Asymptotic and pre-asymptotic tail behavior of a power max-autoregressive model

Marta Ferreira

Universidade do Minho, Dep. Matemática, CMAT

Luísa Canto e Castro

Faculdade de Ciências da Universidade de Lisboa, DEIO-CEAUL

Abstract: Max-autoregressive models have revealed very useful in the analysis of rare events based on time series data, mainly in areas like, hydrology, geophysics and finances. Here we present a power max-autoregressive (pARMAX) process, $\{X_i\}_{i\in\mathbb{Z}}$, defined in such a way that the asymptotic tail dependence coefficient of Ledford and Tawn, for series lag m apart (η_m), exhibits a power decay with m for larger values of c, the main parameter of the process, namely, $\eta_m = c^m$, $c \in (1/2, 1)$. We also look at the threshold-dependent form of the extremal index, which is an important functional when extending discussions of extreme values from independent and identically distributed (i.i.d.) sequences to stationary ones. We state an approach for this functional as well as its connection with coefficient η for the pARMAX process.

Keywords: Markov Chains and Extreme value theory and Dependence conditions and Auto-asymptotic-tail-dependence function and Tail index and Extremal index

1. Introduction

Extreme Value Theory (EVT) became widely used by many researchers in applied sciences when faced with modeling high values of certain phenomena. Ocean wave modeling, wind engineering, thermodynamics of earthquakes, risk assessment on financial markets are some examples. The principal result in EVT is the Extremal Types Theorem which states the possible limit distributions for the maximum of an i.i.d. sequence of r.v.'s, $\widehat{Y}_1$, $\widehat{Y}_2$, ..., $\widehat{Y}_n$. More precisely, let $\widehat{F}$ be the common distribution function (d.f.) of $\{\widehat{Y}_i\}_i$ and denote $\max(\widehat{Y}_1, ..., \widehat{Y}_n) = \bigvee_{i=1}^n \widehat{Y}_i$. We say that $\widehat{F}$ belongs to the max-domain of attraction of the non degenerate extreme value distribution G_γ, (in short $\widehat{F} \in \mathcal{D}(G_\gamma)$), if there exist real sequences, $\{a_n > 0\}$ and $\{b_n\}$, such that,

$$\lim_{n\to\infty} P(\bigvee_{i=1}^n \widehat{Y}_i \le a_n x + b_n) = G_\gamma(x)\,.$$

Function $G_\gamma(x)$, also called, Generalized Extreme Value distribution (GEV), has standard representation given by:

$$G_\gamma(x) = \exp\big\{ -(1+\gamma x)^{-1/\gamma}\big\}\,,\ 1+\gamma x > 0\,.$$

where $G_0(x) = \exp\{-\exp(-x)\}$. The GEV function resumes all the possible limit distributions, i.e., all the max-domains of attraction: Weibull ($\gamma < 0$), Gumbel ($\gamma = 0$) and Fréchet ($\gamma > 0$).

Recently, models for extreme values have been constructed under the more realistic assumption of temporal dependence. Among these models, stationary Markov chains are very interesting, specially because they may have a somewhat simple treatment in what concerns extremal properties. The max-autoregressive moving average processes or MARMA (Davis and Resnick 1989), and also the particular case MARMA(1,0) or ARMAX (Alpuim 1989a, 1989b; Canto e Castro 1994) are some examples. Moreover, they also have similar paths to the ones of heavy tailed ARMA, typically used in modeling stationary data which exhibit sudden large peaks, like telephone signals and stock market prices. Therefore, a satisfactory fit can be achieved with a MARMA model for which deriving extremal features is easier than for heavy tailed ARMA. Here we introduce the power max-autoregressive processs (in short, pARMAX), which can also be an alternative model to fit such data, especially for large values. The pARMAX process satisfies the recursions,

$$X_i = X_{i-1}^c \vee Z_i \ , \ 0 < c < 1, \ i \in \mathbb{Z},$$

where $\{Z_i\}_{i\in\mathbb{Z}}$ is i.i.d. and X_i is independent from Z_j, for all integer $i < j$. We can see in Figure 1 that the corresponding peaks in the pARMAX plots (on the left) and in the AR(1) plots (on the right) are practically the same.

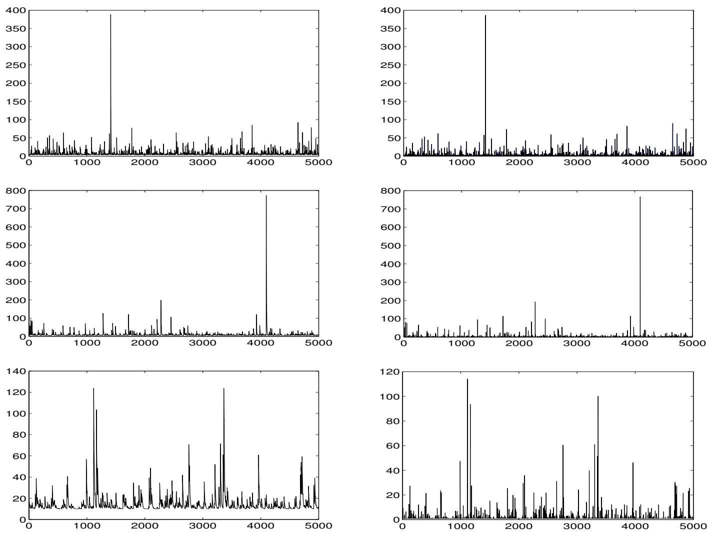

Figure 1. 5000 realizations of pARMAX, $X_i = X_{i-1}^c \ \vee \ Z_i$, on the left, and of AR(1), $X_i = c\,X_{i-1} + Z_i$, on the right, with, from top to bottom, $c = 0.7, 0.8, 0.9$, respectively, and with marginals Pareto(0.7)

We shall prove that pARMAX also possesses easily derived extremal features.

The main differences between pARMAX models and MARMA models respects the bivariate tail dependence of series lag m apart and the extremal index. In MARMA models asymptotic dependence of the tails is observed, whilst pARMAX presents asymptotic independence. More precisely, the coefficient of asymptotic tail dependence of Ledford and Tawn (1996, 1997), denoted by η_m, for MARMA series lag m apart is unitary for all positive integer m, while for pARMAX processes there is a straight relation between the power parameter c and the coefficient of asymptotic tail dependence, when considering series lag m apart (reference to η means η_1). In what respects the extremal index, it is known that for MARMA processes having margins in the Fréchet domain of attraction, it varies with the tail index and with the other parameters of the process, being unitary in the other cases (Canto e Castro, 1992). In the pARMAX process it is unitary if the margins are in the Fréchet domain of attraction, as we shall see.

Although the extremal index is unit, pARMAX models fit well in the formulation proposed by Bortot and Tawn (1998), so it makes sense to consider a threshold-dependent extremal index. Recall that, in order to get an intuitive insight about the relationship between the dependence structure of a process and its extremal behavior, we can look at the extremal index, θ ($\theta \in [0,1]$). In particular, θ measures the tendency of extremes to occur in clusters. A cluster of high levels exceedances is defined to be a set of observations that exceed a threshold u_n within a block of length r_n ($r_n = o(n)$), given that there is at least one exceedance in that block. The extremal index of a general stationary sequence $\{Y_i\}_i$, under a suitable mixing condition, can be obtained like follows:

$$\theta = \lim_{n\to\infty} P(Y_i \leq u_n, 2 \leq i \leq r_n | Y_1 > u_n)\,, \tag{1}$$

where $(u_n)_{n\geq 1}$ is a sequence of normalized levels, i.e., $nP(Y_1 > u_n) = O(1)$ (O'Brien 1987). ($\theta = 0$ corresponds to "pathological" cases which we do not intend to consider here). This parameter plays a very important role when estimating extremal properties of a weakly mixing stationary series with margins in the max-domain of attraction of an extreme value distribution. More precisely, considering that $\{Y_i\}_i$ is weakly mixing, for certain levels $\{u_n\}_{n\geq 1}$, we have,

$$\lim_{n\to\infty} P(\textstyle\bigvee_{i=1}^n Y_i \leq u_n) = \lim_{n\to\infty} \left(P(Y_1 \leq u_n)\right)^{n\theta}. \tag{2}$$

