New Examples of Partial Samples from the Uniform AR(1) Process and Asymptotic Distributions of Extremes

Lenka Glavaš[a]

[a]Faculty of Mathematics, University of Belgrade, Serbia

Abstract. The joint limiting distribution of maximum of the specific sub-sample and maximum of the complete sample from the first-order auto-regressive process with uniform marginal distributions is obtained in this article. There are considered several examples of partial samples, consisted of non-randomly selected terms of the full sample. It is well known that the uniform $AR(1)$ process is strictly stationary random sequence; it doesn't satisfy condition of weak dependency, that prohibits clustering of extremes. As a consequence of this property, some interesting conclusions about joint asymptotic distributions are reached.

1. Introduction and preliminaries

In this paper will be analyzed asymptotic behavior of maxima of some incomplete samples from the uniform $AR(1)$ process.

Generally speaking, limit theorems are essential for understanding the real content of the concept of probability. Extreme value theory is very well-developed branch of the probability theory, dealing with limiting distributions for extremes not only of independent, identically distributed random variables (classical approach), but also of certain stationary random sequences, that are weakly dependent. The interested reader should consult the excellent book [7], for further details.

It turns out that, if a stationary sequence does not satisfy some of the weak dependence conditions, the larger variability of the asymptotic distributional results for extrema arises. Several examples of random sequences, for which one of the conditions of weak dependency is not fulfilled, were studied in the literature, e.g. a uniform $AR(1)$ process [1], [2], and, more recently, a storage process in discrete time with fractional Brownian motion as input [10], introduced in [12].

On the other hand, properties of non-Gaussian first order linear auto-regressive models have been investigated from different points of view. In survey of linear $AR(1)$ models [4] there were described more than 30 different models having linear $AR(1)$ structure, which are very common in many areas of science (such as time series of counts, proportions, binary outcomes, non-negative or heavy-tailed observations etc.).

Here two types of the first-order auto-regressive process with uniform marginal distributions, one with positive and the other with negative lag one correlation, will be considered. It will be shown that extremes

of complete and specific partial samples from each of these two types of uniform $AR(1)$ process manifest different asymptotic behavior. However, the interesting extremal properties of this process, related to the partial samples, are not exhausted in this paper, and could be subject to further consideration, as indicated in conclusion.

The class of first-order auto-regressive processes with uniform marginal distributions - further referred to as uniform $AR(1)$ processes - was originally studied by Chernick [1]. A random process $(X_n)_{n\in\mathbb{N}}$ from this class is defined by recursive formula

$$X_n = \frac{1}{r}X_{n-1} + \epsilon_n, \; n \geq 2, \tag{1.1}$$

where $r \geq 2$ - called parameter of this process - is a positive integer; X_1 is distributed uniformly on the interval $[0,1]$; the ϵ_n's form a sequence of i.i.d. random variables with the discrete uniform distribution on the set $\left\{0, \frac{1}{r}, \ldots, \frac{r-1}{r}\right\}$ and r.v. ϵ_n is independent of X_{n-1} for each n. As a matter of fact, (X_n), defined in this way, is a strictly stationary random sequence with X_n distributed uniformly on the interval $[0,1]$ for each n.

Further properties of a uniform $AR(1)$ process, which were established by Chernick [1], were related to fulfilling of weak dependence conditions - Leadbetter's [5] condition $D(u_n)$ and Loynes' [8] condition $D'(u_n)$. These two conditions are of crucial importance in extreme value theory for strictly stationary random sequences; when they hold it is possible to obtain extremal results similar to those in classical theory for i.i.d. sequences. Namely, he showed that $D(u_n)$ is satisfied, but $D'(u_n)$ fails, with $u_n = 1 - \frac{x}{n}$, $x > 0$.

Due to the fact that a uniform $AR(1)$ process does not satisfy the condition $D'(u_n)$ (under this condition clustering of extremes is restricted) some interesting extremal properties of this process have been expected.

Let M_n be the maximum of the first n terms of the sequence (X_n), i.e. $M_n = \max\{X_1, X_2, \ldots, X_n\}$. Chernick [1] derived a limit theorem for the r.v. M_n. Suppose, further, that some of the random variables $X_1, X_2, \ldots, X_n$ can be observed. Thus shows up partial sample consisted of observed terms amongst first n terms of the sequence (X_n). Let $(c_n)_{n\in\mathbb{N}}$ be a non-random $0-1$-sequence, such that $\lim\limits_{n\to+\infty} \frac{1}{n}\sum\limits_{j=1}^{n} c_j = p$, $0 < p < 1$. This sequence of degenerate random variables is introduced with the purpose to correspond to the sequence (X_n) in the following sense: r.v. X_k is observed if $c_k = 1$, otherwise r.v. X_k is not observed (missing observation). The already mentioned partial sample is simply the set $\{X_k : c_k = 1, \; 1 \leq k \leq n\}$; denote its maximum term by $\widetilde{M}_n$.

