Poisson–Lindley INAR(1) model with applications

M. Mohammadpour[a], Hassan S. Bakouch[b] and M. Shirozhan[a]

[a]*University of Mazandaran*
[b]*Tanta University*

Abstract. The paper focuses on a new stationary integer-valued autoregressive model of first order with Poisson–Lindley marginal distribution. Several statistical properties of the model are established, like spectral density function, multi-step ahead conditional measures, stationarity, ergodicity and irreducibility. We consider several methods for estimating the unknown parameters of the model and investigate properties of the estimators. The performances of these estimators are compared via simulation. The model is motivated by some applications to two real count time series data.

1 Introduction

In the recent years, there has been a growing interest in modeling stationary time series model with discrete marginal distributions, because various kinds of discrete valued time series models are often encountered in practice. If time series with discrete variate do not have a range that is big enough to justify approximation by standard continuous model, then in order to fit and forecast the series it is necessary to use integer valued model. The most widely known models are based on the binomial thinning operator, that was introduced by Steutel and van Harn (1979). In many real life situations, there is a need to model and mathematically describe time series of correlated count observations, so integer valued models have to be used.

Integer-valued time series arise as the number of births at a hospital in successive months, count of accidents, count of patients, count of chromosome interchanges in cells, number of transmitted messages and so on. The first order non-negative integer valued autoregressive (INAR(1)) processes were introduced by McKenzie (1985), Al-Osh and Alzaid (1987).

A significant number of INAR models are either based on the binomial thinning operator or its generalizations. The binomial thinning operator is generated by counting series of independent Bernoulli distributed random variables and it is defined as

$$\alpha \circ X = \sum_{i=1}^{X} W_i, \qquad \alpha \in (0, 1), \tag{1.1}$$

where X is a non negative integer-valued random variable and $\{W_i\}$ is a sequence of independent identically distributed random variables with Bernoulli (α) distribution and is independent of X.

Among the INAR(1) models based on equation (1.1), we cite the Poisson model (Al-Osh and Alzaid (1987), Alzaid and Al-Osh (1988), McKenzie (1988)), the geometric and negative binomial models (Alzaid and Al-Osh (1988), McKenzie (1986)), the generalized Poisson model (Alzaid and Al-Osh (1993)), and the zero-truncated Poisson distribution (Bakouch and Ristic (2010)). Among models based on generalizations of the binomial thinning operator, we cite the negative binomial model (Al-Osh and Aly (1992), Aly and Bouzar (1994b), Ristić, Bakouch and Nastić (2009), Ristić, Nastić and Bakouch (2012)) and the Poisson geometric model of Aly and Bouzar (1994a). Other extensions of the INAR(1) model are surveyed by McKenzie (2003). A significant development of the previous Bernoulli counting series was made by the Ristić, Nastic and Miletic Ilic (2013), where the INAR(1) model with dependent Bernoulli counting series is introduced, as not all counting series random variables are independent. Recently, Schweer and Weiß (2014) introduced a general family of INAR(1) models with compound Poisson innovations.

Count time series data are of main interest in various applied fields, such as medicine, insurance, communications, social sciences, meteorology and ecology. All the above integer valued autoregressive models with different discrete marginal distributions and modified thinning operators have been proposed to improve the quality of fitting and analyzing such data. Sometimes, those models give less accurate analysis because of the nature and properties of the count data that exhibit over-dispersion, skewness and kurtosis. Therefore, there is a need to introduce other integer valued time series models with marginals rather than traditional ones to treat some of their limitations represented in their marginal distributions, for example, Poisson marginal has equidispersion (variance equals to the mean, a situation that may not be consistent with data), zero truncated Poisson marginal is underdispersed (underdispersion occur less frequently than overdispersion), geometric marginal has a constant failure rate (unrealistic feature).

In this paper, the Poisson–Lindley (PL) marginal distribution is used for the first time to model the INAR time series models using the binomial thinning operator with independent Bernoulli counting series random variables. Hereafter, the introduced first-order integer valued autoregressive model with Poisson–Lindley marginal shall be denoted by PLINAR(1).

The discrete Poisson–Lindley distribution of a random variable X has the probability mass function (p.m.f.)

$$f(x, \theta) = \frac{\theta^2(x + \theta + 2)}{(1 + \theta)^{x+3}}, \qquad x = 0, 1, \ldots; \theta > 0,$$

was introduced by Sankaran (1970) to model count data. The PL(θ) distribution belongs to compound Poisson family and has other common properties like unimodality, overdispersion, and infinite divisibility (Ghitany and Al-Mutairi (2009)),

Table 1 *Mean, variance, skewness and kurtosis of the Poisson–Lindley distribution*

Mean	$\frac{\theta+2}{\theta(\theta+1)}$
Variance	$\frac{\theta^3+4\theta^2+6\theta+2}{\theta^2(\theta+1)^2}$
Skewness	$\frac{2(\theta+1)^4(\theta+2)-\theta^3(\theta+2)(\theta+3)}{(2(\theta+1)^3-\theta^2(\theta+2))^{\frac{3}{2}}}$
Kurtosis	$3+\frac{2(\theta+1)^5((\theta+3)^2-3)-\theta^4(\theta+2)((\theta+4)^2-3)}{(2(\theta+1)^3-\theta^2(\theta+2))^2}$

such as the negative binomial distribution. PL(θ) distribution has an increasing hazard rate which is realistic for most of the data. Also, PL(θ) distribution can be viewed as mixture of geometric with parameter $\frac{1}{1+\theta}$ and negative binomial with parameters 2 and $\frac{1}{1+\theta}$ with mixing proportions $\frac{\theta}{1+\theta}$ and $\frac{1}{1+\theta}$, respectively. Moreover, Ghitany and Al-Mutairi (2009) demonstrated that the skewness and kurtosis of the Poisson–Lindley distribution are smaller than the negative binomial distribution, we will further elaborate on this point in Section 5 for real data. Advantages of the PL(θ) distribution than the mentioned counted distributions were the motivation to use the PL(θ) distribution for modeling the integer valued autoregressive model. Further, two real data examples show that the PLINAR(1) model provides a very satisfactory fit and that it is highly competitive with other known and recent INAR models such as Poisson INAR(1), geometric INAR(1), generalized Poisson INAR(1), new geometric INAR(1) and many versions of negative binomial INAR(1).

The corresponding formulas for mean, variance, skewness and kurtosis of the Poisson–Lindley distribution are given in Table 1.

