REVSTAT – Statistical Journal
Volume 4, Number 2, June 2006, 143–151

A NEW DEPENDENCE CONDITION FOR TIME SERIES AND THE EXTREMAL INDEX OF HIGHER-ORDER MARKOV CHAINS

Authors: Helena Ferreira
– Department of Mathematics, University of Beira Interior,
6200 Covilhã, Portugal
ferreira@fenix2.ubi.pt

Received: June 2005 Revised: August 2005 Accepted: August 2005

Abstract:

- We present a new dependence condition for time series and extend the extremal types theorem.

  The dependence structure of a stationary sequence is described by a sequence of extremal functions. Under a stability condition for the sequence of extremal functions, we obtain the asymptotic distribution of the sample maximum.

  As a corollary, we derive a surprisingly simple method for computing the extremal index through a limit of a sequence of extremal coefficients.

  The results may be used to determine the asymptotic distribution of extreme values from stationary time series based on copulas. We illustrate it with the study of the extremal behaviour of d^{th}-order stationary Markov chains in discrete time with continuous state space. For such sequences we present a way to compute the extremal index from the upper extreme value limit for its joint distribution of $d+1$ consecutive variables.

Key-Words:

- *extremal coefficient; dependence; extremes; extremal index; higher-order stationary Markov sequences.*

1. INTRODUCTION

Let $\mathbf{X} = \{X_n\}_{n\geq 1}$ be a stationary sequence with common distribution function F in the domain of attraction of an extreme value distribution G. Therefore there exist real sequences $\mathbf{a} = \{a_n > 0\}_{n\geq 1}$ and $\mathbf{b} = \{b_n\}_{n\geq 1}$ such that

$$F^n\big(u_n(x)\big) \underset{n\to\infty}{\longrightarrow} G(x) , \qquad x \in \mathbb{R} ,$$

where $u_n(x) = a_n x + b_n$.

Let $\{\varepsilon_n^{\mathbf{X}}(.)\}_{n\geq 1}$ be the sequence of functions satisfying

$$P\big(X_1 \leq y, ..., X_n \leq y\big) = F^{\varepsilon_n^{\mathbf{X}}(y)}(y) , \qquad y \in (\alpha_F, \omega_F), \quad n \geq 1 ,$$

where α_F and ω_F denote the left and right end points of F.

This sequence of extremal functions $\{\varepsilon_n^{\mathbf{X}}(.)\}_{n\geq 1}$ associated to $\mathbf{X}$ is inspired by the extremal coefficients considered in Buishand (1984), Tiago de Oliveira (1989) and Smith (1990), among others, to model the dependence of marginals of a multivariate extreme value distribution.

Here we will consider a stability condition for this sequence of extremal functions in order to obtain limiting results for the distribution of maxima $M_n = \max\{X_1, ..., X_n\}$ and the existence of the extremal index of $\mathbf{X}$.

We first point out some properties of $\varepsilon_n^{\mathbf{X}}(.)$ coming directly from the definition.

The Fréchet bounds for $F_n(y) = P(X_1 \leq y, ..., X_n \leq y)$, given by the inequalities $\max\big\{0,\, nF(y) - (n-1)\big\} \leq F_n(y) \leq F(y)$, enables the conclusion that $\varepsilon_n^{\mathbf{X}}(y) \geq 1,\ y \in \mathbb{R}$.

In particular, if $\mathbf{X}$ has a positive dependence structure (Joe (1997)) then

$$F^{\varepsilon_n^{\mathbf{X}}(y)}(y) \geq F^n(y) ,$$

and it would follow that $\varepsilon_n^{\mathbf{X}}(y) \leq n,\ y \in \mathbb{R}$.

Finally, if $(X_1, ..., X_n)$ has a multivariate extreme value distribution then the stability equation for its dependence function D_{F_n} (Deheuvels (1978), Hsing (1989)),

$$D^t_{F_n}(y_1, ..., y_n) = D_{F_n}(y_1^t, ..., y_n^t) ,$$

$t > 0,\ y_1, ..., y_n \in [0, 1]$, leads to $\varepsilon_n^{\mathbf{X}}(y) = \varepsilon_n^{\mathbf{X}},\ y \in \mathbb{R}$. Moreover, this constant $\varepsilon_n^{\mathbf{X}} \in [1, n]$ takes the extreme values 1 or n if and only if F_n has perfect positive dependence or independent marginals, respectively.

