Limiting distribution of the maximum of stationary sequences of stochastic processes

Marta Ferreira

CMAT, University of Minho, Braga, Portugal

Luísa Canto e Castro

CEAUL, Faculty of Sciences, University of Lisbon, Lisbon, Portugal

Abstract

In Extreme Value Theory (EVT) the study of the limit of the normalized maximum of a sequence of stochastic processes has been developed under the independent and identically distributed (i.i.d.) assumption (de Haan and Lin [L. de Haan, T. Lin, On convergence toward an extreme value distribution in $C[0,1]$, Ann. Probab. 29 (2001) 467-483]). Here we drop the independence and present conditions under which the normalized maximum of a stationary sequence still converges to a non-degenerate limit. We will illustrate with some examples.

Key words: Extreme Value Theory in $D[0,1]^d$, max-autoregressive processes
2010 MSC: 60G70, 60J05

1. Introduction

In many processes of science from hydrology, geophysics to finance, often one is confronted with making inference on maxima or exceedances over a high threshold. The first results were developed considering independent random variables but soon models for extreme values have been constructed under the more realistic assumption of temporal dependence. Most often, the processes under study show short term dependence which can conveniently

Email addresses: msferreira@math.uminho.pt (Marta Ferreira), ldloura@fc.ul.pt (Luísa Canto e Castro)

be modeled by a markovian structure. Among these are the autoregressive processes which have been widely studied, both the linear-type ARMA (e.g. Davis and Resnick [3], Leadbetter *et al.* [12]) and the maximal-type MARMA (Davis and Resnick [4]).

Heavy tailed MARMA and ARMA can be good choices for modeling time series data with sudden large peaks, although the former are more convenient for analysis as their finite-dimensional distributions can easily be written explicitly. Actually, MARMA processes and their generalizations have been applied to various phenomena, e.g., a solar thermal energy storage system (Daley and Haslett [2]), the water density in a sill fjord (Helland and Nielsen [10]) or financial series (Zhang and Smith [13]). The more recent power max-autoregressive processes, pARMAX and pRARMAX, which involve a power parameter (Ferreira and Canto e Castro [6], [7]), have proved to be good options for the modeling of financial series (Ferreira and Canto e Castro [7]).

Until now, the whole context of the mentioned works is confined to random variables and random vectors, which we call the classical EVT. The recent work developed by de Haan and Lin [9], Hult and Lindskog [11] and Davis and Mikosch [5] extends the classical EVT to random processes on $[0,1]^d$ with cdlg (right continuous and left-hand limits) sample paths, i.e., belonging to $D[0,1]^d$, the space of cdlg functions $x:[0,1]^d\to\mathbb{R}$ equipped with a metric d_0 which is equivalent to the J_1-*Skorokhod metric* and makes $D[0,1]^d$ complete and separable (the first two cited authors restrict to the case $d=1$ but most of their results are easily extended to $d>1$). There are many applications of this new framework in real life problems. For example, the process $X_i(s)$, $i=1,2,...$ could represent the time series of annual maxima of ozone levels at location s and the problem of interest might be determining the probability that the maximum level over the entire region $[0,1]^2$ does not exceed a certain standard level $f\in D[0,1]^2$ (Davis and Mikosch [5]). Another example presented in de Haan and Lin [9] relates to the probability that the water level $X_i(s)$ on day i at location $s\in[0,1]$ along the Dutch coast will not breach the dykes whose height at any location s is a function $f(s)$, i.e., the probability of interest is

$$P\Big(\bigvee_{i=1}^{n} X_i(s)\le f(s),\ \forall s\in[0,1]\Big). \tag{1}$$

This kind of problem can be solved by establishing a limit theory for the pointwise maximum of i.i.d. random functions (de Haan and Lin [9]). Since bursty phenomena are more appropriately model by heavy-tailed distributions than Gaussian ones, the theory of regular variation provides a convenient and unified background for studying extremes. The papers by de Haan and Lin [9] and Hult and Lindskog [11] state a precise formulation of regular variation of random functions in $C[0,1]$ and $D[0,1]$ and Davis and Mikosch [5] in $D[0,1]^d$. Davis and Mikosch [5] is also a pioneer work in what concerns EVT for random processes when we drop the "independence" assumption. More precisely, based on a point processes approach, they provide approximations for probabilities of some extremal events of the space-time linear process given by,

