CONDITIONAL LEAST SQUARES ESTIMATION OF THE PARAMETERS OF HIGHER ORDER RANDOM ENVIRONMENT INAR MODElS

Petra N. Laketa, Aleksandar S. Nastić

Abstract. Two different random environment INAR models of higher order, precisely RrNGINARmax(p) and RrNGINAR$_1$(p), are presented as a new approach to modeling non-stationary nonnegative integer-valued autoregressive processes. The interpretation of these models is given in order to better understand the circumstances of their application to random environment counting processes. The estimation statistics, defined using the Conditional Least Squares (CLS) method, is introduced and the properties are tested on the replicated simulated data obtained by RrNGINAR models with different parameter values. The obtained CLS estimates are presented and discussed.
Keywords: Random environment; INAR(p); RrNGINAR; negative binomial thinning; geometric marginals; conditional least squares.

1. Introduction

One of the latest and most significant approaches to the modeling of count processes was designed by introducing integer-valued autoregressive (INAR) models almost simultaneously by [7] and [2]. This breakthrough in the analysis of integer-valued time series was a consequence of using a new thinning operator. Namely, the deterministic part of a process random variable was calculated using the realization of a Bernoulli counting sequence limited by the process realization in the preceding moment. This way of modeling was simply more natural and intuitively justified, so it led to much better results in fitting the counting processes than other models known at that time. This was followed by many modifications and generalizations. Some authors considered the thinning operator ([3], [6], [17, 18] and [13]), while others focused on marginal distributions ([8], [1], [4] and [5]). Also, as an alternative to the NGINAR(1) process from [13], a zero-inflated NGINAR(1) process was considered, which is given in [14]. In order to obtain more suitable models for processes of higher correlation between distant elements, INAR models of higher

order were introduced. The most operative approach was developed in [16], where X_n as a process value at time n was defined using p possible preceding random values X_{n-i}, for $i \in \{1, 2, \ldots, p\}$, each with a certain probability. This inspired the construction of models presented in [10] and [9]. So, the evolution of INAR models continued.

All the models listed above corresponded only to stationary counting processes. In many applications, this was found as a frequent limitation. Recently, random environment INAR models, whose marginal distribution depends on random circumstances, have been introduced (more details about these models are given below). However, the conditional least squares (CLS) estimators of random environment INAR models parameters have not been considered so far. Therefore, in this paper, we obtain CLS estimators and test them on the simulated values from the corresponding random INAR model.

Using as a starting point some ideas from [15], [11] defined the r-states random environment integer-valued autoregressive process of order 1, denoted as $(RrINAR(1))$. It is given by

$$X_n(Z_n) = \sum_{i=1}^{X_{n-1}(Z_{n-1})} U_i + \varepsilon_n(Z_{n-1}, Z_n), \ n \in \mathbb{N},$$

where

$$X_n(Z_n) = \sum_{z=1}^{r} X_n(z) I_{\{Z_n=z\}},$$

$$\varepsilon_n(Z_{n-1}, Z_n) = \sum_{z_1=1}^{r} \sum_{z_2=1}^{r} \varepsilon_n(z_1, z_2) I_{\{Z_{n-1}=z_1, Z_n=z_2\}},$$

$\{U_i\}$, $i \in \mathbb{N}$, is a counting sequence of independent and identically distributed (i.i.d.) random variables generating a thinning operator, $\{Z_n\}$, $n \in \mathbb{N}_0$ is an r-states random environment process defined as a Markov chain taking values in $E_r = \{1, 2, \ldots, r\}$. Further, $\{\varepsilon_n(i, j)\}$, $n \in \mathbb{N}_0$, $i, j \in E_r$, are sequences of i.i.d. random variables, for which $\{Z_n\}$, $\{\varepsilon_n(1, 1)\}$, $\{\varepsilon_n(1, 2)\}$, ..., $\{\varepsilon_n(r, r)\}$, are mutually independent, for all $n \in \mathbb{N}_0$, and Z_m and $\varepsilon_m(i, j)$ are independent of $X_n(l)$, for $n < m$ and any $i, j, l \in E_r$. In order to obtain more efficient INAR modeling, a new random environment INAR(1) process with one-step-ahead determined marginal distribution was introduced in [11]. As can be seen, this process is non-stationary, which makes it more applicable in practice. Adapting the process to more dynamical counting data, the authors specify geometric marginals and the negative binomial thinning operator $\alpha*$, which was utilized for construction of the NGINAR(1) model introduced in [13]. This resulted in the r-states random environment $INAR(1)$ process with determined (z_n–guided) geometric marginal distribution based on the negative binomial thinning operator $(RrNGINAR(1))$ given by

