Brazilian Journal of Probability and Statistics
2015, Vol. 29, No. 1, 71–93
DOI: 10.1214/13-BJPS229

First order non-negative integer valued autoregressive processes with power series innovations

Marcelo Bourguignon and Klaus L. P. Vasconcellos

Universidade Federal de Pernambuco

Abstract. In this paper, we introduce a first order non-negative integer valued autoregressive process with power series innovations based on the binomial thinning. This new model contains, as particular cases, several models such as the Poisson INAR(1) model (Al-Osh and Alzaid (*J. Time Series Anal.* **8** (1987) 261–275)), the geometric INAR(1) model (Jazi, Jones and Lai (*J. Iran. Stat. Soc. (JIRSS)* **11** (2012) 173–190)) and many others. The main properties of the model are derived, such as mean, variance and the autocorrelation function. Yule–Walker, conditional least squares and conditional maximum likelihood estimators of the model parameters are derived. An extensive Monte Carlo experiment is conducted to evaluate the performances of these estimators in finite samples. Special sub-models are studied in some detail. Applications to two real data sets are given to show the flexibility and potentiality of the new model.

1 Introduction

In the last three decades, there has been a growing interest in discrete-valued time series models and several models for stationary processes with discrete marginal distributions have been proposed. Al-Osh and Alzaid (1987) proposed the first-order non-negative integer valued autoregressive (INAR(1)) process. Weiß (2009) proposed new autoregressive models for time series of binomial counts. Zhang et al. (2010) introduced pth-order integer valued autoregressive processes with a signed generalized power series thinning operator. Nastić et al. (2012) considered an integer valued autoregressive model of order p with geometric marginal distributions, using the negative binomial thinning. In a very recent paper, Jazi, Jones and Lai (2012b) introduce a new stationary first-order integer valued autoregressive process with zero inflated Poisson innovations.

In general, detailed studies have been conducted not only on the formulation of models but also on properties (Silva and Oliveira, 2004), estimation (Jung, Ronning and Tremayne, 2005), tests (Jung and Tremayne, 2003) and asymptotic distributions of model estimators (Freeland and McCabe, 2005) for different discrete marginal distributions.

Key words and phrases. Conditional maximum likelihood, INAR(1) process, power series distribution.

Received March 2013; accepted August 2013.

Much theoretical work has been concentrated on the use of the Poisson distribution as an integral feature of the model. However, the Poisson distribution is not always suitable for modelling, since its mean and variance are the same and this property may be unacceptable for real data. Furthermore, in many real-life situations there are series which do not contain zeros in a large period of time or are even permanently positive. In these situations, the Poisson distribution is also not suitable for modelling. The situation of zero truncation has been considered by Jazi, Jones and Lai (2012b), who have recently proposed a first-order integer valued AR model with zero inflated Poisson innovations. This has also been studied by Zhu (2012), who has recently proposed integer-valued GARCH models that are based upon the zero inflated Poisson distribution and the zero inflated negative binomial distribution.

There is, therefore, a need to introduce different integer-valued time series models to deal with different particular real situations, like overdispersion or zero-inflation. The idea of considering a distribution for the innovations such that the marginal distribution of the observations will satisfy a given property has been extensively discussed in Weiß (2008), where approaches on how to obtain, for example, the overdispersed negative binomial or generalized Poisson distribution are presented.

In this context, the main purpose of this paper is to propose a new first order non-negative integer valued autoregressive process with power series (PS) innovations based on the binomial thinning operator (Steutel and Van Harn, 1979). The motivation for such a process arises from its potential in modelling and analyzing non-negative integer valued time series when there is an indication of equidispersion, overdispersion, underdispersion or truncated distributions. The use of innovations that come from the PS family of distributions has many advantages, that family of distributions constituting a flexible framework for statistical modelling of discrete data in several real-life situations (Johnson, Kemp and Kotz, 2005).

We consider a sequence of discrete i.i.d. random variables $\{\varepsilon_t; t \in \mathbb{Z}\}$, the distribution of each ε_t being indexed by a parameter θ and defined by the probability mass function

$$\Pr(\varepsilon_t = x) = \frac{a(x)\theta^x}{C(\theta)}, \qquad x \in \mathcal{S}, \tag{1.1}$$

where the support $\mathcal{S}$ of ε_t is a subset of the non-negative integers, $a(x) \geq 0$ depends only on x and there is $s > 0$ such that $C(\theta) = \sum_{x=0}^{\infty} a(x)\theta^x$ is finite for all $\theta \in (0, s)$ (s can be ∞). Although we will always consider θ as a value in $(0, s)$, we will also assume that the power series for $C(\theta)$ converges, in fact, to a finite value for $\theta \in (-s, s)$. If this is the case, then, $C(\theta)$ has derivatives of all orders in $(-s, s)$ and those derivatives can be obtained by differentiating the power series term to term. Also, because $a(x) \geq 0$ for all x, $C(\theta)$ and all its derivatives will be positive in $(0, s)$. For more details on the PS class of distributions, see Noack (1950).

Table 1 *Some distributions in the family* (1)

Distribution	$a(x)$	$C(\theta)$	s	$\mathcal{S}$
1. Bernoulli	1	$1+\theta$	∞	$\{0, 1\}$
2. Binomial	$\binom{n}{x}$	$(1+\theta)^n$	∞	$\{0, 1, \dots, n\}$
3. Geometric	1	$(1-\theta)^{-1}$	1	$\{0, 1, 2, \dots\}$
4. Poisson	$x!^{-1}$	e^{θ}	∞	$\{0, 1, 2, \dots\}$
5. Negative Binomial	$\frac{\Gamma(r+x)}{x!\Gamma(r)}$	$(1-\theta)^{-r}$	1	$\{0, 1, 2, \dots\}$
6. Logarithmic	x^{-1}	$-\log(1-\theta)$	1	$\{1, 2, 3, \dots\}$

Table 1 provides the functions $a(x)$, $C(\theta)$ and the parameter θ corresponding to some special cases of PS distributions such as the Bernoulli, binomial (with n being the integer number of replicas), geometric, Poisson, negative binomial and logarithmic distributions. For the negative binomial, there may exist situations for which we will want r to be integer-valued. In this case, x can be regarded as the random number of failures until exactly r successes are recorded in a sequence of independent trials where the probability of failure is θ. When using the binomial distribution, the value of n may be known in advance or may be estimated; the same holds for the value of r when using the negative binomial distribution.

The paper is structured as follows. The PSINAR(1) (power series INAR(1)) model is formally defined in Section 2 and some of its basic properties are outlined. In Section 3, we propose estimation methods for the model parameters. Three special cases of the proposed model are studied in Section 4. In Section 5, we present some simulation results for the estimation methods. In Section 6, we provide applications to two real data sets. The paper is concluded in Section 7.

2 The PSINAR(1) model

Let Y be a non-negative integer valued random variable and $\alpha \in [0, 1]$. According to Steutel and Van Harn (1979), the binomial thinning operator "$\circ$" is defined as follows

$$\alpha \circ Y = \sum_{j=1}^{Y} Z_j, \tag{2.1}$$

where $\{Z_j\}_{j=1}^{Y}$ are independent and identically distributed (i.i.d.) random variables, independent of Y, with $\Pr(Z_j = 1) = 1 - \Pr(Z_j = 0) = \alpha$, that is, $\{Z_j\}_{j=1}^{Y}$ is an i.i.d. Bernoulli random sequence. Given Y, $\alpha \circ Y$ has a binomial distribution with parameters (Y, α). For an account of the properties of the binomial thinning operator, see Silva and Oliveira (2004). With this operator, the first-order non-negative integer valued autoregressive PS model can be defined. We set

$$Y_t = \alpha \circ Y_{t-1} + \varepsilon_t, \qquad t \in \mathbb{Z}, \tag{2.2}$$

where $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is a sequence of independent and identically distributed integer valued random variables with probability mass function satisfying (1.1), ε_t and y_{t-i} being independent for all $i \geq 1$. Since $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is an i.i.d. sequence with finite mean and variance, we conclude that this sequence is a second-order stationary process. Consequently, the process $\{Y_t\}_{t\in\mathbb{Z}}$ satisfying (2.2) is second-order stationary if $0 \leq \alpha < 1$ (Du and Li, 1991).

We can view a realization of $\{Y_t\}_{t\in\mathbb{Z}}$ as having two components (Freeland and McCabe, 2004): the survivors of elements of Y_{t-1}, each with probability α of survival, and the elements which entered the system in the interval $(t-1, t]$ (the innovation term $\{\varepsilon_t\}_{t\in\mathbb{Z}}$).