Though processes with i.i.d. margins have $\theta = 1$, the converse is false and this can be evidenced through the process pARMAX, which has dependent margins but has unit extremal index, as already mentioned. In fact, there are processes that would be regarded as strongly dependent by other measures, but for which $\theta = 1$. The autoregressive Gaussian processes, for instance, are such an example. Also stationary processes verifying conditions $\mathrm{D}(u_n)$ and $\mathrm{D}'(u_n)$ have $\theta = 1$ (Leadbetter et al. 1983). An unit extremal index means that asymptotic extreme events occur singly. However, it can be possible to observe clustering of exceedances for levels of practical interest (Bortot and Tawn 1998). Based on O'Brien characterization, by taking a large finite n and considering an appropriate upcrossings restriction condition ($\Delta^{(2)}(u_n, r_n)$), the latter authors define an extremal index threshold-form, $\theta(u, r_{[u]})$, which can be used to replace θ in (2) and hence, get some improvement on estimations based on this result, like in high quantiles or return periods.

More recently, some careful attention has been given to the statistical modeling of the tail dependence between consecutive pairs from a stationary first-order Markov chain, since it is important to distinguish asymptotic dependence from asymptotic independence. We classify a Markov chain as asymptotically dependent or asymptotically independent, whenever $b > 0$ or $b = 0$, respectively, in the limit below:

$$\lim_{y\to y^*} P(Y_2 > y | Y_1 > y) = b\,, \tag{3}$$

where y^* is the right-endpoint of Y_1, i.e., $y^* = \sup\{y : P(Y_1 \leq y) < 1\}$ (Bortot and Tawn 1998). For asymptotically independent Markov chains, the degree of dependence between exceedances of y usually decreases as $y \to y^*$, which leads to an extremal feature increasingly resembling an i.i.d. sequence at high levels. In these cases, procedures assuming that the limiting behavior of the chain is exact above a fixed high threshold and the dependence structure between consecutive random variables (r.v.'s) above the threshold can be modeled through a bivariate extreme value distribution (Smith et al. 1997) are not suitable. This problem is overcome by setting the way how $P(Y_2 > y|Y_1 > y)$ converges to zero, as $y \to y^*$, which involves the coefficient of asymptotic tail dependence (Ledford and Tawn 1996, 1997). Based on this approach, a connection between the threshold-dependent extremal index and η is established (Bortot and Tawn 1998).

This paper starts with the settlement of sufficient conditions for the existence of a stationary distribution for the pARMAX process $\{X_i\}_{i\in\mathbb{Z}}$. Then we analyze the marginal max-domain of attraction, the local dependence structure, in particular, we shall prove that $\{X_i\}_{i\in\mathbb{Z}}$ is *β-mixing* and verifies condition $\mathrm{D}''(u_n)$ of Leadbetter and Nandagopalan (1989) for some levels $(u_n)_{n\geq 1}$, which allows an easy way to compute the extremal index. Next we present the joint limiting d.f. of the normalized first passage time of threshold u, $T = \inf\{n \in \mathbb{N} : X_n \geq u\}$, and the corresponding excess, $R_T = X_T - u$, as $u \to \infty$. In Section 3, we compute function ATDF, analogous to the ACF in linear models (Ledford and Tawn 2003), which is based on the coefficient of asymptotic tail dependence, η_m, for the random pair, (X_1, X_{1+m}) and state the already mentioned relationship between η_m and the power parameter, c, of the pARMAX model. Therefore, we derive estimators for c and prove their consistency and asymptotic normality. Section 4 is devoted to the threshold-dependent extremal index, $\theta(u, r_{[u]})$, where we establish a connexion between this functional and η (Bortot and Tawn 1998) in the pARMAX model and derive a better estimate for the return levels.

2. Stationarity and extremal properties of pARMAX

Consider $\{Z_i\}_{i\in\mathbb{Z}}$ a sequence of i.i.d. copies of a r.v., Z, having real nonnegative support and common d.f. F_Z. A sequence $\{X_i\}_{i\in\mathbb{Z}}$ is said to be a pARMAX process if,

$$X_i = X_{i-1}^c \vee Z_i \ , \ 0 < c < 1, \ \ i = 0, \pm 1, \pm 2, ... \tag{4}$$

with X_i independent of Z_j, for all integer $i < j$. The sequence $\{Z_i\}_{i\in\mathbb{Z}}$ is also known as the innovations sequence of the process.

Iterating successively, we have,

$$X_n = X_{n-1}^c \vee Z_n = X_{n-2}^{c^2} \vee Z_{n-1}^c \vee Z_n = ... = X_{n-k}^{c^k} \vee \bigvee_{j=0}^{k-1} Z_{n-j}^{c^j} = \dots . \tag{5}$$

Observe that,

$$X_n = \bigvee_{j=0}^{\infty} Z_{n-j}^{c^j} \tag{6}$$

is a solution of the power max-autoregressive recursion in (4), since

$$X_n = X_{n-1}^c \vee Z_n = \left(\bigvee_{j=0}^{\infty} Z_{n-1-j}^{c^j} \right)^c \vee Z_n = \bigvee_{j=0}^{\infty} Z_{n-j}^{c^j} .$$

Now we will see some sufficient conditions under which the solution given by (6) is well defined, stationary and unique.

Proposition 2.1. *Let $\{Z_i\}_{i\in\mathbb{Z}}$ be an i.i.d. sequence of nonnegative r.v.'s with common d.f. F_Z. Then equations (4) have a stationary solution given by (6) if and only if,*

$$\sum_{j=0}^{\infty} - \log F_Z(x^{1/c^j}) < \infty\,, \ \ \textit{for some } x \geq 0. \tag{7}$$

Proof. We need to show that $\bigvee_{j=0}^{\infty} Z_j^{c^j}$ is almost surely (a.s.) finite if and only if (7) holds. By the Kolmogorov 0-1 law, $P(\bigvee_{j=0}^{\infty} Z_j^{c^j} < \infty) = 0$ or 1, being 1 if $P(\bigvee_{j=0}^{\infty} Z_j^{c^j} \leq x) > 0$ for some nonnegative x.

From the independence hypothesis,

$$P\Big(\bigvee_{j=0}^{\infty} Z_j^{c^j} \leq x \Big) = \prod_{j=0}^{\infty} F_Z(x^{1/c^j})\,, \tag{8}$$

and hence, we have that,

$$P\Big(\bigvee_{j=0}^{\infty} Z_j^{c^j} \leq x \Big) = \exp\Big\{ - \Big(- \log \prod_{j=0}^{\infty} F_Z\big(x^{1/c^j}\big) \Big) \Big\} = \exp\Big\{ - \sum_{j=0}^{\infty} \Big(- \log F_Z\big(x^{1/c^j}\big) \Big) \Big\},$$

which is positive if and only if (7) holds. □

Remark 2.2. *Using convenient representations of domains of attraction of Gumbel and Fréchet(γ), for some $\gamma > 0$, it is easy to prove that condition (7) is valid therein.*

Proposition 2.3. *If (7) holds and Z has support in $[a, \infty[$, $a \geq 1$, then the stationary solution for (4), stated in (6), is unique.*

Proof. In this proof we will follow the same reasoning used for MARMA processes (Davis and Resnick 1989). Suppose that $\{Y_i\}_{i\in\mathbb{Z}}$ is a stationary solution satisfying the power max-autoregressive recursion in (4). Iterating successively, we have,

$$Y_n = Y_{n-1}^c \vee Z_n = Y_{n-2}^{c^2} \vee Z_{n-1}^c \vee Z_n = \ldots = Y_{n-k}^{c^k} \vee \bigvee_{j=0}^{k-1} Z_{n-j}^{c^j} . \tag{9}$$

Since, as $k \to \infty$, $Y_{n-k}^{c^k} \to 1$ in probability and Z is a r.v. with support in $[a, \infty[$ for some $a \geq 1$, by letting $k \to \infty$ in (9), we obtain

$$Y_n = \bigvee_{j=0}^{\infty} Z_{n-j}^{c^j}$$

Hence, $\bigvee_{j=0}^{\infty} Z_{n-j}^{c^j}$ is the unique stationary solution of (4). □

Denoting by $K_i(\cdot)$ the d.f. of the marginal X_i for any integer i, the recurrence (4) and the independence assumptions lead to:

$$K_i(x) = P(X_n \leq x) = P(X_{i-1} \leq x^{1/c}, Z_i \leq x) = K_{i-1}(x^{1/c})F_Z(x)\,.$$

So, any d.f. that satisfies the stationarity equation,

$$K(x) = K(x^{1/c}) F_Z(x)\,, \tag{10}$$

is a stationary distribution of the pARMAX process $\{X_i\}_{i\in\mathbb{Z}}$. It follows an example.

