Statistics and Its Interface Volume 11 (2018) 631–639

Modelling time series of counts with deflation or inflation of zeros

Marcelo Bourguignon

In this paper, we introduce a first order non-negative integer valued autoregressive process with zero-modified geometric innovations based on the binomial thinning operator. This new model will enable one to tackle the problem of deflation or inflation of zeros inherent in the analysis of integer-valued time series data, and contains the INARG(1) model [6] as a particular case. The main properties of the model are derived, such as mean, variance, autocorrelation function, transition probabilities and zero probability. The methods of conditional maximum likelihood, Yule-Walker and conditional least squares are used for estimating the model parameters. A Monte Carlo experiment is conducted to evaluate the performances of these estimators in finite samples. The proposed model is fitted to time series of emergency counts department of a children's hospital and of drugs reselling criminal acts counts illustrating its capabilities in challenging cases of deflated and inflated count data.

AMS 2000 subject classifications: 62M10.
Keywords and phrases: Estimation, INAR(1) process, Integer-valued time series, Zero-modified geometric distribution.

1. INTRODUCTION

McKenzie [12] proposed the first order non-negative integer valued autoregressive [INAR(1)] process based on the binomial thinning operator, i.e., a sequence $\{X_t\}_{t\in\mathbf{Z}}$ is said to be an INAR(1) process if it admits the representation

$$X_t = \alpha \circ X_{t-1} + \epsilon_t, \tag{1}$$

where $0 \le \alpha < 1$, $\{\epsilon_t\}_{t\in\mathbf{Z}}$ is a sequence of independent and identically distributed integer-valued random variables, called innovations, with ϵ_t independent of X_{t-k} for all $k \ge 1$, $\mathrm{E}(\epsilon_t) = \mu_\epsilon$ and $\mathrm{Var}(\epsilon_t) = \sigma_\epsilon^2$, and " $\circ$ " is binomial thinning operator [18], defined by

$$\alpha \circ X_{t-1} = \begin{cases} \sum_{j=1}^{X_{t-1}} Y_j, & X_{t-1} > 0; \\ 0, & X_{t-1} = 0, \end{cases}$$

where the so-called counting series $\{Y_j\}_{j\ge 1}$ is a sequence of independent and identically distributed Bernoulli random variables with $\Pr(Y_j = 1) = 1 - \Pr(Y_j = 0) = \alpha$.

A simple approach based on the binomial thinning is to only change the innovations's distribution. The idea of considering a distribution for the innovations such that the marginal distribution of the observations will satisfy a given property has been discussed in [19]. In this context, [13] considered moment based estimators for the negative binomial INAR(1) process (with negative binomial innovations) based on the ergodicity of the process. [15] proved that if process has a negative binomial distribution, then the innovations have a negative binomial geometric distribution. [6] introduced the INAR(1) process with geometric innovations. [3] used the signed binomial thinning operator to define a first-order process with Skellam-distributed innovations. [5] studied a new stationary INAR(1) process with power series innovations. In a very recent paper, [8] proposed an INAR(1) process with Katz family innovations based on the binomial thinning.

While processes for integer-valued time series are now abundant, there is a shortage of similar processes when the time series refer to data with deflation or inflation of zeros, i.e., processes for modeling count time series with excess (or deficit) of zeros based on thinning operators were discussed by few authors. [7] introduce a new stationary INAR(1) process with zero inflated Poisson innovations. [10] studied the first-order mixed integer-valued autoregressive processes with zero-inflated generalized power series innovations. [11] introduced a zero-inflated Poisson INAR(1) process. Recently, [4] proposed a first-order integer-valued autoregressive process for dealing with count time series with deflation or inflation of zeros [ZMGINAR(1)]. The proposed process has zero-modified geometric marginals; however, the innovation structure form is very complicated. Consequently, the conditional probabilities of this model don't have a simple form. Furthermore, the parameter restrictions aren't liberal.

This paper aims to give a contribution in this direction. The objective of this paper is to propose a new INAR(1) process (1) with zero-modified geometric (ZMG) innovations, denoted by INARZMG(1), based on binomial thinning for modeling nonnegative integer-valued time series with deflation or inflation of zeros (it's possible to extend for the case of zero-modified negative binomial innovations). Advantage of the proposed process is that the ZMG distribution is very flexible. The proposed process is also able to capture equidispersion, underdispersion and overdispersion. Furthermore, the new process has, as a particular case, the INARG(1) process [6]. Additionally, we will provide a comprehensive account of the mathematical properties of the proposed new process. Also, the new process is based on the binomial thinning operator in which some elementary properties are very easy to obtain [16]. Thus, the new process has mathematical simplicity, i.e., in contrast to the ZMGINAR(1) process [4], the proposed process has a very simple innovation structure form and the parameter restrictions are more liberal.

Let $\{\epsilon_t\}_{t\in\mathbf{Z}}$ be a sequence of discrete i.i.d. random variables following a zero-modified geometric (ZMG) distribution with parameters $\mu > 0$ and $\pi \in (-1/\mu, 1)$. More specifically, we here assume that $\{\epsilon_t\}_{t\in\mathbf{Z}}$ has a probability mass function given by

$$\Pr(\epsilon_t = y) = \begin{cases} \dfrac{1+\pi\mu}{1+\mu}, & \text{if } y = 0, \\ (1-\pi)\frac{\mu^y}{(1+\mu)^{y+1}}, & \text{if } y = 1, 2, \ldots, \end{cases} \tag{2}$$

or equivalently,

$$\Pr(\epsilon_t = y) = \mathbf{I}_{\{y=0\}}\,\pi + (1-\pi)\frac{\mu^y}{(1+\mu)^{y+1}}, \quad y = 0, 1, 2, \ldots$$

In short, we name this distribution as the ZMG(π,μ) distribution. The probability generating function (pgf) of $\{\epsilon_t\}_{t\in\mathbf{Z}}$, denoted by $\varphi_\epsilon(s) := \mathrm{E}[s^{\epsilon_t}]$, is given by

$$\varphi_\epsilon(s) = \frac{1+\pi\,\mu(1-s)}{1+\mu(1-s)}, \quad |s| < 1.$$

Remark 1. *Let $Z \sim ZMG(\pi,\mu)$. Then, the moments $E(Z^r)$ are obtained from those of the geometric parent distribution $geom\big(1/(1+\mu)\big)$, say μ_r, by computing $E(Z^r) = (1-\pi)\,\mu_r$.*

The mean and variance of the $\{\epsilon_t\}_{t\in\mathbf{Z}}$ as defined in (2) are

$$\mathrm{E}(\epsilon_t) \equiv \mu_\epsilon = \mu(1-\pi)$$

and

$$\mathrm{Var}(\epsilon_t) \equiv \sigma^2_\epsilon = \mu(1-\pi)[1+\mu(1+\pi)].$$

Thus, the dispersion index, which is the variance-to-mean ratio, will be given by

$$I_\epsilon := \frac{\mu_\epsilon}{\sigma^2_\epsilon} = 1 + \mu(1+\pi);$$

it follows that this model presents equidispersion when $\pi = -1$; underdispersion when $\mu \in (0,1)$ and $\pi \in [-1/\mu, -1)$; and overdispersion when $\pi \in (-1, 1)$.