As remarked earlier, the condition $D'(u_n)$ is not satisfied for uniform $AR(1)$ process (X_n) and that is the main reason why the general result stated in Theorem 3.2. in Mladenović and Piterbarg [10], p. 1981, concerning the limiting distribution of the random vector $(\widetilde{M}_n, M_n)$, as $n \to +\infty$, does not apply. Considering this process Mladenović [9] found out that for some values $u_n < v_n$, events $\{\widetilde{M}_n \leq u_n\}$ and $\{M_n \leq v_n\}$ are asymptotically independent, while for other values $u_n < v_n$ the same events can be asymptotically perfectly dependent; recently, Mladenović and Živadinović [11] also proved that the limiting distribution of the random vector $(\widetilde{M}_n, M_n)$, as $n \to +\infty$, is not uniquely determined by the limit value p.

One of the new results contained in this article, in Section 2, is a limit theorem for the random vector $(\widetilde{M}_n, M_n)$, for three special choices of partial sample.

The other two new limit theorems, that will be presented in Section 3, deal with partial samples from a negatively correlated uniform $AR(1)$ process. This process appeared for the first time in Chernick and Davis [2], where, actually, the existence of negatively correlated uniform sequences was shown.

Let $r \geq 2$ be a positive integer; the class of strictly stationary auto-regressive processes with uniform marginal distributions and lag one correlation equal to $-\frac{1}{r}$ are obtained in the following way. A random

process $(X_n)_{n \in \mathbb{N}}$ from this class is defined by recursive formula

$$X_n = -\frac{1}{r}X_{n-1} + \epsilon_n, \ n \geq 2, \tag{1.2}$$

where X_1 is distributed uniformly on the interval $[0,1]$; ϵ_n's form a sequence of i.i.d. random variables with the discrete uniform distribution on the set $\left\{\frac{1}{r}, \frac{2}{r}, \ldots, 1\right\}$ and r.v. ϵ_n is independent of X_{n-1} for each $n \geq 2$.

Chernick and Davis [2] concluded that for a negatively correlated uniform $AR(1)$ process the condition $D(u_n)$ holds, and, regarding the extremal properties, they derived a limit theorem for the maximum term of the process. From the formulation of Theorem 3.1. [2], p. 87, it is obvious that the maximum of this random sequence behaves as if it corresponds to a uniform $AR(1)$ process with $\frac{1}{r^2}$ as the lag one correlation.

After the theoretical results, in the last part of this article, results of some numerical simulations will be listed (Section 4) and a brief summary of existing results and still open questions will be given (Section 5).

2. The limit theorem - uniform $AR(1)$ process

Let $(X_n)_{n \in \mathbb{N}}$ be the uniform $AR(1)$ process with the parameter r and let $\left(c_n^{(1)}\right)_{n \in \mathbb{N}}$, $\left(c_n^{(2)}\right)_{n \in \mathbb{N}}$, $\left(c_n^{(3)}\right)_{n \in \mathbb{N}}$ be three non-random $0 - 1$-sequences such that:

$c_n^{(1)} = 1$ if $n = 3m$; $c_n^{(1)} = 0$ if $n = 3m - 2$ or $n = 3m - 1$, $m \in \mathbb{N}$;

$c_n^{(2)} = 1$ if $n = 3m - 1$; $c_n^{(2)} = 0$ if $n = 3m - 2$ or $n = 3m$, $m \in \mathbb{N}$;

$c_n^{(3)} = 1$ if $n = 3m - 2$; $c_n^{(3)} = 0$ if $n = 3m - 1$ or $n = 3m$, $m \in \mathbb{N}$;

the maximum of the complete sample $X_1, X_2, \ldots, X_n$ is, as usual, denoted by M_n; the maximum of the sub-sample of the complete sample, determined by the corresponding sequence $\left(c_n^{(i)}\right)$, is denoted by $\widetilde{M}_n^{(i)}$, i.e. $\widetilde{M}_n^{(i)} := \max\left\{X_k : c_k^{(i)} = 1, \ 1 \leq k \leq n\right\}$, $i \in \{1, 2, 3\}$.

Theorem 2.1. *For every $i \in \{1, 2, 3\}$:*

- *if $0 < y \leq \frac{x}{r^2}$, then $\lim\limits_{n \to +\infty} P\left\{\widetilde{M}_n^{(i)} \leq 1 - \frac{x}{n}, M_n \leq 1 - \frac{y}{n}\right\} = e^{-\frac{r^3-1}{3r^3}x}$,*

- *if $0 < \frac{x}{r^2} < y \leq \frac{x}{r}$, then $\lim\limits_{n \to +\infty} P\left\{\widetilde{M}_n^{(i)} \leq 1 - \frac{x}{n}, M_n \leq 1 - \frac{y}{n}\right\} = e^{-\frac{r^2-1}{3r^2}x - \frac{r-1}{3r}y}$,*

- *if $0 < \frac{x}{r} < y \leq x$, then $\lim\limits_{n \to +\infty} P\left\{\widetilde{M}_n^{(i)} \leq 1 - \frac{x}{n}, M_n \leq 1 - \frac{y}{n}\right\} = e^{-\frac{r-1}{3r}x - \frac{2(r-1)}{3r}y}$.*

Proof.
The proof proceeds in the same manner as the one of Theorem 2.1. in Mladenović and Živadinović [11]. Therefore, it is more convenient to omit the proof here. $\square$

