - \epsilon_d(u_n) = L_{d+1}(1) - L_d(1),$$

for $\{u_n\}_{n\geq 1} = \{(1/(1-F))^{\leftarrow}(nx)\}_{n\geq 1}$ and $x > 0$. When $d = 1$ this simplifies to

$$\epsilon_n(u_n)/n \to L_2(1) - 1.$$

"Proof". $F_d \in \mathcal{D}(\mathbf{G}_\gamma)$ implies

$$\lim_{n\to\infty} n\,(1 - F_d(u_n)) = -\log G_0\left(\frac{x^\gamma - 1}{\gamma}\right) = L_d(1/x).$$

Therefore,

$$\epsilon_d(u_n) \sim \frac{\log p_d(u_n)}{\log p_1(u_n)} \sim \frac{1 - p_d(u_n)}{1 - p_1(u_n)} \sim u_n L_d(1/u_n) = L_d(1),$$

the equality by the homogeneity of $\mathbf{L}$: $\mathbf{L}(a\mathbf{x}) = a\mathbf{L}(\mathbf{x})$, for all $a > 0$ and $\mathbf{x} \in (0,\infty)^d$.

The last result motivates the estimator,

$$\widehat{\theta}_L = \widehat{L}_{d+1}(1) - \widehat{L}_d(1). \tag{9}$$

Estimators for L are known from Huang & co-authors.

Or, one can use many of the known relations among the several dependence functions or dependence parameters in e.v.t. Just to mention a few:

- case $d = 1$, $\lambda = 2 - L_2(1)$, with λ the Sibuya's (1960) tail dependence coefficient;

- case $d = 1$, $\chi(t) = \mathbf{L}(t, 1) - t - 1$, with $\chi(t)$, $t > 0$, the dependence function introduced by Sibuya (1960) and Geffroy (1958);

- case $d = 1$, $\mathbf{L}(1, 1) = 2A(1/2)$, where A is the Pickands's (1981) dependence function.

Theorem 2 (Consistency and asymptotic normality of $\widehat{\theta}_L$, F. & F. (2007))**.** *If $F_{d+1} \in \mathcal{D}(\mathbf{G}_\gamma)$ and θ exists,*

1. $\widehat{\theta}_n \to^P \theta, \quad n \to \infty.$

2. *Suppose, additionally, that for some $\alpha > 0$ and for all $x, y > 0$ $(t \to \infty)$*

   $$t\left\{1 - F_{d+1}\left(\left(\frac{1}{1-F}\right)^{\leftarrow}\left(\frac{t}{x}\right)\right)\right\} = L_{d+1}(x) + O(t^{-\alpha}) ,$$

   holds uniformly on $\left\{\sum x_i^2 = 1, x_i \geq 0, i = 1, \ldots, d\right\}$ and L_{d+1} has continuous first order partial derivatives near 1. Then, for $k \to \infty$, $k = o\left(n^{2\alpha/(1+2\alpha)}\right)$,

   $$\sqrt{k}\left(\widehat{\theta}_L - \theta\right) \to^d N\left(0, \sigma_{d+1}^2 + \sigma_d^2\right) ,$$

   where N is a standard normal r.v.,

   $$\sigma_d^2 = L_d(1) + \sum_{i=1}^{d}\left(\left(L_{d+1}^{(i)}(1)\right)^2 - 2L_{d+1}^{(i)}(1)\right) + 2\sum_{i=1}^{d}\sum_{j=1, j\neq i}^{d} L_{d+1}^{(i)}(1) L_{d+1}^{(i)}(1)\left(2 - L_{d+1}^{(i,j)}\right)$$

   and similarly for σ_{d+1}^2.

SIMULATIONS

Considered processes:

- Max-autoregressive process (Markov chain of order 1): W_n i.i.d. unit Fréchet

  $$\begin{cases} X_1 &= W_1/\theta \\ X_n &= \max\{(1-\theta)X_{i-1}, W_n\}, \quad n \geq 2. \end{cases}$$

  Consider $\theta = 0.5$.