Example 2.1 Consider $\{Z_i\}_{i\in\mathbb{Z}}$ with common d.f.,

$$F_Z(x) = \mathbf{1}_{\{x=c\}} + \frac{1-x^{-1/\gamma}}{1-x^{-1/(c\gamma)}}\mathbf{1}_{\{x>1\}}, \tag{11}$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function. Hence, $K(x) = \big(1-x^{-1/\gamma}\big)\mathbf{1}_{\{x\geq 1\}}$, satisfies (10), being, therefore, a stationary distribution for X_i.

Remark 2.4. *From now on, we are going to consider that $\{X_i\}_{i\in\mathbb{Z}}$ is a pARMAX process with innovations $\{Z_i\}_{i\in\mathbb{Z}}$ in the conditions of Proposition 2.3, more precisely, that $\{Z_i\}_{i\in\mathbb{Z}}$ has support in $[1,\infty[$. According to the results above, this means that the process $\{X_i\}_{i\in\mathbb{Z}}$ with representation (6) is the unique stationary solution. Therefore, applying (8), the marginal d.f. of $\{X_i\}_{i\in\mathbb{Z}}$, here denoted by $K(\cdot)$, is given by,*

$$K(x) = \lim_{n\to\infty} P(X_n \leq x) = \prod_{j=0}^{\infty} F_Z\big(x^{1/c^j}\big)\,,$$

which is non degenerate if F_Z is non degenerate.

In fact, the pARMAX process is strictly stationary since it has a markovian structure and stationary margins. The *m-step* transition probability function (t.p.f.) from x to $]-\infty, y]$, is given by:

$$Q^m(x,]-\infty, y]) := P\big(X_{n+m} \leq y | X_n = x\big) = \tfrac{K(y)}{K(y^{1/c^m})}\mathbf{1}_{\big\{x\leq y^{1/c^m}\big\}}, \tag{12}$$

where the last equality is due to (10).

Proposition 2.5. *Let $\{X_i\}_{i\in\mathbb{Z}}$ be a pARMAX process under the conditions of Remark 2.4 with non degenerate innovations $\{Z_i\}_{i\in\mathbb{Z}}$. Then, the following statements hold:*

(i) *The margins of $\{X_i\}_{i\in\mathbb{Z}}$ are in the same max-domain of attraction of the innovations Z, with the same tail index value.*

(ii) *$\{X_i\}_{i\in\mathbb{Z}}$ is β*-mixing.

(iii) *Condition $D''(u_n)$ of Leadbetter and Nandagopalan (1989) holds for $\{X_i\}_{i\in\mathbb{Z}}$, where $(u_n)_{n\geq 1}$ is a real sequence of normalized levels of K, i.e., $n(1-K(u_n)) \to \tau$, as $n\to\infty$, for some $\tau > 0$.*

(iv) *$\{X_i\}_{i\in\mathbb{Z}}$ has unit extremal index whenever Z is in the max-domain of attraction of a Fréchet(γ), for some $\gamma > 0$.*

Proof. (i) The proof is straightforward since, from (5), we have,

$$P\big(\textstyle\bigvee_{i=1}^n X_i \leq x\big) = P\big(X_1 \vee \bigvee_{i=2}^n Z_i \leq x\big). \tag{13}$$

So, if $F_Z \in \mathcal{D}\big(G_\gamma\big)$, then, by the independence assumptions,

$$\lim_{n\to\infty} P\big(\textstyle\bigvee_{i=1}^n X_i \leq a_n x + b_n\big) = \lim_{n\to\infty} K(a_n x + b_n) F_Z^{n-1}(a_n x + b_n) = G_\gamma(x),$$

for conveniently chosen norming constants, $a_n > 0$ and b_n.

(ii) We will show β-*mixing* condition, by proving that $\{X_i\}_{i\in\mathbb{Z}}$ is regenerative and aperiodic (Asmussen 1987).

In what concerns *regeneration*, we will show that $\{X_i\}_{i\in\mathbb{Z}}$ has a regeneration set, that is, a recurrent set R ($K(R)>0$), such that, for some $m>0$, $\epsilon\in(0,1)$ and a probability measure λ, we have,

$$Q^m(x,B)\geq \epsilon\lambda(B),\ x\in R, \tag{14}$$

for all $B\in\mathcal{B}(\mathbb{R})$, where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra in the set of real numbers.

A sufficient condition for *aperiodicity*, is that, besides condition (14), we must also have,

$$Q^{m+1}(x,B)\geq \epsilon^*\lambda(B)\,,\ \ x\in R\,, \tag{15}$$

with the same R and λ, and for some $\epsilon^*\in(0,1)$ (Asmussen 1987).

We are going to prove that $R=]1,r]\subset[1,\infty[$ is a regeneration set. As a subset of the support of X_i ($i\in\mathbb{Z}$) it is obviously recurrent. Let $x\in R$, B a borelian, and let $S=]r,r^{1/c}]$. Now observe that,

$$Q(x,B)=\textstyle\int_B dQ(x,z)\geq \int_{B\cap S} dQ(x,z)$$

and, $\forall x\in R$, $x^c\leq r^c<y$, for any $y\in S$, so that, by (12), the integral in the right hand side becomes

$$\textstyle\int_{B\cap S} dF_Z(z)=P(Z\in B\cap S)=\epsilon\lambda(B)\,.$$

Hence, condition (14) holds for $m=1$, with $\lambda(.)=P(Z\in\ .\cap S)/P(Z\in S)$ and $\epsilon=P(Z\in S)$.

In what respects the aperiodicity, we have that,

$$Q^2(x,B)=\int P(X_{n+2}\in B|X_{n+1}=z)dQ(x,z)\geq\int_S Q(z,B)dQ(x,z)\,.$$

As $S\subset[1,\infty[$, it is a recurrent set. One notices that S is also a regeneration set since, $\forall z\in S$ we have $z^c\leq r<y$ for any $y\in S$, and the same reasoning above leads to $Q(z,B)\geq\epsilon\lambda(B)$, for all borelian B. Therefore, we obtain,

$$Q^2(x,B)\geq\epsilon\lambda(B)\int_S dQ(x,z)=\epsilon\lambda(B)Q(x,S)=\epsilon\lambda(B)P(Z\in S)\,,$$

and so condition (15) holds with $\epsilon^*=\epsilon P(Z\in S)$.