Also, recall that the probability generating function (p.g.f.) of PL distribution is given by

$$\Phi_X(s) = \frac{\theta^2}{1+\theta}\frac{2+\theta-s}{(1+\theta-s)^2}.$$

The paper is organized as follows. In Section 2, we construct the Poisson–Lindley first-order integer-valued autoregressive (PLINAR(1)) model and obtain the p.m.f. of the innovation term of the model. In Section 3, we investigate many properties of this model, such as autocorrelation function, spectral density function, multi-step ahead conditional expectation, variance and partial autocorrelation function. Section 4, deals with estimation of the unknown parameters of the model by using conditional least square estimator, Yule–Walker estimator and maximum likelihood estimator. Also, the asymptotic properties and asymptotic distribution of the estimators are investigated. Some applications of the process for two real data sets are given in Section 5. Some concluding remarks with future issues of the model are given in Section 6.

2 The PLINAR(1) model and its properties

Let us consider the INAR(1) process

$$X_t = \alpha \circ X_{t-1} + \varepsilon_t, \qquad t \geq 1, \tag{2.1}$$

where the binomial thinning operator $\alpha\circ$ defined above, $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables and ε_t is independent of Bernoulli counting process $\{W_i^{(t)}\}$ and X_m for all $m \leq t$. Assume that $\{X_t\}$ is a stationary process with $\mathrm{PL}(\theta)$ distribution, then the innovation process $\{\varepsilon_t\}$ has the p.g.f.

$$\Phi_\varepsilon(s) = \frac{2 + \theta - s}{(1 + \theta - s)^2} \frac{[\theta + \alpha(1 - s)]^2}{1 + \theta + \alpha(1 - s)}. \tag{2.2}$$

Definition 2.1. Let $F(x)$ be a distribution function. We say the distribution function $F(x)$ is a generalized mixture of the distribution functions $F(\,; 1), F(\,; 2), \ldots$ if

$$F(x) = \sum_{i \geq 1} w_i F(x; i),$$

for all x, where the real numbers $w_1, w_2, \ldots$ are such that $\sum_{i\geq 1} w_i = 1$, $\sum_{i\geq 1} |w_i| < \infty$ and for some indices i, $w_i < 0$.

Lemma 2.1. *If $0 < \alpha < 1, \theta \geq 1$, then the generalized mixture*

$$
\begin{aligned}
g(x) = {} & \frac{\theta^2(1-\alpha)^2 + \theta(1-\alpha^2) + 2\alpha}{(\theta(1-\alpha)+1)^2} \frac{\theta}{1+\theta}\left(1 - \frac{\theta}{1+\theta}\right)^x \\
& + \frac{(1-\alpha)}{\theta(1-\alpha)+1}(x+1)\left(\frac{\theta}{1+\theta}\right)^2\left(1 - \frac{\theta}{1+\theta}\right)^x \\
& - \frac{\alpha}{(\theta(1-\alpha)+1)^2}\frac{\theta+1}{\theta+1+\alpha}\left(1 - \frac{\theta+1}{\theta+1+\alpha}\right)^x, \qquad x = 0, 1, \ldots,
\end{aligned}
\tag{2.3}
$$

is a probability mass function.

Proof. The proof is given in the Appendix. □

Theorem 2.1. *Consider the* PLINAR(1) *process defined by equation* (2.1). *The innovation sequence $\{\varepsilon_t\}$ possesses the distribution*

$$f_\varepsilon(x) = \alpha h(x) + (1 - \alpha)g(x),$$

where $h(x)$ is the degenerate distribution function at zero and $g(x)$ is a probability mass function defined by equation (2.3).

Proof. See the Appendix. □

Corollary 2.1. *Based on Lemma* 2.1 *and Theorem* 2.1, *the Poisson–Lindley distribution is discrete self-decomposable* (*see e.g. Steutel and van Harn* (1979)) *for* $\theta \geq 1$.

Applying the result of the Theorem 2.1, the process defined by equation (2.1) can be rewritten as

$$
X_t = \begin{cases} \alpha \circ X_{t-1}, & w.p. \ \alpha, \\ \alpha \circ X_{t-1} + \varepsilon_t, & w.p. \ 1 - \alpha, \end{cases}
$$

where w.p. stands for "with probability". Therefore, we can write the process $\{X_t\}$ as

$$
X_t = \alpha \circ X_{t-1} + I_t H_t, \tag{2.4}
$$

where $P\{I_t = 0\} = 1 - P\{I_t = 1\} = \alpha$, H_t has the probability mass function g in equation (2.3) and X_{t-k} is independent of $I_t H_t$ for $k \geq 1$.

Based on properties of the binomial thinning operator, the marginal distribution of the model (2.4) can be expressed in terms of the innovation sequence $\{I_t H_t\}$ as

$$
X_t \overset{d}{=} \sum_{j=0}^{\infty} \alpha^j \circ (I_{t-j} H_{t-j}).
$$

Remark 2.1. For $\theta \geq 1$, if X_0 is PL(θ), then the process $\{X_t\}$ is Poisson–Lindley for every $t \geq 1$ and strict stationarity ensues by the Markovian property.

Obviously, the autocovariance function of the PLINAR(1) process obtained as

$$
\begin{aligned}
\gamma_k &= \text{Cov}(X_t, X_{t-k}) \\
&= \text{Cov}(\alpha \circ X_{t-1} + I_t H_t, X_{t-k}) \\
&= \alpha \, \text{Cov}(X_{t-1}, X_{t-k}) = \cdots = \alpha^k \gamma_0, \qquad k \geq 0.
\end{aligned}
$$

Then, the autocorrelation function is $\rho_k = \alpha^k$, $\alpha \in (0, 1)$, $k \geq 0$. That is, the autocorrelation function decays exponentially as $k \to \infty$.

By applying the autocovariance function, the spectral density function of the PLINAR(1) process is as follows

$$
\begin{aligned}
f_{xx}(\omega) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \text{Cov}(X_t, X_{t-k}) e^{-i\omega k} \\
&= \frac{(\theta^3 + 4\theta^2 + 6\theta + 2)(1 - \alpha^2)}{2\pi \theta^2 (\theta + 1)^2 (1 + \alpha^2 - 2\alpha \cos(\omega))},
\end{aligned}
$$

where $\omega \in (-\pi, \pi]$.