The question of which distribution to use for the $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ sequence may be rather subjective and it may depend also on the specific situation we are dealing with. For example, in Epidemiology, suppose that a researcher is monitoring the number of individuals in a given population that did not contract a specific disease, that is, suppose Y_t is the number of healthy individuals in the population at time t. Let α be the probability that a healthy individual remains healthy, that is, does not contract the disease, in the next instant of time. Suppose also that, with a very small probability, a sick individual may become cured of this disease, such that, in the next time, we will have no more than one cured individual in the population. Then, the evolution of Y_t may be described by (2.2), using a Bernoulli distribution for ε_t. Suppose now that this same researcher, wishing to observe the evolution of cure in a very specific group, prescribes a given medicine to n sick individuals of that group, such that some of the individuals taking that medicine may become cured. Now the evolution of Y_t may be described by (2.2), using a binomial distribution for ε_t. When treating a serious disease, we can consider that a given individual gets cured or that this same individual ultimately dies. Then, a way of monitoring the efficiency of a given treatment is to observe how many individuals get cured before an individual dies. The evolution of Y_t may then be described by (2.2), with a geometric distribution for ε_t. If we observe how many individuals get cured before m individuals die, then, we can use a negative binomial distribution for ε_t.

A reasonable choice for the distribution of ε_t may also follow from statistical considerations. If it seems reasonable that the mean and variance of the distribution of the observations are equal, then, a simple Poisson model may be adequate. If variance seems to be smaller than the mean, we must discard the geometric and Poisson distributions. A hypothesis test may be used to decide between a geometric and a negative binomial distribution. Also, it may seem reasonable that the observations are necessarily positive, which means that a truncated distribution should be used (see Section 4.3).

From (2.2), it follows that $\{Y_t\}_{t\in\mathbb{Z}}$ is a Markov process. The proofs of the next two propositions can be seen in Appendix A and Appendix B, respectively.

Proposition 1. *For fixed $n \in \mathbb{Z}^+$, the transition probabilities of this process are given by*

$$\Pr(Y_t = k | Y_{t-1} = l) = \begin{cases} \dfrac{1}{C(\theta)} \displaystyle\sum_{i=0}^{\min(l,k-n)} \binom{l}{i} \alpha^i (1-\alpha)^{l-i} \theta^{k-i} a(k-i), \\ \quad \textit{if } \mathcal{S} = \{n, n+1, n+2, \ldots\}, \\ \dfrac{1}{C(\theta)} \displaystyle\sum_{i=\max(0,k-n)}^{\min(l,k)} \binom{l}{i} \alpha^i (1-\alpha)^{l-i} \theta^{k-i} a(k-i), \\ \quad \textit{if } \mathcal{S} = \{0, 1, 2, \ldots, n\}, \end{cases} \tag{2.3}$$

for all $k, l \in \mathbb{Z}^+$, where $(\cdot)$ is the standard combinatorial symbol.

Proposition 2. *The Markov process defined by the transition probabilities above admits a unique stationary distribution.*

The marginal probability function of $\{Y_t\}_{t \in \mathbb{Z}}$ is given by

$$\Pr(Y_t = k) = \begin{cases} \dfrac{1}{C(\theta)} \displaystyle\sum_{l=0}^{\infty} \sum_{i=0}^{\min(l,k-n)} \binom{l}{i} \alpha^i (1-\alpha)^{l-i} \theta^{k-i} a(k-i) \Pr(Y_{t-1} = l), \\ \quad \text{if } \mathcal{S} = \{n, n+1, n+2, \ldots\}, \\ \dfrac{1}{C(\theta)} \displaystyle\sum_{l=0}^{\infty} \sum_{i=\max(0,k-n)}^{\min(l,k)} \binom{l}{i} \alpha^i (1-\alpha)^{l-i} \theta^{k-i} a(k-i) \Pr(Y_{t-1} = l), \\ \quad \text{if } \mathcal{S} = \{0, 1, 2, \ldots, n\}, \end{cases}$$

which is a mixture distribution. It is important to highlight that the support of the distribution of Y_t is *not* a finite set, even if $\mathcal{S}$ is. In fact, it is not difficult to see that, for all $k, l \in \mathbb{Z}^+$, there will exist a positive integer m such that $\Pr(Y_m = k | Y_0 = l) > 0$.

Although we know a unique stationary distribution to exist, the obtention of this stationary distribution from the above equations is in general a difficult task. Alternatively, if Ψ is the probability generating function for this stationary distribution, $\Psi(u) = \mathrm{E}[u^{Y_t}]$, it is not difficult to check that Ψ is that function satisfying $C(\theta)\Psi(u) = C(u\theta)\Psi(\alpha u + (1-\alpha))$, for all u, θ. However, this approach is still not easy. For that very simple situation where $C(\theta) = \mathrm{e}^{\theta}$, which corresponds to the classical Poisson INAR(1) model, we will readily obtain that the stationary distribution is Poisson, its expected value being $\theta/(1-\alpha)$. On the other hand, the general problem of obtaining the stationary distribution of the observations, given a particular $C(\theta)$, seems to be, for most situations, quite difficult.

The moments of the random variable $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ can be easily obtained from the probability generating function $\Psi_{\varepsilon_t}(u) = C(u\theta)/C(\theta)$. The expected value is $\mathrm{E}(\varepsilon_t) = \mu_\varepsilon = \theta G'(\theta)$ and the variance is $\mathrm{Var}(\varepsilon_t) = \sigma_\varepsilon^2 = \theta^2 G''(\theta) + \mu_\varepsilon$, where $G(\theta) = \log[C(\theta)]$, $G'(\theta) = dG(\theta)/d\theta$ and $G''(\theta) = d^2G(\theta)/d\theta^2$. These well-known results can be found, for example, in Johnson, Kemp and Kotz (2005).

The mean and variance of the process $\{Y_t\}_{t\in\mathbb{Z}}$ as defined in (2.2) are

$$\mathrm{E}(Y_t) = \mu = \frac{\theta G'(\theta)}{1-\alpha}$$

and

$$\mathrm{Var}(Y_t) = \sigma^2 = \frac{\theta}{1-\alpha}\left[G'(\theta) + \frac{\theta G''(\theta)}{1+\alpha}\right] = \mu + \frac{\theta^2 G''(\theta)}{1-\alpha^2}.$$

Observe that the variance can be smaller or greater than the mean, depending on the sign of $G''(\theta)$. The dispersion index, which is the variance to mean ratio, will be given by

$$\frac{\sigma^2}{\mu} = 1 + \frac{\theta G''(\theta)}{(1+\alpha)G'(\theta)}.$$

Also, the mean and variance will be equal when G is a linear function, which is the case of the Poisson distribution.

The expressions for the moments of the conditional and unconditional distributions of the observations in a general INAR(1) process can be found, for example, in Rajarshi (2012). For our specific process, we obtain the conditional expectation as

$$\mathrm{E}(Y_t|Y_{t-1}) = \alpha Y_{t-1} + \theta G'(\theta)$$

and the conditional variance as

$$\mathrm{Var}(Y_t|Y_{t-1}) = \alpha(1-\alpha)Y_{t-1} + \theta G'(\theta) + \theta^2 G''(\theta).$$

It is also easy to verify that the autocorrelation function (ACF) at lag k is given by

$$\mathrm{Corr}(Y_t, Y_{t-k}) = \rho(k) = \alpha^k, \qquad k \geq 1, \tag{2.4}$$

which obviously is restricted to be positive.

Next, we consider the problem of estimating the parameters.

3 Estimation of the unknown parameters

This section is concerned with the estimation of the two parameters of interest. We consider three estimation methods, namely, Yule–Walker, conditional least squares and conditional maximum likelihood.