(iii) We must prove that,

$$\lim_{n\to\infty} n\sum_{j=2}^{r_n-1}P\big(X_1>u_n,X_j\leq u_n<X_{j+1}\big)=0\,, \tag{16}$$

where $r_n=[n/k_n]$ is the block length used to define a cluster of high levels exceedances and $\{k_n\}_{n\geq1}$ ($k_n=o(n)$) is some slowly increasing sequence of integers corresponding to the number of blocks. If $j\geq2$, we have,

$$P\big(X_1>u_n,X_j\leq u_n<X_{j+1}\big)=\int_{u_n}^{\infty}\int_{-\infty}^{u_n}P(X_{j+1}>u_n|X_j=z)Q^{j-1}(y,dz)K(dy) \tag{17}$$

Applying (12), the last double integral becomes,

$$\begin{aligned}&\int_{u_n}^{\infty}\int_{-\infty}^{u_n}(1-F_Z(u_n))Q^{j-1}(y,dz)K(dy)\\ =\ &(1-F_Z(u_n))\int_{u_n}^{\infty}Q^{j-1}(y,]-\infty,u_n])K(dy)\\ \leq\ &(1-F_Z(u_n))(1-K(u_n))\,.\end{aligned} \tag{18}$$

Note that $1-F_Z(x)=1-K(x)/K(x^{1/c})\leq 1-K(x)$. Therefore, condition $\mathrm{D}''(u_n)$ in (16) holds, since by (17) and (18), we have, as $n\to\infty$,

$$n\sum_{j=2}^{r_n-1}P\big(X_1>u_n,X_j\leq u_n<X_{j+1}\big)\leq n\frac{n}{k_n}(1-F_Z(u_n))(1-K(u_n))=O(1/k_n).$$

(iv) The assumption is equivalent to state that $1-F_Z$ is regularly varying at infinity with index $-1/\gamma$, i.e.,

$$1-F_Z(x)=x^{-1/\gamma}L_Z(x)\,, \tag{19}$$

where $L_Z(\cdot)$ is a slowly varying function at ∞, and hence, by (i), we also have,

$$1-K(x)=x^{-1/\gamma}L_K(x) \tag{20}$$

for some slowly varying function $L_K(\cdot)$.

Since condition $\mathrm{D}''(u_n)$ holds for some normalized sequence of levels, $(u_n)_{n\geq 1}\equiv (u_n^{(\tau)})_{n\geq 1}$ $(\tau>0)$ (see (iv)) and $\mathrm{D}(u_n)$ also holds (in fact, by (ii), it holds $\forall (u_n)_{n\geq 1}$), if the limit

$$\lim_{n\to\infty}P\big(X_2\leq u_n^{(\tau)}|X_1>u_n^{(\tau)}\big)=\theta\ , \tag{21}$$

exists (i.e. θ is finite) for some of those τ, then it exists for all $\tau>0$ and the process has extremal index θ (Leadbetter and Nandagopalan 1989).

Thus being, we have that,

$$\theta=\lim_{n\to\infty}\frac{P\big(X_2\leq u_n,X_1>u_n\big)}{P\big(X_1>u_n\big)}=\lim_{n\to\infty}\frac{\int_{u_n}^{\infty}Q(y,]-\infty,u_n])K(dy)}{1-K(u_n)}$$

$$=\lim_{n\to\infty}\Big[1-\frac{1-K((u_n)^{1/c})}{1-K(u_n)}\Big],$$

where in the last step we applied the t.p.f. in (12) and $F_Z(u_n)\to 1$, as $n\to\infty$. The result follows from (20).

□

Now we will show that first passage times over high levels are also easily derived for pARMAX processes.

Let $T=T_u=\inf\{n\in\mathbb{N}:X_n\geq u\}$ be the first passage of X_n over threshold u and $R_T=X_T-u$ the respective overshoot. We prove that T and R_T, properly normalized, are asymptotically independent and we calculate their joint limiting distribution as the threshold $u\to\infty$. The convergence in distribution is denoted by "$\overset{d}{\to}$".

Theorem 2.6. *Let $\{X_i\}_{i\in\mathbb{Z}}$ be a pARMAX process under conditions of Proposition 2.5 (iv). Let E and V be independent r.v.'s with distribution functions $\big(1-e^{-x}\big)I_{(0,\infty)}$ and $\big(1-(y+1)^{-1/\gamma}\big)I_{(0,\infty)}$ respectively. Then, when $u\to\infty$,*

$$\big((1-F_Z(u))T,u^{-1}R_T\big)\overset{d}{\longrightarrow}(E,V)\,.$$

Proof. Set $d = 1 - F_Z(u)$. For positive s, v and ω, such that $s < v$,

$$\begin{aligned} P\big(s \le dT \le \omega, R_T > vu\big) &= \sum_j P(R_T > vu, T = j) \\ &= \sum_j P\big(X_1 \le u, ..., X_{j-1} \le u, X_j > (v+1)u\big) \\ &= \sum_j K(u) F_Z^{j-2}(u)(1 - F_Z((v+1)u)\,, \end{aligned}$$

where the summation is over $[s/d] \le j \le [\omega/d]$ ($[\cdot]$ denotes the integer part). By (19), as $u \to \infty$, we have that, $(1 - F_Z((v+1)u))/(1 - F_Z(u)) \sim (v+1)^{-1/\gamma}$ and hence,

$$\begin{aligned} &(1 - F_Z(u)) \textstyle\sum_j F_Z^{j-2}(u) = (1 - F_Z(u))\, \dfrac{1 - F_Z(u)^{[\frac{\omega}{d}] - [\frac{s}{d}] + 1}}{1 - F_Z(u)}\, F_Z(u)^{[\frac{s}{d}] - 2} \\ \sim\;& F_Z(u)^{\frac{s}{d} - 2} - F_Z(u)^{\frac{\omega}{d} - 1} = \big(1 - \tfrac{1}{1/d}\big)^{\frac{s}{d}} F_Z(u)^{-2} - \big(1 - \tfrac{1}{1/d}\big)^{\frac{\omega}{d}} F_Z(u)^{-1} \sim e^{-s} - e^{-\omega}\,. \end{aligned}$$

Since $K(u) \to 1$ as $u \to \infty$, the assertion follows. □

Remark 2.7. *Consider $\{X_i\}_{i\in\mathbb{Z}}$ a pARMAX process under the conditions of Proposition 2.5 (iv). A sequence of normalized levels $(u_n)_{n\ge 1}$ of the marginal d.f., K, is also a sequence of normalized levels of the innovations d.f., $F_Z(.)$. More precisely, by (10), we have,*

$$n(1 - F_Z(u_n)) = \frac{n(1 - K(u_n)) - n(1 - K(u_n^{1/c}))}{K(u_n^{1/c})}\,. \tag{22}$$

If $n(1 - K(u_n)) \to \tau > 0$, as $n \to \infty$, and hence u_n is such that $u_n \sim n^\gamma L_K(u_n)^\gamma/\tau^\gamma$, then,

$$1 - K(u_n^{1/c}) \sim 1 - K\big(K^{-1}(1 - \tau/n)^{1/c}\big) \sim (\tau/n)^{1/c} \mathcal{L}(a_{\tau/n}^{1/c})\,, \tag{23}$$

where $a_t = K^{-1}(1 - t)$ and

$$\mathcal{L}(a_t^{1/c}) = \big[L_{K^{-1}}(t)\big]^{-1/(\gamma c)} L_K\big(a_t^{1/c}\big) \sim \big[L_{K^{-1}}(t)\big]^{-1/(\gamma c)} L_K\big((t^{-\gamma} L_{K^{-1}}(t))^{1/c}\big) \tag{24}$$

is a slowly varying function as $t \downarrow 0$. Therefore, as $n \to \infty$, $n(1 - K(u_n^{1/c})) \sim 0$, and since $K(u_n^{1/c}) \sim 1$, re-taking expression (22), we have,

$$n(1 - F_Z(u_n)) \sim n(1 - K(u_n)) \sim \tau\,,\; n \to \infty\,, \tag{25}$$

3. The function ATDF and its estimation

The function ATDF (Λ_m) "provides a measure of serial dependence between extreme values lag m apart and may be interpreted in a manner that is broadly similar to the ACF" (Ledford and Tawn 2003). More precisely,

$$\Lambda_m = 2\eta_m - 1$$

where η_m is the parameter of tail dependence (Ledford and Tawn 1996, 1997; Draisma et al. 2004) for (X_i, X_{i+m}). Here we name Λ_m as *auto-asymptotic-tail-dependence function* (ATDF).