$$X_t(s) = \sum_{i=0}^{\infty} \psi_i(s) Z_{t-i}(s),\ s \in [0,1]^d \tag{2}$$

where $\{Z_i\}_{i\in\mathbb{Z}}$ is an i.i.d. sequence of random processes with values in $D[0,1]^d$. The coefficients ψ_i are deterministic real-valued processes on $D[0,1]^d$ and the indices s and t refer to the observation of the process at location s at time t. However, they did not established the convergence in distribution of the normalized maximum of $X_1(s), X_2(s), ..., X_n(s)$ and, as far as we know, it has not been established to any non-independent sequence of $D[0,1]^d$-valued random processes. In this work we state the convergence in distribution of the normalized maximum of stationary sequences under some specific conditions (Proposition 2.3). We will show that max-autoregressive processes in general satisfy these conditions. Yet are not the only, as we shall see on the last example.

2. Main Results

Write, for a $x \in D[0,1]^d$, $\delta > 0$ and a set $A \subset [0,1]^d$,

$$w''(x,\delta,A) = \sup_{s_1\le s\le s_2;|s_1-s_2|<\delta; s\in A} \min\{|x(s)-x(s_1)|, |x(s)-x(s_2)|\};\ w''(x,\delta) \text{ if } A=[0,1]^d$$

$$w(x,A) = \sup_{s_1,s_2\in A} |x(s_1)-x(s_2)|,$$

$$|x|_\infty = \sup_{s\in[0,1]^d} |x(s)|.$$

A sequence $\{\zeta_n\}$ of random elements with values in $D[0,1]^d$ is said to be *tight* if and only if the following conditions hold (cf. Billingsley [1]):

(C1) $\forall \epsilon > 0$, $\exists \alpha > 0$, such that, $P(|\zeta_n|_\infty > \alpha) \leq \epsilon$, for all $n \in \mathbb{N}$;

(C2) $\forall \epsilon > 0$ and $\forall \eta > 0$, $\exists 0 < \delta < 1$ and n_0 integer, such that,

(a) $P(w''(\zeta_n, \delta) \geq \eta) \leq \epsilon, \; n \geq n_0$

(b) $P(w(\zeta_n, [0,1]^d \setminus [\delta, 1-\delta]^d) \geq \eta) \leq \epsilon, \; n \geq n_0$

In order to state our main result, we prove the following lemmas.

Lemma 2.1. *If $\{Y_n\}$ and $\{W_n\}$ are tight sequences of random elements of $D[0,1]^d$, and*

$$M_n = Y_n \vee W_n, \tag{3}$$

where operator "$\vee$" means "component-wise maximum", then $\{M_n\}$ is tight.

Proof. We must prove that conditions (C1) and (C2) hold for $\{M_n\}$. In respect to (C1), we have that,

$$P(|M_n|_\infty > \alpha) \;\; \leq \;\; P(|Y_n|_\infty > \alpha) + P(|W_n|_\infty > \alpha) \leq \epsilon,$$

if we take α large enough, as (C1) holds for $\{Y_n\}$ and $\{W_n\}$.
Regarding condition (C2), we verify (a) since (b) is analogous. We have that,

$$w''(M_n, \delta) \;\; \leq \;\; w''(M_n, \delta, A_1) + w''(M_n, \delta, A_1^C)$$

where $A_1 = \{s \in [0,1]^d : Y_n(s) > W(s)\}$. Observe that, $|M_n(s) - M_n(s_i)|$ can be upper bounded either by $|M_n(s) - W_n(s_i)|$ or $|M_n(s) - Y_n(s_i)|$, $i = 1, 2$, and hence,

$$w''(M_n, \delta, A_1) + w''(M_n, \delta, A_1^C) \leq w''(Y_n, \delta, A_1) + w''(W_n, \delta, A_1^C).$$