Remark 1. *In the formulation of the Theorem 2.1, the remaining case when $0 < x < y$ was intentionally omitted, because if $0 < x < y$ then the level $1 - \frac{x}{n}$ is greater than the level $1 - \frac{y}{n}$. Hence, $M_n \leq 1 - \frac{y}{n}$ implies $\widetilde{M}_n^{(i)} \leq 1 - \frac{y}{n} < 1 - \frac{x}{n}$, so that intersection $\left\{\widetilde{M}_n^{(i)} \leq 1 - \frac{x}{n}, M_n \leq 1 - \frac{y}{n}\right\}$ of these two events is just the latter one. Theorem 4.1. in Chernick [1], p. 147, applies.*

Remark 2. *Between these three sequences $\left(c_n^{(i)}\right)$, $i \in \{1, 2, 3\}$, certain relationship can be established: if e.g. $\left(c_n^{(1)}\right)$ is taken as, so to say, the basic one, then terms of the other two sequences can be expressed as: $c_n^{(2)} = c_{n+1}^{(1)}$ and $c_n^{(3)} = c_{n+2}^{(1)}$, for all $n \in \mathbb{N}$.*

3. The limit theorems - negatively correlated uniform $AR(1)$ process

Let $(X_n)_{n\in\mathbb{N}}$ be the negatively correlated uniform $AR(1)$ as defined above by (1.2). Designate, as before, by M_n the maximum of the first n terms of this random sequence and by $\widetilde{M}_n$ the maximum on subset of observed terms among the first n terms. A term X_k, $k \in \mathbb{N}$, is observed if $c_k = 1$; $(c_n)_{n\in\mathbb{N}}$ is a non-random $0-1$-sequence.

The main aim is again to determine the limiting distribution of the random vector $(\widetilde{M}_n, M_n)$ for some particular partial samples. For the same reason indicated in the Remark 1. all considerations will be focused on the case $0 < y \le x$. Otherwise, Theorem 3.1. in Chernick and Davis [2], p. 87, applies.

Theorem 3.1. *Let $(c_n)_{n\in\mathbb{N}}$ be the sequence given by:*

$$c_n = 1 \text{ if } n = 2m; \ c_n = 0 \text{ if } n = 2m - 1, \ m \in \mathbb{N}. \tag{3.1}$$

If $0 < y \le x$, then $\lim_{n\to+\infty} P\left\{\widetilde{M}_n \le 1 - \dfrac{x}{n}, M_n \le 1 - \dfrac{y}{n}\right\} = e^{-\frac{r^2-1}{2r^2}x - \frac{r^2-1}{2r^2}y}.$

Remark 3. *Proofs of the Theorem 3.1. and the next theorem - Theorem 3.2. follow the same reasoning. Therefore, only the latter theorem will be derived in details.*

Theorem 3.2. *Let $(c_n)_{n\in\mathbb{N}}$ be the sequence given by:*

$$c_n = 1 \text{ if } n = 4m - 1 \text{ or } n = 4m; \ c_n = 0 \text{ if } n = 4m - 3 \text{ or } n = 4m - 2, \ m \in \mathbb{N}. \tag{3.2}$$

- *If $0 < y \le \dfrac{x}{r^2}$, then* $\lim_{n\to+\infty} P\left\{\widetilde{M}_n \le 1 - \dfrac{x}{n}, M_n \le 1 - \dfrac{y}{n}\right\} = e^{-\frac{r^4-1}{2r^4}x}$,

- *if $0 < \dfrac{x}{r^2} < y \le x$, then* $\lim_{n\to+\infty} P\left\{\widetilde{M}_n \le 1 - \dfrac{x}{n}, M_n \le 1 - \dfrac{y}{n}\right\} = e^{-\frac{r^2-1}{2r^2}x - \frac{r^2-1}{2r^2}y}$.

Proof.
The following notation will be used:

$$a_m := P\left\{\widetilde{M}_{4m} \le 1 - \frac{x}{n}, M_{4m} \le 1 - \frac{y}{n}\right\}. \tag{3.3}$$

The first step is to determine a_m for $0 < y \le x$, sufficiently large n and some $4m < n$.
Note that the event $\left\{\widetilde{M}_{4m} \le 1 - \dfrac{x}{n}, M_{4m} \le 1 - \dfrac{y}{n}\right\}$ is equivalent to the event (intersection of $4m$ events):

$$\left\{X_{4i-3} \le 1 - \frac{y}{n}, X_{4i-2} \le 1 - \frac{y}{n}, X_{4i-1} \le 1 - \frac{x}{n}, X_{4i} \le 1 - \frac{x}{n}, i = 1, 2, \ldots, m\right\}. \tag{3.4}$$

Use the recurrence relation (1.2) from the definition of the negatively correlated uniform $AR(1)$ process to obtain:

$$X_i = \left(-\frac{1}{r}\right)^{i-1} X_1 + \sum_{j=2}^{i} \left(-\frac{1}{r}\right)^{i-j} \epsilon_j, i \ge 2. \tag{3.5}$$