A special feature of pARMAX processes is the fact that, $\eta_m = \max(c^m, 1/2)$, so that, the ATDF has a power decay as the ACF of AR(1) processes and a cut-off as the MA(p) processes (see Figure 2). Actually, we have the following result.

Proposition 3.1. *Let $\{X_i\}_{i\in\mathbb{Z}}$ be a pARMAX process under conditions of Proposition 2.5 (iv). Then, the random pair (X_1, X_{1+m}) has a coefficient of asymptotic tail dependence $\eta_m = 1/2$ if $c^m \le 1/2$, and c^m if $c^m > 1/2$.*

Proof. Consider notation $a_{xt} = K^{-1}(1-xt)$. In order to obtain η_m, we need to show that,

$$h(x,y) = \lim_{t\downarrow 0} \frac{P\big(X_1 > a_{xt}, X_{1+m} > a_{yt}\big)}{P\big(X_1 > a_t, X_{1+m} > a_t\big)}. \tag{26}$$

is homogeneous (Draisma et al. 2004). Considering the t.p.f. in (12), we develop the expression in numerator:

$$\begin{aligned} P(X_1 > a_{xt}, X_{1+m} > a_{yt}) &= \int_{a_{xt}}^{\infty} 1 - Q^m(u,]-\infty, a_{yt}])K(du) \\ &= 1 - K(a_{xt}) - F_Z(a_{yt})\big[K(a_{yt}^{1/c^m}) - K(a_{yt})\big] \end{aligned} \tag{27}$$

Since, $K \in \mathcal{D}(G_\gamma)$, for some $\gamma > 0$, by (23) and (24), we have that,

$$K\big(a_{yt}^{1/c^m}\big) \sim 1 - (ty)^{1/c^m}\mathcal{L}\big(a_t^{1/c^m}\big).$$

and applying (10), then,

$$\begin{aligned} P(X_1 > a_{xt}, X_{1+m} > a_{yt}) &\sim t(x+y) - 1 + \frac{(1-tx)(1-ty)}{1-(ty)^{1/c^m}\mathcal{L}\big(a_t^{1/c^m}\big)} \\ &\sim \begin{cases} xyt^2 & \text{, if } c^m < 1/2 \\ (yt)^{1/c^m}\mathcal{L}\big(a_t^{1/c^m}\big) & \text{, if } c^m > 1/2\,. \end{cases} \end{aligned} \tag{28}$$

To the denominator in (26) just take $x = y = 1$ in (28), and hence, $h(x,y)$ becomes,

$$h(x,y) = \begin{cases} xy & \text{, if } c^m < 1/2 \\ y^{1/c^m} & \text{, if } c^m > 1/2\,. \end{cases}$$

This is an homogeneous function of order 2 in case $c^m < 1/2$, and of order $1/c^m$ if $c^m > 1/2$, which implies $\eta_m = 1/2$ and $\eta_m = c^m$, respectively. Note that, if $c^m = 1/2$, we have $h(x,y) = y^2$ or $h(x,y) = (xy + k^*y^2)/(1+k^*)$, whenever $\mathcal{L}\big(a_t^2\big) \to \infty$ or $\mathcal{L}\big(a_t^2\big) \to k^*$ ($k^* \ge 0$), respectively, as $t \downarrow 0$, and hence, $\eta_m = 1/2$. □

We conclude that, pARMAX processes are asymptotically independent Markov chains ($\eta < 1$) and, the bigger the value of parameter c, the bigger we must choose lag m in order to get asymptotically independent observations.

For the estimation of ATDF we can use known estimators of the coefficient of asymptotic tail dependence, η_m, (Ledford and Tawn 1996; Peng 1999; Draisma et al. 2004). However the properties of these latter ones were derived under the assumption of independence between random pairs, (X_i, Y_i), $i = 1, ..., n$, i.i.d. copies of (X, Y), while here we must apply to random pairs, (X_i, X_{i+m}), for some m fixed, which are obviously dependent.

Based on (26), parameter η_m can be estimated as the tail index of

$$T_i = \min\left(\frac{1}{1-K(X_i)}, \frac{1}{1-K(X_{i+m})}\right), \tag{29}$$

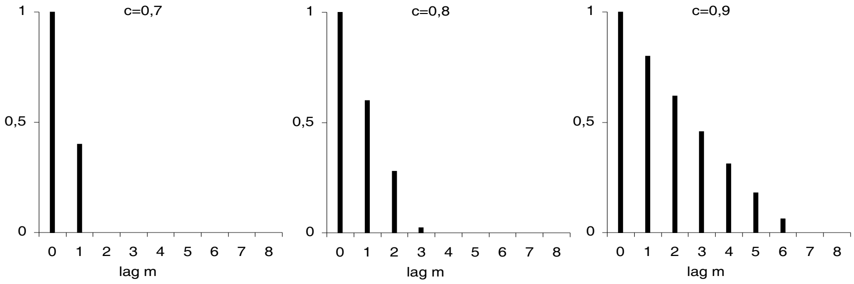

Figure 2. The ATDF (Λ_m) of pARMAX processes with parameters, $c = 0.7$, $c = 0.8$ and $c = 0.9$, respectively, for lags $m = 0, 1, ..., 8$.

in which, replacing the unknown marginal d.f., $K(x)$, by the empirical counterpart, $K_n(x) = (1/n)\sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq x\}}$, leads to,

$$T_i^{(n)} := \min\left(\frac{n+1}{n+1-nK_n(X_i)}, \frac{n+1}{n+1-nK_n(X_{i+m})}\right),\ i = 1, ..., n, \tag{30}$$

(Ledford and Tawn 1996; Draisma et al. 2004). We are going to show that the tail index estimators of the class of Drees, which includes, Hill, maximum likelihood, Pickands, moments and probability weighted moments (Drees 2003), are consistent and asymptotically normal for sequence $\{T_i^{(n)}\}_i$.

By now, observe that,

$$P\big(T_i^{(n)} > x\big) \quad \sim \quad P\Big(X_i > K_n^{-1}\big(1 - \tfrac{1}{x}\big), X_{i+m} > K_n^{-1}\big(1 - \tfrac{1}{x}\big)\Big), \text{ as } n \to \infty. \tag{31}$$

Note also that, the following convergence always hold, for any y in the support of X_i,

$$K_n(y) = \frac{1}{n}\sum_{j=1}^{n} 1_{\{X_j \leq y\}} \xrightarrow{P} K(y)\,, \tag{32}$$

since, $\sum_{j=1}^{n} 1_{\{X_j \leq y\}} \frown Binomial(n, p, \lambda)$, with $p = P(X_j \leq y) = K(y)$, $\lambda = P(X_j \leq y | X_{j-1} \leq y)$, such that, $E(\sum_{j=1}^{n} 1_{\{X_j \leq y\}}) = np$ and

$$Var(\textstyle\sum_{j=1}^{n} 1_{\{X_j \leq y\}}) = np(1-p) + \frac{2p(1-p)(\lambda-p)}{1-\lambda}\left\{(n-1) - \frac{\lambda-p}{1-\lambda}\left[1 - \left(\frac{\lambda-p}{1-p}\right)^n\right]\right\}$$

(Klotz 1973), and hence, by Chebychev inequality,

$$\begin{aligned} P\big(|K_n(y) - p| > \epsilon\big) \quad &\leq \quad \epsilon^{-2} Var\Big(\tfrac{1}{n}\textstyle\sum_{j=1}^{n} 1_{\{X_j \leq y^\alpha\}}\Big) \\ &= \quad \frac{p(1-p)}{\epsilon^2 n} + \frac{2p(1-p)(\lambda-p)}{\epsilon^2(1-\lambda)}\left\{\frac{n-1}{\epsilon^2 n^2} - \frac{\lambda-p}{\epsilon^2(1-\lambda)}\left[\frac{1}{n^2} - \frac{\left(\frac{\lambda-p}{1-p}\right)^n}{n^2}\right]\right\} \\ &\longrightarrow \quad 0, \text{ as } n \to \infty. \end{aligned}$$

