Extended Poisson INAR(1) processes with equidispersion, underdispersion and overdispersion

Marcelo Bourguignon\textsuperscript{a}, Josemar Rodrigues\textsuperscript{b} and Manoel Santos-Neto \textsuperscript{a,c,d}

\textsuperscript{a}Department of Statistics, Federal University of Rio Grande do Norte, Natal, Brazil; \textsuperscript{b}Department of Mathematics and Statistics, University of São Paulo, São Carlos, Brazil; \textsuperscript{c}Department of Statistics, Federal University of São Carlos, São Carlos, Brazil; \textsuperscript{d}Department of Statistics, Federal University of Campina Grande, Campina Grande, Brazil

**ABSTRACT**

Real count data time series often show the phenomenon of the underdispersion and overdispersion. In this paper, we develop two extensions of the first-order integer-valued autoregressive process with Poisson innovations, based on binomial thinning, for modeling integer-valued time series with equidispersion, underdispersion, and overdispersion. The main properties of the models are derived. The methods of conditional maximum likelihood, Yule–Walker, and conditional least squares are used for estimating the parameters, and their asymptotic properties are established. We also use a test based on our processes for checking if the count time series considered is overdispersed or underdispersed. The proposed models are fitted to time series of the weekly number of syphilis cases and monthly counts of family violence illustrating its capabilities in challenging the overdispersed and underdispersed count data.

**ARTICLE HISTORY**

Received 25 September 2017
Accepted 22 March 2018

**KEYWORDS**

Double Poisson distribution; generalized Poisson distribution; INAR(1) process; overdispersion; underdispersion

1. Introduction

McKenzie [30] and Al-Osh and Alzaid [1], independently, introduced the integer-valued autoregressive (INAR) process \( \{X_t\}_{t \in \mathbb{Z}} \) with one lag using binomial thinning operator as follows

\[
X_t = \alpha \circ X_{t-1} + \epsilon_t, \quad t \in \mathbb{Z},
\]

where \( 0 \leq \alpha < 1 \), \( \{\epsilon_t\}_{t \in \mathbb{Z}} \) is a sequence of independent and identically distributed integer-valued random variables, called innovations, with \( \epsilon_t \) independent of \( X_{t-k} \) for all \( k \geq 1 \), \( E(\epsilon_t) = \mu_\epsilon \) and \( \text{Var}(\epsilon_t) = \sigma_\epsilon^2 \). The binomial thinning operator ‘\( \circ \)’ [36] is defined by

\[
\alpha \circ X_{t-1} := \sum_{j=1}^{X_{t-1}} W_j,
\]

where the counting series \( \{W_j\}_{j \geq 1} \) is a sequence of independent and identically distributed Bernoulli random variables with \( \Pr(W_j = 1) = 1 - \Pr(W_j = 0) = \alpha \). From the results of Al-Osh and Alzaid [1], we have that \( \alpha \in [0, 1) \) and \( \alpha = 1 \) are the conditions of (strictly) stationarity and non-stationarity of the process \( \{X_t\}_{t \in \mathbb{Z}} \), respectively. Also, \( \alpha = 0 \) \( (\alpha > 0) \) implies the independence (dependence) of the observations of \( \{X_t\}_{t \in \mathbb{Z}} \).

CONTACT Marcelo Bourguignon  m.p.bourguignon@gmail.com

© 2018 Informa UK Limited, trading as Taylor & Francis Group
The practical motivation for this type of process is the need to model count series with correlated observations. Some examples are daily counts of epileptic seizure in one patient, the number of generics in the pharmaceutical market, the number of guest nights in hotels, the number of different IP addresses, the monthly number of active customers of a mobile phone service provider, daily number of traded stocks in a firm, daily number of visitors to a website, monthly incidence of a disease, and so on. For more details see, for example, Hellström [20], Brännäs et al. [10], Weiß [37,39], Barreto-Souza and Bourguignon [4], Bourguignon and Vasconcellos [8] and Bourguignon and Weiß [9].

If $\varphi_X(s)$ and $\varphi_\epsilon(s)$ denote the probability generating function (pgf) of $\{X_t\}_{t \in \mathbb{Z}}$ and $\{\epsilon_t\}_{t \in \mathbb{Z}}$, respectively, then the stationary marginal distribution of $\{X_t\}_{t \in \mathbb{Z}}$ can be determined from the equation $\varphi_X(s) = \varphi_X(1 - \alpha(1 - s)) \cdot \varphi_\epsilon(s)$, which allows for various types of marginal distributions, including the Poisson [1], geometric [2], generalized Poisson (GP) [38] and Poisson-geometric [6] distributions. Also, the marginal distribution of model (1) may be expressed in terms of arrival process $\{\epsilon_t\}_{t \in \mathbb{Z}}$ [1] as $X_t \overset{d}{=} \sum_{i=0}^{\infty} \alpha^i \circ \epsilon_{t-i}$. Thus, the pgf of $\{X_t\}_{t \in \mathbb{Z}}$ is given by

$$
\varphi_X(s) = \prod_{i=0}^{\infty} \varphi_\epsilon(1 - \alpha^i + \alpha^i s).
$$