Remark 2. *Different values of π lead to different modifications of the zero-modified geometric distribution:*

1. *If $\pi = -1/\mu$, then the distribution (2) becomes the zero-truncated geometric distribution, where the parameter π cancels out and no longer appears as a model parameter, i.e., there is no chance at all of getting a zero observation into the sample;*
2. *For $\pi \in (-1/\mu, 0)$, this yields a zero-deflated geometric distribution. That is, less zeros occur, than expected under the geometric process. Such models are denoted as zero-deflated geometric distribution;*
3. *If $\pi = 0$, than the corresponding distribution is the geometric distribution;*
4. *For $\pi \in (0,1)$, this yields a zero-inflated geometric distribution, which is a geometric process with a proportion of additional zeros;*
5. *If $\pi = 1$, than the corresponding zero-modified distribution is the degenerated at zero one.*

The rest of the paper unfolds as follows. In Section 2, several properties of process are discussed. In Section 3, estimation methods for the model parameters are proposed. Section 4 discusses some simulation results for the estimation methods. In Section 5, we provide applications to two real data sets.

2. PROPERTIES OF THE PROCESS

In this section, we will consider some properties of the new process, such as the mean, variance, conditional moments, the autocorrelation structure and one-step conditional probabilities.

Theorem 2.1. *The process $\{X_t\}_{t\in\mathbf{Z}}$ is an irreducible, aperiodic and positive recurrent (and thus ergodic) Markov chain. Hence, there exists a strictly stationary process satisfying Equation (1) with $\{\epsilon_t\}_{t\in\mathbf{Z}} \sim ZMG(\pi,\mu)$.*

Following the steps of the proof of the Proposition 2 from [5], the proof of our Theorem 1 can be obtained and therefore it is omitted. From Theorem 1, the process $\{X_t\}_{t\in\mathbf{Z}}$ satisfying Equation (1), with $\{\epsilon_t\}_{t\in\mathbf{Z}} \sim$ ZMG(π,μ), is Markovian, stationary, and ergodic. Then, the Markov process admits a unique stationary distribution. From the results of [1], we have that $\alpha \in [0,1)$ and $\alpha = 1$ are the conditions of stationarity and non-stationarity of the process $\{X_t\}_{t\in\mathbf{Z}}$, respectively. Also, $\alpha = 0$ and $\alpha > 0$ respectively imply the independence and dependence of the observations of $\{X_t\}_{t\in\mathbf{Z}}$. Here, we restrict our study to the stationary case.

Theorem 2.2. *The pgf of $\{X_t\}_{t\in\mathbf{Z}}$ can be expressed as*

$$\begin{aligned} \varphi_X(s) &= \prod_{i=0}^{\infty}\varphi_\epsilon(1-\alpha^i+\alpha^i s) = \prod_{i=0}^{\infty}\left[\frac{1+\alpha^i\,\pi\,\mu(1-s)}{1+\alpha^i\,\mu(1-s)}\right] \\ &= \prod_{i=0}^{\infty}\varphi^{(i)}_\epsilon(s), \end{aligned} \tag{3}$$

where $\varphi^{(i)}_\epsilon(s) = \frac{1+\alpha^i\,\pi\,\mu(1-s)}{1+\alpha^i\,\mu(1-s)}$ is the pgf of $\epsilon^{(i)}_t \sim ZMG(\pi, \alpha^i\,\mu)$.

Following the steps of the proof of the Theorem 2 from [7], the proof of our Theorem 1 can be obtained and therefore it is omitted.

Corollary 1. *The marginal distribution of* $\{X_t\}_{t\in\mathbf{Z}}$ *can be expressed as an infinite sum of independent random variables with ZMG*$(\pi, \alpha^i\,\mu)$ *distributions.*

Proposition 1. *If* $\pi = \alpha \in [0, 1)$*, then* $\{X_t\}_{t\in\mathbf{Z}}$ *has geometric marginal distribution with mean* μ*, i.e., the process* $\{X_t\}_{t\in\mathbf{Z}}$ *is reduced to the geometric INAR(1) process introduced by [2].*

Proof. From (3), we have (see [2])

$$\varphi_X(s) = \prod_{i=0}^{\infty}\left[\frac{1+\alpha^i\,\pi\,\mu(1-s)}{1+\alpha^i\,\mu(1-s)}\right] = \prod_{i=0}^{\infty}\left[\frac{1+\alpha^i\,\alpha\,\mu(1-s)}{1+\alpha^i\,\mu(1-s)}\right]$$
$$= \frac{1}{1+\mu(1-s)}, \quad |s| < 1.$$

\# □

Proposition 2. *The proportion of zeros in the INARZMG(1) process is given by*

$$p_0 := Pr(X_t = 0) = \varphi_X(0) = \prod_{i=0}^{\infty}\left[\frac{1+\alpha^i\,\pi\,\mu}{1+\alpha^i\,\mu}\right]$$
$$= Pr(\epsilon_t = 0)\prod_{i=1}^{\infty}\left[\frac{1+\alpha^i\,\pi\,\mu}{1+\alpha^i\,\mu}\right]. \tag{4}$$

For M sufficiently large, the zero probability (4) can be approximated by

$$p_0 \approx \Pr(\epsilon_t = 0)\prod_{i=1}^{M}\left[\frac{1+\alpha^i\,\pi\,\mu}{1+\alpha^i\,\mu}\right]. \tag{5}$$

To analyze the performance of the approximation above, a small simulation study was carried out. For $\alpha = 0.4$, $\mu = 1.0$ and some values of π and M, the values for p_0 were calculated using (5). From Table 1, note that these values for the p_0 are close to $f_0 = \sum_{t=1}^{100}\mathbf{I}_{\{X_t=0\}}/100$.