3.1 Yule–Walker estimation

From a sample $Y_1, \ldots, Y_T$ of a stationary process $\{Y_t\}_{t\in\mathbb{Z}}$, the sample autocorrelation function is given by

$$\widehat{\rho}(k) = \frac{\sum_{t=1}^{T-k}(Y_t - \overline{Y})(Y_{t+k} - \overline{Y})}{\sum_{t=1}^{n}(Y_t - \overline{Y})^2},$$

where $\overline{Y} = 1/T \sum_{t=1}^{T} Y_t$ is the sample mean. The Yule–Walker (YW) estimator of α, based upon the fact that $\rho(k) = \alpha^k$, as in (2.4), is given by

$$\widehat{\alpha} = \widehat{\rho}(1) = \frac{\sum_{t=1}^{T-1}(Y_t - \overline{Y})(Y_{t+1} - \overline{Y})}{\sum_{t=1}^{T}(Y_t - \overline{Y})^2}. \tag{3.1}$$

The first moment of $\{Y_t\}_{t\in\mathbb{Z}}$ is given by $\mathrm{E}(Y_t) = \mu_\varepsilon/(1-\alpha)$. Using this, the estimator of μ_ε is defined as

$$\widehat{\mu}_\varepsilon = (1 - \widehat{\alpha})\overline{Y},$$

where $\widehat{\alpha}$ is given in (3.1). An estimator of the parameter θ can be obtained as the solution of the equation $\widehat{\theta} G'(\widehat{\theta}) = (1 - \widehat{\alpha})\overline{Y}$. The estimator of θ may have closed form, depending on which distribution is being used.

Du and Li (1991) showed that the usual mean estimator, the autocovariance and autocorrelation functions, given by $\overline{Y} = 1/T \sum_{t=1}^{T} Y_t$, $\widehat{\gamma}(k) = 1/T \sum_{t=1}^{T-k}(Y_t - \overline{Y})(Y_{t-k} - \overline{Y})$ and $\widehat{\rho}(k) = \widehat{\gamma}(k)/\widehat{\gamma}(0)$, $0 \le k \le T-1$, respectively, are strongly consistent.

3.2 Conditional least squares estimation

The conditional least squares estimator $\widehat{\boldsymbol{\eta}} = (\widehat{\alpha}, \widehat{\mu}_\varepsilon)^\top$ of $\boldsymbol{\eta} = (\alpha, \mu_\varepsilon)^\top$ is given by

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

where $S_T(\boldsymbol{\eta}) = \sum_{t=2}^{T}[Y_t - g(\boldsymbol{\eta}, Y_{t-1})]^2$ and $g(\boldsymbol{\eta}, Y_{t-1}) = \mathrm{E}(Y_t|Y_{t-1})$. Thus, following Klimko and Nelson (1978), the conditional least squares (CLS) estimators of α and μ_ε can be written in closed form as

$$\widehat{\alpha} = \frac{\sum_{t=2}^{T} Y_t Y_{t-1} - 1/(T-1)\sum_{t=2}^{T} Y_t \sum_{t=2}^{T} Y_{t-1}}{\sum_{t=2}^{T} Y_{t-1}^2 - 1/(T-1)(\sum_{t=2}^{T} Y_{t-1})^2} \tag{3.2}$$

and

$$\widehat{\mu}_\varepsilon = \frac{1}{T-1}\left(\sum_{t=2}^{T} Y_t - \widehat{\alpha}\sum_{t=2}^{T} Y_{t-1}\right),$$

where $\widehat{\alpha}$ is given in (3.2). As in Section 3.1, the estimator of the parameter θ can be obtained by solving the equation $\widehat{\theta} G'(\widehat{\theta}) = \widehat{\mu}_\varepsilon$. The estimator of θ may have closed form, depending on which distribution is being used.

3.3 Conditional maximum likelihood estimation

Suppose that y_1 is fixed. The conditional log-likelihood function for the PSINAR(1) model is given by

$$\ell(\alpha, \theta) = \log\left(\prod_{t=2}^{T} \Pr(Y_t|Y_{t-1})\right) = \sum_{t=2}^{T} \log\big(\Pr(Y_t|Y_{t-1})\big), \tag{3.3}$$

with $\Pr(Y_t|Y_{t-1})$ as in (2.3).

The conditional maximum likelihood (CML) estimators $\widehat{\alpha}$ and $\widehat{\theta}$ of α and θ are defined as the values of α and θ that maximize the conditional log-likelihood function in (3.3). There will be, in general, no closed form for the CML estimates and their obtention will need, in practice, numerical methods.

4 Special cases

In this section, we investigate some special cases of the PSINAR(1) model, giving expressions for mean and variance.

4.1 Geometric INAR(1) model

For $\mathcal{S} = \mathbb{Z}^+$ and $C(\theta) = (1-\theta)^{-1}, \theta \in (0, 1)$ in (1.1), we say that $\{Y_t\}_{t\in\mathbb{Z}}$ is a geometric INAR(1) model. Alzaid and Al-Osh (1988) introduced the INAR(1) process with geometric marginal distribution. Ristić, Bakouch and Nastić (2009) proposed the first-order integer valued autoregressive process with geometric marginal distribution based on negative binomial thinning. Jazi, Jones and Lai (2012a) studied the geometric INAR(1) process.

The transition probabilities of this process are given by

$$\Pr(Y_t = k | Y_{t-1} = l) = (1-\theta) \sum_{i=0}^{\min(k,l)} \binom{l}{i} \alpha^i (1-\alpha)^{l-i} \theta^{k-i}, \qquad 0 < \theta < 1.$$

The mean and variance of $\{Y_t\}_{t\in\mathbb{Z}}$ are

$$E(Y_t) = \mu = \frac{\theta}{(1-\alpha)(1-\theta)} \quad \text{and} \quad \operatorname{Var}(Y_t) = \sigma^2 = \frac{\theta + \alpha\theta(1-\theta)}{(1-\alpha^2)(1-\theta)^2}.$$

The conditional expectation and the conditional variance are given by

$$E(Y_t | Y_{t-1}) = \alpha Y_{t-1} + \frac{\theta}{1-\theta}$$

and

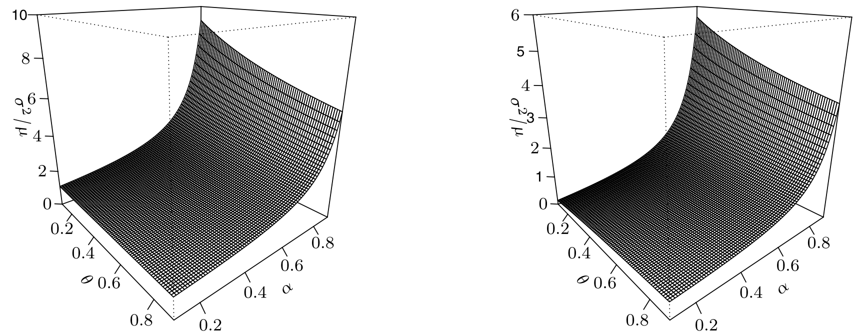
$$\operatorname{Var}(Y_t | Y_{t-1}) = \alpha(1-\alpha) Y_{t-1} + \frac{\theta}{(1-\theta)^2}.$$

Observe that $\mu = (1-\alpha)^{-1}[(1-\theta)^{-1} - 1]$ is an increasing function of α and θ. Also,

$$\sigma^2 = \frac{1}{1-\alpha^2} \cdot \frac{\theta}{(1-\theta)^2} + \frac{\alpha}{1-\alpha^2} \cdot \frac{\theta}{1-\theta}$$

is an increasing function of α and θ. Furthermore, we can easily obtain

$$\frac{\sigma^2}{\mu} = 1 + \frac{\theta}{(1+\alpha)(1-\theta)} > 1.$$



(a) Geometric INAR(1) (b) Logarithmic INAR(1)

Figure 1 *Plot of the variance-to-mean ratio against α and θ.*

The geometric INAR(1) process, therefore, may be used as a model for overdispersed non-negative integer valued time series. From the above expression, we can readily conclude that the variance-mean ratio is an increasing function of θ, but a decreasing function of α.

Figure 1(a) shows how σ^2/μ behaves as a function of α and θ. For more details about the geometric INAR(1) process, see Jazi, Jones and Lai (2012a).

4.2 Poisson INAR(1) model

For $\mathcal{S} = \mathbb{Z}^+$ and $C(\theta) = \mathrm{e}^\theta$ in (1.1), we say that $\{Y_t\}_{t\in\mathbb{Z}}$ is a Poisson INAR(1) model. Al-Osh and Alzaid (1987) proposed and studied the Poisson INAR(1) process. Many new results on it have been obtained in recent years. For example, Hellström (2001) focused on the testing of a unit root, Freeland and McCabe (2005) obtained asymptotic properties of CLS estimators, Weiß (2011) proposed several asymptotic simultaneous confidence regions for the two parameters.