The INAR(1) process is a homogeneous Markov chain and the 1-step transition probabilities of this process are given by

$$
\Pr(X_t = k | X_{t-1} = l) = \sum_{i=0}^{\min(k,l)} \Pr(B_{t-1}^\alpha = i) \cdot \Pr(\epsilon_t = k - i), \quad k, l \geq 0,
$$

where $B_{t-1}^\alpha \sim \text{Binomial}(\alpha, n)$ with $\alpha \in (0, 1)$ and $n \in \mathbb{N}$. The mean and variance of $\{X_t\}_{t \in \mathbb{Z}}$ are given by

$$
\mu_X := \mathbb{E}(X_t) = \frac{\mu_\epsilon}{1 - \alpha} \quad \text{and} \quad \sigma_X^2 := \text{Var}(X_t) = \frac{\alpha \mu_\epsilon + \sigma_\epsilon^2}{1 - \alpha^2},
$$

respectively. A commonly used variability measure of a random variable is the Fisher index of dispersion, defined by $\text{Fl}_X = \text{Var}(X) / \mathbb{E}(X)$, that is a measure of aggregation or disaggregation, for more details see Johnson et al. [24, p. 163]. Thus, the Fisher index of dispersion of $\{X_t\}_{t \in \mathbb{Z}}$ in Equation (1) is given by

$$
\text{Fl}_X = \frac{\text{Fl}_\epsilon + \alpha}{1 + \alpha}, \quad (2)
$$

where $\text{Fl}_\epsilon$ is the Fisher index of dispersion of the innovations $\{\epsilon_t\}_{t \in \mathbb{Z}}$. Furthermore, the autocorrelation function (ACF) at lag $h$ is given by $\rho_X(h) = \alpha^h$, $h \geq 0$, that is, it is of AR(1)-type, but only positive autocorrelation is allowed.

Overdispersion problems arise more frequently, and several solutions were proposed to overcome the additional variability in the data. We found various applications in the literature, with most of them analyzing data from veterinary or medical studies. Also, a reason for overdispersion, reported in the literature, is the presence of a positive correlation between the monitored events [40]. Meanwhile, the underdispersion describes
deficient variation in count data and can be caused by model over-fitting or seen in data sets with small sample values [35]. Moreover, underdispersion is a phenomenon only observed on rare occasions. The situation of smaller variability inside the data than assumed through a theoretical distribution model are found in some textbooks on biology and parasitology (see [14, 19]). Finally, we can see underdispersed political science data in King [28].

Equation (2) shows that the dispersion behavior of the observations \( \{X_t\}_{t \in \mathbb{Z}} \) is controlled by that of the innovations \( \{\epsilon_t\}_{t \in \mathbb{Z}} \), that is, Equation (2) implies that we obtain an INAR(1) process with the distribution of the \( \{X_t\}_{t \in \mathbb{Z}} \) being equidispersion, underdispersion or overdispersion iff the distribution of the innovations is chosen to be equidispersion, underdispersion or overdispersion, respectively. Thus, a simple approach is to only change the innovations distribution in such a way that the marginal distribution of the process is underdispersed or overdispersed.

In this context, based on binomial thinning operator, Jazi et al. [22] introduced the INAR(1) process with geometric innovations. Jazi et al. [21] discussed an INAR(1) process with zero-inflated Poisson innovations. Weiß [41] studied a stationary INAR(1) process with Good and certain types of power law weighted Poisson innovations. However, the applications focus on underdispersion, but Good and power law weighted Poisson distributions can also deal with equidispersion and overdispersion, see Kemp [26] and Del Castillo and Pérez-Casany [15]. Schweer and Weiß [33] introduced a first-order non-negative integer-valued autoregressive process with compound Poisson innovations. Bourguignon and Vasconcellos [7] studied a new stationary INAR(1) process with power series innovations. Andersson and Karlis [3] used the signed binomial thinning operator to define a first-order process with Skellam-distributed innovations. Fernández-Fontelo et al. [18] introduced a generalization of the classical Poisson-based INAR models whose innovations follow a Hermite distribution. Kim and Lee [27] considers the INAR(1) process with Katz family innovations. This paper aims to give a contribution in this direction.

The main objective of this paper is to propose two new binomial thinning INAR(1) processes with double Poisson (DP) and GP innovations, denoted by INARDP(1) and INARGP(1), respectively, for modeling non-negative integer-valued time series with equidispersion, underdispersion or overdispersion. We are choosing these distributions (double Poisson and GP) for the \( \{\epsilon_t\}_{t \in \mathbb{Z}} \) as ideal option among other candidates with overdispersion or underdispersion relative to Poisson distribution, because, the other options have only one type of dispersion or the probability mass function of distribution and moments of innovations are very complicated. We also propose a test (using the test provided by Schweer and Weiß [34]) based on our models for checking if the count time series considered is overdispersed or underdispersed. Additionally, we will provide a comprehensive account of the mathematical properties of these two new processes which are very easy to obtain [16] and the parameter restrictions are liberal \( [\alpha \in (0, 1)] \). Furthermore, the proposed processes have, as a particular case, the Poisson INAR(1) [INARP(1)] process [1].

The article is organized as follows. In Section 2, we construct two new models to properly capture different types of dispersions, and some of its properties are outlined. Section 3 discusses some simulation results for the estimation methods. Two applications with the real data sets are presented in Section 4.
2. Extended Poisson INAR(1) processes

In this section, we propose two extensions of the Poisson INAR(1) process in Equation (1) to deal with equidispersion, underdispersion and overdispersion problems. For better presentation, this section is divided into two subsections. The INARDP(1) and INARGP(1) processes are presented in Sections 2.1 and 2.2, respectively.