In this way all events in (3.4) can be expressed in terms of the r.v. X_1 and random variables from the sequence (ϵ_n), so inserting the equations (3.5) to the expression a_m and doing some rearrangements leads to:

$$a_m = P\left\{X_1 \le 1 - \frac{y}{n}, X_1 \ge -r + r\epsilon_2 + \frac{ry}{n}, X_1 \le r^2 + r\epsilon_2 - r^2\epsilon_3 - \frac{r^2 x}{n}, X_1 \ge -r^3 + r\epsilon_2 - r^2\epsilon_3 + r^3\epsilon_4 + \frac{r^3 x}{n}, \right.$$

$$X_1 \le r^{4i-4} + \sum_{j=2}^{4i-3} (-1)^j r^{j-1} \epsilon_j - \frac{r^{4i-4}y}{n},$$

$$X_1 \ge -r^{4i-3} + \sum_{j=2}^{4i-2} (-1)^j r^{j-1} \epsilon_j + \frac{r^{4i-3}y}{n},$$

$$X_1 \le r^{4i-2} + \sum_{j=2}^{4i-1} (-1)^j r^{j-1} \epsilon_j - \frac{r^{4i-2}x}{n},$$

$$X_1 \ge r^{4i-1} + \sum_{j=2}^{4i} (-1)^j r^{j-1} \epsilon_j + \frac{r^{4i-1}x}{n}, i = 2, 3, \dots, m\Big\}.$$

After that the law of total probability can be applied, conditioning on $(4m-1)$ random variables $\epsilon_2, \dots, \epsilon_{4m}$; assumptions made about these random variables should be exploited (recall (ϵ_n) is sequence of i.i.d. random variables with the discrete uniform distribution). Here it is convenient to take notice of the following, of course when n is large enough so that $\frac{r^{4m}}{x}, \frac{r^{4m-2}}{y} \in (0, 1)$:

– the event $A_i := \left\{ X_1 \le r^{4i-4} + \sum_{j=2}^{4i-3} (-1)^j r^{j-1} \epsilon_j - \frac{r^{4i-4}y}{n} \right\}$ is non-trivial when all random variables with even indexes, $\epsilon_2, \epsilon_4, \dots, \epsilon_{4i-4}$, take value $\frac{1}{r}$, and, at the same time, all random variables with odd indexes, $\epsilon_3, \epsilon_5, \dots, \epsilon_{4i-3}$, take value 1 - if this is the case A_i becomes $\left\{ X_1 \le 1 - \frac{r^{4i-4}y}{n} \right\}, 2 \le i \le m$;

– the event $B_i := \left\{ X_1 \ge -r^{4i-3} + \sum_{j=2}^{4i-2} (-1)^j r^{j-1} \epsilon_j + \frac{r^{4i-3}y}{n} \right\}$ is non-trivial when all random variables with even indexes, $\epsilon_2, \epsilon_4, \dots, \epsilon_{4i-2}$, take value 1, and, at the same time, all random variables with odd indexes, $\epsilon_3, \epsilon_5, \dots, \epsilon_{4i-3}$, take value $\frac{1}{r}$ - if this is the case B_i becomes $\left\{ X_1 \ge \frac{r^{4i-3}y}{n} \right\}, 1 \le i \le m$;

– the event $C_i := \left\{ X_1 \le r^{4i-2} + \sum_{j=2}^{4i-1} (-1)^j r^{j-1} \epsilon_j - \frac{r^{4i-2}x}{n} \right\}$ is non-trivial when all random variables with even indexes, $\epsilon_2, \epsilon_4, \dots, \epsilon_{4i-2}$, take value $\frac{1}{r}$, and, at the same time, all random variables with odd indexes, $\epsilon_3, \epsilon_5, \dots, \epsilon_{4i-1}$, take value 1 - if this is the case C_i becomes $\left\{ X_1 \le 1 - \frac{r^{4i-2}x}{n} \right\}, 1 \le i \le m$;

– the event $D_i := \left\{ X_1 \ge r^{4i-1} + \sum_{j=2}^{4i} (-1)^j r^{j-1} \epsilon_j + \frac{r^{4i-1}x}{n} \right\}$ is non-trivial when all random variables with even indexes, $\epsilon_2, \epsilon_4, \dots, \epsilon_{4i}$, take value 1, and, at the same time, all random variables with odd indexes, $\epsilon_3, \epsilon_5, \dots, \epsilon_{4i-1}$, take value $\frac{1}{r}$ - if this is the case D_i becomes $\left\{ X_1 \ge \frac{r^{4i-1}x}{n} \right\}, 1 \le i \le m$.

Depending on which values the random variables $\epsilon_2, \dots, \epsilon_{4m}$ take, it can be proceeded with reducing the number of events of the form $\{X_1 \le u_n\}$ or $\{X_1 \ge v_n\}$ within the intersections and thus simplifying, as much as possible, the expression a_m. At this point it should be obvious that two separate cases have to be distinguished, because a_m will be different if $0 < r^2 y \le x$ or $0 < y \le x < r^2 y$. Finally, in both cases all

probabilities will be calculated using the fact that X_1 is distributed uniformly on $[0, 1]$ - the required concise equation for a_m in each case is obtained.