The transition probabilities of this process are given by

$$\Pr(Y_t = k | Y_{t-1} = l) = \mathrm{e}^{-\theta} \sum_{i=0}^{\min(k,l)} \binom{l}{i} \alpha^i (1-\alpha)^{l-i} \frac{\theta^{k-i}}{(k-i)!}, \qquad \theta > 0.$$

The mean and variance of $\{Y_t\}_{t\in\mathbb{Z}}$ are

$$\mathrm{E}(Y_t) = \mu = \frac{\theta}{1-\alpha} \quad \text{and} \quad \mathrm{Var}(Y_t) = \sigma^2 = \frac{\theta}{1-\alpha}.$$

The mean and variance are equal for this model, increasing both with θ and α. For more details about the Poisson INAR(1) process, see Al-Osh and Alzaid (1987).

4.3 Truncated models

Truncated Poisson and negative binomial models have been discussed, among others, by Creel and Loomis (1990) and Grogger and Carson (1991).

Table 2 *Some distributions truncated at zero in the family* (1)

Distribution	$a(x)$	$C(\theta)$	s	$\mathcal{S}$
1. Binomial	$\binom{n}{x}$	$(1+\theta)^n - 1$	∞	$\{1, 2, \dots, n\}$
2. Geometric	1	$\theta(1-\theta)^{-1}$	1	$\{1, 2, 3, \dots\}$
3. Poisson	$x!^{-1}$	$e^{\theta} - 1$	∞	$\{1, 2, 3, \dots\}$

The PSINAR(1) model defined by (2.1) has the flexibility of modelling data by a Markovian process for which the state space is some proper subset of the nonnegative integers. This can be achieved, for example, by considering truncated distributions for the innovations.

Table 2 provides the functions $a(x)$ and $C(\theta)$ corresponding to some special cases of PS distributions truncated at zero.

The logarithmic INAR(1) can also be a model for a series of counts where zeros are not observed. For $\mathcal{S} = \{1, 2, 3, \dots\}$ and $C(\theta) = -\log(1-\theta)$, $\theta \in (0, 1)$ in (1.1), we define $\{Y_t\}_{t\in\mathbb{Z}}$ as the logarithmic INAR(1) model. The transition probabilities of this process are given by

$$\Pr(Y_t = k | Y_{t-1} = l) = -\frac{1}{\log(1-\theta)} \sum_{i=0}^{\min(k-1,l)} \binom{l}{i} \alpha^i (1-\alpha)^{l-i} \frac{\theta^{k-i}}{k-i}.$$

The mean and variance of $\{Y_t\}_{t\in\mathbb{Z}}$ are

$$\mathrm{E}(Y_t) = \mu = \frac{a\theta}{(1-\alpha)(1-\theta)}$$

and

$$\mathrm{Var}(Y_t) = \sigma^2 = \frac{a\theta[\alpha(1-\theta) + (1-a\theta)]}{(1-\alpha^2)(1-\theta)^2} = \frac{\mu[\alpha(1-\theta) + (1-a\theta)]}{(1+\alpha)(1-\theta)},$$

where $a = -1/\log(1-\theta)$. The conditional expectation and the conditional variance are given by

$$\mathrm{E}(Y_t | Y_{t-1}) = \alpha Y_{t-1} + \frac{a\theta}{1-\theta}$$

and

$$\mathrm{Var}(Y_t | Y_{t-1}) = \alpha(1-\alpha) Y_{t-1} + \frac{a\theta(1-a\theta)}{(1-\theta)^2}.$$

Both mean and variance are increasing functions of α and θ. The variance-mean ratio can be easily seen to be

$$\frac{\sigma^2}{\mu} = 1 + \frac{\theta(1-a)}{(1+\alpha)(1-\theta)};$$

it follows that this model presents

- equidispersion when $\theta = 1 - \mathrm{e}^{-1}$,
- overdispersion when $\theta > 1 - \mathrm{e}^{-1}$,
- underdispersion when $\theta < 1 - \mathrm{e}^{-1}$.

Figure 1(b) shows how σ^2/μ behaves as a function of α and θ.

5 Monte Carlo simulation study

The performances of the YW, CLS and CML estimators for a sample size of T observed values of $\{Y_t\}$ is the motivation of this section. Some numerical results for different values of the parameters α and θ are presented in Tables 3, 4 and 5.

Table 3 *Bias and MSE (in parentheses) of the parameters in geometric INAR(1)*

Sample size	Parameters	CLS	YW	CML
		$\alpha = 0.3, \theta = 0.3$		
$T = 100$	$\widehat{\alpha}$	−0.0216 (0.0120)	−0.0245 (0.0119)	−0.0097 (0.0078)
	$\widehat{\theta}$	0.0044 (0.0023)	0.0054 (0.0023)	0.0013 (0.0019)
$T = 200$	$\widehat{\alpha}$	−0.0086 (0.0064)	−0.0101 (0.0064)	−0.0027 (0.0038)
	$\widehat{\theta}$	0.0011 (0.0012)	0.0017 (0.0012)	−0.0003 (0.0009)
$T = 300$	$\widehat{\alpha}$	−0.0056 (0.0042)	−0.0065 (0.0042)	−0.0023 (0.0026)
	$\widehat{\theta}$	0.0002 (0.0008)	0.0005 (0.0008)	−0.0001 (0.0006)
		$\alpha = 0.7, \theta = 0.3$		
$T = 100$	$\widehat{\alpha}$	−0.0362 (0.0079)	−0.0435 (0.0085)	−0.0065 (0.0025)
	$\widehat{\theta}$	0.0166 (0.0042)	0.0215 (0.0043)	−0.0009 (0.0020)
$T = 200$	$\widehat{\alpha}$	−0.0168 (0.0036)	−0.0204 (0.0037)	−0.0042 (0.0012)
	$\widehat{\theta}$	0.0079 (0.0020)	0.0104 (0.0020)	0.0007 (0.0010)
$T = 300$	$\widehat{\alpha}$	−0.0100 (0.0024)	−0.0123 (0.0024)	−0.0007 (0.0008)
	$\widehat{\theta}$	0.0052 (0.0015)	0.0068 (0.0015)	−0.0005 (0.0006)
		$\alpha = 0.3, \theta = 0.7$		
$T = 100$	$\widehat{\alpha}$	−0.0232 (0.0097)	−0.0262 (0.0097)	0.0003 (0.0027)
	$\widehat{\theta}$	0.0024 (0.0013)	0.0033 (0.0013)	−0.0027 (0.0008)
$T = 200$	$\widehat{\alpha}$	−0.0095 (0.0048)	−0.0108 (0.0047)	0.0024 (0.0014)
	$\widehat{\theta}$	−0.0022 (0.0008)	−0.0020 (0.0008)	−0.0020 (0.0005)
$T = 300$	$\widehat{\alpha}$	−0.0077 (0.0033)	−0.0087 (0.0033)	−0.0005 (0.0009)
	$\widehat{\theta}$	0.0011 (0.0005)	0.0014 (0.0005)	−0.0008 (0.0003)
		$\alpha = 0.7, \theta = 0.7$		
$T = 100$	$\widehat{\alpha}$	−0.0335 (0.0070)	−0.0411 (0.0076)	−0.0016 (0.0009)
	$\widehat{\theta}$	0.0121 (0.0030)	0.0169 (0.0029)	−0.0024 (0.0010)
$T = 200$	$\widehat{\alpha}$	−0.0163 (0.0035)	−0.0195 (0.0036)	−0.0003 (0.0004)
	$\widehat{\theta}$	0.0053 (0.0017)	0.0073 (0.0017)	−0.0014 (0.0005)
$T = 300$	$\widehat{\alpha}$	−0.0099 (0.0019)	−0.0122 (0.0019)	0.0009 (0.0002)
	$\widehat{\theta}$	0.0042 (0.0010)	0.0056 (0.0010)	−0.0009 (0.0002)