Therefore, we can state,

$$\begin{aligned} P(w''(M_n, \delta) \geq \eta) \;\; &\leq \;\; P(w''(Y_n, \delta) + w''(W_n, \delta) \geq \eta) \\ &\leq \;\; P(w''(Y_n, \delta) \geq \eta/2) + P(w''(W_n, \delta) \geq \eta/2) \leq \epsilon, \end{aligned}$$

for small enough δ and large n, as condition (C2) holds for $\{Y_n\}$ and $\{W_n\}$. $\square$

Lemma 2.2. *If $\{Y_n\}$ and $\{a_n^{-1} W_n\}$ are tight sequences of random elements of $D[0,1]^d$, where (a_n) is a real sequence satisfying $0 < a_n \to \infty$, and*

$$S_n = Y_n + W_n, \tag{4}$$

then $\{a_n^{-1} S_n\}$ is tight.

Proof. The proof follows by an adaptation of the proof in Lemma 5.1 in Davis and Mikosch [5].

In respect to (C1), just observe that,

$$\begin{aligned} P(|a_n^{-1} S_n|_\infty > \alpha) &\le P(a_n^{-1}|Y_n + W_n|_\infty > \alpha) \\ &\le P(|a_n^{-1} Y_n|_\infty > \alpha/2) + P(|a_n^{-1} W_n|_\infty > \alpha/2) \le \epsilon, \end{aligned}$$

if we choose α large enough, since (C1) holds for $\{Y_n\}$ and $\{a_n^{-1} W_n\}$. Regarding condition (C2), we verify (a) since (b) is analogous. We consider the following decomposition for any $\lambda > 0$:

$$\begin{aligned} P(w''(S_n,\delta) \ge a_n \eta) &= P(w''(S_n,\delta) \ge a_n\eta, |Y_n|_\infty > a_n\lambda) \\ &\quad + P(w''(S_n,\delta) \ge a_n\eta, |Y_n|_\infty \le a_n\lambda) \\ &= I_1 + I_2 \end{aligned}$$

Note that, $I_1 \le P(|Y_n|_\infty > a_n\lambda) \le \epsilon/2$, considering large enough n and applying (C1). For I_2 we have,

$$w''(S_n,\delta) \le w''(W_n,\delta) + 2|Y_n|_\infty \le w''(W_n,\delta) + 2a_n\lambda$$

and hence, for $\eta > 2\lambda$, sufficiently small δ and large n,

$$I_2 \le P(w''(W_n,\delta) > a_n(\eta - 2\lambda)) \le \epsilon/2. \ \square$$

Proposition 2.3. *Sequences $\{M_n\}$ and $\{S_n\}$ under conditions of Lemma 2.1 and Lemma 2.2, respectively, whose finite-dimensional distributions converge to the respective of some $D[0,1]^d$-valued random process, converge in distribution in $D[0,1]^d$.*

2.1. Examples

We will see that the normalized maxima of any stationary MARMA random process with values in $D[0,1]^d$ and regularly-varying innovations, converges to a non-degenerate process. For simplicity we consider the MAR(1) case, usually called ARMAX, but the procedure and the result holds exactly for general MARMA.

We define the $D[0,1]^d$-valued ARMAX random process under similar assumptions considered for ARMAX in a random variables setting, the same procedure that is adopted by Davis and Mikosch [5] with linear processes (2). Consider $\{Z_i\}_{i\in\mathbb{Z}}$ an i.i.d. sequence of copies of a non-negative $D[0,1]^d$-valued random process Z. The ARMAX random process $\{X_i\}_{i\in\mathbb{Z}}$ satisfies the following recursion equations:

$$X_i(s) = c(s)X_{i-1}(s) \vee Z_i(s), i = 0, \pm 1, \pm 2, ..., \tag{5}$$

with $0 < c(s) < 1$ deterministic on $D[0,1]^d$ and Z_i independent of X_j for $j < i$. We assume that, for some sequence of normalizing constants $0 < a_n \to \infty$,