2.1. DP INAR(1) model

Efron [17] proposed, based on the double exponential family, the DP distribution. This model is indexed by two parameters \( \mu > 0 \) and \( \phi > 0 \). The probability mass function (pmf) is given by

\[
p(y; \mu, \phi) \equiv \Pr(Y = y) = Z(\mu, \phi) \cdot \tilde{p}(y; \mu, \phi) = Z(\mu, \phi) \phi^{1/2} e^{-\phi \mu} \left( \frac{e \mu}{y!} \right)^{\phi y},
\]

\[
y = 0, 1, 2, \ldots,
\]

where \( Z(\mu, \phi)^{-1} = \sum_{y=0}^{\infty} \tilde{p}(y; \mu, \phi) = \phi^{1/2} e^{-\phi \mu} \sum_{y=0}^{\infty} \left[ \frac{e^{-\phi y}}{y!} \right] (e \mu)^{\phi y} \approx 1 + \frac{1 - \phi}{12 \mu \phi} \) (1 + \( \frac{1}{\mu \phi} \)) is the normalizing constant. Efron [17] obtained this closed form approximation to the infinite series \( \sum_{y=0}^{\infty} \tilde{p}(y; \mu, \phi) \) and showed that \( \tilde{p}(y; \mu, \phi) \) is reasonably good when \( \mu \) is large (\( \mu = 10 \)). However, they are highly unreliable when \( \mu \) is small. Zhu [42] used the DP distribution without the normalizing constant providing a good fit for the mean and variance. Zou et al. [43] suggested the sum of the first \( k \) terms as an approximation to the sum \( \sum_{y=0}^{\infty} \tilde{p}(y; \mu, \phi) \), and recommended the value \( k \) to be at least twice as large as the sample mean.

The mean and the variance of the pmf \( p(y; \mu, \phi) \) are given by (approximation)

\[
E(Y) \approx \mu \quad \text{and} \quad \text{Var}(Y) \approx \frac{\mu}{\phi}.
\]

Thus, the DP distribution allows for both overdispersion (\( \phi < 1 \)) and underdispersion (\( \phi > 1 \)). If \( \phi = 1 \), the DP distribution collapses to the Poisson distribution. It is possible to show that the pgf is given by

\[
\varphi_Y(s) = \frac{Z(\mu s, \phi)}{Z(\mu, \phi)}. \tag{4}
\]

Now, let \( \{\epsilon_t\}_{t \in \mathbb{Z}} \) be a sequence of discrete i.i.d. random variables following a DP distribution with pmf given in Equation (3). In short, we name this process as the INARDP(1) process. Thus, the transition probabilities of this process are given by

\[
\Pr(X_t = k | X_{t-1} = l) = Z(\mu, \phi) \phi^{1/2} e^{-\phi \mu} \sum_{i=0}^{\min(k,l)} \binom{l}{i} \alpha^i (1 - \alpha)^{l-i} 
\]

\[
\times \frac{e^{-(k-i)} (k-i)^{(k-i)}}{(k-i)!} \cdot \left( \frac{e \mu}{k-i} \right)^{\phi (k-i)}. \tag{4}
\]

The mean, variance and Fisher index of dispersion of \( \{X_t\}_{t \in \mathbb{Z}} \) are given by

\[
\mu_X = \frac{\mu}{(1 - \alpha)}, \quad \sigma_X^2 = \frac{\mu (1 + \alpha \phi)}{\phi (1 - \alpha^2)} \quad \text{and} \quad \text{Fl}_X = \frac{1 + \alpha \phi}{\phi + \alpha \phi}.
\]
Table 1. Dispersion index of INARDP(1) model for various values of $\alpha$ and $\phi$.

| $\phi$ | $\alpha \rightarrow$ | $\alpha = 0.3$ | $\alpha = 0.5$ | $\alpha = 0.7$ |
|--------|---------------------|----------------|----------------|----------------|
| 0.3    | 2.7949              | 2.5556         | 2.3725         |
| 0.5    | 1.7692              | 1.6667         | 1.5882         |
| 0.7    | 1.3297              | 1.2857         | 1.2521         |
| 1.3    | 0.8225              | 0.8462         | 0.8642         |
| 1.5    | 0.7436              | 0.7778         | 0.8039         |
| 1.7    | 0.6833              | 0.7255         | 0.7578         |

It follows that this process shows equidispersion for $\phi = 1$, while we have underdispersion for $\phi > 1$, and overdispersion for $\phi < 1$. Note that the parameter $\mu$ does not change the dispersion index of the process. Table 1 contains the dispersion index of the INARDP(1) model for various parameter values.

The conditional expectation and the conditional variance are given, respectively, by

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

and

$$\text{Var}(X_t|X_{t-1}) = \alpha(1-\alpha)X_{t-1} + \mu/\phi.$$

In practice, the true values of the model parameters of the process are not known but have to be estimated from a given realization $X_1, \ldots, X_T$ of the process. We consider three estimation methods, namely, conditional least squares (CLS), Yule–Walker (YW) and conditional maximum likelihood (CML).

### 2.1.1. CLS estimation

The CLS estimator $\hat{\eta} = (\hat{\alpha}_{CLS}, \hat{\mu}_{CLS}, \hat{\phi}_{CLS})^T$ of $\eta = (\alpha, \mu, \phi)^T$ is given by

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

where $S_T(\eta) = \sum_{t=2}^T [X_t - \alpha X_{t-1} - \mu]^2$. Then, the CLS estimators of $\alpha$ ($\hat{\alpha}_{CLS}$), and $\mu$ ($\hat{\mu}_{CLS}$), can be written in closed form as

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

and

$$\hat{\mu}_{CLS} = \frac{\sum_{t=2}^T X_t - \hat{\alpha}_{CLS} \sum_{t=2}^T X_{t-1}}{T-1}.$$  