Case $0 < r^2 y \leq x$.
For such values x and y and $1 \leq i \leq m - 1$:

$$\left\{ X_1 \leq 1 - \frac{r^{4i-2}x}{n} \right\} \subseteq \left\{ X_1 \leq 1 - \frac{r^{4i}y}{n} \right\}, \text{ since } 1 - \frac{r^{4i}y}{n} \geq 1 - \frac{r^{4i-2}x}{n},$$

$$\left\{ X_1 \geq \frac{r^{4i-1}x}{n} \right\} \subseteq \left\{ X_1 \geq \frac{r^{4i+1}y}{n} \right\}, \text{ since } \frac{r^{4i-1}x}{n} \geq \frac{r^{4i+1}y}{n}.$$

Therefore, after grouping:

$$a_m = \frac{r^{4m-3}(r(r-1)-1)}{r^{4m-1}} P\left\{ X_1 \leq 1 - \frac{y}{n} \right\}$$

$$+ \frac{r(r^4-1)}{r^{4m-1}} \sum_{i=1}^{m-1} r^{4(m-1-i)} P\left\{ X_1 \leq 1 - \frac{r^{4i-2}x}{n} \right\} + \frac{r}{r^{4m-1}} P\left\{ X_1 \leq 1 - \frac{r^{4m-2}x}{n} \right\}$$

$$+ \frac{r^{4m-4}(r^2-1)}{r^{4m-1}} P\left\{ X_1 \leq 1 - \frac{y}{n}, X_1 \geq \frac{ry}{n} \right\}$$

$$+ \frac{r^4-1}{r^{4m-1}} \sum_{i=1}^{m-1} r^{4(m-1-i)} P\left\{ X_1 \leq 1 - \frac{y}{n}, X_1 \geq \frac{r^{4i-1}x}{n} \right\} + \frac{1}{r^{4m-1}} P\{ X_1 \leq 1 - \frac{y}{n}, X_1 \geq \frac{r^{4m-1}x}{n} \},$$

and further:

$$a_m = \frac{r^2-r-1}{r^2} \left(1 - \frac{y}{n} \right) + \frac{r^4-1}{r^2} \sum_{i=1}^{m-1} \frac{1}{r^{4i}} \left(1 - \frac{r^{4i-2}x}{n} \right) + \frac{1}{r^{4m-2}} \left(1 - \frac{r^{4m-2}x}{n} \right)$$

$$+ \frac{r^2-1}{r^3} \left(1 - \frac{y}{n} - \frac{ry}{n} \right) + \frac{r^4-1}{r^3} \sum_{i=1}^{m-1} \frac{1}{r^{4i}} \left(1 - \frac{y}{n} - \frac{r^{4i-1}x}{n} \right) + \frac{1}{r^{4m-1}} \left(1 - \frac{y}{n} - \frac{r^{4m-1}x}{n} \right).$$

Computing partial sums in the previous expression leads to:

$$a_m = 1 - \left(\frac{2(r^4-1)}{r^4}(m-1) + 2 \right) \cdot \frac{x}{n} - \frac{2(r^2-1)}{r^2} \cdot \frac{y}{n};$$

if $m \to +\infty$ and $k := \left[\frac{n}{4m} \right] \to +\infty$, as $n \to +\infty$, then the following equality holds:

$$\lim_{n \to +\infty} a_m^k = e^{-\frac{r^4-1}{2r^4}x}. \tag{3.6}$$

Case $0 < y \leq x < r^2 y$.
For such values x and y and $1 \leq i \leq m - 1$:

$$\left\{ X_1 \leq 1 - \frac{r^{4i+2}x}{n} \right\} \subseteq \left\{ X_1 \leq 1 - \frac{r^{4i}y}{n} \right\} \subseteq \left\{ X_1 \leq 1 - \frac{r^{4i-2}x}{n} \right\},$$

$$\left\{ X_1 \geq \frac{r^{4i+1}x}{n} \right\} \subseteq \left\{ X_1 \geq \frac{r^{4i+1}y}{n} \right\} \subseteq \left\{ X_1 \geq \frac{r^{4i-1}x}{n} \right\}.$$

Therefore, after grouping:

$$a_m = \frac{r^{4m-3}(r(r-1)-1)}{r^{4m-1}} P\left\{ X_1 \leq 1 - \frac{y}{n} \right\} + \frac{r(r^2-1)}{r^{4m-1}} \sum_{i=1}^{m-1} r^{4(m-1-i)} P\left\{ X_1 \leq 1 - \frac{r^{4i}y}{n} \right\}$$

$$+\frac{r^3(r^2-1)}{r^{4m-1}}\sum_{i=1}^{m-1}r^{4(m-1-i)}P\Big\{X_1\le 1-\frac{r^{4i-2}x}{n}\Big\}+\frac{r}{r^{4m-1}}P\Big\{X_1\le 1-\frac{r^{4m-2}x}{n}\Big\}$$

$$+\frac{r^4(r^2-1)}{r^{4m-1}}\sum_{i=1}^{m}r^{4(m-1-i)}P\Big\{X_1\le 1-\frac{y}{n},X_1\ge\frac{r^{4i-3}y}{n}\Big\}$$

$$+\frac{r^2(r^2-1)}{r^{4m-1}}\sum_{i=1}^{m-1}r^{4(m-1-i)}P\Big\{X_1\le 1-\frac{y}{n},X_1\ge\frac{r^{4i-1}x}{n}\Big\}+\frac{1}{r^{4m-1}}P\{X_1\le 1-\frac{y}{n},X_1\ge\frac{r^{4m-1}x}{n}\Big\},$$

and proceeding similarly as above leads to:

$$a_m = 1 - \left(\frac{2(r^2-1)}{r^2}(m-1)+2\right)\cdot\frac{x}{n}-\frac{2(r^2-1)}{r^2}m\cdot\frac{y}{n};$$

if $m\to+\infty$ and $k:=\left[\dfrac{n}{4m}\right]\to+\infty$, as $n\to+\infty$, then the following equality holds:

$$\lim_{n\to+\infty}a_m^k = e^{-\frac{r^2-1}{2r^2}x-\frac{r^2-1}{2r^2}y}. \tag{3.7}$$