$$(1.1) \qquad X_n(z_n) = \alpha * X_{n-1}(z_{n-1}) + \varepsilon_n(z_{n-1}, z_n), \ n \in \mathbb{N},$$

where $\alpha \in (0,1)$, the counting sequence $\{U_i\}$, $i \in \mathbb{N}$, incorporated in $\alpha*$, makes a sequence of i.i.d. random variables with the probability mass function (pmf) given by

$$P(U_i = u) = \frac{\alpha^u}{(1+\alpha)^{u+1}} , \ u \in \mathbb{N}_0,$$

and finally the process pmf is defined as

$$(1.2) \qquad P(X_n(z_n) = x) = \frac{\mu_{z_n}^x}{(1+\mu_{z_n})^{x+1}} , \ x \in \mathbb{N}_0,$$

where $\mu_{z_n} \in \{\mu_1, \mu_2, \ldots, \mu_r\}$ and $r \in \mathbb{N}$.

1.1. Interpretation of the random environment INAR processes of higher order

Continuing the efforts towards the optimal fitting of the counting processes, models of higher order were introduced in [12]. Two approaches were used, which we discuss in what follows.

Definition 1. *Let z_n be the realization of a random environment process $\{Z_n\}$ at the moment $n \geqslant 0$. We say that $\{X_n(z_n)\}_{n \in \mathbb{N}_0}$ is an INAR process with r-states random environment guided geometric marginals based on the negative binomial thinning operator of maximal order p ($RrNGINARmax(p)$), $p \in \mathbb{N}$, if the random variable $X_n(z_n)$ is defined as*

$$(1.3) \qquad X_n(z_n) = \begin{cases} \alpha * X_{n-1}(z_{n-1}) + \varepsilon_n(z_{n-1}, z_n), & w.p. \ \phi_1^{(p_n)}, \\ \alpha * X_{n-2}(z_{n-2}) + \varepsilon_n(z_{n-2}, z_n), & w.p. \ \phi_2^{(p_n)}, \\ \quad \vdots & \quad \vdots \\ \alpha * X_{n-p_n}(z_{n-p_n}) + \varepsilon_n(z_{n-p_n}, z_n), & w.p. \ \phi_{p_n}^{(p_n)}, \end{cases}$$

for $n \geqslant 1$, where

$$p_n = \begin{cases} p, & p_n^* \geq p, \\ p_n^*, & p_n^* < p, \end{cases}$$

$p_n^* = \max\{i \in \{1, 2, \ldots, n\} : z_{n-1} = z_{n-2} = \cdots = z_{n-i}\}$ *and the following conditions are satisfied:*

1. $\phi_i^{(p_n)} \geqslant 0$, $i \in \{1, 2, \ldots, p_n\}$, $\sum_{i=1}^{p_n} \phi_i^{(p_n)} = 1$,

2. $\alpha \in (0,1)$ and the counting sequence $\{U_i\}_{i \in \mathbb{N}}$ of the negative binomial thinning operator $\alpha*$ has pmf $P(U_i = u) = \frac{\alpha^u}{(1+\alpha)^{u+1}}$, $u \in \{0, 1, 2, \ldots\}$,

3. $P(X_n(z_n) = x) = \frac{\mu_{z_n}^x}{(1+\mu_{z_n})^{x+1}}$, $x \in \{0, 1, 2, \ldots\}$, where $\mu_{z_n} \in \{\mu_1, \mu_2, \ldots, \mu_r\}$, $\mu_i > 0$, $i \in \{1, 2, \ldots, r\}$ and $r \in \mathbb{N}$ is the number of states of the random environment process $\{Z_n\}$,

4. for fixed $i, j \in E_r = \{1, 2, .., r\}$, $\{\varepsilon_n(i, j)\}_{n \in \mathbb{N}}$ is a sequence of i.i.d. random variables,

5. $\{Z_n\}, \{\varepsilon_n(1,1)\}, \{\varepsilon_n(1,2)\}, \ldots, \{\varepsilon_n(r,r)\}$ *are mutually independent sequences of random variables,*

6. $X_n(l)$ *is independent of* Z_m *and* $\varepsilon_m(i,j)$, *for* $0 \leq n < m$ *and any* $i, j, l \in E_r$.