The mean and variance of $\{X_t\}_{t\in\mathbf{Z}}$ are given, respectively, by

$$\mu_X := \mathrm{E}(X) = \frac{\mu(1-\pi)}{1-\alpha}$$

and

$$\sigma_X^2 := \mathrm{Var}(X) = \frac{\mu(1-\pi)[1+\mu(1+\pi)+\alpha]}{1-\alpha^2}.$$

Table 1. Simulated values of p_0 *for some values of* M*,* $\alpha = 0.4$ *and* $\mu = 1.0$

π	f_0	$M = 10$	$M = 100$	$M = 1000$
−0.4	0.1251	0.12535	0.12534	0.12534
−0.2	0.1928	0.19345	0.19343	0.19343
0.2	0.3797	0.37898	0.37895	0.37895
0.4	0.5003	0.50002	0.5	0.5

Thus, the dispersion index, which is the variance-to-mean ratio, will be given by

$$\mathrm{I}_X := \frac{\sigma_X^2}{\mu_X} = 1 + \frac{\mu(1+\pi)}{1+\alpha},$$

it follows that this model presents equidispersion when $\pi = -1$; underdispersion when $\mu \in (0, 1)$ and $\pi \in [-1/\mu, -1)$; and overdispersion when $\pi \in (-1, 1)$.

The conditional distribution of $\alpha \circ X_{t-1}$ given X_{t-1} is binomial with parameters α and X_{t-1}. Then, the conditional pgf of X_t given X_{t-1} becomes

$$\varphi_{X_t|X_{t-1}}(s) = \frac{(1-\alpha+\alpha\,s)^{X_{t-1}}[1+\pi\,\mu(1-s)]}{1+\mu(1-s)}, \quad |s| < 1.$$

The 1-step ahead conditional expectation and the conditional variance are given by

$$\mathrm{E}(X_{t+1}|X_t) = \alpha\,X_t + \mu(1-\pi)$$

and

$$\mathrm{Var}(X_{t+1}|X_t) = \alpha(1-\alpha)X_t + \mu(1-\pi)[1+\mu(1+\pi)].$$

The transition probabilities of this process are given by

$$\Pr(X_t = k|X_{t-1} = l)$$
$$= \sum_{i=0}^{\min(k,l)}\binom{l}{i}\alpha^i(1-\alpha)^{l-i}$$
$$\times\left[\pi\,\mathbf{I}_{\{0\}}(k-i) + (1-\pi)\frac{\mu^{k-i}}{(1+\mu)^{k-i+1}}\right], \tag{6}$$

$k, l \geq 0$, where $(\cdot)$ is the standard combinatorial symbol.

Thus, using (6), we obtain that the transition probability from zero to non-zero and zero to zero are

$$\theta = \Pr(X_t \neq 0|X_{t-1} = 0) = \frac{\mu(1-\pi)}{1+\mu}$$

and

$$1-\theta = \Pr(X_t = 0|X_{t-1} = 0) = \frac{1+\pi\,\mu}{1+\mu},$$

respectively. The run length of zeros in the process, N, follows a geometric distribution with termination probability θ, i.e., $\Pr(N = n) = \theta(1-\theta)^{n-1}, n \geq 1$. Thus, the average run length of zeros in the process is given by

$$\mathrm{E}(N) = \frac{1+\mu}{\mu(1-\pi)}.$$

The expected run length of zeros for the INARG(1) process ($\pi = 0$) [6] is $\mathrm{E}(N_0) = (1+\mu)/\mu$. Note that $\mathrm{E}(N) = (1-\pi)^{-1}\mathrm{E}(N_0)$. Thus, $\mathrm{E}(N) \geq \mathrm{E}(N_0)$ for $\pi \in [0, 1)$ and $\mathrm{E}(N) < \mathrm{E}(N_0)$ for $\pi \in (-1/\mu, 0)$.

The autocorrelation function (ACF) at lag h is given by

$$\operatorname{Corr}(X_t, X_{t-h}) = \rho_X(h) = \alpha^h, \quad h \geq 0.$$

That is, the autocorrelation function decays exponentially as $h \to \infty$.

The spectral density function $f(\omega)$ of any order contains important information about the properties of the process. The spectral density of the INARZMG(1) process is

$$f(\omega) = \frac{\mu(1-\pi)[1+\mu(1+\pi)+\alpha]}{2\pi[1+\alpha^2-2\alpha\cos(\omega)]}, \quad \omega \in (-\pi, \pi].$$

Next section, we consider the problem of estimating the parameters of the INARZMG(1) process.

3. ESTIMATION AND INFERENCE OF THE UNKNOWN PARAMETERS

In practice, the true values of the model parameters α, μ and π are not known but have to be estimated from given time series data. This section is concerned with the estimation of the three parameters of interest. Here, we consider three estimation methods: Yule-Walker, conditional least squares and conditional maximum likelihood.

3.1 Yule–Walker estimation

The Yule-Walker (YW) estimators of α, λ and μ are based upon the sample autocorrelation function $\widehat{\rho}(k)$, using that $\rho_X(1) = \alpha$, the first moment, and the dispersion index of X_t, which are $\mathrm{E}(X_t) = \mu(1-\pi)/(1-\alpha)$, and $\mathrm{I}_X = 1+\mu(1+\pi)/(1+\alpha)$, respectively. Let $X_1, X_2, \ldots, X_T$ be a random sample of size T from the INARZMG(1) process. Then, YW estimators of α, μ and π are given by

$$\widehat{\alpha}_{\mathrm{YW}} = \frac{\sum\limits_{t=1}^{T-1}(X_t-\overline{X})(X_{t+1}-\overline{X})}{\sum_{t=1}^{T}(X_t-\overline{X})^2},$$

$$\widehat{\mu}_{\mathrm{YW}} = \frac{1}{2}\left[(1+\widehat{\alpha}_{\mathrm{YW}})(\widehat{I}-1)+\overline{X}(1-\widehat{\alpha}_{\mathrm{YW}})\right]$$

and

$$\widehat{\pi}_{\mathrm{YW}} = 1 - \frac{\overline{X}(1-\widehat{\alpha}_{\mathrm{YW}})}{\widehat{\mu}_{\mathrm{YW}}},$$

where $\overline{X} = (1/T)\sum_{t=1}^{T} X_t$, $\widehat{I} = \widehat{\gamma}(0)/\overline{X}$ and $\widehat{\gamma}(k) = (1/T)\sum_{t=1}^{T-k}(X_t-\overline{X})(X_{t-k}-\overline{X})$.