- Markov chain of order 1 with copula (Kimeldorf and Sampson (1975)),

  $$C(u,v) = u + v - 1 + ((1-u)^{-1} + (1-v)^{-1} - 1)^{-1},$$

  for $u, v \in (0,1)$; $\theta = .5$;

- Markov Gaussian process,

  $$\begin{cases} X_1 &= (1-\alpha^2)^{-1/2}\epsilon_1 \\ X_n &= \alpha X_{n-1} + \epsilon_n, \quad n \geq 2, \end{cases}$$

  where $\{\epsilon_n\}_{n\geq 1}$ are i.i.d. $N(0, 1-\alpha^2)$ r.v.'s; take $\alpha = 0.5$; $\theta = 1$.

Considered estimators:

$$\begin{aligned}
\widehat{\theta}_n^A &= \frac{\sum_{i=1}^k 1_{\left\{M_r^{(i)} > X_{n-\tau+1,n}\right\}}}{\tau} \\
\widehat{\theta}_n^B &= \frac{n \log\left(k^{-1}\sum_{i=1}^k 1_{\left\{M_r^{(i)} \le X_{n-\tau+1,n}\right\}}\right)}{-r\tau} \\
\widehat{\theta}_{SW} &= \frac{\log\left(k^{-1}\sum_{i=1}^k 1_{\left\{M_r^{(i)} \le X_{n-\tau+1,n}\right\}}\right)}{r \log\left(\frac{n-\tau+1}{n}\right)}
\end{aligned}$$

(e.g. Smith and Weissman, 1994)

$$\widehat{\theta}_R = \frac{\sum_{i=1}^{n-r} 1_{\{X_i > X_{n-\tau+1,n}\}} 1_{\{X_{i+1} \le X_{n-\tau+1,n}\}} \cdots 1_{\{X_{i+r} \le X_{n-\tau+1,n}\}}}{\tau},$$

(cf. Hsing, 1991,1993; Smith and Weissman, 1994; Weissman and Novak, 1998)

$$\widehat{\theta}_L = \widehat{L}_2(1) - 1 = \frac{1}{k}\sum_{i=1}^n 1_{\left\{X_1^{(i)} \ge X_1^{(n-k+1,n)} \text{ or } X_2^{(i)} \ge X_2^{(n-k+1,n)}\right\}} - 1.$$

Copula $\mathbf{C}$:

$$p_d(x) = P(X_1 \leq x, \ldots, X_d \leq x) = \mathbf{C}(F(x), \ldots, F(x)),$$

i.e., denote $c_d(.) = p_d\left(F^{\leftarrow}(.)\right)$, the arrow meaning the inverse function.

Lemma 1. *For $\{X_n\}_{n\geq 1}$ stationary Markov chain of order d,*

$$\theta = \lim_{x\to 1} \frac{c_d(x) - c_{d+1}(x)}{1-x} \tag{10}$$

in case one of the two exist. In particular, for $d = 1$,

$$\theta = \lim_{x\to 1} \frac{x - c_2(x)}{1-x} = c_2'(1) - 1. \tag{11}$$

Example.

Stationary Markov chain of order 1 with copula (Kimeldorf and Sampson (1975))

$$C(u,v) = u + v - 1 + ((1-u)^{-1} + (1-v)^{-1} - 1)^{-1},$$

$u, v \in (0,1)$. Then $c_2(x) = 2x^2(1+x)^{-1}$, hence

$$\lim_{x\to 1} \frac{x - c_2(x)}{1-x} = \lim_{x\to 1} \frac{x}{1+x} = \frac{1}{2}$$

i.e. $\theta = 0.5$.