Definition 2. *Let* z_n *be the realization of a random environment process* $\{Z_n\}$ *at the moment* $n \geqslant 0$. *We say that* $\{X_n(z_n)\}_{n \in \mathbb{N}_0}$ *is an INAR process with* r-*states random environment guided geometric marginals based on the negative binomial thinning operator of order* p ($RrNGINAR_1(p)$) *if the random variable* $X_n(z_n)$ *is defined as*

$$
X_n(z_n) = \begin{cases}
\alpha * X_{n-1}(z_{n-1}) + \varepsilon_n(z_{n-1}, z_n), & w.p.\ \phi_1^{(p_n)}, \\
\alpha * X_{n-2}(z_{n-2}) + \varepsilon_n(z_{n-2}, z_n), & w.p.\ \phi_2^{(p_n)}, \\
\quad\vdots & \quad\vdots \\
\alpha * X_{n-p_n}(z_{n-p_n}) + \varepsilon_n(z_{n-p_n}, z_n), & w.p.\ \phi_{p_n}^{(p_n)},
\end{cases}
$$

for $n \geqslant 1$, *where*

$$
p_n = \begin{cases}
p, & p_n^* \geq p, \\
1, & p_n^* < p,
\end{cases}
$$

$p_n^* = \max\{i \in \{1, 2, \ldots, n\} : z_{n-1} = z_{n-2} = \cdots = z_{n-i}\}$ *and conditions* $1 - 6$ *from Definition* 1 *are satisfied.*

Since the distribution parameter values of the processes may vary over time, it could happen that each of the equations (1.3) and (2), at a certain moment, contains differently distributed X_n random variables, which would make the models pretty complicated to work with. In order to avoid this, each of these models is defined with the ability of changing the number of possibilities (possible expressions) on the right side of the equation. So, the process introduced by Definition 1 has a fully variable order, possibly taking all the values from 1 to p. When the process random state changes, then the order of the process becomes equal to 1 and then starts rising successively, until it reaches p (when the process takes shape of the model of fixed order), or the state changes again. However, for the process given by Definition 2, the order takes one of two possible values. Namely, every time the state changes, the order becomes equal to 1 and it remains the same until there is a series of enough (p) previous process elements corresponding to the same state, when the order becomes equal to p. By virtue of these qualities, these processes are the most suitable for counting, for example, some elements of the observed unstable system or some random events recorded in a variable environment. In each case, certain area conditions or random circumstances may affect the dynamics of the interactions in the observed populations, which further affects the values of counts. So, the finite number of possible combinations of circumstances in which the population is observed is represented by the finite number (r) of random states and is modeled by the Markov process $\{Z_n\}$. Its realization $\{z_n\}$ directly determines the value of the selected marginal distribution. Hence, while being in the same state z_n, the process behaves as a stationary one with the marginal parameter value μ_{z_n}.

Nevertheless, its non-stationarity comes from changing its mean parameter value μ_{z_n}, which is directly guided by $\{z_n\}$. So, the counting process is basically piece-by-piece stationary, where each piece is as long as the random process $\{Z_n\}$ remains in the same state, i.e. the population circumstances do not change.

2. Conditional least squares estimators

Let $\{X_n(z_n)\}$ be the RrNGINARmax(p) or RrNGINAR$_1$(p) time series model. In order to apply Theorem 2 from [12] we have to suppose conditions from that theorem. Let $\mu_1 > 0$, $\mu_2 > 0$, ..., $\mu_r > 0$ and let us suppose that $0 \leq \alpha \leq \min\left\{\frac{\mu_l}{1+\mu_k}, k, l \in E_r\right\}$, $z_n = j$ and $z_{n-1} = i$, for $i, j \in E_r$. Now, recalling the mentioned theorem, the conditional expectation of the random variable X_n for given $X_{n-1}, X_{n-2}, ..., X_{n-p_n}$ is

$$E(X_n|H_{n-1}) = \mu_j - \alpha\mu_i + \alpha \sum_{l=1}^{p_n} \phi_l^{(p_n)} X_{n-l},$$

where H_{n-1} represents σ-algebra generated by $X_{n-1}, X_{n-2}, ...$. Now, if we define new parameters as $\theta_l^{(z_n)} = \alpha\phi_l^{(p_n)}$, for $l \in \{1, 2, ..., p_n\}$, then $\alpha = \sum_{l=1}^{p_n} \theta_l^{(p_n)}$ and consequently

$$
\begin{aligned}
E(X_n|H_{n-1}) &= \mu_j - \alpha\mu_i + \theta_1^{(p_n)} X_{n-1} + \theta_2^{(p_n)} X_{n-2} + ... + \theta_{p_n}^{(p_n)} X_{n-p_n} \\
&= \mu_j - \sum_{l=1}^{p_n} \theta_l^{(p_n)} \mu_i + \theta_1^{(p_n)} X_{n-1} + \theta_2^{(p_n)} X_{n-2} + ... + \theta_{p_n}^{(p_n)} X_{n-p_n}.
\end{aligned}
$$