Also, the joint probability generating function of the random variables X_t and X_{t-1} is obtained as follows

$$\Phi_{x_t,x_{t-1}}(s_1, s_2) = \alpha \Phi_{\alpha \circ x_{t-1},x_{t-1}}(s_1, s_2) + (1-\alpha) \Phi_{\alpha \circ x_{t-1}+\varepsilon_t,x_{t-1}}(s_1, s_2)$$

$$= \alpha \Phi_{x_t}(s_2(1-\alpha+\alpha s_1))$$

$$+ (1-\alpha)\left[\Phi_{\varepsilon_t}(s_1)\Phi_{x_t}(s_2(1-\alpha+\alpha s_1))\right]$$

$$= \alpha \frac{\theta^2}{1+\theta}\left(\frac{2+\theta-(s_2(1-\alpha+\alpha s_1))}{(1+\theta-(s_2(1-\alpha+\alpha s_1)))^2}\right)$$

$$+ (1-\alpha)\left[\frac{2+\theta-s_1}{(1+\theta-s_1)^2}\frac{(\theta+\alpha(1-s_1))^2}{1+\theta+\alpha(1-s_1)}\right.$$

$$\times \left.\frac{\theta^2}{1+\theta}\frac{2+\theta-(s_2(1-\alpha+\alpha s_1))}{(1+\theta-(s_2(1-\alpha+\alpha s_1)))^2}\right].$$

It is obvious that the PLINAR(1) process is not time reversible as $\Phi_{x_t,x_{t-1}}$ is not symmetric in (s_1, s_2).

3 Conditional properties

Let us provide the $(k+1)$-step ahead conditional mean and conditional variance of the PLINAR(1) process. Note that unconditional mean and variance of the model (2.4) are

$$E(X_t) = E(H_t) = \frac{\theta+2}{\theta(\theta+1)}, \qquad \mathrm{Var}(X_t) = \frac{\theta^3+4\theta^2+6\theta+2}{\theta^2(\theta+1)^2},$$

respectively.

Using the definition of the PLINAR(1) process and induction method, the $(k+1)$-step ahead conditional mean obtained as

$$E(X_{t+k} \mid X_{t-1} = x) = \alpha^{k+1}x + \left(1-\alpha^{k+1}\right)\frac{\theta+2}{\theta(\theta+1)}, \qquad (3.1)$$

and if $k \longrightarrow \infty$ then $E(X_{t+k} \mid X_{t-1} = x) \longrightarrow \frac{\theta+2}{\theta(\theta+1)}$, which is the unconditional mean of the process.

Using the model definition and its variance, we get

$$E(H_t^2) = \frac{\theta^3+5\theta^2+10\theta-2\alpha+6}{\theta^2(\theta+1)^2},$$

and noting that

$$\mathrm{Var}(I_t H_t) = (1-\alpha)E(H_t^2) - (1-\alpha)^2(E(H_t))^2,$$

implies

$$\text{Var}(I_t H_t) = \frac{(1 - \alpha)}{\theta^2 (\theta + 1)^2} (\theta^3 + 4\theta^2 + 6\theta + 2 + \alpha(\theta^2 + 4\theta + 2)).$$

All the above implies that the $(k + 1)$-step ahead conditional variance, $\text{Var}(X_{t+k} \mid X_{t-1} = x)$, is

$$\alpha^{k+1}(1 - \alpha^{k+1})x + \frac{1 - \alpha^{2(k+1)}}{1 - \alpha^2} \text{Var}(I_t H_t) + \frac{(1 - \alpha^k)(\alpha - \alpha^{k+2})}{1 - \alpha^2} E(I_t H_t).$$

When $k \longrightarrow \infty$, we find that

$$\text{Var}(X_{t+k} \mid X_{t-1} = x) \to \frac{\text{Var}(I_t H_t)}{1 - \alpha^2} + \frac{\alpha E(I_t H_t)}{1 - \alpha^2}$$

$$= \frac{\theta^3 + 4\theta^2 + 6\theta + 2}{\theta^2 (\theta + 1)^2},$$

which is the unconditional variance of the process.

The partial autocorrelation function at lag $h > 1$ by using equation (3.1) is

$$\beta(h) = \text{Corr}(X_{h+1} - E(X_{h+1} \mid X_2, \ldots, X_h), X_1)$$

$$= \text{Corr}\left(X_{h+1} - \alpha X_h + (1 - \alpha)\frac{\theta + 2}{\theta(\theta + 1)}, X_1\right) = \alpha^h - \alpha(\alpha^{h-1}) = 0,$$

where $\text{Corr}(\cdot)$ stands for correlation. From this and noting that $\beta(1) = 1$, we conclude that for PLINAR(1) the partial autocorrelation cuts off at lag 1 and the PLINAR(1) model possesses an autoregressive nature.

Now, let us turn to transition probabilities. As the PLINAR(1) process is Markovian, transition probabilities play a major role in determining some characteristics of the process. Based on equation (2.4), the transition probabilities of the process are

$$p_{ij} = p(X_t = j \mid X_{t-1} = i)$$

$$= \sum_{k=0}^{\min(i,j)} \binom{i}{k} \alpha^k (1 - \alpha)^{i-k} p(I_t H_t = j - k),$$

(3.2)

where the process $I_t H_t$ is defined in equation (2.4). Equation (3.2) and the assumption $f_\varepsilon > 0$ (Theorem 2.1) imply that the transition probabilities $p_{ij} > 0$ and hence the process $\{X_t\}$ is irreducible and aperiodic Markov chain. Thus, it is either positive recurrent (and hence ergodic) or $\lim_{n\to\infty} p_{ij}^n = 0$.

4 Estimation and inference

Many features of the model depend on its parameters, so estimation of the model parameters is an important issue. For this purpose, we consider several methods

for estimation the unknown parameters from the realization $X_1, \ldots, X_n$ of the PLINAR(1) process defined by equation (2.4). These estimators are compared via Monte Carlo simulations in terms of their mean and standard deviations.

4.1 Conditional least square estimation

The conditional least squares estimators of the parameters α and μ are obtained by minimizing the function

$$Q_n = \sum_{t=2}^{n}(X_t - \alpha X_{t-1} - \mu(1-\alpha))^2,$$

where $\mu = E(X_t) = \frac{\theta+2}{\theta(\theta+1)}$. The estimators are given by

$$\hat{\alpha}_{\text{CLS}} = \frac{(n-1)\sum_{t=2}^{n}(X_{t-1}X_t) - \sum_{t=2}^{n}X_t \sum_{t=2}^{n}X_{t-1}}{(n-1)\sum_{t=2}^{n}X_{t-1}^2 - (\sum_{t=2}^{n}X_{t-1})^2},$$

$$\frac{\hat{\theta}_{\text{CLS}}+2}{\hat{\theta}_{\text{CLS}}(\hat{\theta}_{\text{CLS}}+1)} = \hat{\mu}_{\text{CLS}} = \frac{\sum_{t=2}^{n}X_t - \hat{\alpha}_{\text{CLS}}\sum_{t=2}^{n}X_{t-1}}{(n-1)(1-\hat{\alpha}_{\text{CLS}})},$$

and hence estimator of the parameter θ is

$$\hat{\theta}_{\text{CLS}} = \frac{-(\hat{\mu}_{\text{CLS}}-1) + \sqrt{(\hat{\mu}_{\text{CLS}}-1)^2 + 8\hat{\mu}_{\text{CLS}}}}{2\hat{\mu}_{\text{CLS}}}.$$