Thus being, assuming that K is strictly increasing at $K^{-1}(q)$, we have, for $0 < q < 1$,

$$K_n^{-1}(q) \xrightarrow{P} K^{-1}(q). \tag{33}$$

On the theoretical result, usually attributed to Slutsky, in which, any given random elements, Y_n, Y and Y_n', such that $Y_n \xrightarrow{d} Y$ and $Y_n' \xrightarrow{P} a$, then, $(Y_n, Y_n') \xrightarrow{d} (Y, a)$, we have that, for each i, and by the theorem of continuous application,

$$\begin{aligned} &P\Big(X_i > K_n^{-1}\big(1-\tfrac{1}{x}\big), X_{i+m} > K_n^{-1}\big(1-\tfrac{1}{x}\big)\Big) \\ \underset{n\to\infty}{\longrightarrow}\ &P\Big(X_i > K^{-1}\big(1-\tfrac{1}{x}\big), X_{i+m} > K^{-1}\big(1-\tfrac{1}{x}\big)\Big) \end{aligned} \tag{34}$$

Therefore, given (31), we conclude that, $T_i^{(n)} \stackrel{d}{\longrightarrow} T_i$. Similarly is proof that,

$$P\big(T_i^{(n)} > x, T_{i+k}^{(n)} > y\big) \stackrel{d}{\longrightarrow} P(T_i > x, T_{i+k} > y). \tag{35}$$

Now we shall see that sequence $\{T_i\}_i$ (and therefore $\{T_i^{(n)}\}_i$ given (35)) is under Drees conditions (Drees 2003). More precisely, $\{T_i\}_i$ must have a β-*mixing* dependence structure (immediate from Proposition 2.5 (ii)), and there must exist $\epsilon > 0$ and functions h_m $\forall m \in \mathbb{N}$, such that,

$$\lim_{n\to\infty} \frac{n}{k_n} P\Big(T_1 > K^{-1}\Big(1-\frac{k_n}{n}x\Big), T_{1+m} > K^{-1}\Big(1-\frac{k_n}{n}y\Big)\Big) \to h_m(x,y), \tag{36}$$

and, considering the extreme interval $I_n(x,y) = \big]K^{-1}(1-yk_n/n), K^{-1}(1-xk_n/n)\big]$,

$$\tfrac{n}{k_n} P\big(T_1 \in I_n(x,y), T_{1+m} \in I_n(x,y)\big) \le (y-x)\Big(\tilde\rho(m) + D_1\Big(\tfrac{k_n}{n}\Big)^{\alpha}\Big), \tag{37}$$

with $0<\alpha<1$, $\forall m \in \mathbb{N}$, $0 < x, y \le 1+\epsilon$, $D_1 \ge 0$, where $\tilde\rho(m)$ is a sequence satisfying, $\sum_{m=1}^{\infty} \tilde\rho(m) < \infty$, and $\{k_n\}$ is a convenient positive integers sequence such that $k_n \to \infty$, $k_n/n \to 0$, as $n \to \infty$.

Observe that, by Proposition 3.1, $F_T \in \mathcal{D}(G_{\gamma^*})$, with $\gamma^* = \max(1/2, c^m)$ and, for the sake of simplicity, it is assumed that, for some real constants, $d > 0$ and $d^* > 0$,

$$F_T^{-1}(1-t) \sim d^* t^{-\gamma_*} \quad \text{and} \quad K^{-1}(1-t) \sim dt^{-\gamma}, \ \text{ as } t \downarrow 0. \tag{38}$$

Denoting the quantile, $K^{-1}\big(1 - 1/F_T^{-1}(1-\frac{k_n}{n}x)\big) = a_{n,x}$, we have,

$$\begin{aligned} &P\big(T_1 > F_T^{-1}(1-\tfrac{k_n}{n}x), T_{1+m} > F_T^{-1}(1-\tfrac{k_n}{n}y)\big) \\ &= P(X_1 > a_{n,x}, X_2 > a_{n,x}, X_{1+m} > a_{n,y}, X_{2+m} > a_{n,y}) \end{aligned} \tag{39}$$

Note that, by (38), $a_{n,x} \sim \big(\frac{k_n}{n}x\big)^{-\gamma^*\gamma} d(d^*)^{\gamma}$ and hence,

$$K\big(a_{n,x}^{1/c^j}\big) \sim 1 - \big(\tfrac{k_n}{n}x\big)^{\gamma^*} d^{-1/(\gamma c^j)} (d^*)^{-1/c^j} d^{1/\gamma} \sim 1 - \big(\tfrac{k_n}{n}x\big)^{\gamma^*} A^{1/c^j} d^{1/\gamma}, \ \ \forall j \ge 0, \tag{40}$$

where $A = d^{-1/\gamma}(d^*)^{-1}$.

Considering (38) and (39), the same reasoning of Proposition 3.1 leads to,

$$\begin{aligned} &P\big(T_1 > F_T^{-1}(1-\tfrac{k_n}{n}x), T_{1+m} > F_T^{-1}(1-\tfrac{k_n}{n}y)\big) \\ \underset{t\downarrow 0}{\sim}\ &\big(y\tfrac{k_n}{n}\big)^{\frac{\gamma^*}{c^{m+1}}} d^{1/\gamma} A^{1/c^{m+1}} + (xy)^{2\gamma^*}\big(\tfrac{k_n}{n}\big)^{4\gamma^*} d^{1/\gamma} A^4 + (xy)^{\frac{\gamma^*}{c}}\big(\tfrac{k_n}{n}\big)^{2\frac{\gamma^*}{c}} d^{1/\gamma} A^{2/c} \\ \underset{n\to\infty}{\sim}\ &\begin{cases} xy\big(\tfrac{k_n}{n}\big)^2 d^{1/\gamma} A^4\big(1+\mathbf{1}_{\{c=1/2\}}\big) & , c \le 1/2 \\ xy\big(\tfrac{k_n}{n}\big)^2 d^{1/\gamma} A^{2/c} + \mathbf{1}_{\{c^m=1/2\}}\big(y\tfrac{k_n}{n}\big)^2 d^{1/\gamma} A^{1/c^{m+1}} & , 1/2 < c \le 1/2^{1/m} \\ \big(y\tfrac{k_n}{n}\big)^{1/c^m} d^{1/\gamma} A^{1/c^{m+1}} & , c > 1/2^{1/m}. \end{cases} \end{aligned}$$

Thus being, (36) holds for $\{T_i\}_i$, and hence by (35) also holds for $\{T_i^{(n)}\}_i$, with $h_m(x,y)=0$, $\forall m\geq 1$.

With respect to (37), observe that, we can state,

$$\frac{n}{k_n}P\Big(T_1\in I_n(x,y),T_{1+m}\in I_n(x,y)\Big)$$

$$\leq \frac{n}{k_n}\Big[P\Big(a_{n,y}<X_1\leq a_{n,x},X_1>a_{n,y}^{1/c^m}\Big)+P\Big(a_{n,y}<X_1\leq a_{n,x},\bigvee_{j=0}^{m-1}(Z_{m-j+1}^{c^j})>a_{n,y}\Big)\Big]$$

$$+\frac{n}{k_n}\Big[P\Big(a_{n,y}<X_2\leq a_{n,x},X_2>a_{n,y}^{1/c^m}\Big)+P\Big(a_{n,y}<X_2\leq a_{n,x},\bigvee_{j=0}^{m-1}(Z_{m-j+2}^{c^j})>a_{n,y}\Big)\Big].$$

By the independence assumptions, the last expression becomes,

$$\frac{n}{k_n}\Big\{2\Big[K(a_{n,x})-K\big(a_{n,y}^{1/c^m}\big)\Big]+2\Big[K(a_{n,x})-K(a_{n,y})\Big]\cdot\Big[1-\prod_{j=0}^{m-1}F_Z\big(a_{n,y}^{1/c^j}\big)\Big]\Big\}$$

$$\leq \frac{n}{k_n}\Big\{2\Big[K(a_{n,x}^{1/c^m})-K(a_{n,y}^{1/c^m})\Big]+2\Big[K(a_{n,x})-K(a_{n,y})\Big]\Big[1-K(a_{n,y})\Big]\Big\}$$