**Proposition 2.1:** The estimators $\hat{\alpha}_{CLS}$ and $\hat{\mu}_{CLS}$ given in Equation (5) are strongly consistent and satisfy the asymptotic normality

$$\sqrt{T}[(\hat{\alpha}_{CLS}, \hat{\mu}_{CLS})^T - (\alpha, \mu)^T] \xrightarrow{d} N_2((0,0)^T, \Sigma),$$

as $T \to \infty$, where the asymptotic covariance matrix $\Sigma$ is given by

$$\Sigma = \begin{pmatrix}
\frac{\gamma \alpha (1-\alpha)^3 (1+\alpha)^2 \phi^2}{\mu^2 (1+\alpha \phi)^2} + 1 - \alpha^2 & \alpha (1-\alpha) - \mu (1+\alpha) & \frac{\gamma \alpha (1-\alpha)^2 \phi^2}{\mu (1+\alpha \phi)^2} \\
\frac{\alpha (1-\alpha) - \mu (1+\alpha)}{\mu (1+\alpha \phi)^2} & \frac{\gamma \alpha (1-\alpha) (1+\alpha)^2 \phi^2}{1+\alpha \phi} & \mu \cdot \frac{1+\alpha \phi}{1-\alpha} + \frac{\phi}{\alpha} \mu \\
\frac{-\gamma \alpha (1-\alpha)^2 \phi^2}{\mu (1+\alpha \phi)^2} & \frac{\gamma \alpha (1-\alpha) (1+\alpha)^2 \phi^2}{1+\alpha \phi} & \phi - \alpha \mu
\end{pmatrix},$$

with $\gamma = \mu_{X,3} - 3\mu_X \cdot \sigma^2 - \mu_X^3, \mu_{X,3} = E(X_t^3)$. 

The proof of the proposition above can be found in the appendix. For estimation of the parameter $\phi$, we will use the two-step CLS estimation methods proposed by Karlsen and Tjostheim [25] which consists in minimizing the function

$$Q_T(\eta) = \sum_{t=2}^{T} \left[ (X_t - E(X_t|X_{t-1})]^2 - \text{Var}(X_t|X_{t-1}) \right]^2$$

$$= \sum_{t=2}^{T} \left[ (X_t - \alpha X_{t-1} - \mu)^2 - \alpha(1 - \alpha)X_{t-1} + \mu/\phi) \right]^2.$$ 

Thus, solving the equation $\partial Q_T(\eta)/\partial \phi = 0$ and replacing $\alpha$ and $\mu$ with the appropriate CLS estimates, we obtain that the CLS estimate of the parameter $\phi$ is given by

$$\hat{\phi}_{\text{CLS}} = \frac{\sum_{t=2}^{T} X_t - \hat{\alpha}_{\text{CLS}} \sum_{t=2}^{T} X_{t-1}}{\sum_{t=2}^{T} [(X_t - \hat{\alpha}_{\text{CLS}}X_{t-1} - \hat{\mu}_{\text{CLS}})^2 - \hat{\alpha}_{\text{CLS}}(1 - \hat{\alpha}_{\text{CLS}})X_{t-1}]}.$$ 

### 2.1.2. YW estimation

Additionally, we propose the method of moments based on the sample quantities of $\rho_X(1)$, $E(X_t)$ and $\text{Var}(X_t)$ whose estimators of $\alpha$, $\mu$ and $\phi$ are given by

$$\hat{\alpha}_{\text{YW}} = \frac{\sum_{t=1}^{T-1} (X_t - \bar{X})(X_{t+1} - \bar{X})}{\sum_{t=1}^{T} (X_t - \bar{X})^2}, \quad \hat{\mu}_{\text{YW}} = (1 - \hat{\alpha}_{\text{YW}})\bar{X}$$

and

$$\hat{\phi}_{\text{YW}} = \frac{\bar{X}}{\hat{\gamma}(0)(1 + \hat{\alpha}_{\text{YW}}) - \bar{X}\hat{\alpha}_{\text{YW}}},$$

respectively, where $\bar{X} = (1/T) \sum_{t=1}^{T} X_t$.

### 2.2. GP INAR(1) model

A random variable $Y$ is said to have a GP distribution [13] if its pmf is given by

$$\Pr(Y = y) = \frac{\mu (\mu + y \phi)^{y-1} e^{-(\mu + y \phi)}}{y!}, \quad y = 0, 1, \ldots,$$

where $\Pr(Y = y) = 0$, for $y > m$ if $\phi < 0$ and $\mu > 0$, $\max(-1, -\mu/m) < \phi < 1$. The value $m$ is the largest positive integer for which $\mu + \phi m > 0$ when $\phi < 0$. According to Consul and Famoye [12], when $\phi < 0$, the GP distribution includes a truncation due to $\Pr(Y = y) = 0$ for all $y \geq m$ and the sum $\sum_{y=0}^{m} \Pr(Y = y)$ is usually a little less than unity. However, this truncation error is less than 0.5% when $m \geq 4$ implying that the truncation error does not make any difference in practical applications [42] (see Section 4). It is obvious that the Poisson distribution with parameter $\mu$ is a special case when $\phi = 0$. Joe and Zhu [23] proved that the GP distribution is a mixture of Poisson distribution. Then, for $0 < \phi < 1$, the pmf is given by
Table 2. Dispersion index of INARGP(1) model for various values of $\alpha$ and $\phi$.

| $\phi$ ↓ | $\alpha \rightarrow$ | $\alpha = 0.3$ | $\alpha = 0.5$ | $\alpha = 0.7$ |
|---------|---------------------|----------------|----------------|----------------|
| $\phi = -0.7$ | $0.4969$ | $0.5640$ | $0.6153$ |
| $\phi = -0.5$ | $0.5726$ | $0.6296$ | $0.6732$ |
| $\phi = -0.3$ | $0.6859$ | $0.7278$ | $0.7598$ |
| $\phi = 0.3$ | $1.8006$ | $1.6939$ | $1.6122$ |
| $\phi = 0.5$ | $3.3077$ | $3.2764$ | $3.1764$ |
| $\phi = 0.7$ | $8.7778$ | $7.7407$ | $6.9477$ |

the corresponding pgf (see [11]) is given by

$$\varphi_Y(s) = e^{\mu[u(s,\phi) - 1]},$$

where $u(s, \phi)$ is the smaller root of the equation $u = se^{\phi(u-1)}$. For $\phi > 0$, the GP distribution belongs to the class of discrete self-decomposable (DSD) distributions [38], but, for $\phi < 0$, the GP distribution loses its DSD property [24, p. 336].