3.2 Conditional least squares estimation

The conditional least squares (CLS) estimator $\widehat{\boldsymbol{\eta}} = (\widehat{\alpha}, \widehat{\mu}, \widehat{\pi})^{\mathsf{T}}$ of $\boldsymbol{\eta} = (\alpha, \mu, \pi)^{\mathsf{T}}$ is given by

$$\widehat{\boldsymbol{\eta}} = \arg\min_{\boldsymbol{\eta}}(S_T(\boldsymbol{\eta})),$$

where $S_T(\boldsymbol{\eta}) = \sum_{t=2}^{T}[X_t - g(\boldsymbol{\eta}, X_{t-1})]^2$ and $g(\boldsymbol{\eta}, X_{t-1}) = \mathrm{E}(X_t|X_{t-1}) = \alpha X_{t-1} + \mu(1-\pi)$. However, note that $\alpha X_{t-1} + \mu(1-\pi)$ depends on μ and π only through $\mu(1-\pi)$. Therefore, we cannot define estimators of μ and π (together) with this technique in our case. Thus, here we use the CLS method to find estimators for α and π and assume that μ is known. Then, the CLS estimators of α and π can be written in closed form as

$$\widehat{\alpha}_{\mathrm{CLS}} = \frac{(T-1)\sum_{t=2}^{T} X_t Z_{t-1} - \sum_{t=2}^{T} X_t \sum_{t=2}^{T} X_{t-1}}{(T-1)\sum_{t=2}^{T} X_{t-1}^2 - \left(\sum_{t=2}^{T} X_{t-1}\right)^2} \tag{7}$$

and

$$\widehat{\pi}_{\mathrm{CLS}} = 1 - \frac{\sum_{t=2}^{T} X_t - \widehat{\alpha}\sum_{t=2}^{T} X_{t-1}}{\mu(1-\widehat{\alpha})}, \tag{8}$$

where μ will be replaced by some consistent estimator $\widehat{\mu}$. The parameter μ can be estimated as in the previous subsection, i.e., the estimator of the parameter μ is given by $\widehat{\mu}_{\mathrm{CLS}} = \widehat{\mu}_{\mathrm{YW}}$.

Proposition 3. *The estimators $\widehat{\alpha}_{\mathrm{CLS}}$ and $\widehat{\pi}_{\mathrm{CLS}}$ given in (7) and (8) are strongly consistent for estimating α and π, respectively, and satisfy the asymptotic normality*

$$\sqrt{T}[(\widehat{\alpha}_{\mathrm{CLS}}, \widehat{\pi}_{\mathrm{CLS}})^{\top} - (\alpha, \pi)^{\top}] \xrightarrow{d} N_2((0,0)^{\top}, \mathbf{V^{-1}WV^{-1}}),$$

where

$$\mathbf{V} = \begin{pmatrix} \mu_{X,2} & -\mu\,\mu_X \\ -\mu\,\mu_X & \mu^2 \end{pmatrix}$$

and

$$\mathbf{W} = \begin{pmatrix} \alpha(1-\alpha)\mu_{X,3} + \sigma_\epsilon^2\mu_{X,2} & -\mu[\alpha(1-\alpha)\mu_{X,2} + \sigma_\epsilon^2\mu_X] \\ -\mu[\alpha(1-\alpha)\mu_{X,2} + \sigma_\epsilon^2\mu_X] & \mu^2[\alpha(1-\alpha)\mu_X + \sigma_\epsilon^2\mu_X] \end{pmatrix},$$

with $\mu_{X,r} = E(X_t^r)$.

3.3 Conditional maximum likelihood estimation

The conditional log-likelihood function for the INARZMG(1) model is given by

$$\ell(\alpha, \theta) = \log\left[\prod_{t=2}^{T}\Pr(X_t|X_{t-1})\right] = \sum_{t=2}^{T}\log[\Pr(X_t|X_{t-1})], \tag{9}$$

with $\Pr(X_t|X_{t-1})$ as in (6). The conditional maximum likelihood (CML) estimators $\widehat{\alpha}, \widehat{\mu}$ and $\widehat{\pi}$ of α, μ and π are defined as the values of α, μ and π that maximize the conditional log-likelihood function in (9). There will be no closed form for the CML estimates and their obtention will need, in practice, numerical methods. As starting values for the algorithm, we have used the estimates obtained by the YW method. Since Fisher information matrix is not available, the standard errors are obtained as the square roots of the elements in the diagonal of the inverse of the negative of the Hessian of the conditional log-likelihood calculated at the conditional maximum likelihood estimates.

In next section, a Monte Carlo simulation experiment will be conducted to evaluate the performance of the estimators discussed in this section.

4. MONTE CARLO SIMULATION STUDY

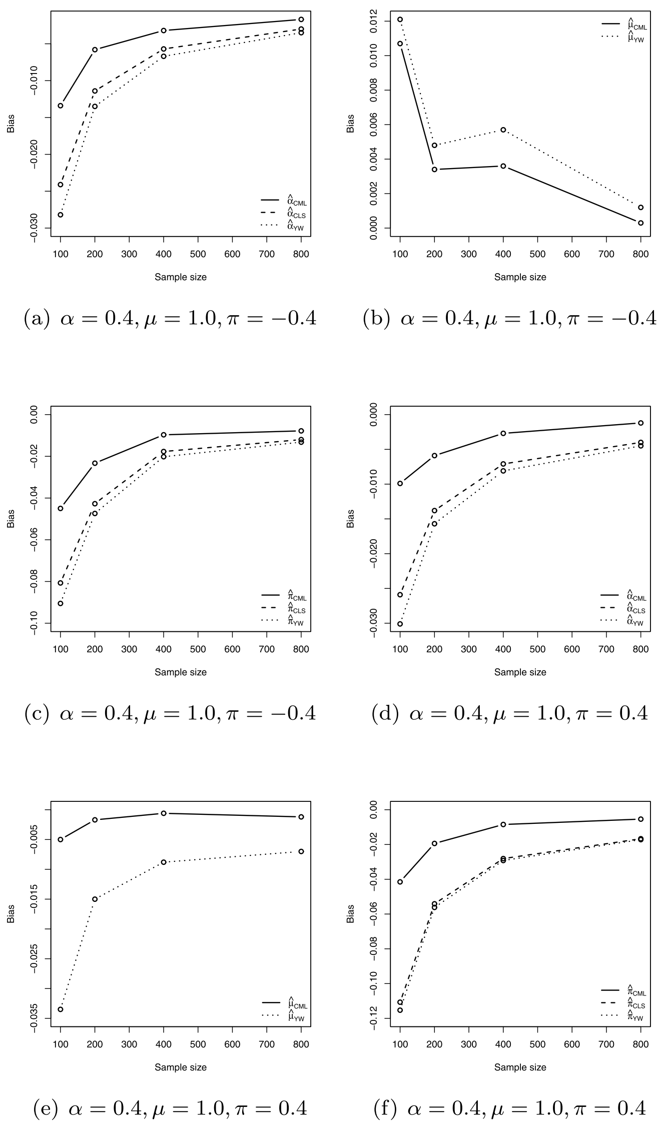
In this section, the properties of the estimators discussed in the previous sections are now investigated for finite sample sizes $T = 100, 200, 400$ and 800 from INARZMG(1) series with $\alpha = 0.4$, $\{\epsilon_t\}_{t\in\mathbf{Z}}$ being an i.i.d. ZMG sequence with $\mu = 1$ and $\pi = -0.4, -0.2, 0.2$ and 0.4. For each different situation, we have estimated the empirical mean and the mean squared error (MSE). All simulations were carried out using the R programming language [14]. The number of Monte Carlo replications was 5000.