GENERALIZATIONS

Define

$$p_r^{(i,j)}(v_n) = P(M_r^{(i)} \le v_n, M_r^{(j)} \le v_n)$$

and

$$\epsilon_r^{(i,j)}(v_n) = \frac{\log p_r^{(i,j)}(v_n)}{\log p_1(v_n)}, \qquad i, j = 1, \dots, k.$$

Theorem 3 (Consistency of $\widehat{\theta}_n$). *Suppose for some sequences $\{v_n\}_{n\ge1}$ and $r = \{r_n\}_{n\ge1}$,*

$$\frac{\epsilon_r(v_n)}{r} \to \theta \in [0,1], \qquad n \to \infty. \tag{12}$$

1. *For $p_r \to p_0 \in (0,1)$, take $k = \{k_n\}_{n\ge1} \to \infty$ and*

$$\sup_{1\le i,j\le k} \left(\epsilon_r^{(i,j)} - 2\epsilon_r\right) \log p_1 \to 0, \qquad n \to \infty. \tag{13}$$

2. *For $p_r \to 1$, take $k = \{k_n\}_{n\ge1}$ with $k(1-p_r) \to \infty$ and*

$$\sup_{1\le i,j\le k} \left(\epsilon_r^{(i,j)} - 2\epsilon_r\right) \frac{\log p_1}{(1-p_r)^2} \to 0, \qquad n \to \infty. \tag{14}$$

Then, for $\log \widehat{P(X_1 \le v_n)}$ *any "consistent"estimator of* $\log p_1$,

$$\widehat{\theta}_n \to^P \theta, \quad n \to \infty.$$

Example. Max-autoregressive process: W_n i.i.d. unit Fréchet

$$\begin{cases} X_1 &= W_1/\theta \\ X_n &= \max\{(1-\theta)X_{i-1}, W_n\}, \quad n \geq 2 \end{cases}$$

$$\begin{aligned} p_1(x) &= e^{-1/(\theta x)}, \\ p_r(x) &= e^{-(1+\theta(r-1))/(\theta x)}, \end{aligned}$$

$x > 0$, $\theta \in (0, 1]$; hence for all $x > 0$,

$$\epsilon_n(x) = 1 + \theta(n-1),$$

$$\theta = \lim_{n\to\infty} \frac{\epsilon_n}{n},$$

and the extremal functions are stable, e.g.

$$\Delta_k = 2k - 1.$$

SOME REFERENCES

Ancona-Navarrete, M.A. and Tawn, J.A. (2000). A comparison of methods for estimating the extremal index. *Extremes* **3**, 5-38.

Ferreira, H. (2005). A new dependence condition for time series and the extremal index of higher-order Markov chains. *Revstat*; accepted for publication.

Ferreira, A. and Ferreira, H. (2007) Extremal functions, extremal index and Markov chains. Submitted (Technical report: Notas e comunicações CEAUL 12/2007).

Geffroy, J. (1958). Contributions á la theorie des valeurs extrémes. *Publ. Inst. Statist. Univ. Paris 7* **8**, 37-185.

Huang, X. (1992). Statistics of Bivariate Extreme Values *Tinbergen Institute Research Series, Ph.D. thesis.*.

Hsing, T. (1991). Estimating the parameters of rare events. *Stochast. Process. Appl.* **37**, 117-139.

Hsing, T. (1993). Extremal index estimation for a weakly dependent stationary sequence. *Ann. Statist.* **21**, 2043-2071.

Hsing, T., Husler, J. and Leadbetter, M.R. (1988). On the exceedence point process for stationary sequence. *Probab. Theory Related Fields* **78**, 97-112.

Kimeldorf, G. and Sampson, A.R. (1975) Uniform representations of bivariate distributions. *Comm. Statist.* **4**, 617-627.

Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes* Springer, Berlin.

Pickands III, J. (1981) Multivariate Extreme Value Distributions. *Proceedings: 43rd Session of the International Statistical Institute. Book 2, Buenos Aires, Argentina*, 859-878.

Sibuya, M. (1960). Bivariate extreme statistics. *Ann. Inst. Statist. Math. Tokyo*, **11**, 195-210.

Smith, R.L. and Weissman, I. (1994). Estimating the extremal index. *J. R. Statist. Soc. B* **56**, 515-528.