The mean and variance are given by

$$E(Y) = \frac{\mu}{(1 - \phi)}$$

and

$$Var(Y) = \frac{\mu^2}{(1 - \phi)^3},$$

and the third non-central moment is

$$E(Y^3) = \frac{\mu^2}{(1 - \phi)^3}[\mu^2(1 - \phi)^2 + 3\mu(1 - \phi) - 2(1 - \phi) + 3].$$

The $FI_Y$ is $1/(1 - \phi)^2$, soon when $\phi \in (0, 1)$ the GP distribution possesses the property of overdispersion. For $\phi \in (-1, 0)$ the model possesses the property underdispersion.

Let $\{\epsilon_t\}_{t \in \mathbb{Z}}$ be a sequence of discrete i.i.d. random variables following a GP distribution with parameters $\mu$ and $\phi$ with pmf given in Equation (6). In short, we name this process as the INARGP(1) process.

The conditional expectation, the conditional variance and the transition probabilities of this process are given by

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

$$Var(X_t|X_{t-1}) = \alpha(1 - \alpha)X_{t-1} + \mu/(1 - \phi)^3,$$

and

$$Pr(X_t = k|X_{t-1} = l) = \sum_{i=0}^{\min(k,l)} \binom{l}{i} \alpha^i (1 - \alpha)^{l-i} \frac{\mu^{k-i}e^{-\mu + (k-i)\phi}}{(k-i)!},$$

respectively.
Now, we introduce some properties and we will estimate the unknown parameters \( \alpha \), \( \mu \) and \( \phi \) of the INARGP(1) process. The YW estimates are briefly discussed. Also, CLS estimators will be derived, and their asymptotic properties will be considered.

### 2.2.1. YW estimation

The YW estimators of \( \alpha \), \( \mu \) and \( \phi \) are based upon the sample ACF \( \hat{\rho}(k) \) with \( \rho_X(1) = \alpha \) and the first moment and the dispersion index of \( X_t \) given by \( \mathbb{E}(X_t) = \mu / \left( (1 - \alpha)(1 - \phi) \right) \), and \( \text{FI}_X = \frac{1 + \alpha(1 - \phi)^2}{(1 + \alpha)(1 - \phi)^2} \), respectively. Let \( X_1, X_2, \ldots, X_T \) be a random sample of size \( T \) from the INARPG(1) process. Then, the YW estimators of \( \alpha \), \( \mu \) and \( \phi \) are given by

\[
\hat{\alpha}_{YW} = \frac{\sum_{t=1}^{T-1} (X_t - \bar{X})(X_{t+1} - \bar{X})}{\sum_{t=1}^{T} (X_t - \bar{X})^2}, \quad \hat{\mu}_{YW} = (1 - \hat{\alpha}_{YW})(1 - \hat{\phi}_{YW})\bar{X},
\]

and

\[
\hat{\phi}_{YW} = \frac{\hat{\alpha}_{YW} \cdot \text{FI}_X - \sqrt{\hat{\alpha}_{YW} \cdot \text{FI}_X - \hat{\alpha}_{YW} + \text{FI}_X - \hat{\alpha}_{YW} + \text{FI}_X}}{\hat{\alpha}_{YW} \cdot \text{FI}_X - \hat{\alpha}_{YW} + \text{FI}_X}.
\]

### 2.2.2. CLS estimation

The CLS estimator \( \hat{\eta} = (\hat{\alpha}_{CLS}, \hat{\mu}_{CLS}, \hat{\phi}_{CLS})^T \) of \( \eta = (\alpha, \mu, \phi)^T \) is given by

\[
\hat{\eta} = \arg \min_{\eta} S_T(\eta),
\]

where \( S_T(\eta) = \sum_{t=2}^{T} [X_t - g(\eta, X_{t-1})]^2 \) and \( g(\eta, X_{t-1}) = \mathbb{E}(X_t \mid X_{t-1}) = \alpha X_{t-1} + \mu / (1 - \phi) \). However, note that \( \alpha X_{t-1} + \mu / (1 - \phi) \) depends on \( \mu \) and \( \phi \) only through \( \mu / (1 - \phi) \) implying that it is not possible to obtain the estimators of \( \mu \) and \( \phi \). Thus, we use the CLS method to find estimators for \( \alpha \) and \( \phi \) assuming that \( \mu \) is known. Then, in this case the CLS estimators of \( \alpha \) and \( \phi \) can be written in closed form as

\[
\hat{\alpha}_{CLS} = \frac{(T - 1) \sum_{t=2}^{T} X_t X_{t-1} - \sum_{t=2}^{T} X_t \sum_{t=2}^{T} X_{t-1}}{(T - 1) \sum_{t=2}^{T} X_t^2 - (\sum_{t=2}^{T} X_{t-1})^2},
\]

and

\[
\hat{\phi}_{CLS} = 1 - \frac{\mu (T - 1)}{\sum_{t=2}^{T} X_t - \hat{\alpha} \sum_{t=2}^{T} X_{t-1}}, \quad (8)
\]

where \( \mu \) will be replaced by some consistent estimator \( \hat{\mu} \) as the one in the previous subsection given by \( \hat{\mu}_{CLS} = \hat{\mu}_{YW} \).
Proposition 2.2: The estimators $\hat{\alpha}_{CLS}$ and $\hat{\phi}_{CLS}$ given in Equation (8) are strongly consistent for estimating $\alpha$ and $\phi$, respectively, and satisfy the asymptotic normality