The last step consists in using standard arguments as in [9], p. 1419, and the fact that two level mixing condition is satisfied, and the statement of the Theorem 3.2. follows straightforward. □

4. Numerical simulations

After having obtained the preceding theoretical results computer simulations were performed. For this purpose functions, written in statistical software, were executed for different suitable arguments.

There were made NS simulations of the first n terms of the random sequence (X_n) and counted the number of the realizations of the event $\Big\{\widetilde{M}_n\le 1-\dfrac{x}{n},M_n\le 1-\dfrac{y}{n}\Big\}$, for some values of x and y. Hence, it was possible to get an estimate $\widehat{P}=\widehat{P}(x,y)$ of the probability of this event ($\widehat{P}$ was computed simply as a ratio of the number of realizations of this event and NS). Described procedure was used several times so several estimates $\widehat{P}_j=\widehat{P}_j(x,y)$ were obtained.

Of course, the direct comparison of estimated probabilities $\widehat{P}_j(x,y)$ to those calculated by using the equations from the limit theorems - further labeled with $G(x,y)$, was possible, as shown in tables below. The following three tables are associated with examples of partial samples, that appear in the Theorem 2.1, 3.1, 3.2, respectively.

In order to illustrate the quality of results obtained from simulation runs the maximum absolute differences of limiting and estimated probabilities are computed.

Table 1: The examples of partial samples determined by the sequences $\left(c_n^{(i)}\right)$, $i \in \{1, 2, 3\}$, as in the Theorem 2.1.

						$r = 2$, $NS = 5000$, $n = 12000$					
x	y	$G(x, y)$	$\widehat{P}_1^{(1)}$	$\widehat{P}_2^{(1)}$	$\widehat{P}_3^{(1)}$	$\widehat{P}_1^{(2)}$	$\widehat{P}_2^{(2)}$	$\widehat{P}_3^{(2)}$	$\widehat{P}_1^{(3)}$	$\widehat{P}_2^{(3)}$	$\widehat{P}_3^{(3)}$
1	0.8	0.6483443	0.6492	0.6484	0.6556	0.6532	0.6552	0.6486	0.6292 [3]	0.6440	0.6518
1	0.6	0.6930406	0.6912	0.6944	0.6938	0.6938	0.6778 [2]	0.6918	0.6872	0.6850	0.7088
1	0.4	0.7285736	0.7284	0.7304	0.7262	0.7352	0.7308	0.7366	0.7264	0.7112	0.7246
1	0.3	0.7408182	0.7456	0.7474	0.7434	0.7332	0.7360	0.7342	0.7540	0.7412	0.7284
1	0.2	0.7470175	0.7394	0.7344 [1]	0.7388	0.7548	0.7400	0.7424	0.7484	0.7352	0.7394
1	0.1	0.7470175	0.7550	0.7436	0.7518	0.7502	0.7396	0.7498	0.7410	0.7438	0.7360

The maximum absolute differences (that appear for the values in shaded cells of the table) are:

$$^1 \max_{x,y,j} \left| G(x, y) - \widehat{P}_j^{(1)} \right| = 0.0126175$$

$$^2 \max_{x,y,j} \left| G(x, y) - \widehat{P}_j^{(2)} \right| = 0.0152406$$

$$^3 \max_{x,y,j} \left| G(x, y) - \widehat{P}_j^{(3)} \right| = 0.0191443.$$

Table 2: The example of partial sample determined by the sequence (c_n), defined by (3.1)

				$r = 2$, $NS = 5000$, $n = 10000$			
x	y	$G(x, y)$	$\widehat{P}_1$	$\widehat{P}_2$	$\widehat{P}_3$	$\widehat{P}_4$	$\widehat{P}_5$
1	0.8	0.5091564	0.5134	0.5056	0.5028	0.5144	0.4946
1	0.4	0.5915554	0.5898	0.5800	0.5970	0.5918	0.6026
1	0.2	0.6376282	0.6408	0.6324	0.6346	0.6422	0.6308
1	0.1	0.6619932	0.6468	0.6604	0.6568	0.6676	0.6696
2	1.6	0.2592403	0.2642	0.2590	0.2580	0.2494	0.2662
2	0.8	0.3499377	0.3536	0.3394	0.3446	0.3664 [1]	0.3456
2	0.4	0.4065697	0.4142	0.3950	0.4174	0.4122	0.4210
2	0.2	0.4382350	0.4310	0.4332	0.4340	0.4304	0.4290