In this paper we will only assume that the sequences $\left\{\varepsilon_n^{\mathbf{X}}(u_n(x))\right\}_{n\geq 1}$, $x \in \mathbb{R}$, satisfy a stability condition introduced in section 2. Such condition is sufficient to conclude that if $F_n(u_n(x))$ converges to a non degenerate distribution G_* then G_* is in the class of max-stable distributions.

Moreover, we recall the definition of extremal index θ and prove that it can be computed from the limit of $\varepsilon_n^{\mathbf{X}}(u_n^{\tau_0})/n$, for some $\tau_0 > 0$, where $\{u_n^{\tau}\}_{n\geq 1}$ denotes a real sequence such that $n(1-F(u_n^{\tau})) \xrightarrow[n\to\infty]{} \tau > 0$.

In section 3 we apply the results to Markov chains in discrete time with continuous state space. After the calculation of the extremal index of a Markov chain of order 1 based on a given bivariate dependence (copula) function, we demonstrate a sufficient condition for the existence of extremal index of a d^{th}-order Markov chain and compute its value. For such sequences, when the distribution of $d{+}1$ consecutive variables is in the domain of attraction of a $(d{+}1)$-multivariate extreme distribution H_{d+1}, it holds

$$\theta \;=\; -\ln D_{H_{d+1}}(e^{-1},...,e^{-1}) + \ln D_{H_d}(e^{-1},...,e^{-1}) \ ,$$

where $D_{H_{d+1}}$, D_{H_d} denote the dependence functions of the multivariate distribution functions H_{d+1}, H_d, respectively, and

$$H_d(y_1,...,y_d) = H_{d+1}(y_1,...,y_d,+\infty) \ .$$

The notation introduced in this paragraph will be used throughout the paper.

2. STABLE EXTREMAL FUNCTIONS

We now introduce a stability condition for the sequence $\{\varepsilon_n^{\mathbf{X}}(.)\}_{n\geq 1}$ under which we can, asymptotically, relate the dependence measure $\varepsilon_n^{\mathbf{X}}(.)$ for $(X_1,...,X_n)$ to the analogous measure $\varepsilon_{[n/k]}^{\mathbf{X}}(.)$ for $\left(X_{(i-1)[n/k]+1},...,X_{i[n/k]}\right)$, $1\leq i\leq k$.

Definition. The sequence $\{\varepsilon_n^{\mathbf{X}}(.)\}_{n\geq 1}$ is stable over the real sequence $\{u_n\}_{n\geq 1}$ if, for each $k\geq 1$, it holds

$$\left|\varepsilon_n^{\mathbf{X}}(u_n) - k\,\varepsilon_{[n/k]}^{\mathbf{X}}(u_n)\right| \xrightarrow[n\to\infty]{} \varepsilon_k \geq 0 \ . \tag{2.1}$$

We shall pursue the direction of this dependence condition and extend the extremal types theorem (Leadbetter *et al.* (1983)). Although the dependence between X_i and X_j does not necessarily fall off when $|i-j|$ increases, as occurs in the condition $D(u_n)$ of Leadbetter (1974), the condition (2.1) is still appropriate for the argument of extremes.

Proposition 2.1. *Let* $\mathbf{X}=\{X_n\}_{n\geq 1}$ *be a stationary sequence with common distribution function* F *and* $\mathbf{a}=\{a_n>0\}_{n\geq 1}$, $\mathbf{b}=\{b_n\}_{n\geq 1}$ *real sequences such that* $F^n(u_n(x)) \underset{n\to\infty}{\longrightarrow} G(x)$, $x\in\mathbb{R}$, *where* $u_n(x)=a_n x+b_n$ *and* G *is a non degenerate distribution function.*