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

where

$$V^{-1} = \begin{pmatrix}
\frac{\mu \epsilon \sigma_\epsilon^2}{\mu \epsilon \sigma_\epsilon^2 \mu X_2 - (1 - \alpha)^2 \mu_X^4} & \frac{(1 - \alpha)\mu_X^2}{\mu \epsilon \sigma_\epsilon^2 \mu X_2 - (1 - \alpha)^2 \mu_X^4} \\
\frac{\mu \epsilon \sigma_\epsilon^2 \mu X_2 - (1 - \alpha)^2 \mu_X^4}{\mu \epsilon \sigma_\epsilon^2 \mu X_2 - (1 - \alpha)^2 \mu_X^4} & \frac{\mu_X^2}{\mu \epsilon \sigma_\epsilon^2 \mu X_2 - (1 - \alpha)^2 \mu_X^4}
\end{pmatrix},$$

and

$$W = \begin{pmatrix}
\frac{\alpha (1 - \alpha) \mu X_3 + \sigma_\epsilon^2 \mu X_2}{\mu \epsilon \sigma_\epsilon^2 (\alpha \mu \epsilon + \sigma_\epsilon^2)} & \frac{\mu \epsilon \sigma_\epsilon^2 (\alpha \mu \epsilon + \sigma_\epsilon^2)}{\lambda [\alpha (1 - \alpha) \mu X_2 + \mu X \sigma_\epsilon^2] / (1 - \theta)^2} \\
\frac{\mu \epsilon \sigma_\epsilon^2 (\alpha \mu \epsilon + \sigma_\epsilon^2)}{\lambda [\alpha (1 - \alpha) \mu X_2 + \mu X \sigma_\epsilon^2] / (1 - \theta)^2} & \frac{\mu \epsilon \sigma_\epsilon^2 (\alpha \mu \epsilon + \sigma_\epsilon^2)}{\lambda [\alpha (1 - \alpha) \mu X_2 + \mu X \sigma_\epsilon^2] / (1 - \theta)^2}
\end{pmatrix},$$

with $\mu X_2 = \sigma_\epsilon^2 + \mu_X^2$ and $\mu X_3 = E(X^3_t)$.

2.2.3. Conditional maximum likelihood

Let $X_1, X_2, \ldots, X_T$, with $X_1$ fixed, be a random sample of size $T$ from a stationary INARDP(1) or INARGP(1) process with vector parameters $\eta$. The conditional log-likelihood function for the INARDP(1) or INARGP(1) process is given by

$$\ell(\eta) = \sum_{t=2}^{T} \log[\Pr(X_t = k | X_{t-1} = l)],$$

with $\Pr(X_t = k | X_{t-1} = l)$ as in Equations (4) or (7). CML estimates $\hat{\eta}_{CML}$ for $\eta$ are obtained by maximizing $\ell(\eta)$. In practical scenery there will be no closed form for the CML estimates and numerical methods need to be necessary. As starting values for the algorithm, we have used the estimates obtained by the YW or CLS methods. Since the 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 CML estimates.

Remark 2.1 (tsinteger R package): The theoretical results of this paper has been implemented into a piece of statistical software: the tsinteger package for R [31,32]. To install this package, the R code below must be used.

devtools::install_github("projecttsinteger/tsintegerpackage")

This package contains a collection of utilities for analyzing data from INAR(1) processes. Some of the functions are: epoinar(), est.inar(), epoinar.sim() and equi.test().

3. Experimental evaluation

This section contains results from a simulation study that illustrates the performances of the different methods of estimation for parameters of the models described in the previous
Table 3. Empirical bias and MSE (in parentheses) of estimators of $\alpha$, $\mu$, $\phi$.

| $T$  | $\hat{\alpha}_{\text{CLS}}$ | $\hat{\alpha}_{\text{YW}}$ | $\hat{\alpha}_{\text{CML}}$ | $\hat{\mu}_{\text{CLS}}$ | $\hat{\mu}_{\text{YW}}$ | $\hat{\mu}_{\text{CML}}$ | $\hat{\phi}_{\text{CLS}}$ | $\hat{\phi}_{\text{YW}}$ | $\hat{\phi}_{\text{CML}}$ |
|------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| 100  | -0.0188        | -0.0218        | -0.0168        | 0.0855         | 0.1060         | 0.0998         | 0.0302         | 0.0246         | 0.0245         |
|      | (0.0101)       | (0.0101)       | (0.0090)       | (0.5900)       | (0.5842)       | (0.5502)       | (0.0119)       | (0.0112)       | (0.0123)       |
| 200  | -0.00103       | -0.0118        | -0.0079        | 0.0295         | 0.0398         | 0.0493         | 0.0141         | 0.0114         | 0.0107         |
|      | (0.0048)       | (0.0048)       | (0.0041)       | (0.2766)       | (0.2754)       | (0.2489)       | (0.0052)       | (0.0050)       | (0.0053)       |
| 400  | -0.0051        | -0.0059        | -0.0043        | -0.0105        | -0.0053        | 0.0236         | 0.0669         | 0.0055         | 0.0056         |
|      | (0.0023)       | (0.0023)       | (0.0019)       | (0.1365)       | (0.1356)       | (0.1159)       | (0.0025)       | (0.0025)       | (0.0026)       |
| 800  | -0.0017        | -0.0020        | -0.0013        | -0.0324        | -0.0298        | 0.0063         | 0.0022         | 0.0015         | 0.0020         |
|      | (0.0012)       | (0.0012)       | (0.0010)       | (0.0704)       | (0.0701)       | (0.0585)       | (0.0012)       | (0.0012)       | (0.0012)       |