Now we derive the asymptotic properties of the estimators $\hat{\alpha}_{\text{CLS}}$ and $\hat{\mu}_{\text{CLS}}$. Since the process $\{X_t\}$ is strictly stationary and ergodic, by Theorems 3.1 and 3.2 in Tjøstheim (1986) it follows that the estimators $\hat{\alpha}_{\text{CLS}}$ and $\hat{\mu}_{\text{CLS}}$ are strongly consistent estimators for the parameters α and μ and $\sqrt{n}(\hat{\alpha}_{\text{CLS}} - \alpha, \hat{\mu}_{\text{CLS}} - \mu)'$ converges in distribution to the bivariate normal distribution with mean zero vector and covariance matrix given by

$$G = \begin{bmatrix} \frac{\gamma\sigma^2 + \alpha(1-\alpha)(\mu_3 - \mu^3 - 2\mu\sigma^2)}{\sigma^4} & \alpha \\ \alpha & \frac{\alpha(1-\alpha)\mu + \gamma}{(1-\alpha)^2} \end{bmatrix},$$

where $\sigma^2 = \text{Var}(X_t)$, $\gamma = \text{Var}(I_t H_t)$ and

$$\mu_3 = E(X_t^3) = \frac{(\theta^5 + 10\theta^4 + 41\theta^3 + 80\theta^2 + 72\theta + 24)}{\theta^3(\theta+1)^3}.$$

4.2 Yule–Walker estimation

Let $\hat{\gamma}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(X_t - \overline{X})(X_{t+k} - \overline{X})$, $0 \le k < n$, be the sample auto covariance function of X, where $\overline{X} = \frac{1}{n}\sum_{t=1}^{n}X_t$ is the sample mean. Since $\alpha = \frac{\gamma_1}{\gamma_0}$ and $\mu =$

$E(X_t)$, the Yule–Walker estimators of α and μ are

$$\hat{\alpha}_{\text{YW}} = \frac{\hat{\gamma}_1}{\hat{\gamma}_0} = \frac{\sum_{t=2}^{n}(X_t - \overline{X})(X_{t-1} - \overline{X})}{\sum_{t=1}^{n}(X_t - \overline{X})^2},$$

$$\hat{\mu}_{\text{YW}} = \overline{X} = \frac{\hat{\theta} + 2}{\hat{\theta}(\hat{\theta} + 1)}.$$

So the explicit estimator for parameter θ is

$$\hat{\theta}_{\text{YW}} = \frac{-(\overline{X} - 1) + \sqrt{(\overline{X} - 1)^2 + 8\overline{X}}}{2\overline{X}}.$$

Asymptotic properties of Yule–Walker estimators is given by the following theorem.

Theorem 4.1. *Conditional least squares and Yule–Walker estimators of the* PLINAR(1) *process are asymptotically equivalent.*

Proof. The proof follows by showing the two conditions.

(i)

$$\hat{\mu}_{\text{CLS}} - \hat{\mu}_{\text{YW}} = o\big(n^{-\frac{1}{2}}\big),$$

(ii)

$$\hat{\alpha}_{\text{CLS}} - \hat{\alpha}_{\text{YW}} = o\big(n^{-\frac{1}{2}}\big).$$

From the definitions of $\hat{\mu}_{\text{CLS}}$ and $\hat{\mu}_{\text{YW}}$, we find that

$$\lim_{n \to \infty} \sqrt{n-1}(\hat{\mu}_{\text{CLS}} - \hat{\mu}_{\text{YW}}) = \lim_{n \to \infty} \frac{\alpha}{(1 - \alpha)\sqrt{n-1}}(X_n - X_1) = 0,$$

which gives condition (i).

Definitions of $\hat{\alpha}_{\text{CLS}}$ and $\hat{\alpha}_{\text{YW}}$ imply

$$\sqrt{n}(\hat{\alpha}_{\text{CLS}} - \hat{\alpha}_{\text{YW}}) = \frac{A}{[(n-1)\sum_{t=2}^{n} X_{t-1}^2 - (\sum_{t=2}^{n} X_{t-1})^2][\sum_{t=1}^{n}(X_t - \overline{X})^2]},$$

where A is equal to

$$\left[(n-1)\sum_{t=2}^{n} X_t X_{t-1} - \sum_{t=2}^{n} X_t \sum_{t=1}^{n} X_{t-1}\right]\left[\sum_{t=1}^{n}(X_t^2 - 2\overline{X}X_t + \overline{X}^2)\right]$$

$$- \left[\sum_{t=2}^{n} X_t X_{t-1} - \overline{X}\sum_{t=2}^{n}(X_t + X_{t-1}) + (n-1)\overline{X}^2\right]$$

$$\times \left[(n-1)\sum_{t=2}^{n} X_{t-1}^2 - \left(\sum_{t=2}^{n} X_{t-1}\right)^2\right].$$

Dividing nominator and denominator by $(n-1)^2$ and using $n \approx (n-1)$, for large n, we obtain

$$\sqrt{n}(\hat{\alpha}_{\text{CLS}} - \hat{\alpha}_{\text{YW}})$$

$$= \frac{X_n - \overline{X}}{\sqrt{n}} \frac{\frac{1}{n-1}\sum_{t=2}^{n} X_t X_{t-1} - (\frac{1}{n-1}\sum_{t=2}^{n} X_t)(\frac{1}{n-1}\sum_{t=1}^{n} X_{t-1})}{[\frac{1}{n-1}\sum_{t=2}^{n} X_{t-1}^2 - (\frac{1}{n-1}\sum_{t=2}^{n} X_{t-1})^2][\frac{1}{n}\sum_{t=1}^{n} X_t^2 - \overline{X}^2]}.$$

Due to stationarity and ergodicity of the sequence $\{X_t\}$, it holds

$$\lim_{n \to \infty} \sqrt{n}(\hat{\alpha}_{\text{CLS}} - \hat{\alpha}_{\text{YW}}) = 0,$$

then we get condition (ii). $\qquad \square$

4.3 Maximum likelihood estimation

Finally, we derive the maximum likelihood estimators (MLEs) of the unknown parameters α and θ. The MLEs of the parameters are obtained by maximization of the log-likelihood function

$$\ln L(X_1, \ldots, X_n; \theta, \alpha) = 2 \ln \theta - (X_1 + 3) \ln(1 + \theta)$$

$$+ \ln(2 + \theta + X_1) + \sum_{l=1}^{n} \ln p(X_l \mid X_{l-1}; \theta, \alpha),$$

where $p(X_l \mid X_{l-1}; \theta, \alpha)$ is given by equation (3.2).