If $F_n(u_n(x)) \underset{n\to\infty}{\longrightarrow} G_*(x)$, $x\in\mathbb{R}$, *for some non degenerate distribution* G_* *and* $\{\varepsilon_n^{\mathbf{X}}(.)\}_{n\geq 1}$ *is stable over the real sequence* $\{u_n(x)\}_{n\geq 1}$, *for all* $x\in\mathbb{R}$, *then* G_* *is of extreme value type.*

Proof: Since every max-stable distribution is of extreme value type, it is sufficient to prove that there are real sequences $\{\alpha_n>0\}_{n\geq 1}$ and $\{\beta_n\}_{n\geq 1}$ such that

$$G_*^n(\alpha_n x+\beta_n) = G_*(x)\ , \qquad n\geq 1\ . \tag{2.2}$$

We follow essentially the proof of Theorem 1.3.1 of Leadbetter *et al.* (1983): if $F_n(u_{nk}(x)) \underset{n\to\infty}{\longrightarrow} G_*^{1/k}(x)$, $x\in\mathbb{R}$, $k\geq 1$, then (2.2) holds. To obtain this last convergence we note that $F_{nk}(u_{nk}(x)) \underset{n\to\infty}{\longrightarrow} G_*(x)$ and

$$\begin{aligned}
\Big|F_{nk}\big(u_{nk}(x)\big) - F_n^k\big(u_{nk}(x)\big)\Big| \;&=\\
&= \Big|F^{\varepsilon_{nk}^{\mathbf{X}}(u_{nk}(x))}\big(u_{nk}(x)\big) - F^{k\varepsilon_n^{\mathbf{X}}(u_{nk}(x))}\big(u_{nk}(x)\big)\Big|\\
&= F^{\varepsilon_{nk}^{\mathbf{X}}(u_{nk}(x))}\big(u_{nk}(x)\big)\Big|1 - F^{k\varepsilon_n^{\mathbf{X}}(u_{nk}(x))-\varepsilon_{nk}^{\mathbf{X}}(u_{nk}(x))}\big(u_{nk}(x)\big)\Big| \;=\; o(1)\ ,
\end{aligned}$$

by applying (2.1). □

The proof points out that the convergence in (2.1) can be weakned. The result holds for bounded sequences $\big|\varepsilon_n^{\mathbf{X}}(u_n(x)) - k\,\varepsilon_{[n/k]}^{\mathbf{X}}(u_n(x))\big|$, $x\in\mathbb{R}$, $k\geq 1$.

As a corollary we provide a relation between the sequence of extremal coefficients $\{\varepsilon_n^{\mathbf{X}}(u_n^\tau)\}_{n\geq 1}$ and the extremal index θ of $\mathbf{X}$.

Specifically, $\mathbf{X}$ has extremal index θ (Leadbetter *et al.* (1983)) if, for each $\tau>0$, there exists $\{u_n^\tau\}_{n\geq 1}$ such that $\lim\limits_{n\to+\infty} n(1-F(u_n^\tau))=\tau$ and $\lim\limits_{n\to+\infty} F_n(u_n^\tau)=e^{-\theta\tau}$. If θ exists then is given by

$$\theta \;=\; \frac{\ln \lim\limits_{n\to+\infty} F_n(u_n^\tau)}{\ln \lim\limits_{n\to+\infty} F^n(u_n^\tau)}\ .$$

Proposition 2.2. *Let* $\mathbf{X}$ *be a stationary sequence with common distribution function* F *such that, for each* $\tau>0$, *there exists* $\{u_n^\tau\}_{n\geq 1}$ *satisfying* $n(1-F(u_n^\tau)) \underset{n\to\infty}{\longrightarrow} \tau>0$. *If, for each* $\tau>0$, $\{\varepsilon_n^{\mathbf{X}}(.)\}_{n\geq 1}$ *is stable over* $\{u_n^\tau\}_{n\geq 1}$ *then:*

(**i**) *there are constants* θ' *and* θ'' *satisfying* $\liminf_{n\to+\infty} F_n(u_n^\tau) = e^{-\theta'\tau}$ *and* $\limsup_{n\to+\infty} F_n(u_n^\tau) = e^{-\theta''\tau}$, *for all* $\tau>0$;