$$a_n^{-1} \bigvee_{i=1}^{n} Z_i \xrightarrow{d} \xi \tag{6}$$

where ξ is some non-degenerate $D[0,1]^d$-valued random process. Observe that, this is equivalent to assume that Z is an α-regularly varying process for some $\alpha > 0$, if a $C[0,1]^d$-valued version of ξ exists (Lemma 2.2 in Davis and Mikosch [5]). We are going to prove that we also have

$$a_n^{-1} \bigvee_{i=1}^{n} X_i \xrightarrow{d} \xi. \tag{7}$$

We start by noting that, using the recursion (5) and the independence assumptions, we have for $s \in [0,1]^d$ and $y > 0$,

$$P(\textstyle\bigvee_{i=1}^n X_i(s) \leq y) = P(X_1(s) \leq y) P(\bigvee_{i=2}^n Z_i(s) \leq y) \tag{8}$$

and hence, $\bigvee_{i=1}^n X_i \overset{d}{=} X_1 \vee \bigvee_{i=1}^{n-1} Z_{i+1}$. Therefore in order to prove (7) we can show that $a_n^{-1} X_1 \vee \bigvee_{i=1}^{n-1} Z_{i+1} \xrightarrow{d} \xi$.

With regard to the finite-dimensional convergence, for $s_i \in [0,1]^d$, $y_i > 0$, $i = 1, ..., k$, and using (6),

$$\begin{aligned} & \lim_{n\to\infty} P(X_1(s_j) \leq a_n y_j,\ j = 1, ..., k) P(\textstyle\bigvee_{i=1}^{n-1} Z_{i+1}(s_j) \leq a_n y_j,\ j = 1, ..., k) \\ = & \lim_{n\to\infty} P(\textstyle\bigvee_{i=1}^{n} Z_i(s_j) \leq a_n y_j,\ j = 1, ..., k) \\ = & P(\xi(s_1) \leq y_1, ..., \xi(s_k) \leq y_k). \end{aligned}$$

In what respects *tightness*, it suffices to note that $a_n^{-1} X_1 \vee \bigvee_{i=1}^{n-1} Z_{i+1}$ is a random element of type M_n in (3) under the conditions of Lemma 2.1 (take $Y_n = a_n^{-1} X_1$ and $W_n = a_n^{-1} \bigvee_{i=1}^{n-1} Z_{i+1}$): the convergence assumption (6) for innovations $\{Z_i\}_{i\in\mathbb{Z}}$ implies $a_n^{-1} \bigvee_{i=1}^{n-1} Z_{i+1}$ is *tight* and the *tightness* of $a_n^{-1} X_1$, with $0 < a_n \to \infty$, is straightforward from the *tightness* of the single element X_1.

Other examples for which (8) holds, and therefore all the same procedure, are listed bellow:

- The random coefficient version of (5), i.e., $X_i(s) = U_i(s) X_{i-1}(s) \vee Z_i(s)$, where $\{U_i\}_{i\in\mathbb{Z}}$ is an i.i.d. sequence of copies of a random process U independent of Z, also known as RARMAX, which is widely applicable as a real-valued process (see for instance, Daley and Haslett [2] and Helland and Nielsen [10]).

- The power max-autoregressive processes pARMAX and pRARMAX (Ferreira and Canto e Castro [6], [7]), $X_i(s) = U_i(s) X_{i-1}(s)^{c(s)} \vee Z_i(s)$, with $0 < c(s) < 1$ deterministic on $D[0,1]^d$ and $\{Z_i\}_{i\in\mathbb{Z}}$ and $\{U_i\}_{i\in\mathbb{Z}}$ under the same assumptions of the previous example (when $U_i(s) = 1$ we have pARMAX). As real-valued processes, these are good alternatives to heavy-tailed ARMA usually used on the modeling of asymptotic tail independent phenomena.