Let $k \in E_r$, $p_n = p$ and $J_k = \{n \in \mathbb{N}|X_n, X_{n-1}, ..., X_{n-p_k} \in U^{(k)}\}$, where $U^{(k)}$ represents the process subsample which consists of all the elements corresponding to the same state k. In conducting the conditional least squares (CLS) estimation, the aim is to minimize the following sum of squares
(2.1)

$$Q_N^{(k)}(\mathbf{a}) = \sum_{n \in J_k} \left(X_n - \mu_j - \sum_{l=1}^{p} \theta_l^{(p)} \mu_i - \theta_1^{(p)} X_{n-1} - \theta_2^{(p)} X_{n-2} - ... - \theta_p^{(p)} X_{n-p}\right)^2,$$

with respect to the vector $\mathbf{a} = (\theta_1^{(p)}, \theta_2^{(p)}, ..., \theta_p^{(p)}, \mu_k)'$. This is achieved by solving the system $\frac{\partial Q_N}{\partial \theta_1^{(p)}} = 0$, $\frac{\partial Q_N}{\partial \theta_2^{(p)}} = 0, ..., \frac{\partial Q_N}{\partial \theta_p^{(p)}} = 0$, $\frac{\partial Q_N}{\partial \mu_k} = 0$. Since the summation in the previous expression is over the set J_k, it holds that $X_n, X_{n-1}, ..., X_{n-p} \in U^{(k)}$ and $z_n = z_{n-1} = ... = z_{n-p} = k$. So, considering the process on the subsample $U^{(k)}$, we deal with the CGINAR(p) model introduced in [10]. Therefore, the corresponding results and equations obtained for the CGINAR(p) model can be used here. Thus, we have

(2.2)
$$\mu_{k,p} = \frac{1}{1 - \sum_{i=1}^{p} \theta_{i,p}^{(k)}} \left(\overline{X}^{(0)} - \sum_{i=1}^{p} \theta_{i,p}^{(k)} \overline{X}^{(i)}\right),$$

where

$$\overline{X}^{(i)} = \frac{1}{|J_k|} \sum_{n \in J_k} X_{n-j}, \quad j \in \{0, 1, ..., p\}.$$

Replacing (2.2) in (2.1) the system becomes

$$(2.3) \qquad \sum_{j=1}^{p} \theta_j^{(p)} \widehat{\gamma}^*(|l - j|) = \widehat{\gamma}^*(l), \quad l = 1, 2, ..., p,$$

where

$$\widehat{\gamma}^*(|l - j|) = \frac{1}{|J_k|} \sum_{n \in J_k} X_{n-l} X_{n-j} - \overline{X}^{(l)} \overline{X}^{(j)}.$$

Solving it gives us $\widehat{\theta}_j^{(p)} = \frac{D_j^*}{D^*}$, $j = 1, 2, ..., p$, where D_j^* and D^* are the appropriate determinants from Kramer's method. Substituting the last equations in (2.2) we get

$$\widehat{\mu}_k^{CLS} = \frac{1}{1 - \sum_{i=1}^{p} \frac{D_i^*}{D^*}} \left(\frac{1}{|J_k|} \sum_{n \in J_k} X_n - \sum_{j=1}^{p} \frac{D_i^*}{D^*} \cdot \frac{1}{|J_k|} \sum_{n \in J_k} X_{n-j} \right).$$

Therefore,

$$(2.4) \qquad \widehat{\alpha}^{(k),CLS} = \frac{\sum_{j=1}^{p} D_j^*}{D^*},$$

$$\widehat{\phi}_{i,p}^{(k),CLS} = \frac{D_i^*}{\sum_{j=1}^{p} D_j}, \quad i \in \{1, 2, ..., p\}.$$

Finally, using the preceding results for each $k \in \{1, 2, ..., r\}$, it is only left to calculate the weighted thinning parameter and the weighted probabilities, respectively, as

$$(2.5) \qquad \widehat{\alpha}^{CLS} = \frac{\sum_{k=1}^{r} |J_k| \widehat{\alpha}^{(k),CLS}}{\sum_{k=1}^{r} |J_k|},$$

$$(2.6) \qquad \widehat{\phi}_{i,p}^{CLS} = \frac{\sum_{k=1}^{r} |J_k| \widehat{\phi}_{i,p}^{(k),CLS}}{\sum_{k=1}^{r} |J_k|},$$

which represent the required estimators.