(ii) *the convergence of* $\{F_n(u_n^{\tau_0})\}_{n\geq 1}$*, for some* $\tau_0 > 0$*, implies* $\theta' = \theta''$ *and* $\lim_{n\to+\infty} F_n(u_n^\tau) = e^{-\theta\tau}$*, for all* $\tau > 0$. □

We omit the proof since it follows the same discussion used in Theorem 3.7.1 of Leadbetter *et al.* (1983) from the result

$$\left|F_n(u_n^\tau) - F_{[n/k]}^k(u_n^\tau)\right| = o(1) .$$

Since $\lim_{n\to+\infty} F_n(u_n^\tau) = \lim_{n\to+\infty} F^{\varepsilon_n^{\mathbf{X}}(u_n^\tau)}(u_n^\tau)$ and $\lim_{n\to+\infty} F^n(u_n^\tau) = e^{-\tau}$ the second statement of the above result can be rewritten as follows.

Corollary 2.1. *Let* $\mathbf{X}$ *be a stationary sequence with common distribution function* F *such that, for each* $\tau > 0$*, there exists* $\{u_n^\tau\}_{n\geq 1}$ *satisfying* $n(1-F(u_n^\tau)) \underset{n\to\infty}{\longrightarrow} \tau > 0$.

If, for each $\tau > 0$, $\{\varepsilon_n^{\mathbf{X}}(.)\}_{n\geq 1}$ *is stable over* $\{u_n^\tau\}_{n\geq 1}$ *then* $\mathbf{X}$ *has extremal index* θ *if and only if* $\theta = \lim_{n\to+\infty} \frac{\varepsilon_n^{\mathbf{X}}(u_n^{\tau_0})}{n}$*, for some* $\tau_0 > 0$. □

This surprisingly simple result presents a new method for computing the extremal index, through a limit of a sequence of extremal coefficients, and relates the extremal index with the dependence structure of $\mathbf{X}$.

3. CALCULATING THE EXTREMAL INDEX OF MARKOV CHAINS

The stationary Markov chains are important both from the applied and theoretical points of view and a sizeable literature on its extremal behaviour is available. There are stationary Markov sequences for which the condition $D(u_n^\tau)$ fails and, in general, it is not easy to show directly from the functional form of its distributions that $D(u_n^\tau)$ holds for each $\tau > 0$. O'Brien (1987) and Rootzen (1988) propose instead a general method by considering $\mathbf{X}$ as a measurable function of a Harris chain.

Since

$$\frac{\varepsilon_n^{\mathbf{X}}(u_n^\tau)}{n} = \frac{\ln D_{F_n}\big(F(u_n^\tau), ..., F(u_n^\tau)\big)}{\ln F^n(u_n^\tau)} ,$$

the above corollary seems to be suitable for the computation of θ in stationary sequences constructed from a given dependence function and a univariate margin.

We will apply the previous results to Markov models which can be defined from families of dependence functions. We start by illustrating the results with a Markov chain of order 1.

Example 3.1. Let $\mathbf{X} = \{X_n\}_{n\geq 1}$ be a stationary Markov chain of order 1 with common distribution function F such that, for each $\tau > 0$, there exists $\{u_n^\tau\}_{n\geq 1}$ satisfying $n(1-F(u_n^\tau)) \xrightarrow[n\to\infty]{} \tau > 0$.

Suppose that the dependence function D_{F_2} of (X_1, X_2) is defined (Kimeldorf and Sampson (1975)) by

$$D_{F_2}(u,v) \;=\; u+v-1+\Big((1-u)^{-1}+(1-v)^{-1}-1\Big)^{-1}\;, \qquad u,\, v \in [0,1]\;.$$