The next example is an autoregressive process whose maximum can be represented as a random element S_n given in Lemma 2.2.

Consider a sequence $\{Z_i\}_{i\in\mathbb{Z}}$ of i.i.d. copies of a random process Z satisfying (6) and let V be some random process, all with values in $D[0,1]^d$, such that, $X_i = Z_i + V$ (Ferreira and de Haan [8]). It is straightforward that $\bigvee_{i=1}^n X_i \stackrel{d}{=} V + \bigvee_{i=1}^n Z_i$ and hence, for $s_i \in [0,1]^d$, $y_i > 0$, $i = 1, ..., k$, and applying (6), the finite-dimensional convergence of normalized maxima is given by,

$$
\begin{aligned}
& \lim_{n\to\infty} P(a_n^{-1}V(s_j) + a_n^{-1}\textstyle\bigvee_{i=1}^n Z_i(s_j) \leq y_j,\ j = 1, , ..., k) \\
= \ & \lim_{n\to\infty} P(a_n^{-1}\textstyle\bigvee_{i=1}^n Z_i(s_j) \leq y_j,\ j = 1, ..., k) \\
= \ & P(\xi(s_1) \leq y_1, ..., \xi(s_k) \leq y_k).
\end{aligned}
$$

The *tightness* of the normalized maxima is straightforward. Observe that $a_n^{-1}V + a_n^{-1}\bigvee_{i=1}^n Z_i$ is as S_n in (4) under the conditions of Lemma 2.2 (just take $Y_n = X_1$ and $W_n = \bigvee_{i=1}^{n-1} Z_{i+1}$), since the single element V is tight, as well is $a_n^{-1}\bigvee_{i=1}^n Z_{i+1}$ by the convergence assumption (6).

References

[1] P. Billingsley, Convergence of Probability Measures. Wiley, New York, 1968.

[2] D.J. Daley, J. Haslett, A thermal energy storage process with controlled input, Adv. Appl. Probab. 14 (1982) 257-271

[3] R.A. Davis, S. Resnick, Limit theory for moving averages of random variables with regularly varying tail probabilities, Ann. Probab. 13 (1985) 179-195.

[4] R.A. Davis, S. Resnick, Basic properties and prediction of max-ARMA processes, Adv. Appl. Probab. 21 (1989) 781-803.

[5] R.A. Davis, T. Mikosch, Extreme value theory for space-time processes with heavy-tailed distributions. Stochastic Process. Appl. 118(4) (2008) 560-584.

[6] M. Ferreira, L. Canto e Castro, Tail and dependence behavior of levels that persist for a fixed period of time, Extremes 11 (2008) 113-133.

[7] M. Ferreira, L. Canto e Castro, Modeling rare events through a pRARMAX process, Tech. Rep. of Center of Statistics and Applications of University Lisbon (CEAUL, 2009); http://www.ceaul.fc.ul.pt/notas.html?ano=2009.

[8] A. Ferreira, L. de Haan, Extreme Value Theory: An Introduction, Springer-Verlag, 2006.

[9] L. de Haan, T. Lin, On convergence toward an extreme value distribution in $C[0,1]$, Ann. Probab. 29 (2001) 467-483.

[10] I.S. Helland, T.S. Nilsen, On a general random exchange model, J. Appl. Probab. 13 (1976) 781-790.

[11] H. Hult, F. Lindskog, Extremal behavior of regularly varying stochastic processes, Stochastic Process. Appl. 115(2) (2005) 249-274.

[12] M.R. Leadbetter, G. Lindgren, H. Rootzén, Extremes and Related Properties of Random Sequences and Processes, Springer-Verlag, New-York, Heidelberg-Berlin, 1983.

[13] Z. Zhang, R.L. Smith, Modelling Financial Time Series Data as Moving Maxima Processes, Tech. Rep. Dept. Stat. (Univ. North Carolina, Chapel Hill, NC, 2001); http://www.stat.unc.edu/faculty/rs/ papers/RLS_Papers.html.