Table 4 *Bias and MSE (in parentheses) of the parameters in Poisson INAR(1)*

Sample size	Parameters	CLS	YW	CML
		$\alpha = 0.3, \theta = 1.0$		
$T = 100$	$\widehat{\alpha}$	−0.0236 (0.0106)	−0.0265 (0.0105)	−0.0133 (0.0087)
	$\widehat{\theta}$	0.0303 (0.0282)	0.0348 (0.0284)	0.0160 (0.0251)
$T = 200$	$\widehat{\alpha}$	−0.0094 (0.0052)	−0.0109 (0.0052)	−0.0061 (0.0043)
	$\widehat{\theta}$	0.0119 (0.0146)	0.0143 (0.0145)	0.0072 (0.0128)
$T = 300$	$\widehat{\alpha}$	−0.0063 (0.0035)	−0.0073 (0.0035)	−0.0022 (0.0028)
	$\widehat{\theta}$	0.0098 (0.0096)	0.0114 (0.0096)	0.0038 (0.0080)
		$\alpha = 0.7, \theta = 1.0$		
$T = 100$	$\widehat{\alpha}$	−0.0317 (0.0068)	−0.0389 (0.0072)	−0.0050 (0.0022)
	$\widehat{\theta}$	0.0926 (0.0800)	0.1161 (0.0852)	0.0037 (0.0276)
$T = 200$	$\widehat{\alpha}$	−0.0155 (0.0033)	−0.0189 (0.0034)	−0.0024 (0.0011)
	$\widehat{\theta}$	0.0502 (0.0367)	0.0613 (0.0378)	0.0070 (0.0136)
$T = 300$	$\widehat{\alpha}$	−0.0090 (0.0019)	−0.0113 (0.0020)	−0.0004 (0.0007)
	$\widehat{\theta}$	0.0335 (0.0236)	0.0410 (0.0242)	0.0045 (0.0089)
		$\alpha = 0.3, \theta = 2.0$		
$T = 100$	$\widehat{\alpha}$	−0.0160 (0.0103)	−0.0187 (0.0102)	−0.0056 (0.0094)
	$\widehat{\theta}$	0.0445 (0.0965)	0.0531 (0.0960)	0.0147 (0.0886)
$T = 200$	$\widehat{\alpha}$	−0.0074 (0.0049)	−0.0088 (0.0049)	−0.0028 (0.0041)
	$\widehat{\theta}$	0.0194 (0.0482)	0.0236 (0.0481)	0.0065 (0.0418)
$T = 300$	$\widehat{\alpha}$	−0.0061 (0.0034)	−0.0072 (0.0034)	−0.0013 (0.0028)
	$\widehat{\theta}$	0.0178 (0.0312)	0.0207 (0.0312)	0.0038 (0.0266)
		$\alpha = 0.7, \theta = 2.0$		
$T = 100$	$\widehat{\alpha}$	−0.0335 (0.0075)	−0.0405 (0.0079)	−0.0029 (0.0022)
	$\widehat{\theta}$	0.2270 (0.3462)	0.2721 (0.3648)	0.0219 (0.1024)
$T = 200$	$\widehat{\alpha}$	−0.0164 (0.0033)	−0.0199 (0.0034)	−0.0026 (0.0011)
	$\widehat{\theta}$	0.1064 (0.1516)	0.1299 (0.1569)	0.0149 (0.0533)
$T = 300$	$\widehat{\alpha}$	−0.0111 (0.0021)	−0.0134 (0.0021)	−0.0009 (0.0007)
	$\widehat{\theta}$	0.0708 (0.0926)	0.0867 (0.0950)	0.0023 (0.0306)

The sample sizes considered were $T = 100$, 200 and 300. The Monte Carlo simulation experiments were performed using the R programming language; see http://www.r-project.org. The number of Monte Carlo replications was 1000. The CML estimates of α and θ are obtained by maximizing the conditional log-likelihood function using the BFGS quasi-Newton nonlinear optimization algorithm with numerical derivatives. For each different situation, we have estimated the bias and the mean squared error (MSE). The YW and CLS estimates for θ are not obtained directly for the logarithmic INAR(1) model. In this case, the estimate of θ is obtained as the value of θ that minimizes $g(\theta) = \{\theta/[(1-\theta)\log(1-\theta)] - \widehat{\mu}\}^2$.

Table 5 *Bias and MSE (in parentheses) of the parameters in logarithmic INAR(1)*

Sample size	Parameters	CLS	YW	CML
		$\alpha = 0.3, \theta = 0.3$		
$T = 100$	$\widehat{\alpha}$	−0.0171 (0.0112)	−0.0198 (0.0111)	−0.0007 (0.0026)
	$\widehat{\theta}$	−0.0211 (0.0562)	0.0147 (0.0537)	−0.0105 (0.0075)
$T = 200$	$\widehat{\alpha}$	−0.0100 (0.0061)	−0.0116 (0.0061)	−0.0010 (0.0013)
	$\widehat{\theta}$	−0.0085 (0.0250)	−0.0052 (0.0245)	−0.0064 (0.0040)
$T = 300$	$\widehat{\alpha}$	−0.0079 (0.0040)	−0.0088 (0.0039)	−0.0001 (0.0008)
	$\widehat{\theta}$	−0.0044 (0.0149)	−0.0025 (0.0131)	−0.0040 (0.0026)
		$\alpha = 0.7, \theta = 0.3$		
$T = 100$	$\widehat{\alpha}$	−0.0340 (0.0076)	−0.0408 (0.0081)	−0.0002 (0.0008)
	$\widehat{\theta}$	0.0169 (0.1167)	0.0466 (0.1039)	−0.0126 (0.0072)
$T = 200$	$\widehat{\alpha}$	−0.0160 (0.0033)	−0.0195 (0.0035)	−0.0001 (0.0004)
	$\widehat{\theta}$	0.0059 (0.0588)	0.0212 (0.0553)	−0.0066 (0.0035)
$T = 300$	$\widehat{\alpha}$	−0.0134 (0.0021)	−0.0156 (0.0022)	0.0001 (0.0002)
	$\widehat{\theta}$	0.0032 (0.0343)	0.0080 (0.0331)	−0.0046 (0.0024)
		$\alpha = 0.3, \theta = 0.7$		
$T = 100$	$\widehat{\alpha}$	−0.0219 (0.0113)	−0.0251 (0.0113)	−0.0006 (0.0021)
	$\widehat{\theta}$	−0.0149 (0.0154)	−0.0119 (0.0171)	−0.0091 (0.0029)
$T = 200$	$\widehat{\alpha}$	−0.0104 (0.0052)	−0.0119 (0.0052)	0.0001 (0.0010)
	$\widehat{\theta}$	−0.0054 (0.0040)	−0.0042 (0.0039)	−0.0045 (0.0014)
$T = 300$	$\widehat{\alpha}$	−0.0068 (0.0035)	−0.0078 (0.0035)	−0.0001 (0.0007)
	$\widehat{\theta}$	−0.0033 (0.0024)	−0.0026 (0.0024)	−0.0022 (0.0009)
		$\alpha = 0.7, \theta = 0.7$		
$T = 100$	$\widehat{\alpha}$	−0.0354 (0.0081)	−0.0433 (0.0088)	−0.0004 (0.0006)
	$\widehat{\theta}$	−0.0276 (0.1968)	−0.0103 (0.1397)	−0.0072 (0.0028)
$T = 200$	$\widehat{\alpha}$	−0.0143 (0.0031)	−0.0180 (0.0032)	−0.0002 (0.0003)
	$\widehat{\theta}$	−0.0050 (0.0380)	0.0088 (0.0371)	−0.0035 (0.0014)
$T = 300$	$\widehat{\alpha}$	−0.0087 (0.0020)	−0.0109 (0.0020)	0.0001 (0.0002)
	$\widehat{\theta}$	0.0016 (0.0100)	−0.0047 (0.0122)	−0.0020 (0.0010)

Tables 3, 4 and 5 present the biases and MSE's (given in parentheses) of the different estimators for geometric INAR(1), Poisson INAR(1) and logarithmic INAR(1) models, respectively. It is noteworthy that the CML estimators of the parameters α and θ display biases and MSE's that are much smaller than those of the corresponding YW and CLS for almost all sample sizes considered in the experiment. Note that as the sample size increases, the bias tends to zero in all three cases, confirming that the estimators are asymptotically unbiased.

It is expected for the CML estimator to have the best performance, since it uses the whole information of the distribution. The empirical investigation presented here suggests that, generally speaking, the CML is, in fact, much better than the

YW and CLS. Thus, we recommend the use of the CML method to estimate the model parameters of an INAR(1) process with PS innovation.

6 Applications to real data

We assess the efficiency of the proposed model in an analysis of real data. The first data set is obtained from the crime data section of the forecasting principles site (http://www.forecastingprinciples.com). This data series represents the counting of sex offences reported in the 21st police car beat in Pittsburgh, during one month. The data consist of 144 observations, starting in January 1990 and ending in December 2001. These data were previously studied by Ristić, Bakouch and Nastić (2009) and are listed in Table 6. The required numerical evaluations are implemented using the R software.