Considering (40) and conditions in (37), then,

$$2\Big[K(a_{n,x}^{1/c^m})-K(a_{n,y}^{1/c^m})\Big]+2\Big[K(a_{n,x})-K(a_{n,y})\Big]\Big[1-K(a_{n,y})\Big]$$

$$=2\Big[\big(\tfrac{k_n}{n}\big)^{\gamma^*/c^m}d^{1/\gamma}A^{1/c^m}\big(y^{\gamma^*/c^m}-x^{\gamma^*/c^m}\big)\Big]+2\Big[\big(\tfrac{k_n}{n}\big)^{\gamma^*}d^{1/\gamma}A(y^{\gamma^*}-x^{\gamma^*})\Big]\big(\tfrac{k_n}{n}y\big)^{\gamma^*}d^{1/\gamma}A$$

$$\leq (y-x)\frac{k_n}{n}\Big(2d^{1/\gamma}A^{1/c^m}\Big[\frac{k_n}{n}(1+\epsilon)\Big]^{\gamma^*/c^m-1}\frac{1+\epsilon}{\delta}+\big(\tfrac{k_n}{n}\big)^{2\gamma^*-1}2\frac{(1+\epsilon)^{2\gamma^*}}{\delta}d^{2/\gamma}A^2\Big)$$

$$\leq (y-x)\frac{k_n}{n}\Big(\widetilde{\rho}(m)+\big(\tfrac{k_n}{n}\big)^{\alpha}D_1\Big).$$

where $\delta=y-x$, $\widetilde{\rho}(m)=2d^{1/\gamma}A^{1/c^m}\Big[\frac{k_n}{n}(1+\epsilon)\Big]^{\gamma^*/c^m-1}\frac{1+\epsilon}{\delta}$, $D_1=2\frac{(1+\epsilon)^{2\gamma^*}}{\delta}d^{2/\gamma}A^2>0$ and $\alpha=2\gamma^*-1$. Note that, $\sum_{m=0}^{\infty}\widetilde{\rho}(m)<\infty$ since, from some order n,

$$\lim_{m\to\infty}\widetilde{\rho}(m+1)/\widetilde{\rho}(m)\sim\Big\{A\Big[\frac{k_n}{n}(1+\epsilon)\Big]^{\gamma^*}\Big\}^{\frac{1}{c^m}(1/c-1)}<1.$$

Given (35), for large enough n, we have,

$$\frac{n}{k_n}P\big(T_1^{(n)}>x,T_{1+m}^{(n)}>y\big)\sim\frac{n}{k_n}P\big(T_1>x,T_{1+m}>y\big)\leq(y-x)\Big(\widetilde{\rho}(m)+\big(\tfrac{k_n}{n}\big)^{\alpha}D_1\Big).$$

and so, (37) also holds for sequence $\{T_i^{(n)}\}_i$.

Example 3.1

Consider pARMAX series from Example 2, plotted in Figure 1, that is, the marginals are Pareto(0.7) and innovations have d.f. (11) with $\gamma=0.7$ and $c=0.7,0.8,0.9$, respectively (sample size $n=5000$). For each of these series, estimates of η_m are obtained, for lags $m = 1, ..., 6$, by the Hill estimator based on the largest $k+1$ order statistics of $T_i^{(n)}$, defined as follows:

$$\widehat{\eta}_m = \frac{1}{k}\sum_{i=1}^{k} \log T_{n-i+1:n}^{(n)} - \log T_{n-k:n}^{(n)}.$$

Results in Table 1 show that the estimates are, in general, very close to their true values. The values of k were chosen in a range of stability of Hill's estimates.

Table 1. Estimates of η_m ($m = 1, ..., 6$) obtained by Hill estimator based on the largest $k = 200, 500, 1000$ order statistics of $T_i^{(n)}$. The true values for each lag m are on the lines beginning with "η_m"

$c = 0.7$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$
η_m	0.7	0.5	0.5	0.5	0.5	0.5
$k = 200$	0.69	0.54	0.49	0.48	0.48	0.48
$k = 500$	0.68	0.52	0.48	0.48	0.48	0.48
$k = 1000$	0.66	0.49	0.46	0.48	0.48	0.47
$c = 0.8$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$
η_m	0.8	0.64	0.51	0.5	0.5	0.5
$k = 200$	0.81	0.67	0.56	0.51	0.49	0.49
$k = 500$	0.79	0.66	0.56	0.5	0.49	0.49
$k = 1000$	0.8	0.62	0.53	0.47	0.46	0.49
$c = 0.9$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$
η_m	0.9	0.81	0.73	0.66	0.59	0.53
$k = 200$	0.89	0.79	0.7	0.63	0.58	0.54
$k = 500$	0.88	0.79	0.69	0.62	0.56	0.53
$k = 1000$	0.87	0.76	0.66	0.58	0.52	0.48

4. Estimation of return levels

In this section we will use relation (2) to estimate return levels in pARMAX processes. However, since θ is unitary in these cases, there is an advantage (for inferential purposes) if θ is replaced by the threshold-dependent extremal index (Bortot and Tawn 1998). More precisely, for a sequence $\{Y_i\}_i$ as in (1), the authors cited considered,

$$\begin{aligned}\theta(u, r_{[u]}) &:= P(Y_i \le u, 2 \le i \le r_{[u]} | Y_1 > u) \\ &= 1 - P(Y_2 > u | Y_1 > u) - \sum_{j=3}^{r_{[u]}} P\big(\textstyle\bigvee_{i=2}^{j-1} Y_i \le u, Y_j > u | Y_1 > u\big),\end{aligned} \tag{41}$$

which is an approximation of the limit (1) for n fixed, $u_n \equiv u$ and $r_{[u]}$ the block length. Moreover, condition $\Delta^{(2)}(u_n, r_n)$ used by the same authors is also valid for pARMAX processes, as will be seen in the next proposition, so that,

$$1 - \theta(u, r_{\lfloor u \rfloor}) \sim P(Y_2 > u | Y_1 > u) \sim (t(u))^{1-1/\eta} L(t(u)), \text{ with } t(u) = (P(Y_1 > u))^{-1}. \tag{42}$$

Therefore, taking normalized levels u_n of pARMAX, i.e., $\lim_{n\to\infty} n(1-K(u_n))=\tau$, with $\tau>0$ fixed, using (13) and (23), we derive an analogous result to m-dependent Gaussian stationary sequences (Rootzén 1983), that is,

$$P(\bigvee_{i=1}^n X_i \le u_n) - K^n(u_n) \sim e^{-\tau} n^{1-1/c} \mathcal{L}(a_{\tau/n}^{1/c})\tau^{1/c}, \tag{43}$$

which, with a simple modification, as stated in Bortot and Tawn (1998), leads to,

$$P(\bigvee_{i=1}^n X_i \le u_n) - K^{n\theta(u_n)}(u_n) = o\big(n^{1-1/c}\mathcal{L}(a_{\tau/n}^{1/c})\big), \tag{44}$$

where $\theta(u)\equiv\theta(u,r_{[u]})$ is given in (42).