Based on Lemma 6, from [10], the estimators $\widehat{\alpha}^{CLS}$, $\widehat{\mu}_k^{CLS}$ and $\widehat{\phi}_{i,p}^{CLS}$ are asymptotically almost surely equivalent to the corresponding Yule-Walker estimators. So, the strong consistency of the Yule-Walker estimators, proved in [12], implies the strong consistency of the here observed CLS estimators.

3. Simulation results

In this section we try to confirm the correctness of the introduced CLS estimators. With that in mind, we have simulated 100 replicates of realizations of the processes RrNGINARmax(p) and RrNGINAR$_1$(p), each of size 10000. Parameter values for $\alpha, p, r, \boldsymbol{\mu}, \mathbf{p}_{mat}$ and $\boldsymbol{\phi}$ are chosen and then the corresponding models are simulated. The transition probability matrix of the random environment process is denoted by $\mathbf{p}_{mat}$, and $\boldsymbol{\mu}$ is a vector of means. In the case of RrNGINARmax(p) model, the p_nth row, $p_n \in \{2, \ldots, p\}$, of the matrix $\boldsymbol{\phi}$ contains probabilities $\phi_i^{(p_n)}$, $i \in \{1, 2, \ldots, p_n\}$ and in the case of RrNGINAR$_1$(p) model, the last row represents probabilities $\phi_i^{(p)}$, $i \in \{1, 2, \ldots, p\}$. The simulated realization of random environment process, $\{z_n\}$, is obtained using $\mathbf{p}_{mat}$ and then the sequence $\{p_n\}$ is specified based on the corresponding definition. We have considered six different cases of chosen parameter values and presented all the results in the appropriate tables. Also, we have decided for the same parameter values as in the case of Yule-Walker parameter estimation discussed in [12]. There are three tables. In the first one we have $p = 2, r = 2$, in the second $p = 3, r = 2$ and in the last $p = 3, r = 3$. In the first table, for $r = p = 2$ we considered different choices of other parameters. The larger α gives better estimates for probabilities $\phi_i^{(p_n)}$. The higher diagonal values of p_{mat} ensures longer subsamples and, consequently, better results. Also, the higher values of p and r implies more subsamples and, therefore, a larger number of them and smaller sizes, which gives us worse results for the same samples size. Finally, for the small sample sizes it is possible to have very small subsamples and to get bad results.

Table 3.1: $r = 2$, $p = 2$

True values $\boldsymbol{\mu} = (1, 2)$, $\alpha = 0.3$, $\boldsymbol{\phi} = \begin{bmatrix} 1 & 0 \\ 0.6 & 0.4 \end{bmatrix}$, $\mathbf{p}_{mat} = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$

n	$\hat{\mu}_1^{CLS}$	$\hat{\mu}_2^{CLS}$	$\hat{\alpha}^{CLS}$	$\hat{\phi}_{2,1}^{CLS}$	$\hat{\phi}_{2,2}^{CLS}$	$\hat{\alpha}^{CLS}$	$\hat{\phi}_{2,1}^{CLS}$	$\hat{\phi}_{2,2}^{CLS}$
500	1.0100	1.9964	0.3359	0.6900	0.3100	0.2963	0.6207	0.3793
SE	0.1195	0.2214	0.1751	0.2451	0.2451	0.1557	0.3836	0.3836
1000	1.0119	1.9976	0.3307	0.6248	0.3752	0.2896	0.6127	0.3873
SE	0.0797	0.1373	0.1176	0.1364	0.1364	0.1187	0.1229	0.1229
5000	1.0024	2.0047	0.3026	0.6048	0.3952	0.2978	0.5984	0.4016
SE	0.0354	0.0600	0.0478	0.0595	0.0595	0.0565	0.0579	0.0579
10000	1.0016	2.0072	0.3020	0.5990	0.4010	0.2956	0.6029	0.3971
SE	0.0249	0.0429	0.036	0.0386	0.0386	0.0406	0.0393	0.0393

Table 3.2: $r = 2$, $p = 2$

True values $\boldsymbol{\mu} = (1,2)$, $\alpha = 0.15$, $\boldsymbol{\phi} = \begin{bmatrix} 1 & 0 \\ 0.5 & 0.5 \end{bmatrix}$, $\mathbf{p}_{mat} = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$