Table 7 displays some descriptive statistics. We see that the data set assumes the value 0. Thus, the logarithmic INAR(1) model is not appropriate. Furthermore, the sample variance is much larger than the sample mean, hence, the data seem to be overdispersed. Consequently, a Poisson marginal distribution would not be appropriate. The series and its sample autocorrelation are displayed in Figure 2.

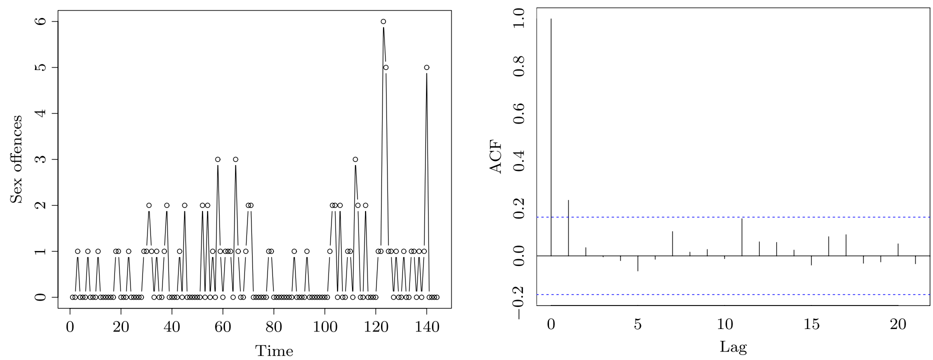
Analyzing Figure 2 we conclude that first order autoregressive models may be appropriate for the given data series. The behavior of the series indicates that it may

Table 6 *Sex offences*

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1990	0	0	1	0	0	0	1	0	0	0	1	0
1991	0	0	0	0	0	1	1	0	0	0	1	0
1992	0	0	0	0	1	1	2	1	0	1	0	0
1993	1	2	0	0	0	0	1	0	2	0	0	0
1994	0	0	0	2	0	2	0	1	0	3	1	0
1995	1	1	1	0	3	1	0	0	1	2	2	0
1996	0	0	0	0	0	1	1	0	0	0	0	0
1997	0	0	0	1	0	0	0	0	1	0	0	0
1998	0	0	0	0	0	1	2	2	0	2	0	0
1999	1	1	0	3	2	0	0	2	0	0	0	0
2000	1	1	6	5	1	1	0	1	0	0	1	0
2001	0	1	1	0	1	0	1	5	0	0	0	0

Table 7 *Descriptive statistics*

Min.	Q_1	Q_2	Mean	$\widehat{\rho}(1)$	Q_3	Max.	Var.
0.0000	0.0000	0.0000	0.5903	0.2348	1.0000	6.0000	1.0268



(a) Series (b) Sample autocorrelation

Figure 2 *Counting of sex offences in Pittsburgh with sample ACF.*

Table 8 *Estimated parameters (with corresponding standard errors in parentheses), AIC, RMS and MA*

Model	CML estimates	AIC	RMS	MA
Geometric INAR(1)	$\widehat{\alpha} = 0.1143$ (0.0754) $\widehat{\theta} = 0.3449$ (0.0364)	302.57	0.9913	0.7270
Negative Binomial INAR(1)	$\widehat{\alpha} = 0.2021$(0.0660) $\widehat{\theta} = 0.0794$ (0.0103) $\widehat{r} = 5.4993$ (0.0001)	102.15	0.9842	0.7237
NGINAR(1)	$\widehat{\alpha} = 0.1660$ (0.0965) $\widehat{\mu} = 0.5929$ (0.0958)	301.75	0.9862	0.7235

be a mean stationary time series. We compared the geometric INAR(1) with the Negative Binomial INAR(1) (corresponding to $C(\theta) = (1-\theta)^{-r}$) and also with the geometric first-order integer valued autoregressive (NGINAR(1)) model with geometric marginal distribution (Ristić, Bakouch and Nastić, 2009). Table 8 provides the CML estimates (with corresponding standard errors in parentheses) of the model parameters and three goodness-of-fit statistics: AIC (Akaike information criterion), RMS (root mean square of differences between observations and predicted values) and MA (absolute mean of differences between observations and predicted values). 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.

From Table 8, we conclude that the geometric INAR(1) and the NGINAR(1) models are competitive. Also, we can compare the AIC's to conclude that the proposed negative binomial INAR(1) model produces much better fits to the data. The

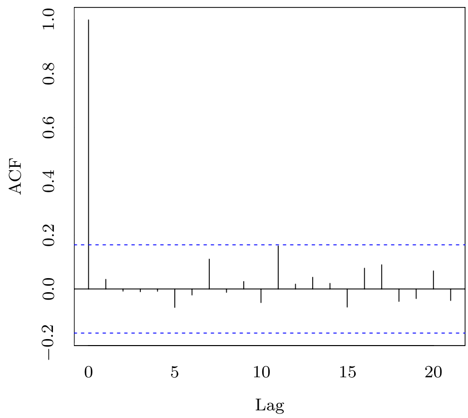


Figure 3 *Sample autocorrelations of the residuals obtained from Negative Binomial INAR*(1) *model.*

estimated model is

$$Y_t = 0.20 \circ Y_{t-1} + \varepsilon_t,$$

where $\varepsilon_t \sim \text{Negative Binomial}(5.50, 0.92)$. The sample autocorrelations of the residuals can be seen in Figure 3.

The second data set is given by Bakouch and Ristić (2010) as an application of their ZTPINAR(1) model, for which the marginal distribution of the observations is a zero truncated Poisson. Their original data counts family violence in the 11th police car beat in Pittsburgh, during one month. The data set is obtained from the crime data section of the forecasting principles site (http://www.forecastingprinciples.com). It consists of 144 observations, starting in January 1990 and ending in December 2001. In order to use their zero truncated Poisson model, the authors transformed the series, adding 1 to each observation. The transformed data are listed in Table 9.

Table 10 displays some descriptive statistics. We see that the transformed data set does not assume the value 0. Thus, the logarithmic INAR(1) and truncated Poisson INAR(1) may be also appropriate. The transformed series and its sample autocorrelations are displayed in Figure 4.

Analyzing Figure 4, we conclude that the first order autoregressive models may be appropriate for the given data series. The behavior of the series indicates that it may be a mean stationary time series. We compared the logarithmic INAR(1) and the truncated Poisson INAR(1) (corresponding to $C(\theta) = e^{\theta} - 1$) fittings with that of the ZTPINAR(1).

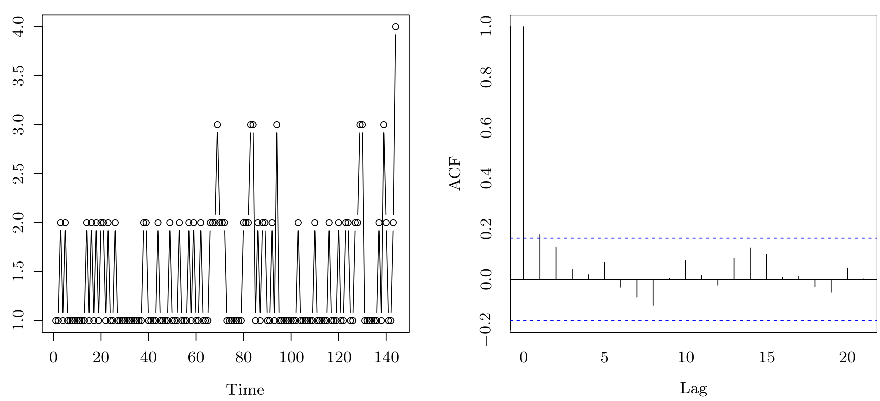
Table 11 provides the CML estimates (with corresponding standard errors in parentheses) of the model parameters and goodness of-fit statistics. From this table, we observe that the three models are competitive, the first two being only

Table 9 *Family violences*

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1990	1	1	2	1	2	1	1	1	1	1	1	1
1991	1	2	1	2	1	2	1	2	2	1	2	1
1992	1	2	1	1	1	1	1	1	1	1	1	1
1993	1	2	2	1	1	1	1	2	1	1	1	1
1994	2	1	1	1	2	1	1	1	2	1	2	1
1995	1	2	1	1	1	2	2	2	3	2	2	2
1996	1	1	1	1	1	1	1	2	2	2	3	3
1997	1	2	1	2	2	1	1	2	1	3	1	1
1998	1	1	1	1	1	1	2	1	1	1	1	1
1999	1	2	1	1	1	1	1	2	1	1	1	2
2000	1	1	2	2	1	1	2	2	3	3	1	1
2001	1	1	1	1	2	1	3	2	1	1	2	4

Table 10 *Descriptive statistics*

Min.	Q_1	Q_2	Mean	$\widehat{\rho}(1)$	Q_3	Max.	Var.
1.0000	1.0000	1.000	1.403	0.1770	2.0000	4.0000	0.3821



(a) Series (b) Sample autocorrelation

Figure 4 *The sample path of the second time series and the autocorrelation function.*

marginally better. The estimated truncated Poisson INAR(1) model, which is only marginally better than the other two, is

$$Y_t = 0.21 \circ Y_{t-1} + \varepsilon_t,$$

where $\varepsilon_t \sim$ Truncated Poisson(0.24). The sample autocorrelations of the residuals are shown in Figure 5.