The MLEs can be easily computed by using the function nlm from statistical package R, taking the conditional least squares estimators as initial values of the function nlm. A numerical investigation of the MLE estimates is assessed and compared to CLS and YW estimates via a Monte Carlo simulation in the next subsection.

4.4 Some simulation results

To examine the performance of the CLS, YW and ML estimators, a Monte Carlo simulation was conducted for different sample sizes ($n = 50, 100, 500, 1000, 5000, 10,000$). We simulated 10,000 samples from the PLINAR(1) process for true parameter values $\alpha = 0.1, \theta = 1; \alpha = 0.3, \theta = 2$ and $\alpha = 0.5, \theta = 3$. Table 2 provides the mean and standard deviation (in brackets) of the estimates for different values of the parameters α and θ with different sample sizes.

From Table 2, we observe that the estimates obtained from the three estimation methods are convergent in their values. The CLS and YW estimates have similar numerical values. Also, increasing the sample size implies smaller standard deviation and MLEs converge faster to the true parameter values. Further, we conclude that the MLEs have the smallest standard deviations than the others, hence the MLEs give the best performance than the CLS and YW estimates.

Table 2 *Mean and standard deviation (in brackets) of the estimators for different values of the parameters α and θ*

n	$\hat{\alpha}_{\text{CLS}}$	$\hat{\theta}_{\text{CLS}}$	$\hat{\alpha}_{\text{YW}}$	$\hat{\theta}_{\text{YW}}$	$\hat{\alpha}_{\text{ML}}$	$\hat{\theta}_{\text{ML}}$
			$(\alpha, \theta) = (0.1, 1)$			
50	0.0994102	1.149081	0.0983538	1.150767	0.0993817	1.142961
	(0.0113089)	(0.0134960)	(0.0112335)	(0.0143792)	(0.0107299)	(0.0132052)
100	0.1006051	1.145384	0.1001877	1.148386	0.0997218	1.138520
	(0.0108718)	(0.0132487)	(0.0110591)	(0.0138757)	(0.0105309)	(0.0131841)
500	0.0998751	1.138529	0.0994009	1.145684	0.0997259	1.106318
	(0.0105092)	(0.0129523)	(0.0107118)	(0.0134780)	(0.0104294)	(0.0117204)
1000	0.0998039	1.135861	0.0995280	1.140273	0.0999521	1.083892
	(0.0100442)	(0.0123955)	(0.0104629)	(0.0129511)	(0.0082953)	(0.0093295)
5000	0.0999793	1.076428	0.0998529	1.106299	0.0999962	1.028725
	(0.0095922)	(0.0082024)	(0.0088207)	(0.0108362)	(0.0038261)	(0.0058263)
10,000	0.0999974	1.014721	0.999693	1.084037	0.1000007	1.000481
	(0.0038574)	(0.0052983)	(0.0049302)	(0.0087206)	(0.0006294)	(0.0009173)
			$(\alpha, \theta) = (0.3, 2)$			
50	0.3186633	2.495004	0.3254938	2.503447	0.299762	2.439466
	(0.012787)	(0.056107)	(0.014497)	(0.059660)	(0.012173)	(0.0527319)
100	0.3092002	2.493216	0.3147419	2.497559	0.3032524	2.327441
	(0.012187)	(0.052288)	(0.0130745)	(0.0549282)	(0.0104201)	(0.0450027)
500	0.30814	2.378716	0.3102891	2.384481	0.301836	2.185692
	(0.011622)	(0.0497849)	(0.0123052)	(0.0506455)	(0.0083709)	(0.0272195)
1000	0.302746	2.157266	0.306215	2.261185	0.301052	2.106573
	(0.0098937)	(0.010586)	(0.010067)	(0.031839)	(0.0058401)	(0.006834)
5000	0.300639	2.061905	0.303988	2.138359	0.300088	2.00396
	(0.0047439)	(0.005722)	(0.007628)	(0.0076394)	(0.0007219)	(0.002849)
10,000	0.300096	2.010522	0.300692	2.025630	2.999999	2.000594
	(0.0008375)	(0.0017544)	(0.0011784)	(0.004204)	(0.000063)	(0.0002893)
			$(\alpha, \theta) = (0.5, 3)$			
50	0.486303	4.182591	0.512058	4.202716	0.491703	4.011847
	(0.013681)	(0.145623)	(0.015136)	(0.158446)	(0.012834)	(0.114824)
100	0.5076018	4.116091	0.509162	4.152312	0.497109	3.963158
	(0.013418)	(0.134507)	(0.014995)	(0.138476)	(0.0086253)	(0.107127)
500	0.4979793	4.026146	0.500327	4.056057	0.499162	3.571908
	(0.0084954)	(0.107728)	(0.011643)	(0.110977)	(0.0061081)	(0.068135)
1000	0.499527	3.832826	0.496101	3.940611	0.499425	3.28083
	(0.008289)	(0.087229)	(0.010832)	(0.088392)	(0.0024507)	(0.019230)
5000	0.499863	3.298065	0.499362	3.480159	0.500206	3.072011
	(0.001840)	(0.007284)	(0.002372)	(0.030719)	(0.0006827)	(0.0056093)
10,000	0.499911	3.061087	0.500485	3.170842	0.5000082	3.006235
	(0.000635)	(0.004972)	(0.000904)	(0.008865)	(0.000192)	(0.000274)

5 Real data analysis

In this section, we discuss some possible applications of the PLINAR(1) model for two real count time series data.

The data give numbers of submissions to animal health laboratories, monthly from January 2003 to December 2009, from a region in New Zealand (Aghababaei Jazi, Jones and Lai (2012)). The submissions contain many categories for presenting symptoms. The first data set is skin lesions, given in Table 3, and the second data set is anorexia, given in Table 4. The two data series are empirically overdispersed with dispersion indices $\hat{I}_x = \frac{S_x^2}{\bar{X}} = 2.34$ and $\hat{I}_x = 3.53$, respectively. We apply the overdispersion test described in Schweer and Weiß (2014) with significance level $\alpha = 0.05$. The critical values for the two data series are 1.26 and 1.32, respectively. The observed value of the index of dispersion, $\hat{I}_x$, exceeds the critical value completely, hence the data series do not stem from an equidispersed Poisson INAR(1) process. Therefore, the Poisson–Lindley or negative binomial could appear to be more appropriate than the Poisson model for the two series. Further, using the MLE $\hat{\theta}$ of "i.i.d. Poisson Lindley" and MLEs $\hat{n}$ and $\hat{p}$ of "i.i.d. Negative Binomial" at Tables 5 and 6, we get the skewness and kurtosis