The maximum absolute difference (that appears for the value in shaded cell of the table) is:

$$^1 \max_{x,y,j} \left| G(x, y) - \widehat{P}_j \right| = 0.0164623.$$

Table 3: The example of partial sample determined by the sequence (c_n), defined by (3.2)

$r = 2, NS = 5000, n = 10000$							
x	y	$G(x, y)$	$\widehat{P_1}$	$\widehat{P_2}$	$\widehat{P_3}$	$\widehat{P_4}$	$\widehat{P_5}$
1	0.8	0.5091564	0.5060	0.5170	0.4946	0.4996	0.5000
1	0.4	0.5915554	0.5922	0.6012	0.5964	0.5944	0.5930
1	0.2	0.6257840	0.6272	0.6282	0.6130	0.6202	0.6264
1	0.1	0.6257840	0.6236	0.6154	0.6278	0.6318	0.6300
2	1.6	0.2592403	0.2548	0.2628	0.2382 [1]	0.2716	0.2612
2	0.8	0.3499377	0.3620	0.3438	0.3570	0.3436	0.3626
2	0.4	0.3916056	0.3966	0.4042	0.3954	0.3954	0.3874
2	0.2	0.3916056	0.3978	0.3912	0.3836	0.3994	0.3772

The maximum absolute difference (that appears for the value in shaded cell of the table) is:

[1] $\max\limits_{x,y,j} \left| G(x, y) - \widehat{P_j} \right| = 0.0210403.$

5. Discussion and conclusions

1. The Theorem 2.1. shows that the limiting distribution of random vector, whose components are maximum of partial (aforementioned - specific, regular) sample and maximum of the whole sample, remains the same even if the indexes of observed random variables are, so to say, translated by one or two places (see Remark 2).

2. The non-random $0 - 1$-sequences $(c_n)_{n \in \mathbb{N}}$, defined by (3.1) and (3.2), that determine partial samples from negatively correlated uniform $AR(1)$ process in the Theorem 3.1. and the Theorem 3.2, have already been used in Theorem 3.1. [9], p. 1415, and in Theorem 2.1. [11], respectively, also for determining partial samples, but from ordinary uniform $AR(1)$ process. Comparing the new limit results for the random vector $(\widetilde{M}_n, M_n)$ with those previous ones, the following can be noted:

⋄ if (c_n) is defined by (3.1) and

  – the underlying random sequence (X_n) is negatively correlated uniform $AR(1)$ process - for all values x and y such that $0 < y \le x$, the limiting distribution of $(\widetilde{M}_n, M_n)$ is the one from the Theorem 3.1; the events $\left\{ \widetilde{M}_n \le 1 - x/n \right\}$ and $\left\{ M_n \le 1 - y/n \right\}$ are asymptotically independent

  – the underlying random sequence (X_n) is ordinary uniform $AR(1)$ process - the region in $\mathbb{R}^2$ that consists of all the points (x, y), such that $0 < y \le x$, splits in two parts - in this parts appear different expressions for the limiting probability; the events $\left\{ \widetilde{M}_n \le 1 - x/n \right\}$ and $\left\{ M_n \le 1 - y/n \right\}$ are asymptotically independent if $0 < x/r < y \le x$, and if $0 < y \le x/r$ these two events are asymptotically perfectly dependent;

⋄ if (c_n) is defined by (3.2) and

  – the underlying random sequence (X_n) is negatively correlated uniform $AR(1)$ process - the limiting distribution of $(\widetilde{M}_n, M_n)$ is the one from the Theorem 3.2; the events $\left\{ \widetilde{M}_n \le 1 - x/n \right\}$ and $\left\{ M_n \le 1 - y/n \right\}$ are asymptotically independent if $0 < x/r^2 < y \le x$, and if $0 < y \le x/r^2$ these two events are asymptotically perfectly dependent

– the underlying random sequence (X_n) is ordinary uniform $AR(1)$ process - the region in $\mathbb{R}^2$ that consists of all the points (x, y), such that $0 < x/r^2 < y \leq x$, splits in two parts - in this parts and in the region where $0 < y \leq x/r^2$ appear different expressions for the limiting probability; the events $\left\{\widetilde{M}_n \leq 1 - x/n\right\}$ and $\left\{M_n \leq 1 - y/n\right\}$ are asymptotically independent if $0 < x/r^2 < y \leq x$, and if $0 < y \leq x/r^2$ these two events are asymptotically perfectly dependent.

This is shown on the illustrations below. In the regions labeled with D_1 Chernick's limit results apply; in the regions labeled with D_3 the asymptotic perfect dependence is present while the asymptotic independence appears in the regions labeled with D_2; the region where the events are also asymptotically independent, but the expression for the limiting probability differs from the one in the region labeled with D_2, is labeled with D_4.
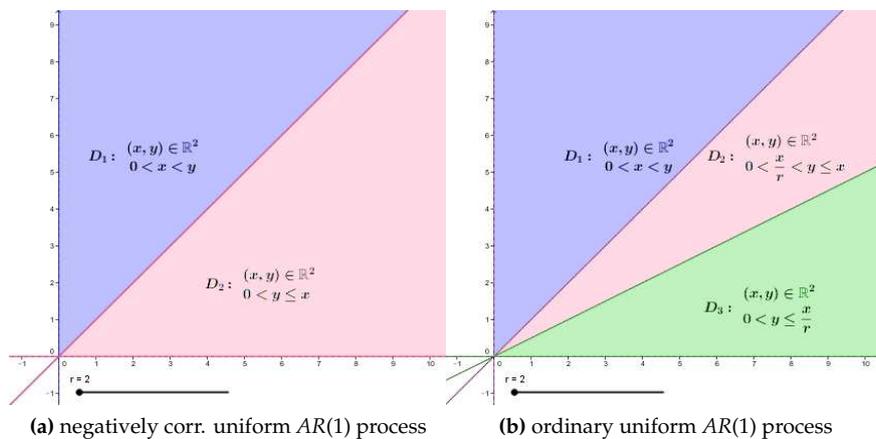