Table 2 presents the empirical mean and MSE's (given in parentheses) of the different estimators for INARZMG(1) process. Note that as the sample size increases, the bias tends to zero in all three cases, confirming that the estimators are asymptotically unbiased.

Figure 1 shows the bias of the simulated estimates of α, μ and π. From this figure, we observe that all estimators (for $\pi = 0.4$) of the parameters are positively biased, i.e. the estimators don't exceed the true value of the parameters. Figure 2 shows the MSE of the simulated estimates of α, μ and π. From this figure, we observe that the CLS and YW estimators have almost the same MSE. The empirical investigation presented here suggests that the performance of the CML estimator is superior to those of the YW and CLS estimators. Therefore, we recommend the use CML estimator as

Table 2. Empirical means and mean squared errors (in parentheses) of the estimates of the parameters for $(\alpha, \mu) = (0.4, 1)$ and some values of π and T

		α			μ		π		
T	π	$\widehat{\alpha}_{\mathrm{YW}}$	$\widehat{\alpha}_{\mathrm{CLS}}$	$\widehat{\alpha}_{\mathrm{CML}}$	$\widehat{\mu}_{\mathrm{YW}}$	$\widehat{\mu}_{\mathrm{CML}}$	$\widehat{\pi}_{\mathrm{YW}}$	$\widehat{\pi}_{\mathrm{CLS}}$	$\widehat{\pi}_{\mathrm{CML}}$
100	−0.4	0.3718	0.3759	0.3866	1.0121	1.0107	−0.4905	−0.4807	−0.4450
		(0.0104)	(0.0103)	(0.0066)	(0.0485)	(0.0408)	(0.1071)	(0.1055)	(0.0614)
	−0.2	0.3718	0.3758	0.3843	1.0017	1.0023	−0.3025	−0.2940	−0.2584
		(0.0102)	(0.0101)	(0.0068)	(0.0577)	(0.0457)	(0.1105)	(0.1092)	(0.0651)
	0.2	0.3737	0.3775	0.3895	0.9928	1.0056	0.0961	0.1014	0.1559
		(0.0106)	(0.0106)	(0.0053)	(0.0909)	(0.0643)	(0.0956)	(0.0947)	(0.0444)
	0.4	0.3699	0.3741	0.3901	0.9665	0.9950	0.2847	0.2893	0.3585
		(0.0114)	(0.0114)	(0.0054)	(0.1193)	(0.0812)	(0.0865)	(0.0853)	(0.0330)
200	−0.4	0.3865	0.3886	0.3942	1.0048	1.0034	−0.4474	−0.4427	−0.4233
		(0.0051)	(0.0051)	(0.0032)	(0.0245)	(0.0200)	(0.0548)	(0.0546)	(0.0294)
	−0.2	0.3858	0.3877	0.3925	1.0004	1.0015	−0.2520	−0.2480	−0.2267
		(0.0052)	(0.0052)	(0.0032)	(0.0289)	(0.0226)	(0.0547)	(0.0544)	(0.0296)
	0.2	0.3876	0.3896	0.3958	0.9893	0.9957	0.1450	0.1478	0.1775
		(0.0054)	(0.0054)	(0.0025)	(0.0447)	(0.0304)	(0.0437)	(0.0433)	(0.0181)
	0.4	0.3843	0.3862	0.3941	0.9850	0.9983	0.3438	0.3459	0.3806
		(0.0058)	(0.0058)	(0.0025)	(0.0596)	(0.0393)	(0.0360)	(0.0357)	(0.0132)
400	−0.4	0.3933	0.3943	0.3968	1.0057	1.0036	−0.4202	−0.4177	−0.4097
		(0.0025)	(0.0025)	(0.0016)	(0.0124)	(0.0096)	(0.0277)	(0.0276)	(0.0140)
	−0.2	0.3925	0.3934	0.3961	0.9979	0.9984	−0.2288	−0.2269	−0.2158
		(0.0026)	(0.0026)	(0.0016)	(0.0142)	(0.0111)	(0.0265)	(0.0263)	(0.0141)
	0.2	0.3929	0.3940	0.3976	0.9952	0.9992	0.1691	0.1706	0.1882
		(0.0027)	(0.0027)	(0.0013)	(0.0242)	(0.0154)	(0.0220)	(0.0219)	(0.0088)
	0.4	0.3919	0.3929	0.3973	0.9912	0.9994	0.3708	0.3719	0.3915
		(0.0030)	(0.0029)	(0.0012)	(0.0319)	(0.0200)	(0.0169)	(0.0168)	(0.0061)
800	−0.4	0.3965	0.3970	0.3983	1.0012	1.0003	−0.4132	−0.4120	−0.4078
		(0.0012)	(0.0012)	(0.0008)	(0.0060)	(0.0047)	(0.0138)	(0.0138)	(0.0069)
	−0.2	0.3969	0.3974	0.3981	1.0010	1.0008	−0.2125	−0.2115	−0.2069
		(0.0013)	(0.0013)	(0.0007)	(0.0076)	(0.0056)	(0.0137)	(0.0137)	(0.0065)
	0.2	0.3966	0.3971	0.3990	0.9997	0.9985	0.1880	0.1886	0.1948
		(0.0014)	(0.0014)	(0.0006)	(0.0120)	(0.0080)	(0.0104)	(0.0104)	(0.0041)
	0.4	0.3955	0.3960	0.3988	0.9930	0.9988	0.3828	0.3833	0.3946
		(0.0015)	(0.0015)	(0.0006)	(0.0157)	(0.0099)	(0.0080)	(0.0080)	(0.0029)