Table 3 *First time series*: *Skin-lesions (Mean* = 1.43, *Variance* = 3.36 *and* 1*st order sample autocorrelation* = 0.24)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2003	2	5	0	0	1	0	1	3	0	3	0	1
2004	3	3	6	3	1	0	0	0	0	0	0	1
2005	0	0	1	3	0	1	0	0	0	0	2	1
2006	3	1	1	2	3	1	0	2	2	1	6	0
2007	1	0	0	1	0	2	0	0	0	2	3	0
2008	2	4	1	1	0	0	1	1	1	8	1	3
2009	2	4	9	3	4	2	0	1	0	0	0	0

Table 4 *Second time series*: *Anorexia (Mean* = 0.82, *Variance* = 2.90 *and* 1*st order sample autocorrelation* = 0.49)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2003	0	1	3	1	4	1	1	4	11	2	1	1
2004	2	2	0	0	0	0	0	0	0	0	0	0
2005	0	0	0	0	0	0	0	0	0	0	0	1
2006	0	0	0	0	0	0	3	5	6	3	2	1
2007	0	0	0	0	0	0	1	0	2	0	0	0
2008	0	0	0	0	0	0	0	2	4	0	1	0
2009	1	0	0	0	2	1	0	0	0	0	0	0

Table 5 *Estimated parameters, AIC and RMS for the first time series*

Model	MLE	AIC	LL	BIC	CAIC	RMS
i.i.d. Poisson	$\hat{\lambda} = 1.43$	311.44	-154.72	316.30	311.58	
i.i.d. Geometric	$\hat{\mu} = 1.43$	278.4	-138.2	283.26	278.54	
i.i.d. Negative Binomial	$\hat{n} = 1.063$					
	$\hat{p} = 0.43$	280.4	-138.2	285.26	280.54	
i.i.d. Poisson–Lindley	$\hat{\theta} = 1.04$	278.6	-138.3	283.46	278.74	
PINAR(1)	$\hat{\lambda} = 1.18$					
	$\hat{\alpha} = 0.17$	306.22	-151.11	311.08	306.36	1.78
GINAR(1)	$\hat{p} = 0.58$					
	$\hat{\alpha} = 0.12$	277.72	-136.86	282.58	277.86	1.79
NGINAR(1)	$\hat{p} = 1.41$					
	$\hat{\alpha} = 0.17$	277.12	-136.56	281.98	277.26	1.78
NBRCINAR(1)	$\hat{n} = 1.2$					
	$\hat{p} = 0.45$					
	$\hat{\rho} = 0.17$	279.93	-136.96	284.79	280.07	1.78
NBIINAR(1)	$\hat{n} = 1.01$					
	$\hat{p} = 1.03$					
	$\hat{\rho} = 0.31$	276.14	-135.07	281.00	276.28	1.78
GPQINAR(1)	$\hat{\lambda} = 0.81$					
	$\hat{\theta} = 0.33$					
	$\hat{\rho} = 0.16$	280.43	-137.21	285.29	280.57	1.79
PLINAR(1)	$\hat{\theta} = 1.05$					
	$\hat{\alpha} = 0.25$	223.81	-109.9	228.67	223.95	1.42

for both negative binomial and Poisson Lindley of the two data series. As we can see from Table 7, the skewness and kurtosis of the Poisson–Lindley distribution are smaller than those of the negative binomial distribution for both data series. Also, the sample skewness and sample kurtosis of the two data series, written in parentheses, are (1.87, 7.06) and (3.38, 17.7), respectively. Hence, we conclude that the Poisson–Lindley distribution offers more flexibility for modeling the data series than the negative binomial. The sample paths and partial autocorrelation functions (PACFs) of the two series are shown in Figures 1 and 2, respectively. The figures suggest that first order autoregression models are appropriate for both data series. Also, we expect some positive correlation in the data series because symptoms of the disease changing gradually over time. This expectation was confirmed by getting the sample autocorrelations of the two data series.

For comparison purposes, we use the two data series and compare the PLINAR(1) model to the following INAR models: PINAR(1) (Al-Osh and Alzaid (1987)), NBRCINAR(1) (Weiß (2008)), NBIINAR(1) (Al-Osh and Aly (1992)), GPQINAR(1) (Alzaid and Al-Osh (1993)), GINAR(1) (Alzaid and Al-Osh (1988)), NGINAR(1) (Ristić, Bakouch and Nastić (2009)). For each INAR

Table 6 *Estimated parameters, AIC and RMS for the Second time series*

Model	MLE	AIC	LL	BIC	CAIC	RMS
i.i.d. Poisson	$\hat{\lambda} = 0.82$	264.4	−131.2	269.26	264.54	
i.i.d. Geometric	$\hat{\mu} = 0.82$	212.62	−105.31	217.48	212.76	
i.i.d. Negative Binomial	$\hat{n} = 0.32$					
	$\hat{p} = 0.28$	203.9	−99.95	208.76	204.04	
i.i.d. Poisson–Lindley	$\hat{\theta} = 1.67$	215.42	−106.71	220.28	215.56	
PINAR(1)	$\hat{\lambda} = 0.5$					
	$\hat{\alpha} = 0.38$	229.05	−112.52	233.91	229.19	1.51
GINAR(1)	$\hat{p} = 0.48$					
	$\hat{\alpha} = 0.38$	189.38	−92.67	194.24	189.52	1.49
NGINAR(1)	$\hat{p} = 0.76$					
	$\hat{\alpha} = 0.66$	183.97	−89.98	188.83	184.11	1.5
NBRCINAR(1)	$\hat{n} = 0.58$					
	$\hat{p} = 0.4$					
	$\hat{\rho} = 0.34$	193.77	−93.88	198.63	193.91	1.49
NBIINAR(1)	$\hat{n} = 0.26$					
	$\hat{p} = 0.95$					
	$\hat{\rho} = 0.65$	178.6	−86.3	183.46	178.74	1.5
GPQINAR(1)	$\hat{\lambda} = 0.34$					
	$\hat{\theta} = 0.38$					
	$\hat{\rho} = 0.35$	195.44	−94.72	200.30	195.58	1.51
PLINAR(1)	$\hat{\theta} = 1.71$					
	$\hat{\alpha} = 0.49$	174.19	−85.45	179.05	174.33	1.48

Table 7 *Skewness and kurtosis for both Negative Binomial and Poisson Lindley of the three data series*