We get, for each $\tau > 0$,

$$\begin{aligned}\varepsilon_n^{\mathbf{X}}(u_n^\tau) \;&=\; \frac{\ln D_{F_2}^{n-1}\big(F(u_n^\tau), F(u_n^\tau)\big) - \ln F^{n-2}(u_n^\tau)}{\ln F(u_n^\tau)} \\ &=\; \frac{\ln F^n(u_n^\tau) - \ln\Big(\frac{1+F(u_n^\tau)}{2}\Big)^{n-1}}{\ln F(u_n^\tau)}\end{aligned}$$

and, for each $k \geq 1$,

$$\lim_{n\to+\infty}\Big|\varepsilon_{nk}^{\mathbf{X}}(u_{nk}^\tau) - k\,\varepsilon_n^{\mathbf{X}}(u_{nk}^\tau)\Big| \;=\; \lim_{n\to+\infty}\frac{(k-1)\,\ln\Big(\frac{1+F(u_{nk}^\tau)}{2}\Big)}{\ln F(u_{nk}^\tau)} \;=\; \frac{k-1}{2}\;.$$

Therefore, for each $\tau > 0$, $\{\varepsilon_n^{\mathbf{X}}(.)\}_{n\geq 1}$ is stable over $\{u_n^\tau\}_{n\geq 1}$ and

$$\theta \;=\; \lim_{n\to+\infty}\frac{\ln F^n(u_n^\tau) - \ln\Big(\frac{1+F(u_n^\tau)}{2}\Big)^{n-1}}{\ln F^n(u_n^\tau)} \;=\; \frac{1}{2}\;.$$

The following result is a contribution to compute θ for the special cases where the dependence struture of $\mathbf{X}$ is given. Smith (1992) and Perfekt (1994), among others, present a technique for calculating the extremal index of Markov chains under the assumption that a multivariate extreme limit distribution exists for the joint distribution of sucessive variables and suitable conditions on the transition probabilities.

We also assume here that the joint distribution of $d+1$ consecutive variables is in the domain of attraction of some multivariate extreme value distribution H_{d+1} and prove that this is sufficient for the stability condition to hold and compute θ from H_{d+1}.

Proposition 3.1. *Let $\mathbf{X}$ be a d^{th} order stationary Markov chain with the joint distribution F_{d+1} of $d+1$ sucessive variables in the domain of attraction of a $(d+1)$-multivariate extreme value distribution H_{d+1}. Then:*

(i) $\{\varepsilon_n^{\mathbf{X}}(.)\}_{n\geq 1}$ *is stable over* $\{u_n^\tau\}_{n\geq 1}$, *for each* $\tau > 0$;

(ii) $\mathbf{X}$ *has extremal index* $\theta = -\ln D_{H_{d+1}}(e^{-1}, ..., e^{-1}) + \ln D_{H_d}(e^{-1}, ..., e^{-1})$.

Proof: We first note that if F_{d+1} is in the domain of attraction of an extreme value distribution then the same holds for the common distribution F of variables in $\mathbf{X}$ and for each $\tau>0$ there exists $\{u_n^\tau\}_{n\geq 1}$ satisfying $n(1-F(u_n^\tau)) \underset{n\to\infty}{\longrightarrow} \tau > 0$.

It follows from the Markov property (Joe, 1997) that

$$\lim_{n\to+\infty}\left|\varepsilon_{nk}^{\mathbf{X}}(u_{nk}^\tau) - k\,\varepsilon_n^{\mathbf{X}}(u_{nk}^\tau)\right| =$$

$$= \lim_{n\to+\infty}\left|\frac{\ln D_{F_{d+1}}^{(k-1)d}\big(F(u_{nk}^\tau),...,F(u_{nk}^\tau)\big) - \ln D_{F_d}^{(k-1)(d+1)}\big(F(u_{nk}^\tau),...,F(u_{nk}^\tau)\big)}{\ln F(u_{nk}^\tau)}\right|$$

$$= \lim_{n\to+\infty}(k-1)\left|\frac{d\ln D_{F_{d+1}}^{nk}\big(F(u_{nk}^\tau),...,F(u_{nk}^\tau)\big) - (d+1)\ln D_{F_d}^{nk}\big(F(u_{nk}^\tau),...,F(u_{nk}^\tau)\big)}{-\tau}\right| .$$