Proposition 4.1. *Consider $(u_n)_{n\ge1}$ a sequence of normalized levels of K and $\{r_n\}_{n\ge1}$ a nondecreasing integers sequence as stated in (1). Then, condition $\Delta^{(2)}(u_n,r_n)$ holds for the pARMAX process.*

Proof. We must prove that,

$$\sum_{j=3}^{r_n} \frac{P(X_1>u_n, \bigvee_{i=2}^{j-1} X_i \le u_n, X_j>u_n)}{P(X_1>u_n, X_2>u_n)} \to 0\,,\ \text{as } n\to\infty\,, \tag{45}$$

(Bortot and Tawn 1998). Observe that,

$$P\Big(\bigvee_{i=2}^{j-1} X_i \le u, X_j>u, X_1>u\Big) = P\Big(\bigvee_{i=2}^{j-1} X_i \le u, X_1>u\Big) - P\Big(\bigvee_{i=2}^{j} X_i \le u, X_1>u\Big)$$

According to (13), $P\big(\bigvee_{i=2}^{j} X_i \le u, X_1>u\big) = \big[K(u^{1/c})-K(u)\big]\big[F_Z(u)\big]^{j-1}$, so that,

$$P\big(\bigvee_{i=2}^{j-1} X_i \le u, X_j>u, X_1>u\big) = \big[K(u^{1/c})-K(u)\big]\big[F_Z(u)\big]^{j-2}\big[1-F_Z(u)\big]. \tag{46}$$

Applying (27), (46) and simple calculations, expression in (45) becomes,

$$\frac{[K(u_n^{1/c})-K(u_n)]F_Z(u_n)[1-(F_Z(u_n))^{r_n-2}]}{1-K(u_n)-F_Z(u_n)[K(u_n^{1/c})-K(u_n)]}\,.$$

Considering the first order of Taylor's approach of $(F_Z(u_n))^{r_n-2}$, as $n\to\infty$, we have, $(F_Z(u_n))^{r_n-2}\sim 1-(r_n-2)(1-F_Z(u_n))$. Since $(u_n)_{n\ge1}$ is a sequence of normalized levels of K, by (23), (25) and given (24), we have successively,

$$\sum_{j=3}^{r_n} \frac{P\big(X_1>u_n, \bigvee_{i=2}^{j-1} X_i \le u_n, X_j>u_n\big)}{P(X_1>u_n, X_2>u_n)} \sim \frac{\big[\frac{\tau}{n}-\big(\frac{\tau}{n}\big)^{1/c}\mathcal{L}(u_n^{1/c})\big]\big(1-\frac{\tau}{n}\big)\frac{\tau}{n}(r_n-2)}{\frac{\tau}{n}-\big(1-\frac{\tau}{n}\big)\big[\frac{\tau}{n}-\big(\frac{\tau}{n}\big)^{1/c}\mathcal{L}(u_n^{1/c})\big]}$$

$$\sim \frac{\big[\big(\frac{\tau}{n}\big)^2-\big(\frac{\tau}{n}\big)^{1/c+1}\mathcal{L}(u_n^{1/c})-\big(\frac{\tau}{n}\big)^3+\big(\frac{\tau}{n}\big)^{1/c+2}\mathcal{L}(u_n^{1/c})\big](r_n-2)}{\big(\frac{\tau}{n}\big)^{1/c}\mathcal{L}(u_n^{1/c})+\big(\frac{\tau}{n}\big)^2-\big(\frac{\tau}{n}\big)^{1/c+1}\mathcal{L}(u_n^{1/c})}$$

$$\sim \begin{cases} r_n-2 & ,\text{if } c\le 1/2 \\ \big(\frac{\tau}{n}\big)^{2-1/c}\dfrac{r_n-2}{\mathcal{L}(u_n^{1/c})} & ,\text{if } c>1/2 \end{cases}\quad ,\ n\to\infty\,.$$

Hence, when $c\le1/2$, $\Delta^{(2)}(u_n,r_n)$ holds for $r_n=2$ and, when $c>1/2$, the condition holds for any r_n such that, $r_n=o\big(n^{2-1/c}\big)$. □

Example 4.1

When there is evidence of asymptotic independence, which means that cluster characteristics, in particular the extremal index, change with threshold, return level estimation can be improved if the approximation, $P(\bigvee_{i=1}^{n} X_i \leq u) \approx K(u)^{n\theta}$, is used with θ replaced by $\theta(u)$ (see (43) and (44)). Considering again the pARMAX processes of Example 3, we can see somewhat significantly different return levels estimates, if we take $\theta = \theta(u)$ (model I) and $\theta = 1$ (model II) in the previous approach. Moreover, differences become larger with higher values of parameter c (Figure 3). The estimates obtained for the 100 year return level of pARMAX (considering, for instance, $n = 250$ observations per year) are, using model I, 1126, 1184 and 893, with $c = 0.7$, $c = 0.8$ and $c = 0.9$, respectively, and 1195 considering model II.

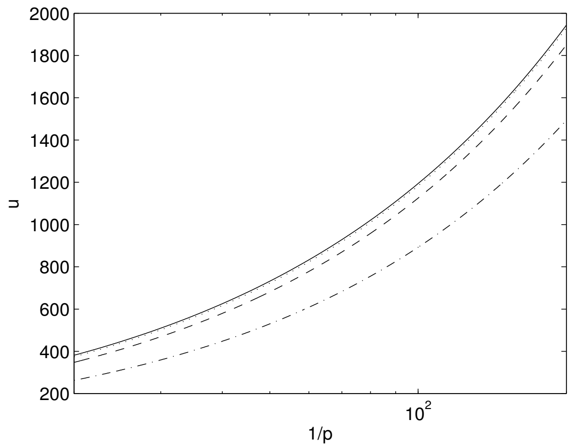

Figure 3. Return level estimates for pARMAX process against $1/p$, ($p = P(\bigvee_{i=1}^{n} X_i > u)$), on a logarithmic scale: solid line taking $\theta = 1$ (model II); dotted line, dashed line and dotted-dashed line taking $\theta = \theta(u)$ in (42) (model I), with $c = 0.7$, $c = 0.8$ and $c = 0.9$, respectively

References

[1] Alpuim MT (1989a) Contribuições à teoria de extremos em sucessões dependentes. Ph. D. Thesis, FCUL

[2] Alpuim MT(1989b) An extremal markovian sequence. J. Appl. Probab. 26:219-232

[3] Asmussen S(1987) Applied Probability and Queues. John Wiley & Sons

[4] Bortot P, J. Tawn (1998) Models for the extremes of Markov Chains. Biometrika 85:851-867

[5] Canto e Castro L (1992) Sobre a Teoria Assintótica de Extremos. Ph. D. Thesis, FCUL

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

[7] Draisma G, Drees H, Ferreira A, de Haan L (2004) Bivariate tail estimation: dependence in asymptotic independence. Bernoulli 10:251-280

[8] Drees H (2003) Extreme quantile estimation for dependent data with applications to finance. Bernoulli 9:617-657

[9] Klotz J (1973) Statistical Inference in Bernoulli Trials with dependence. Ann. Statist. 1:373-379

[10] Leadbetter MR, Lindgren G, Rootzén H (1983) Extremes and Related Properties of Random Sequences and Processes. Springer-Verlag, New York, Heidelberg-Berlin

[11] Leadbetter MR , Nandagopalan S (1989) On Exceedance Point Processes for Stationary Sequences under Mild Oscillation Restrictions. In: Hüsler, J. and Reiss R.-D. (ed) Extreme Value Theory. Springer-Verlag, New York, pp 69-80

[12] Ledford A, Tawn JA (1996) Statistics for near independence in multivariate extreme values. Biometrika 83:169-187

[13] Ledford A, Tawn JA (1997) Modelling Dependence within joint tail regions. J. R. Stat. Soc. Ser. B Stat. Methodol. 59:475-499

[14] Ledford A, Tawn JA (2003) Diagnostics for dependence within time series. J. R. Stat. Soc. Ser. B Stat. Methodol. 65:521-543

[15] O'Brien GL (1987) Extreme values for stationary and Markov sequences. Ann. Prob. 15:281-291

[16] Peng L (1999) Estimation of the coefficient of the tail dependence in bivariate extremes. Statist. Probab. Lett. 43:399-409

[17] Rootzén H (1983) The rate of convergence of extremes of stationary normal sequences. Adv. in Appl. Probab. 15:54-80

[18] Smith RL, Tawn JA, Coles SG (1997) Markov chain models for threshold exceedances. Biometrika 84:249-68