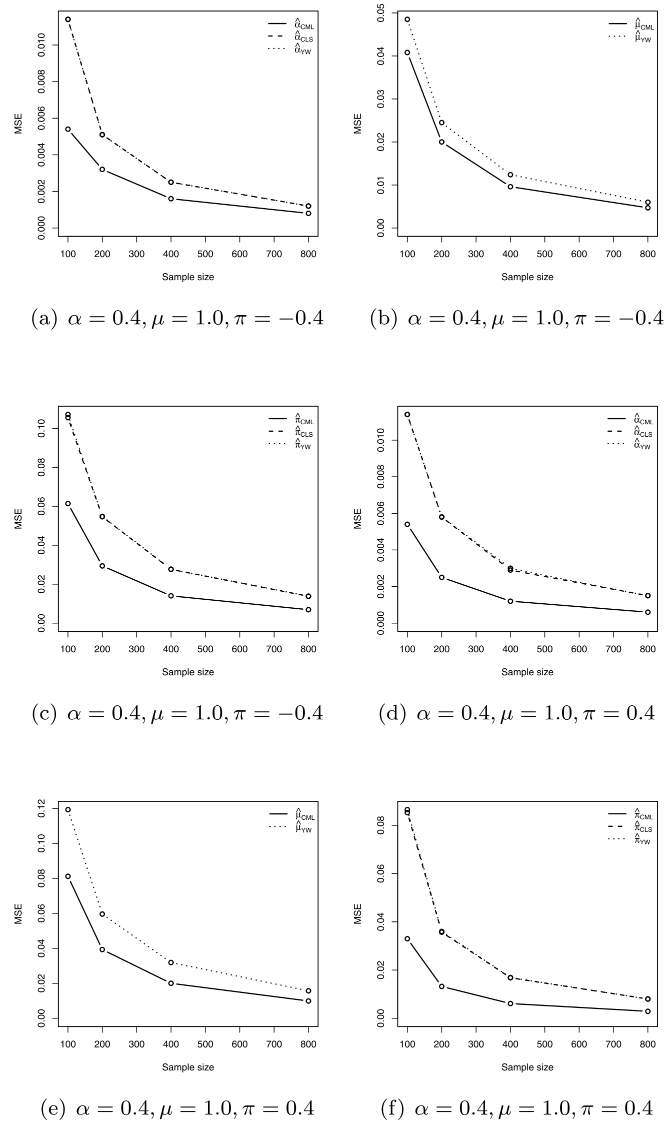


(a) $\alpha = 0.4, \mu = 1.0, \pi = -0.4$ (b) $\alpha = 0.4, \mu = 1.0, \pi = -0.4$

(c) $\alpha = 0.4, \mu = 1.0, \pi = -0.4$ (d) $\alpha = 0.4, \mu = 1.0, \pi = 0.4$

(e) $\alpha = 0.4, \mu = 1.0, \pi = 0.4$ (f) $\alpha = 0.4, \mu = 1.0, \pi = 0.4$

Figure 1. Bias from 5000 simulated estimates of α, μ and π.



(a) $\alpha = 0.4, \mu = 1.0, \pi = -0.4$ (b) $\alpha = 0.4, \mu = 1.0, \pi = -0.4$

(c) $\alpha = 0.4, \mu = 1.0, \pi = -0.4$ (d) $\alpha = 0.4, \mu = 1.0, \pi = 0.4$

(e) $\alpha = 0.4, \mu = 1.0, \pi = 0.4$ (f) $\alpha = 0.4, \mu = 1.0, \pi = 0.4$

Figure 2. MSE from 5000 simulated estimates of α, μ and π.

the estimators for the parameters of a INARZMG(1) model, having a good performance in terms of bias and MSE.

5. REAL DATA EXAMPLES

In this section, we compare the proposed process by means of two real data sets (with deflation and inflation of zeros) for illustrative purposes. We compared the INARZMG(1) process with the INARG(1) process (special case) [7] and ZINAR(1) process [6]. In order to estimate the parameters of these processes, we adopt the CML method (as discussed in Section 3) and all the computations were done using the R programming language [14].

5.1 Modelling deflation of zeros

In the first application, we consider a count data time series about the utilization of the examination room of the emergency department of a children's hospital. The data were monitored on Thursday, 16 July 2009. Within time intervals of 10-min length (between 08:00:00 and 23:59:59), the number of patients between the call for examination and the first treatment were determined. The proportion of zeros in the series considered is 4.17%. Then, we have evidence that there is deflation of zeros in the emergency counts. Thus, the use of the INARZMG(1) model for fitting this data set seems justified. Table 3 displays some descriptive statistics of the emergency counts. The time series data and their sample autocorrelation and partial autocorrelation functions are displayed in the Figure 3.

Analyzing Figure 3, we conclude that a first order autoregressive model may be appropriate for the given data series, because of the clear cut-off after lag 1 in the partial autocorrelations.

Table 3. Descriptive statistics for the emergency counts

Minimum	Median	Mean	Variance	Maximum	T
0	2.5	2.5625	1.8697	7	96

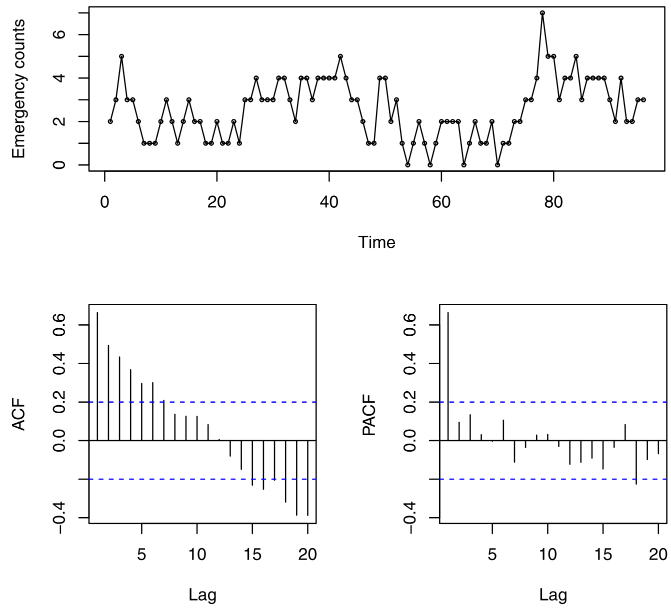


Figure 3. Plots of the time series, autocorrelation and partial autocorrelation functions for the emergency counts.

Table 4. Estimates of the parameters (standard errors in parentheses) and goodness-of-fit statistics for the emergency counts

Model		CML Estimates	AIC	BIC
INARZMG(1)	$\widehat{\alpha}$	0.6344 (0.0659)	275.97	283.66
	$\widehat{\mu}$	0.2344 (0.0867)		
	$\widehat{\pi}$	−3.0346 (1.3683)		
ZINAR(1)	$\widehat{\alpha}$	0.6321 (0.0684)	276.54	284.24
	$\widehat{\lambda}$	0.4566 (0.1544)		
	$\widehat{\rho}$	−1.0839 (0.6445)		
INARG(1)	$\widehat{\alpha}$	0.7559 (0.0379)	290.11	295.24
	$\widehat{p}$	0.3883 (0.0464)		

Table 4 provides the estimates (with standard errors in parentheses) of the model parameters and two goodness-of-fit statistics: Akaike information criterion (AIC) and Bayesian information criterion (BIC). In general, it is expected that the better model to fit the data presents the smaller values for these quantities. From this table, we observe that the proposed model being better.