(a) negatively corr. uniform $AR(1)$ process   **(b)** ordinary uniform $AR(1)$ process

Figure 1: The example of partial sample determined by the sequence (c_n), defined by (3.1); $r = 2$

(a) negatively corr. uniform $AR(1)$ process   **(b)** ordinary uniform $AR(1)$ process
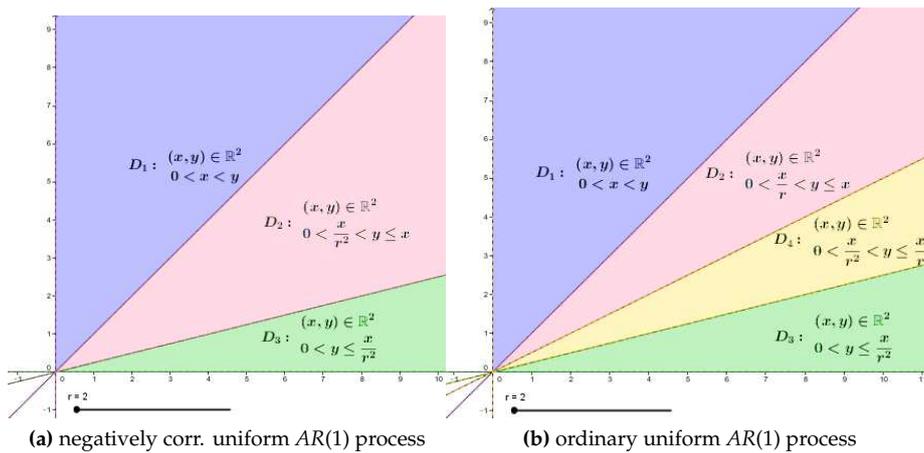
Figure 2: The example of partial sample determined by the sequence (c_n), defined by (3.2); $r = 2$

3. It turned out that computer simulations provide quite good approximations of the limiting probability $\lim_{n \to +\infty} P\left\{\widetilde{M}_n \leq 1 - x/n, M_n \leq 1 - y/n\right\}$. At the same time, writing functions (in statistical software) for this

purpose is not too complicated. Therefore, numerical simulations may prove very useful and, in the situations in which theoretical results are not available, even indispensable.

4. The question about determining (a class of) limiting distributions of the random vector $(\widetilde{M}_n, M_n)$, when partial sample of observed random variables from the uniform $AR(1)$ process is determined by the sequence (I_n) of indicator random variables instead by the non-random $0 - 1$-sequence (c_n), is still open. At this point just approximations of limiting probabilities can be made, using computer simulations.

Acknowledgements

References

[1] M. R. Chernick, *A limit theorem for the maximum of autoregressive processes with uniform marginal distributions*, Annals of Probability **9** (1981), 145-149

[2] M. R. Chernick, R. A. Davis, *Extremes in autoregressive processes with uniform marginal distributions*, Statistics and Probability Letters **1** (1982), 85-88

[3] R. A. Davis, *Maxima and minima of stationary sequences*, Annals of Probability **7** (1979), 453-460

[4] G. K. Grunwald, R. J. Hyndman, L. Tedesco, R. L. Tweedie, *A unified view of linear AR(1) models*, Clayton, Australia: Monash University (1997)

[5] M. R. Leadbetter, *On extreme values in stationary sequences*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **28** (1974), 289-303

[6] M. R. Leadbetter, *Extremes and local dependence in stationary sequences*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **65** (1983), 291-306

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

[8] R. M. Loynes, *Extreme values in uniformly mixing stationary stochastic process*, Annals of Mathematical Statistics **36** (1965), 993-999

[9] P. Mladenović, *Maximum of a partial sample in the uniform AR(1) processes*, Statistics and Probability Letters **79** (2009), 1414-1420

[10] P. Mladenović, V. Piterbarg, *On asymptotic distribution of maxima of complete and incomplete samples from stationary sequences*, Stochastic Processes and their Applications **116** (2006), 1977-1991

[11] P. Mladenović, L. Živadinović, *Uniform AR(1) processes and maxima on partial samples*, Communications in Statistics - Theory and Methods (2013), DOI: 10.1080/03610926.2013.786785

[12] I. Norros, *A storage model with self-similar input*, Queueing Systems **16** (1994), 387-396

[13] G. L. O'Brien, *Limit theorems for the maximum term of a stationary process*, Annals of Probability **2** (1974), 540-545