n	$\widehat{\mu}_1^{CLS}$	$\widehat{\mu}_2^{CLS}$	$\widehat{\alpha}^{CLS}$	$\widehat{\phi}_{2,1}^{CLS}$	$\widehat{\phi}_{2,2}^{CLS}$	$\widehat{\alpha}^{CLS}$	$\widehat{\phi}_{2,1}^{CLS}$	$\widehat{\phi}_{2,2}^{CLS}$
500	0.99609	2.0089	0.1735	-0.3945	1.3945	0.1679	0.2434	0.7566
SE	0.1014	0.1588	0.1451	15.064	15.064	0.1389	2.2148	2.2148
1000	0.9977	2.0143	0.1475	0.5418	0.4582	0.1547	0.3885	0.6115
SE	0.0666	0.1259	0.0966	0.4849	0.4849	0.0878	0.9751	0.9751
5000	1.0045	1.9993	0.1505	0.4970	0.5030	0.1508	0.4893	0.5107
SE	0.0360	0.0618	0.037	0.1008	0.1008	0.0384	0.1113	0.1113
10000	1.0024	1.9981	0.1494	0.5039	0.4961	0.1514	0.4964	0.5036
SE	0.0252	0.0425	0.027	0.0702	0.0702	0.0297	0.0682	0.0682

Table 3.3: $r = 2$, $p = 2$

True values $\boldsymbol{\mu} = (1,2)$, $\alpha = 0.3$, $\boldsymbol{\phi} = \begin{bmatrix} 1 & 0 \\ 0.6 & 0.4 \end{bmatrix}$, $\mathbf{p}_{mat} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$

n	$\widehat{\mu}_1^{CLS}$	$\widehat{\mu}_2^{CLS}$	$\widehat{\alpha}^{CLS}$	$\widehat{\phi}_{2,1}^{CLS}$	$\widehat{\phi}_{2,2}^{CLS}$	$\widehat{\alpha}^{CLS}$	$\widehat{\phi}_{2,1}^{CLS}$	$\widehat{\phi}_{2,2}^{CLS}$
500	0.9957	2.0003	0.3322	0.7480	0.252	0.3075	0.6919	0.3081
SE	0.0996	0.1919	0.1192	2.8876	2.8876	0.0988	0.5416	0.5416
1000	0.9928	2.0004	0.3108	0.6208	0.3792	0.3036	0.6556	0.3444
SE	0.0732	0.1345	0.0892	0.5292	0.5292	0.0837	0.3017	0.3017
5000	1.0019	2.0008	0.3037	0.5973	0.4027	0.2976	0.5931	0.4069
SE	0.0414	0.06231	0.0380	0.0894	0.0894	0.0387	0.0818	0.0818
10000	0.9993	2.0030	0.3020	0.5904	0.4096	0.2985	0.5929	0.4071
SE	0.0245	0.0418	0.0264	0.0702	0.0702	0.0284	0.0633	0.0633

Table 3.4: $r = 2$, $p = 2$

True values $\boldsymbol{\mu} = (4,5)$, $\alpha = 0.5$, $\boldsymbol{\phi} = \begin{bmatrix} 1 & 0 \\ 0.6 & 0.4 \end{bmatrix}$, $\mathbf{p}_{mat} = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$

n	$\widehat{\mu}_1^{CLS}$	$\widehat{\mu}_2^{CLS}$	$\widehat{\alpha}^{CLS}$	$\widehat{\phi}_{2,1}^{CLS}$	$\widehat{\phi}_{2,2}^{CLS}$	$\widehat{\alpha}^{CLS}$	$\widehat{\phi}_{2,1}^{CLS}$	$\widehat{\phi}_{2,2}^{CLS}$
500	3.9973	5.0266	0.5277	0.6094	0.3906	0.5151	0.6108	0.3892
SE	0.4207	0.5014	0.1589	0.1595	0.1595	0.1364	0.1420	0.1420
1000	3.9776	5.0367	0.5166	0.5973	0.4027	0.5109	0.5927	0.4073
SE	] 0.3344	0.3312	0.1009	0.0975	0.0975	0.1100	0.0935	0.0935
5000	3.9923	5.0179	0.4960	0.5944	0.4056	0.5031	0.5867	0.4133
SE	0.1340	0.1635	0.0559	0.0417	0.0417	0.0638	0.0513	0.0513
10000	3.9947	5.0157	0.4997	0.5931	0.4069	0.5050	0.5935	0.4065
SE	0.0985	0.1157	0.0398	0.0285	0.0285	0.0428	0.0391	0.0391