Table 11 *Estimated parameters (with corresponding standard errors in parentheses), AIC, RMS and MA*

Model	CML estimates	AIC	RMS	MA
Logarithmic INAR(1)	$\widehat{\alpha} = 0.2199$ (0.0447) $\widehat{\theta} = 0.1727$ (0.0798)	233.21	0.6061	0.5205
Truncated Poisson INAR(1)	$\widehat{\alpha} = 0.2045$ (0.0569) $\widehat{\theta} = 0.2356$ (0.1378)	232.87	0.6059	0.5214
ZTPINAR(1)	$\widehat{\alpha} = 0.4202$ (0.0878) $\widehat{\lambda} = 0.7450$ (0.1142)	233.46	0.6061	0.5221

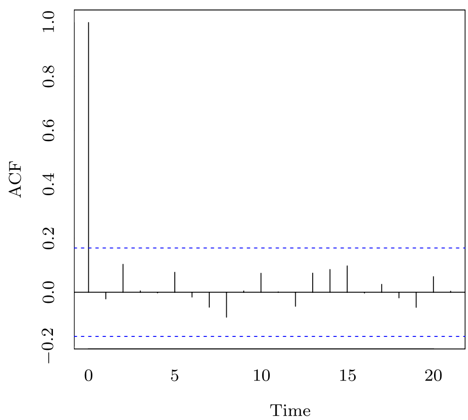


Figure 5 *Sample autocorrelations of the residuals obtained from truncated Poisson INAR(1) model.*

7 Concluding remarks

In this paper, we introduce first order non-negative integer valued autoregressive processes with power series innovations based on binomial thinning. The main properties of the model are derived, such as the mean, variance, autocorrelation function and transition probabilities. Three methods for estimating the model parameters are considered. Special sub-models (Geometric INAR(1), Poisson INAR(1) and Logarithmic INAR(1) models) are studied in some detail. We observe that the use of innovations that come from a PS distribution has many advantages, and allows us to create processes for modelling series of counts in several real-life situations. Indeed, the general Markovian process that is studied in this paper offers many modelling options to the user depending on the characteristics of the data. In the simulation study, we compare YW, CLS and CML estimators. The simulation results show that the YW and CLS methods produce estimators with similar performances and that CML is much better. Thus, we recommend the use of the CML method to estimate the model parameters of an INAR(1) process with PS innovations. Finally, we fitted PS models to two real data sets to show the potential of the new proposed model. These applications also demonstrate the practical relevance of the new model.

Appendix A: Proof of Proposition 1

For the PSINAR(1) process with binomial thinning operation, the conditional distribution of Y_t given Y_{t-1} is the convolution of the binomial distribution of the result of the thinning operation, $\alpha \circ Y_{t-1}$, with the PS distribution of the innovation process, ε_t (Sprott, 1983). Thus, let $\bullet$ denote convolution, let

$$f_1(i) = \binom{Y_{t-1}}{i} \alpha^i (1-\alpha)^{Y_{t-1}-i}, \qquad i = 0, 1, 2, \ldots, Y_{t-1} \quad \text{and}$$

$$f_2(i) = \frac{\theta^i a(i)}{C(\theta)}, \qquad i \in S.$$

Then,

$$\Pr(Y_t = k | Y_{t-1} = l) = f_1 \bullet f_2 = \sum_i f_1(i) f_2(k-i).$$

If $S = \{n, n+1, n+2, \ldots\}$ for fixed $n \in \mathbb{Z}^+$, then

$$0 \le i \le l \quad \text{and} \quad k - i \ge n \quad \Rightarrow \quad 0 \le i \le \min(l, k-n),$$

thus

$$\Pr(Y_t = k | Y_{t-1} = l) = \sum_{i=0}^{\min(l,k-n)} f_1(i) f_2(k-i).$$

If $S = \{0, 1, 2, \ldots, n\}$ for fixed $n \in \mathbb{Z}^+$, then

$$0 \le i \le l \quad \text{and} \quad 0 \le k - i \le n \quad \Rightarrow \quad \max(0, k-n) \le i \le \min(l, k),$$

thus

$$\Pr(Y_t = k | Y_{t-1} = l) = \sum_{i=\max(0,k-n)}^{\min(l,k)} f_1(i) f_2(k-i).$$

Appendix B: Proof of Proposition 2

Since $\Pr(Y_t = k | Y_{t-1} = l) > 0$, for all k, l, our process is an *irreducible* process, in the sense that every $k \in \mathcal{S}$ can be reached from every $l \in \mathcal{S}$. It also has stationary transition probabilities, in the sense that these transition probabilities do not involve t. Let $P^t(l,k) = \Pr(Y_t = k|Y_0 = l)$. Following Hoel, Port and Stone (1972), the existence of a stationary distribution is equivalent to

$$\lim_{t\to\infty} \frac{1}{t} \sum_{m=1}^{t} P^m(x,x) > 0, \qquad x \in \mathcal{S}.$$

From Hoel, Port and Stone (1972), if the inequality above is valid for a particular $y \in \mathcal{S}$, it will then, because our process is irreducible, be valid for all $x \in \mathcal{S}$. We will suppose, without loss of generality, that the smallest element of $\mathcal{S}$ is zero, in the sense that $a(0) > 0$ in the power series expansion of $C(\theta)$. Then, we will prove that

$$\lim_{t\to\infty} \frac{1}{t} \sum_{m=1}^{t} P^m(0,0) > 0.$$

This will be proved if we show that $\lim_{m\to\infty} P^m(0,0)$ exists and it is positive. Equivalently, we will show that $\lim_{m\to\infty} \log(P^m(0,0))$ exists and is finite.

We begin by showing by induction that

$$P^m(x,0) = \frac{a(0)(1-\alpha^m)^x}{C(\theta)^m} \prod_{i=1}^{m-1} C\big(\theta(1-\alpha^i)\big).$$

For $m = 1$, the above expression is reduced to $\Pr(Y_1 = 0|Y_0 = x) = a(0)(1-\alpha)^x/C(\theta)$, which is trivially true. Suppose it is valid for a given m. Then

$$\begin{aligned}
P^{m+1}(x,0) &= \sum_{z=0}^{\infty} P^1(x,z)P^m(z,0) \\
&= \frac{a(0)}{C(\theta)^{m+1}} \left(\prod_{i=1}^{m-1} C\big(\theta(1-\alpha^i)\big)\right) \\
&\quad \times \sum_{z=0}^{\infty} \sum_{i=0}^{\min(x,z)} (1-\alpha^m)^z \binom{x}{i} \alpha^i (1-\alpha)^{x-i} \theta^{z-i} a(z-i) \\
&= \frac{a(0)}{C(\theta)^{m+1}} \left(\prod_{i=1}^{m-1} C\big(\theta(1-\alpha^i)\big)\right) \\
&\quad \times \sum_{i=0}^{x} \sum_{z=i}^{\infty} (1-\alpha^m)^z \binom{x}{i} \alpha^i (1-\alpha)^{x-i} \theta^{z-i} a(z-i) \\
&= \frac{a(0)}{C(\theta)^{m+1}} \left(\prod_{i=1}^{m-1} C\big(\theta(1-\alpha^i)\big)\right)
\end{aligned}$$