The mean, variance and dispersion index within the estimated model [INARZMG(1)] are given by, respectively, $\widehat{\mu}_X = 2.5867$, $\widehat{\sigma}^2_X = 1.8319$ and $\widehat{I}_X = 0.7082$. Note that these values for the fitted model are close to the corresponding empirical values. The proportions of zeros based on the estimated INARZMG(1) model is 0.0366 (using (5) with $M = 1000$), the empirical one 0.0417. Furthermore, analyzing Table 4, note that $\widehat{\pi} < 0$, which implies that the model presents deflation of zeros. We test the null hypothesis H_0 : INARG(1) against the alternative hypothesis H_1 : INARZMG(1), i.e., $H_0 : \pi = 0$ against $H_1 : \pi \neq 0$ (with a significance level at 5%). The conditional likelihood ratio (LR) statistic to test the hypothesis is 16.15 (p-value < 0.01). Thus, we reject the null hypothesis in favor of the INARZMG(1) model using any usual significance level. Also, the residuals (from INARZMG(1) model) are not correlated.

5.2 Modelling inflation of zeros

The second data set is obtained from the crime data section of the forecasting principles site (http://www.forecastingprinciples.com). This data series represents the drug reselling criminal acts which took place in the area of Pittsburgh which is under the jurisdiction of the 56th police car beat. The data consist of 144 observations, starting in January 1990 and ending in December 2001. The series has a large proportion of zero values. The proportion of zeros in the Drugs-56 series is 47%. This basically means that in these years there were no crimes or the offenders were not arrested; however, there are not such long periods of lack of crimes, i.e. there are no such long runs of zeros. Then, we have evidence that there is inflation of zeros in the Drugs-56 series. Thus, the use of the INARZMG(1) model for fitting this data set appears justified.

The series, autocorrelation and partial autocorrelation functions are displayed in Figure 4. Examining the Figure 4 we conclude that a first-order autoregressive model may be appropriate for the given data series, given the pattern of the sample partial autocorrelation function and the clear cut-off.

Table 5 displays some descriptive statistics for the Drugs-56 counts data. We see that the sample variance is larger than the sample mean. Table 6 gives the CML estimates (with standard errors in parentheses), AIC and BIC for the fitted models. Since the values of the AIC and BIC are smaller for the INARZMG(1) process compared to those values of the ZINAR(1) and INARG(1) models, the new model seems a competitive model for these data.

The mean, variance and dispersion index within the estimated model [INARZMG(1)] are given by, respectively, $\widehat{\mu}_X = 1.722$, $\widehat{\sigma}^2_X = 5.7827$ and $\widehat{I}_X = 3.3584$. Note that these values for the fitted model are close to the corresponding empirical values. The proportion of zeros within the fitted INARZMG(1) model equals 0.4219 (using (5) with $M = 1000$), the empirical one 0.4653. Thus, the proposed model works well for capturing the proportion of zeros in this application. Furthermore, analyzing Table 6, note

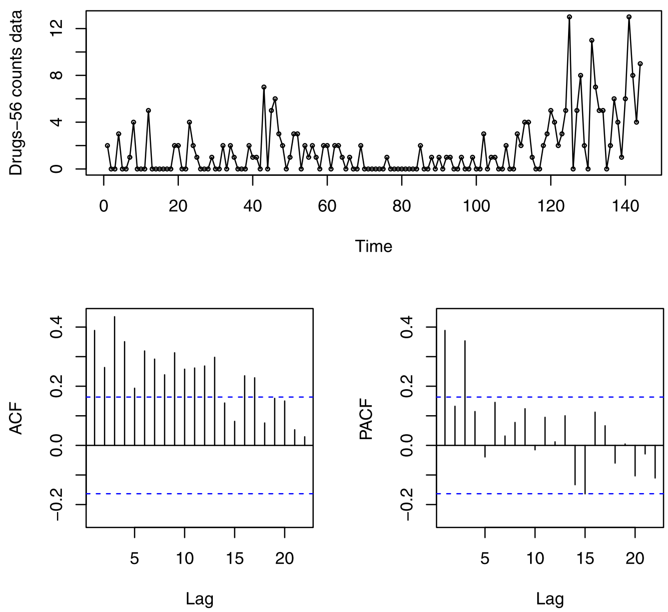


Figure 4. Plots of the time series, autocorrelation and partial autocorrelation functions for the Drugs-56 counts data.

Table 5. Descriptive statistics for the Drugs-56 counts data

Minimum	Median	Mean	Variance	Maximum	T
0	1	1.7153	6.4289	13	144

Table 6. Estimates of the parameters (standard errors in parentheses) and goodness-of-fit statistics for the Drugs-56 counts data

Model		CML Estimates		AIC	BIC
	$\widehat{\alpha}$	0.1482	(0.0686)	502.64	511.55
INARZMG(1)	$\widehat{\mu}$	2.0873	(0.3281)		
	$\widehat{\pi}$	0.2973	(0.0808)		
	$\widehat{\alpha}$	0.2156	(0.0535)	527.20	536.11
ZINAR(1)	$\widehat{\lambda}$	2.9131	(0.2604)		
	$\widehat{\rho}$	0.5351	(0.0502)		
INARG(1)	$\widehat{\alpha}$	0.1148	(0.0729)	512.49	518.43
	$\widehat{p}$	0.6035	(0.0316)		

that $\widehat{\pi} > 0$, which implies that the model presents inflation of zeros. Testing the null hypothesis H_0 : INARG(1) against the alternative hypothesis H_1 : INARZMG(1), we find that the LR statistic has the value 11.85 (p-value < 0.01). We conclude that the INARZMG(1) model is significantly better than the INARG(1) model. Furthermore, the sample autocorrelations of the residuals obtained from the INARZMG(1) model do not show any significant value.

APPENDIX A. APPENDIX SECTION

A.1 Proof of Proposition 3

Let $X_1, \ldots, X_T$ be a sample of an INARZMG(1) process. It can be verified that the regularity conditions given in Theorem 3.2 of [9], p. 634, are satisfied by INARZMG(1) process.