Skin lessions	Negative Binomial	Poisson Lindley
skewness	2.59	1.8
kurtosis	8.98	7.59
Anorexia	Negative Binomial	Poisson Lindley
skewness	3.62	1.99
kurtosis	22.3	8.51

model, we obtain the maximum likelihood estimates, log-likelihood function (LL), Akaike information criterion (AIC), Bayesian information criterion (BIC), consistent Akaike information criterion (CAIC), and the root mean squares of differences of observations and predicted values RMS. The obtained results, for both data series, are presented in Tables 5 and 6. As we can see from Tables 5 and 6, the
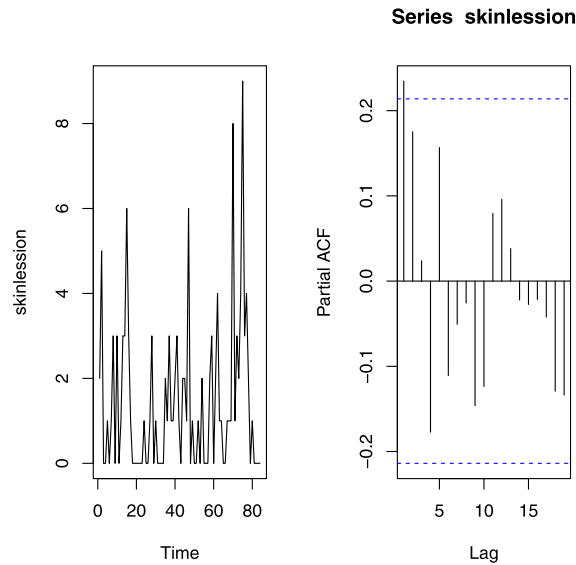
Series skinlession

Figure 1   *The sample path and sample partial autocorrelation function of the first time series.*
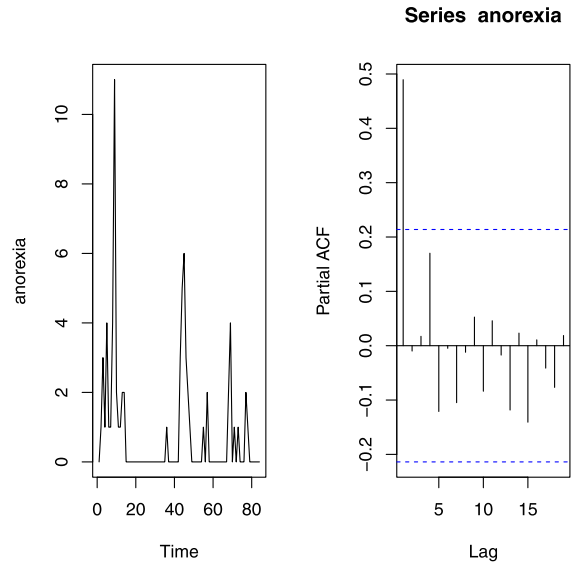
Series anorexia

Figure 2   *The sample path and sample partial autocorrelation function of the third time series.*

smallest values of the AIC, BIC, CAIC and RMS and the largest value of LL are obtained for the PLINAR(1) model. Thus, we can conclude that the PLINAR(1) model provides the best fit among the other INAR models. Also, it is worth to find that the PLINAR(1) model works well than other models with more two pa-

Table 8 *Prediction k months of out-sample time series*

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$	$k = 9$	$k = 10$
First data	1.062718	1.328397	1.394817	1.411422	1.415573	1.416611	1.416871	1.416935	1.416952	1.416956
Second data	0.408299	0.608366	0.706398	0.754434	0.777972	0.789505	0.795157	0.797926	0.799283	0.799948

rameters based on the value of LL that, in most, becomes larger for models with extra parameters. Table 8 displays the kth month ahead forecasting from January to October of the year 2010 for the two data sets via equation (3.1), where $\hat{\alpha}, \hat{\theta}$ of the data series are given in Tables 5, 6. From Table 8, we observe that:

- The conditional forecasts converge to the unconditional mean of the PLINAR(1) process for both data sets at $k = 10$, where unconditional mean for first and second data are $E(X_t) = 1.416957$, $E(X_t) = 0.800587$, respectively.
- The forecasted values of the data provide real values in general and thus they are reasonable to integer-valued nature of the data, so taking their greatest integers gives the forecasted values of the two data sets from January to October of the year 2010 which correspond $k = 1$ to $k = 10$. Therefore, the forecasted values for skin lesions data will be 1 and for Anorexia data will be 0 for January and 1 for other months in the mentioned period.

6 Conclusions

A new stationary first-order integer-valued autoregressive model with Poisson–Lindley marginal is introduced. We get the probability mass function of the innovation term of the process as a generalized mixture of geometric, negative binomial and degenerate distributions. Many properties of the model are obtained such as autocorrelation function, spectral density function, multi-step ahead conditional expectation, variance and partial autocorrelation function. The unknown parameters of the model are estimated using the methods of conditional least squares, Yule–Walker and maximum likelihood, and performance of the estimates of all methods is investigated via simulation. The model is fitted to two data sets taken from the New Zealand animal health laboratories and a small analysis of such data is justified under this model. It is shown that, the model is the best fit among the compared INAR(1) models based on some goodness-of-fit statistics among them log-likelihood function, Akaike information criterion (AIC), Bayesian information criterion (BIC), consistent Akaike information criterion (CAIC). Predictive capacity of the model for ten months of out-sample data are checked via the k-step ahead conditional mean. Finally, we hope that this model will be able to attract wider applicability in count time series data.

Some issues of future research may be represented in extending the results to bivariate case, censored time series data and Bayesian estimation.

Appendix section

In this appendix, we give the following proofs.

Proof of Lemma 2.1. The function $g(x)$ is a generalized mixture of $\text{Geo}(\frac{\theta}{1+\theta})$, $\text{NB}(2, \frac{\theta}{1+\theta})$ and $\text{Geo}(\frac{\theta+1}{\theta+1+\alpha})$, where Geo and NB denote the Geometric and Negative Binomial distributions, respectively. Since $\frac{\theta^2(1-\alpha)^2+\theta(1-\alpha^2)+2\alpha}{(\theta(1-\alpha)+1)^2} + \frac{(1-\alpha)}{\theta(1-\alpha)+1} + \frac{-\alpha}{(\theta(1-\alpha)+1)^2} = 1$, it follows that $\sum_{x=0}^{\infty} g(x) = 1$.