$$
\begin{aligned}
&\times \sum_{i=0}^{x}\big(1-\alpha^m\big)^i \binom{x}{i}\alpha^i (1-\alpha)^{x-i} \sum_{z=i}^{\infty}\big(1-\alpha^m\big)^{z-i}\theta^{z-i}a(z-i) \\
&= \frac{a(0)}{C(\theta)^{m+1}}\left(\prod_{i=1}^{m-1} C\big(\theta\big(1-\alpha^i\big)\big)\right) \\
&\quad \times C\big(\theta\big(1-\alpha^m\big)\big)\sum_{i=0}^{x}\big(1-\alpha^m\big)^i \binom{x}{i}\alpha^i (1-\alpha)^{x-i} \\
&= \frac{a(0)}{C(\theta)^{m+1}}\left(\prod_{i=1}^{m} C\big(\theta\big(1-\alpha^i\big)\big)\right)\sum_{i=0}^{x}\left(\frac{1-\alpha^m}{1-\alpha}\right)^i \binom{x}{i}\alpha^i (1-\alpha)^{x} \\
&= \frac{a(0)(1-\alpha)^x}{C(\theta)^{m+1}}\left(\prod_{i=1}^{m} C\big(\theta\big(1-\alpha^i\big)\big)\right) \\
&\quad \times \sum_{i=0}^{x}\binom{x}{i}\big(1+\alpha+\cdots+\alpha^{m-1}\big)^i \alpha^i \\
&= \frac{a(0)(1-\alpha)^x}{C(\theta)^{m+1}}\left(\prod_{i=1}^{m} C\big(\theta\big(1-\alpha^i\big)\big)\right)\sum_{i=0}^{x}\binom{x}{i}\big(\alpha+\alpha^2+\cdots+\alpha^m\big)^i \\
&= \frac{a(0)(1-\alpha)^x}{C(\theta)^{m+1}}\left(\prod_{i=1}^{m} C\big(\theta\big(1-\alpha^i\big)\big)\right)\big(1+\alpha+\cdots+\alpha^m\big)^x \\
&= \frac{a(0)(1-\alpha^{m+1})^x}{C(\theta)^{m+1}}\prod_{i=1}^{m} C\big(\theta\big(1-\alpha^i\big)\big).
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
P^{m+1}(0,0) &= \sum_{z=0}^{\infty} P^1(0,z)P^m(z,0) \\
&= \frac{a(0)}{C(\theta)^{m+1}}\left(\prod_{i=1}^{m-1} C\big(\theta\big(1-\alpha^i\big)\big)\right)\sum_{z=0}^{\infty} a(z)\theta^z\big(1-\alpha^m\big)^z \\
&= \frac{a(0)}{C(\theta)^{m+1}}\left(\prod_{i=1}^{m-1} C\big(\theta\big(1-\alpha^i\big)\big)\right)C\big(\theta\big(1-\alpha^m\big)\big) \\
&= \frac{a(0)}{C(\theta)^{m+1}}\prod_{i=1}^{m} C\big(\theta\big(1-\alpha^i\big)\big) \\
&= \frac{a(0)}{C(\theta)}\prod_{i=1}^{m}\left(\frac{C(\theta(1-\alpha^i))}{C(\theta)}\right).
\end{aligned}
$$

Observe that for $m=0$ we get $P^1(0,0)=a(0)/C(\theta)$, which is trivially true. Now, we have

$$\log\bigl(P^{m+1}(0,0)\bigr)=\log\bigl(a(0)\bigr)-\log\bigl(C(\theta)\bigr)-\sum_{i=1}^{m}\bigl[\log\bigl(C(\theta)\bigr)-\log\bigl(C\bigl(\theta\bigl(1-\alpha^i\bigr)\bigr)\bigr)\bigr].$$

Let $G(\theta)=\log(C(\theta))$. From the intermediate value theorem, we can, for each i, obtain a value $\theta_i\in((1-\alpha^i)\theta,\theta)$ such that $G(\theta)-G(\theta(1-\alpha^i))=G'(\theta_i)\theta\alpha^i$. Now,

$$\log\bigl(P^{m+1}(0,0)\bigr)=\log\bigl(a(0)\bigr)-G(\theta)-\theta\sum_{i=1}^{m}G'(\theta_i)\alpha^i.$$

Because $G'=C'/C$ is positive, the sum above is a sum of positive terms. It remains, then, to show that the infinite series converges if we let $m\to\infty$. But this is immediate. Observe that $\theta_i\in((1-\alpha^i)\theta,\theta)\subset[(1-\alpha)\theta,\theta]$. Since G' has a derivative, $G''=(C''/C)-(G')^2$, G' must be continuous. Let $M(\theta)$ be the maximum value of G' in $[(1-\alpha)\theta,\theta]$. Then

$$0\le\sum_{i=1}^{\infty}G'(\theta_i)\alpha^i\le M(\theta)\sum_{i=1}^{\infty}\alpha^i=\frac{\alpha M(\theta)}{1-\alpha}<\infty.$$

Acknowledgments

We gratefully acknowledge the financial support from the Brazilian governmental institutions CAPES and CNPq. We are also grateful to the associate editor and two anonymous referees for helpful comments and suggestions.

References

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

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

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

Creel, M. D. and Loomis, J. B. (1990). Theoretical and empirical advantages of truncated count data estimators for analysis of deer hunting in California. *American Journal of Agricultural Economics* **72**, 434–441. DOI:10.2307/1242345.

Du, J. and Li, Y. (1991). The integer valued autoregressive (INAR(p)) model. *Journal of Time Series Analysis* **12**, 129–142. MR1108796

Freeland, R. K. and McCabe, B. P. M. (2004). Analysis of low count time series data by Poisson autoregression. *Journal of Time Series Analysis* **25**, 701–722. MR2089191

Freeland, R. K. and McCabe, B. P. M. (2005). Asymptotic properties of CLS estimators in the Poisson AR(1) model. *Statistcs & Probability Letters* **73**, 147–153. MR2159250

Grogger, J. T. and Carson, R. T. (1991). Models for truncated counts. *Journal of Applied Econometrics* **6**, 225–238. DOI:10.1002/jae.3950060302.

Hellström, J. (2001). Unit root testing in integer valued AR(1) models. *Economics Letters* **70**, 9–14. DOI:10.1016/S0165-1765(00)00344-X.
Hoel, P. G., Port, S. C. and Stone, C. J. (1972). *Introduction to Stochastic Processes*. Long Grove, IL: Waveland Press.
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
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. MR2991911
Johnson, N. L., Kemp, A. K. and Kotz, S. (2005). *Univariate Discrete Distributions* 3rd ed. Hoboken, NJ: Wiley. MR2163227
Jung, R. C. and Tremayne, A. R. (2003). Testing for serial dependence in time series models of counts. *Journal of Time Series Analysis* **24**, 65–84. MR1960383
Jung, R. C., Ronning, G. and Tremayne, A. R. (2005). Estimation in conditional first order autoregression with discrete support. *Statistical Papers* **46**, 195–224. MR2155712
Klimko, L. and Nelson, P. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistic* **6**, 629–642. MR0494770
Nastić, A. S., Ristić, M. M. and Bakouch, H. S. (2012). A combined geometric INAR(p) model based on negative binomial thinning. *Mathematical and Computer Modelling* **55**, 1665–1672. MR2899113
Noack, A. (1950). A class of random variables with discrete distributions. *Annals of Mathematical Statistics* **21**, 127–132. MR0033472
Rajarshi, M. B. (2012). *Statistical Inference for Discrete Time Stochastic Processes*. New Delhi: Springer. MR2976569
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
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. MR2062676
Sprott, D. A. (1983). Estimating the parameters of a convolution by maximum likelihood. *Journal of the American Statistical Association* **78**, 457–460. MR0711123
Steutel, F. W. and Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability* **7**, 893–899. MR0542141
Weiß, C. H. (2008). Thinning operations for modelling time series of counts—a survey. *Advances in Statistical Analysis* **92**, 319–341. MR2426093
Weiß, C. H. (2009). A new class of autoregressive models for time series of binomial counts. *Communications in Statistics—Theory and Methods* **38**, 447–460. MR2510797
Weiß, C. H. (2011). Simultaneous confidence regions for the parameters of a Poisson INAR(1) model. *Statistical Methodology* **8**, 517–527. MR2834037
Zhang, H., Wang, D. and Zhu, F. (2010). Inference for INAR(p) processes with signed generalized power series thinning operator. *Journal of Statistical Planning and Inference* **140**, 667–683. MR2558396
Zhu, F. (2012). Zero inflated Poisson and negative binomial integer-valued GARCH models. *Journal of Statistical Planning and Inference* **142**, 826–839. MR2863870

Departamento de Estatística
Universidade Federal de Pernambuco
Cidade Universitária
50740-540, Recife, PE
Brazil
E-mail: m.p.bourguignon@gmail.com
klaus@de.ufpe.br