Since

$$D_{F_{d+1}}^{nk}\big(F(u_{nk}^\tau),...,F(u_{nk}^\tau)\big) = D_{F_{d+1}^{nk}}\big(F^{nk}(u_{nk}^\tau),...,F^{nk}(u_{nk}^\tau)\big)$$

converges to $D_{H_{d+1}}(e^{-\tau},...,e^{-\tau})$, we find

$$\begin{aligned}\varepsilon_k &= (k-1)\left|\frac{d\,\ln D_{H_{d+1}}(e^{-\tau},...,e^{-\tau}) - (d+1)\,\ln D_{H_d}(e^{-\tau},...,e^{-\tau})}{-\tau}\right| \\ &= (k-1)\big(-d\,\ln D_{H_{d+1}}(e^{-1},...,e^{-1})\big) + (d+1)\,\ln D_{H_d}(e^{-1},...,e^{-1}) .\end{aligned}$$

Then, by applying the corollary 2.1, we get

$$\begin{aligned}\theta &= \lim_{n\to+\infty}\frac{\ln D_{F_{d+1}}^{n-d}\big(F(u_n^\tau),...,F(u_n^\tau)\big) - \ln D_{F_d}^{n-d-1}\big(F(u_n^\tau),...,F(u_n^\tau)\big)}{\ln F^n(u_n^\tau)} \\ &= \lim_{n\to+\infty}\frac{\ln D_{F_{d+1}^n}\big(F^n(u_n^\tau),...,F^n(u_n^\tau)\big) - \ln D_{F_d^n}\big(F^n(u_n^\tau),...,F^n(u_n^\tau)\big)}{-\tau} \\ &= -\ln D_{H_{d+1}}(e^{-1},...,e^{-1}) + \ln D_{H_d}(e^{-1},...,e^{-1}) .\end{aligned}$$

□

One can easily construct examples to illustrate the result. We note instead that F_2 in the previous example defined by

$$\begin{aligned}F_2(x,y) &= D_{F_2}\big(F(x),F(y)\big) \\ &= F(x)+F(y)-1+\Big(\big(1-F(x)\big)^{-1}+\big(1-F(y)\big)^{-1}-1\Big)^{-1},\end{aligned}$$

$x,y\in\mathbb{R}$, is in the domain of attraction of

$$\begin{aligned}H_2(x,y) &= D_{H_2}\big(G(x),G(y)\big) \\ &= G(x)\,G(y)\,\exp\Big(\big(-\ln G(x)\big)^{-1}+\big(-\ln G(y)\big)^{-1}\Big)^{-1} ,\end{aligned}$$

where $G(x)=H_2(x,+\infty)=H_2(+\infty,x)$ (Joe (1997)).

Therefore, we can apply directly (ii) above and find $\theta=-\ln D_{H_2}(e^{-1},e^{-1})-1 = \frac{3}{2}-1=\frac{1}{2}$.

ACKNOWLEDGMENTS

I am grateful to the referee for his corrections and rigorous report.

REFERENCES

[1] Buishand, T.A. (1984). Bivariate extreme-value data and the station-year method, *J. Hydrol.*, **69**, 77–95.

[2] Deheuvels, P. (1978). Caractérisation complète des lois extrèmes multivariées et de la convergence aux types extrèmes, *Publ. Inst. Statist. Univ. Paris*, **23**.

[3] Hsing, T. (1989). Extreme value theory for multivariate stationary sequences, *J. Multivariate Anal.*, **29**, 274–291.

[4] Joe, H. (1997). *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.

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

[6] Leadbetter, M.R. (1974). On extreme values in stationary sequences, *Z. Wahrsch. Verw. Gebiete*, **28**, 289–303.

[7] Leadbetter, M.R.; Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, Berlin.

[8] O'Brien, G.L. (1987). Extreme values for stationary and Markov sequences, *Ann. Probab.*, **15**, 281–291.

[9] Perfekt, R. (1994). Extremal behaviour of stationary Markov chains with applications, *Ann. Appl. Probab.*, **4**, 529–548.

[10] Rootzén, H. (1988). Maxima and exceedances of stationary Markov chains, *Adv. in Appl. Probab.*, **20**, 371–390.

[11] Smith, R.L. (1992). The extremal index for a Markov chain, *J. Appl. Probab.*, **29**, 37–45.

[12] Tiago de Oliveira, J. (1962/63). Structure theory of bivariate extremes, extensions, *Est. Mat., Estat. e Econ.*, **7**, 165–195.