$\alpha = 0.3$, $\mu = 5.0$ and $\phi = 0.5$ (overdispersed case)

| $T$  | $\hat{\alpha}_{\text{CLS}}$ | $\hat{\alpha}_{\text{YW}}$ | $\hat{\alpha}_{\text{CML}}$ | $\hat{\mu}_{\text{CLS}}$ | $\hat{\mu}_{\text{YW}}$ | $\hat{\mu}_{\text{CML}}$ | $\hat{\phi}_{\text{CLS}}$ | $\hat{\phi}_{\text{YW}}$ | $\hat{\phi}_{\text{CML}}$ |
|------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| 100  | -0.0193        | -0.0223        | -0.0185        | 0.1399         | 0.1613         | 0.1289         | 0.1348         | 0.0926         | 0.1263         |
|      | (0.0097)       | (0.0096)       | (0.0094)       | (0.5002)       | (0.4974)       | (0.4865)       | (0.2972)       | (0.2649)       | (0.2800)       |
| 200  | -0.0110        | -0.0124        | -0.0105        | 0.0814         | 0.0916         | 0.0729         | 0.0671         | 0.0474         | 0.0622         |
|      | (0.0048)       | (0.0048)       | (0.0046)       | (0.2521)       | (0.2518)       | (0.2417)       | (0.1199)       | (0.1134)       | (0.1135)       |
| 400  | -0.0054        | -0.0061        | -0.0051        | 0.0421         | 0.0472         | 0.0344         | 0.0299         | 0.0204         | 0.0240         |
|      | (0.0024)       | (0.0024)       | (0.0023)       | (0.1265)       | (0.1264)       | (0.1217)       | (0.0549)       | (0.0535)       | (0.0508)       |
| 800  | -0.0021        | -0.0025        | -0.0021        | 0.0191         | 0.0217         | 0.0134         | 0.0175         | 0.0128         | 0.0125         |
|      | (0.0012)       | (0.0012)       | (0.0012)       | (0.0631)       | (0.0631)       | (0.0606)       | (0.0276)       | (0.0272)       | (0.0259)       |

$\alpha = 0.3$, $\mu = 5.0$ and $\phi = 2.0$ (underdispersed case)

sections. The simulation study was carried out to compare the estimates obtained from the YW, CLS and CML methods. These three methods of estimation was based on their empirical bias and mean square error (MSE).

In our simulation study, a random sample of size $T = 100, 200, 400$ and 800 was generated and values of $X_1$ were independently drawn from the DP (or GP) with corresponding values of $X_t$ given by

$$X_t = 0.3 \circ X_{t-1} + \epsilon_t,$$

where $\epsilon_t$ were set to be independently drawn from the following two studied distributions: (a) DP with parameters $\mu = 5.0$ and $\phi = 0.5$ and 2.0; and (b) GP with parameters $\mu = 1.0$ and $\phi = -0.5$ and 0.5. The simulation process was replicated 5,000 times. The empirical means and mean square error of the three methods of estimation were then computed. All simulations were accomplished by using R software. To simulate the innovations process, we simulate a DP and a GP one with R’s `rdpois` (rmutil package) and `rgenpois` (HMMpa package) functions, respectively.

Tables 3 and 4 show the empirical bias and mean square error of the estimators obtained from the YW, CLS and CML methods under INARDP(1) and INARGP(1) models, respectively. The results set out in Table 3, under INARDP(1) model, show that the CML estimators has the best performance on empirical bias and MSE compared with the YW and CLS estimators. For the estimators of $\alpha$ and different values of $\phi$, we notice that the MSE of both methods are very much similar. Note that for large sample size both methods given a smaller bias and MSE (very close to zero) for estimates of $\alpha$, $\mu$ and $\phi$. The bias of the estimators of $\alpha$ are negative and does not depend on the values of other parameters considered.

However, the results displayed in Table 4, under INARGP(1) model, reveal that for the overdispersed case the estimator $\hat{\alpha}_{\text{CML}}$ turns out to be better than the other methods as
Table 4. Empirical bias and MSE (in parentheses) of estimators of $\alpha$, $\mu$, $\phi$.

| $T$ | $\hat{\alpha}_{\text{CLS}}$ | $\hat{\alpha}_{\text{YW}}$ | $\hat{\alpha}_{\text{CML}}$ | $\mu_{\text{YW}}$ | $\mu_{\text{CML}}$ | $\phi_{\text{CLS}}$ | $\phi_{\text{YW}}$ | $\phi_{\text{CML}}$ |
|-----|-----------------|-----------------|-----------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| 100 | $-0.0263$       | $-0.0289$       | $0.1994$        | $-0.0461$           | $-0.3057$           | $0.1458$           | $0.1524$           | $-0.0785$           |
|     | (0.0116)        | (0.0116)        | (0.0977)        | (0.0413)            | (0.1580)            | (0.0503)           | (0.0518)           | (0.1509)            |
| 200 | $-0.0182$       | $-0.0197$       | $0.1943$        | $-0.0659$           | $-0.3141$           | $0.1520$           | $0.1552$           | $-0.0465$           |
|     | (0.0063)        | (0.0063)        | (0.0860)        | (0.0255)            | (0.1364)            | (0.0433)           | (0.0441)           | (0.1397)            |
| 400 | $-0.0047$       | $-0.0054$       | $0.1034$        | $-0.1054$           | $-0.2483$           | $0.2057$           | $0.2073$           | $0.1300$            |
|     | (0.0032)        | (0.0032)        | (0.0429)        | (0.0211)            | (0.0840)            | (0.0508)           | (0.0514)           | (0.1172)            |
| 800 | $-0.0037$       | $-0.0041$       | $0.0748$        | $-0.1102$           | $-0.2296$           | $0.2134$           | $0.2142$           | $0.1092$            |
|     | (0.0017)        | (0.0017)        | (0.0285)        | (0.0175)            | (0.0675)            | (0.0511)           | (0.0514)           | (0.1038)            |

$\alpha = 0.3$, $\mu = 1.0$ and $\phi = -0.5$ (underdispersed case)

$\alpha = 0.3$, $\mu = 1.0$ and $\phi = 0.5$ (overdispersed case)

it gives in all cases lower bias (MSE) than the YW and CLS methods. However, for the underdispersed case we have a reverse scenery. For example, for the estimator of $\alpha$, the CLS method has the best performance on empirical bias. However, when estimating $\phi$, the CML method shows the best performance on empirical bias. Finally, we notice that the CLS and YW methods present similar MSE behaviors.