It remains to show that $g(x) \geq 0$ for $x = 0, 1, \ldots$. The function $g(x)$ in equation (2.3) can be written as

$$g(x) = \left(\frac{1}{1+\theta}\right)^x r(x),$$

where

$$r(x) = \left[A\frac{\theta}{1+\theta} + B\left(\frac{\theta}{1+\theta}\right)^2 (x+1) + C\frac{\theta+1}{\theta+1+\alpha}\left(\frac{\alpha(1+\theta)}{\theta+1+\alpha}\right)^x\right]$$

and

$$A = \frac{\theta^2(1-\alpha)^2 + \theta(1-\alpha^2) + 2\alpha}{(\theta(1-\alpha)+1)^2},$$

$$B = \frac{(1-\alpha)}{\theta(1-\alpha)+1}, \qquad C = \frac{-\alpha}{(\theta(1-\alpha)+1)^2}.$$

It is easy to verify $(r(x))' > 0$, for $x = 0, 1, \ldots$ and $\lim_{x \to \infty} r(x) = +\infty$. So the positivity of the function $g(x)$ follows by proving that $r(0) \geq 0$,

$$r(0) = \left[A\frac{\theta}{1+\theta} + B\left(\frac{\theta}{1+\theta}\right)^2 + C\frac{\theta+1}{\theta+1+\alpha}\right].$$

It suffices to notice that

$$r(0) \geq \frac{A}{2} + \frac{B}{4} + C,$$

and

$$\frac{A}{2} + \frac{B}{4} + C = \frac{\theta^2(1-\alpha)^2 + \theta(1-\alpha^2) + 2\alpha}{2(\theta(1-\alpha)+1)^2}$$

$$+ \frac{1-\alpha}{4(\theta(1-\alpha)+1)} - \frac{\alpha}{(\theta(1-\alpha)+1)^2}$$

$$= (1-\alpha)\frac{2\theta^2(1-\alpha) + \theta(3+\alpha) + 1}{4(\theta(1-\alpha)+1)^2} \geq 0$$

which completes the proof. $\qquad\square$

Proof of Theorem 2.1. By applying equation (2.2), the p.g.f. of ε can be written as

$$\Phi_\varepsilon(s) = \alpha + (1-\alpha)$$
$$\times \left[\left(s^2(\theta\alpha - \alpha) + s\left(-\theta^2(1+\alpha) - 2\theta\alpha + 2\alpha\right) \right. \right.$$
$$\left. \left. + \theta^3 + \theta^2(\alpha+2) + \theta\alpha - \alpha \right) / \left((1+\theta-s)^2\left(1+\theta+\alpha(1-s)\right)\right) \right].$$

Now using partial fraction decomposition, the last equation becomes

$$\alpha + (1-\alpha)\left[A\frac{\theta}{1+\theta-s} + B\frac{\theta^2}{(1+\theta-s)^2} + C\frac{(\theta+1)}{1+\theta+\alpha(1-s)} \right],$$

where A, B and C defined in Lemma 2.1. So the random variable ε is a mixture of discrete component 0 with probability α and a generalized mixture of $\mathrm{Geo}(\frac{\theta}{1+\theta})$, $\mathrm{NB}(2, \frac{\theta}{1+\theta})$, $\mathrm{Geo}(\frac{\theta+1}{\theta+1+\alpha})$ with probability $1-\alpha$. $\qquad\square$

Acknowledgements

References

Aghababaei Jazi, M., Jones, G. and Lai, C.-D. (2012). First-order integer valued AR processes with zero inflated Poisson innovations. *Journal of Time Series Analysis* **33**, 954–963. MR2991911

Al-Osh, M. A. and Aly, E. E. A. A. (1992). First order autoregressive time series with negative binomial and geometric marginals. *Communications in Statistics: Theory and Methods* **21**, 2483–2492. MR1186065

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

Aly, E-E. A. A. and Bouzar, N. (1994a). Explicit stationary distributions for some Galton-Watson processes with immigration. *Communications in Statistics Stochastic Models* **10**, 499–517. MR1268562

Aly, E. E. A. A. and Bouzar, N. (1994b). On some integer-valued autoregressive moving average models. *Journal of Multivariate Analysis* **50**, 132–151. MR1292612

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

Alzaid, A. A. and Al-Osh, M. A. (1993). Some autoregressive moving average processes with generalized Poisson marginal distributions. *Annals of the Institute of Statistical Mathematics* **45**, 223–232. MR1232490

Bakouch, H. S. and Ristic, M. M. (2010). Zero truncated Poisson integer-valued AR(1) model. *Metrika* **72**, 265–280. MR2725101

Ghitany, M. E. and Al-Mutairi, D. K. (2009). Estimation methods for the discrete Poisson–Lindley distribution. *Journal of Statistical Computation and Simulation* **79**, 1–9. MR2655672

McKenzie, E. (1985). Some simple models for discrete variate time series. *Water Resources Bulletin* **21**, 645–650.

McKenzie, E. (1986). Autoregressive moving-average processes with negative binomial and geometric distributions. *Advances in Applied Probability* **18**, 679–705. MR0857325

McKenzie, E. (1988). Some ARMA models for dependent sequences of Poisson counts. *Advances in Applied Probability* **20**, 822–835. MR0968000

McKenzie, E. (2003). Discrete variate time series. In *Stochastic Processes: Modelling and Simulation* (D. N. Shanbhag et al., eds.), *Amsterdam, North-Holland, Handbook of Statistics* **21**, 573–606. MR1973555

Ristić, M. M., Bakouch, H. S. and Nastić, A. S. (2009). A new geometric first-order integer-valued autoregressive (NGINAR(1)) process. *Journal of Statistical Planning and Inference* **139**, 2218–2226. MR2507983

Ristić, M. M., Nastić, A. S. and Bakouch, H. S. (2012). Estimation in an integer-valued autoregressive process with negative binomial marginals (NBINAR(1)). *Communications in Statistics: Theory and Methods* **41**, 606–618. MR2874378

Ristić, M. M., Nastic, S. A. and Miletic Ilic, V. A. (2013). A geometric time series model with dependent Bernoulli counting series. *Journal of Time Series Analysis* **34**, 466–476. MR3070869

Sankaran, M. (1970). The discrete Poisson–Lindley distribution. *Biometrics* **26**, 145–149.

Schweer, S. and Weiß, C. H. (2014). Compound Poisson INAR(1) processes: Stochastic properties and testing for overdispersion. *Computational Statistics and Data Analysis* **77**, 267–284. MR3210062

Steutel, F. W. and van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Annals of Probability* **7**, 893–899. MR0542141

Tjøstheim, D. (1986). Estimation in nonlinear time series models. *Stochastic Processes and Their Applications* **21**, 251–273. MR0833954

Weiß, C. H. (2008). Thining operations for modeling time series of count—a survey. *AStA Advances in Statistical Analysis* **92**, 319–341. MR2426093

M. Mohammadpour
M. Shirozhan
Faculty of Mathematical Sciences
University of Mazandaran
Babolsar
Iran
E-mail: m.mohammadpour@umz.ac.ir
         M.shirozhan@gmail.com

H. S. Bakouch
Department of Mathematics
Faculty of Science
Tanta University
Tanta
Egypt
E-mail: hnbakouch@yahoo.com
         hassan.bakouch@science.tanta.edu.eg