Table 3.5: $r = 2$, $p = 3$

True values $\boldsymbol{\mu} = (1, 2)$, $\alpha = 0.3$, $\boldsymbol{\phi} = \begin{bmatrix} 1 & 0 & 0 \\ 0.6 & 0.4 & 0 \\ 0.5 & 0.3 & 0.2 \end{bmatrix}$, $\mathbf{p}_{mat} = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$

n	$\hat{\mu}_1^{CLS}$	$\hat{\mu}_2^{CLS}$	$\hat{\alpha}^{CLS}$	$\hat{\phi}_{2,1}^{CLS}$	$\hat{\phi}_{2,2}^{CLS}$	$\hat{\phi}_{3,1}^{CLS}$	$\hat{\phi}_{3,2}^{CLS}$	$\hat{\phi}_{3,3}^{CLS}$	$\hat{\alpha}^{CLS}$	$\hat{\phi}_{3,1}^{CLS}$	$\hat{\phi}_{3,2}^{CLS}$	$\hat{\phi}_{3,3}^{CLS}$
500	1.0025	1.9650	0.3139	0.5793	0.4207	0.6083	0.2945	0.0972	0.3022	0.0761	0.3122	0.6118
SE	0.1186	0.1981	0.171	1.5608	1.5608	0.7284	0.2673	0.7066	0.1032	3.5502	1.1115	3.6095
1000	1.0023	2.0057	0.3094	0.6659	0.3341	0.5140	0.3125	0.1735	0.2958	0.5485	0.2607	0.1908
SE	0.0808	0.1338	0.0988	0.5715	0.5715	0.1731	0.1701	0.1455	0.0735	0.2608	0.2818	0.1948
5000	0.9951	2.0011	0.3058	0.6155	0.3845	0.4902	0.3026	0.2072	0.2985	0.4941	0.3095	0.1964
SE	0.0335	0.0652	0.0477	0.1239	0.1239	0.0669	0.0610	0.0677	0.0347	0.0729	0.0751	0.0591
10000	0.9995	2.0019	0.3009	0.5924	0.4076	0.4970	0.2972	0.2058	0.2983	0.4943	0.3113	0.1944
SE	0.0257	0.0461	0.0329	0.0787	0.0787	0.0506	0.0460	0.0514	0.0248	0.0500	0.0503	0.0434

Table 3.6: $r = 3$, $p = 3$

True values $\boldsymbol{\mu} = (1, 1.5, 2)$, $\alpha = 0.3$, $\boldsymbol{\phi} = \begin{bmatrix} 1 & 0 & 0 \\ 0.6 & 0.4 & 0 \\ 0.5 & 0.3 & 0.2 \end{bmatrix}$, $\mathbf{p}_{mat} = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0.1 & 0.2 & 0.7 \end{bmatrix}$

n	$\hat{\mu}_1^{CLS}$	$\hat{\mu}_2^{CLS}$	$\hat{\mu}_3^{CLS}$	$\hat{\alpha}^{CLS}$	$\hat{\phi}_{2,1}^{CLS}$	$\hat{\phi}_{2,2}^{CLS}$	$\hat{\phi}_{3,1}^{CLS}$	$\hat{\phi}_{3,2}^{CLS}$	$\hat{\phi}_{3,3}^{CLS}$	$\hat{\alpha}^{CLS}$	$\hat{\phi}_{3,1}^{CLS}$	$\hat{\phi}_{3,2}^{CLS}$	$\hat{\phi}_{3,3}^{CLS}$
500	0.9749	1.5187	2.0151	0.3331	0.9811	0.0189	0.9537	0.3139	-0.2675	0.3027	0.5286	0.2473	0.2241
SE	0.1527	0.1659	0.2519	0.1341	2.1042	2.1042	2.8729	1.6246	4.3460	0.0990	0.5116	0.8627	0.7784
1000	0.9886	1.5182	1.9855	0.3143	0.7626	0.2374	0.4260	0.6499	-0.0759	0.3010	0.5560	0.2774	0.1666
SE	0.1051	0.1161	0.1819	0.1000	0.8830	0.8830	0.7415	2.3627	1.9789	0.0638	0.5080	0.4014	0.5072
5000	1.0025	1.5043	1.9918	0.3047	0.6003	0.3997	0.5133	0.2923	0.1944	0.3018	0.5003	0.3050	0.1947
SE	0.0516	0.0572	0.0785	0.0458	0.1328	0.1328	0.1031	0.0999	0.0982	0.0271	0.0970	0.1020	0.1070
10000	1.0038	1.4999	1.9961	0.2988	0.5984	0.4016	0.4998	0.2970	0.2032	0.3032	0.4955	0.3087	0.1958
SE	0.0335	0.0390	0.0562	0.0290	0.0874	0.0874	0.0572	0.0635	0.0561	0.0191	0.0714	0.0678	0.0629

4. Conclusion

Varying the sizes of the simulated samples, we have noticed quite a similar behavior of the here obtained estimates compared to those obtained by the Yule-Walker statistics, thus confirming the asymptotical equivalence mentioned at the end of Section 2. Also, the convergence of the obtained estimations to the real parameter values, which is easy to observe in all the following tables, confirms the strong consistency of the conditional least squares estimators.