Consider the following quantities $\mathrm{E}_{t|t-1} \equiv \mathrm{E}(X_t|X_{t-1}) = \alpha X_{t-1} + \mu(1-\pi)$ and $\mathrm{d}_{t|t-1} = \mathrm{Var}(X_t|X_{t-1}) = \alpha(1-\alpha)X_{t-1} + \sigma_\epsilon^2$, and calculate

$$\frac{\partial \mathrm{E}_{t|t-1}}{\partial\alpha} = X_{t-1}, \quad \frac{\partial \mathrm{E}_{t|t-1}}{\partial\pi} = -\mu, \quad \frac{\partial^2 \mathrm{E}_{t|t-1}}{\partial\alpha^2} = 0,$$
$$\frac{\partial^2 \mathrm{E}_{t|t-1}}{\partial\pi^2} = 0, \quad \frac{\partial^2 \mathrm{E}_{t|t-1}}{\partial\pi\partial\alpha} = 0.$$

Define the 2×2 matrix $\mathbf{V}$ according to Equation (3.2) in [9] as

$$\mathbf{V} = \mathrm{E}\left(\begin{bmatrix} \frac{\partial \mathrm{E}_{t|t-1}}{\partial\alpha} \\ \frac{\partial \mathrm{E}_{t|t-1}}{\partial\pi}\end{bmatrix}\begin{bmatrix} \frac{\partial \mathrm{E}_{t|t-1}}{\partial\alpha} & \frac{\partial \mathrm{E}_{t|t-1}}{\partial\pi}\end{bmatrix}\right)$$
$$= \begin{pmatrix} \mathrm{E}(X_{t-1}^2) & -\mu\,\mathrm{E}(X_{t-1}) \\ -\mu\,\mathrm{E}(X_{t-1}) & \mu^2 \end{pmatrix} = \begin{pmatrix} \mu_{X,2} & -\mu\,\mu_X \\ -\mu\,\mu_X & \mu^2\end{pmatrix}$$

and the 2×2 matrix $\mathbf{W}$ according to Equation (3.5) in [9] as

$$\mathbf{W} = \mathrm{E}\left(\begin{bmatrix} \frac{\partial \mathrm{E}_{t|t-1}}{\partial\alpha} \\ \frac{\partial \mathrm{E}_{t|t-1}}{\partial\pi}\end{bmatrix}\mathrm{d}_{t|t-1}\begin{bmatrix} \frac{\partial \mathrm{E}_{t|t-1}}{\partial\alpha} & \frac{\partial \mathrm{E}_{t|t-1}}{\partial\pi}\end{bmatrix}\right)$$
$$= \begin{pmatrix} \alpha(1-\alpha)\mu_{X,3} + \sigma_\epsilon^2\mu_{X,2} & -\mu[\alpha(1-\alpha)\mu_{X,2} + \sigma_\epsilon^2\mu_X] \\ -\mu[\alpha(1-\alpha)\mu_{X,2} + \sigma_\epsilon^2\mu_X] & \mu^2[\alpha(1-\alpha)\mu_X + \sigma_\epsilon^2\mu_X]\end{pmatrix}.$$

Hence, the estimators $\widehat{\alpha}_{\mathrm{CLS}}$ and $\widehat{\pi}_{\mathrm{CLS}}$ CLS of α and π have the following asymptotic distribution:

$$\sqrt{T}[(\widehat{\alpha}_{\mathrm{CLS}}, \widehat{\pi}_{\mathrm{CLS}})^\top - (\alpha,\pi)^\top] \xrightarrow{d} \mathrm{N}_2((0,0)^\top, \mathbf{V^{-1}WV^{-1}}).$$

ACKNOWLEDGEMENTS

We thank the associate editor and the reviewer for the constructive comments.

Received 22 June 2017

REFERENCES

[1] Al-Osh, M. A. and Alzaid, A. A. (1987). First-order integer valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis*, **8** 261–275. MR0903755

[2] Alzaid, A. A. and Al-Osh, M. A. (1988). First-order integer valued autoregressive (INAR(1)) process: Distributional and regression properties. *Statistica Neerlandica*, **42** 53–61. MR0959714

[3] Andersson, J. and Karlis, D. (2014). A parametric time series model with covariates for integers in **Z**. *Statistical Modelling*, **14** 135–56. MR3192556

[4] Barreto-Souza, W. (2015). Zero-Modified Geometric INAR(1) Process for Modelling Count Time Series with Deflation or Inflation of Zeros. *Journal of Time Series Analysis*, **36** 839–852.

[5] Bourguignon, M., Vasconcellos, K. L. (2015) First order non-negative integer valued autoregressive processes with power series innovations. *Brazilian Journal of Probability and Statistics*, **29** 71–93. MR3299108

[6] Jazi, M. A., Jones, G. and Lai, C. D. (2012a). Integer valued AR(1) with geometric innovations. *Journal of the Iranian Statistical Society*, **11** 173–190. MR3010343

[7] Jazi, M. A., Jones, G. and Lai, C. D. (2012b). First-order integer valued AR processes with zero inflated Poisson innovations. *Journal of Time Series Analysis*, **33** 954–963.

[8] Kim, H., Lee, S. (2017) On first-order integer-valued autoregressive process with Katz family innovations. *Journal of Statistical Computation and Simulation*, **87** 546–562.

[9] Klimko, L. A. and Nelson, P. I. (1978). On conditional least squares estimation for stochastic processes. *Ann. Statist.*, **6** 629–642.

[10] Li, C., Dehui Wang, D. and Zhang, H. (2015). First-order mixed integer-valued autoregressive processes with zero-inflated generalized power series innovations. *Journal of the Korean Statistical Society*, **44** 232–246.

[11] Maiti, R., Biswas, A., Das, S. (2015). Time Series of Zero-Inflated Counts and their Coherent Forecasting. *Journal of Forecasting*, **34** 694–707.

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

[13] Leonenko N., Savani, V., Zhigljavsky, A. (2007) Autoregressive negative binomial processes. *Annales de l'I.S.U.P*, **LI** 25–47.

[14] R Core Team (2016) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org

[15] Savani, V., Zhigljavsky, A. (2007) Efficient parameter estimation for independent and INAR(1) negative binomial samples. *Metrika*, **65** 207–225.

[16] Silva, M. E. and Oliveira, V. L. (2004). Difference equations for the higher-order moments and cumulants of the INAR(1) model. *Journal of Time Series Analysis*, **25** 317–333.

[17] Silva, I. and Silva, M. E. (2006). Asymptotic distribution of the Yule-Walker estimator for INAR(p) processes. *Statistics and Probability Letters*, **76** 1655–1663. MR2248854

[18] Steutel, F. W. and Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability*, **7** 893–899.

[19] Weiss, C. H. (2008) Thinning operations for modelling time series of counts — a survey. *AStA Advances in Statistical Analysis*, **92** 319–341.

Marcelo Bourguignon
Universidade Federal do Rio Grande do Norte
Departamento de Estatística
Natal, RN
Brazil
E-mail address: m.p.bourguignon@gmail.com