Some negative values for $\widehat{\phi}_{3,3}^{CLS}$ are obtained when the sample size is small, which is induced by the model properties. Namely, $\phi_{3,3}$ represents the probability that $X_n(z_n)$ will depend on $X_{n-3}(z_{n-3})$. In this case $\phi_{3,3} = 0.2$, so the portion of the data from which we can obtain $\widehat{\phi}_{3,3}^{CLS}$ is approximately 0.2. However, another "reduction" of the data occurs since all estimators are defined on the subsamples with the same state. So, in this case, the subsample is too small to get a good result. By enlarging the data size, $\widehat{\phi}_{3,3}^{CLS}$ converge to the true value. This effect of the small subsample also results in the large values of standard deviations for the small sample size.

Acknowledgements

REFERENCES

1. AL-OSH, M.A., ALY, E.E.A.A.: *First order autoregressive time series with negative binomial and geometric marginals*, Commun. Statist. Theory Meth.(1992) 21, 2483–2492.

2. AL-OSH, M.A., ALZAID, A.A.: *First-order integer-valued autoregressive (INAR(1)) process*, J. Time Ser. Anal. (1987) 8, 261–275.

3. ALY, E.E.A.A., BOUZAR, N.: *On Some Integer-Valued Autoregressive Moving Average Models*, Journal of Multivariate Analysis (1994) 50, 132–151.

4. ALZAID, A. A., AL-OSH, M.A.: *Some autoregressive moving average processes with generalized Poisson marginal distributions*, Ann. Inst. Statist. Math. (1993) 45, 223–232.

5. BAKOUCH, H.S., RISTIĆ, M.M. *Zero Truncated Poisson Integer Valued AR(1) Model*, Metrika (2010) 72(2), 265-280.

6. LATOUR, A.: *Existence and stochastic structure of a non-negative integer-valued autoregressive process*, J. Time Ser. Anal. (1998) 19, 439–455

7. McKenzie, E. : *Some simple models for discrete variate time series*, Water Resour. Bull.(1985) 21, 645–650.

8. McKenzie, E.: *Autoregressive moving-average processes with negative binomial and geometric distributions*, Adv. Appl. Prob. (1986) 18, 679–705.

9. Nastić, A.S., Ristić, M.M.: *Some geometric mixed integer-valued autoregressive (INAR) models*, Statistics and Probability Letters (2012) 82, 805-811.

10. Nastić, A.S., Ristić, M.M., Bakouch, H.S.: *A combined geometric INAR(p) model based on negative binomial thinning*, Mathematical and Computer Modelling (2012) 55, 1665–1672.

11. Nastić, A.S., Laketa, P.N., Ristić, M.M.: *Random Environment Integer-Valued Autoregressive process*, J. Time Ser. Anal.(2016) 37, 267–287.

12. Nastić, A.S., Laketa, P.N., Ristić, M.M.: *Random environment INAR models of higher order*, RevStat: Statistical Journal (2017)Volume 17, Number 1, January 2019

13. Ristić, M.M., Bakouch, H.S., Nastić, A.S.: *A new geometric first-order integer-valued autoregressive (NGINAR(1)) process*, J. Stat. Plan. Inference (2009) 139, 2218–2226.

14. Ristić, M.M., Bourguignon, M., Nastić, A.S.: *Zero-Inflated NGINAR(1) process*, J. Communications in Statistics - Theory and Methods (2018)

15. Tang, M., Wang, Y.: *Asymptotic Behaviour of Random Coefficient INAR Model under Random Environment defined by Difference Equation, Advances in Difference Equations*, 2014(1), 1–9.

16. Weiss, C.H.: *The combined INAR(p) models for time series of counts*, Statist. Probab. Lett. (2008) 72, 1817–1822.

17. Zheng, H., Basawa, I.V., Datta, S.: *Inference for pth-order random coefficient integer-valued autoregressive processes*, J. Time Ser. Anal. (2006) 27, 411–440.

18. Zheng, H., Basawa, I.V., Datta, S.: *First-order random coefficient integer-valued autoregressive processes*, J. Stat. Plann. Inf. (2007) 137, 212–229.

Petra N. Laketa
University of Niš,
Faculty of Sciences and Mathematics
Department of Mathematics
18000 Niš, Serbia
petra.laketa@pmf.edu.rs

Aleksandar S. Nastić
University of Niš,
Faculty of Sciences and Mathematics
Department of Mathematics
18000 Niš, Serbia
anastic78@gmail.com