4. Real data examples

To illustrate the applications of the proposed models, we consider in this section two real data sets with overdispersion and underdispersion. We compared the proposed processes with the INARP($1$) process (special case). In order to estimate the parameters of these processes, we adopt the CML method and all the computations were done using the tsinteger package.

Remark 4.1 (Detecting overdispersion or underdispersion): For testing the null hypothesis $H_0: X_1, \ldots, X_T$ stem from an equidispersed Poisson INAR($1$) process ($\text{FI}_X = 1$) against the alternative of an overdispersed (or underdispersed) marginal distribution, we suggest to use the the following test provided by Schweer and Weiß [33]. Let $z_{1-\beta}$ be the quantile of the $(1-\beta)$-quantile of the N(0,1)-distribution, that is, $\Phi(z_{1-\beta}) = 1 - \beta$, for $\beta \in (0,1)$, where $\Phi(\cdot)$ is the distribution function of a N(0,1)-distribution. We reject the null hypothesis $H_0: \phi = 1$ or $\phi = 0$ (equidispersion) in favor of alternative hypothesis $H_1: \phi < 1$ or $\phi > 0$ (overdispersion) if

$$\hat{\text{FI}}_X > z_{1-\beta} \sqrt{\frac{2(1 + \alpha^2)}{T(1 - \alpha^2)}},$$
where $\hat{F}_I X := \sum_{t=1}^{T} (X_t - \bar{X})^2 / \sum_{t=1}^{T} X_t$ with $\bar{X} := (1/T) \sum_{t=1}^{T} X_t$. Furthermore, if the alternative hypothesis of interest is $\mathcal{H}_1 : \phi > 1$ or $\phi < 0$ (underdispersion), we reject $\mathcal{H}_0$ in favor of an alternative hypothesis if

$$\hat{F}_I X < z_\beta \sqrt{\frac{2(1 + \alpha^2)}{T(1 - \alpha^2)}}.$$ 

### 4.1. Overdispersed data: weekly number of syphilis cases

As a first example, we consider the data set consisting of the weekly number of syphilis cases in the United States from 2007–2010 in Mid-Atlantic states given in \texttt{tsinteger} package available for download at \texttt{data(syphilis)}. The data consist of 209 observations, and they were already analyzed by Borges et al. [5].

The sample mean is 24.63, the sample variance is 105.68, and the first-order autocorrelation is 0.2322. The empirical Fisher index of dispersion is 4.29. The sample variance is much larger than the sample mean, hence, the data seems to be overdispersed. The equidispersion test [33] rejected the null hypothesis of equidispersion, the $p$-value for the test being $< 0.01$. Consequently, a Poisson marginal distribution seems to not be appropriate.

The series together with its sample autocorrelation and partial ACFs are displayed in Figure 1. Analyzing Figure 1 we conclude that a first-order autoregressive model may be appropriate for the given data series, given the pattern of the sample partial ACF and the clear cut-off.

Table 5 gives the CML estimates (with corresponding standard errors in parentheses), Akaikie information criterion (AIC) and Bayesian information criterion (BIC) for the fitted models. Since the values of the AIC and BIC are smaller for the INARGP(1) and INARDP(1) models compared to those values of the INARP(1) model. The likelihood ratio (LR) statistic to test the hypothesis $\mathcal{H}_0 : \text{INARP}(1)$ against the alternative hypothesis $\mathcal{H}_1 : \text{INARGP}(1)$, that is, $\mathcal{H}_0 : \phi = 0$ against $\mathcal{H}_1 : \phi \neq 0$, is 403.39 ($p$-value $< 0.01$). Furthermore, the LR statistic to test the hypothesis $\mathcal{H}_0 : \text{INARP}(1)$ against the alternative hypothesis $\mathcal{H}_1 : \text{INARDP}(1)$, that is, $\mathcal{H}_0 : \phi = 1$ against $\mathcal{H}_1 : \phi \neq 1$, is 453.04 ($p$-value $< 0.01$). Thus, we reject the null hypothesis in favor of the ONARGP(1) and NARDP(1) models using any usual significance level. Therefore, the INARGP(1) and INARDP(1) models are significantly better than the INARP(1) model based on the LR statistic.

We note that $\hat{\phi}_{\text{CML}} > 0$ and $\hat{\phi}_{\text{CML}} < 1$ for the INARGP(1) and INARDP(1) models, respectively, which implies that the dispersion index of these models is greater than 1 in accordance with equidispersion test [33]. From the figures of Table 5 and according to LR tests, the INARDP(1) model fits the current data better than other models, that is, these values indicate that the null hypothesis is strongly rejected for the INARP(1) model. These results illustrate the potentiality of the INARGP(1) and INARDP(1) models and the importance of the additional parameter [in INARP(1) model]. Also, the residuals are not correlated.

### 4.2. Underdispersed data: family violence counts

As the second example, we consider the series of monthly counts of family violences in the 11th police car beat in Pittsburgh, during one month. It consists of 143 observations,
Figure 1. Plots of the time series, autocorrelation and partial ACFs for the number of syphilis cases.

Table 5. Estimates of the parameters, MSE (in parentheses), AIC, BIC and estimated quantities for the number of syphilis cases.

| Model   | Parameter | CML estimate    | AIC     | BIC     | $\mu_X$ | $\sigma^2_X$ | $F_X$ |
|---------|-----------|-----------------|---------|---------|---------|-------------|-------|
| INARGP(1) | $\alpha$  | 0.0798 (0.0497) | 1615.15 | 1625.18 | 24.72   | 137.04      | 5.54  |
|          | $\mu$     | 9.3614 (0.8164) |         |         |         |             |       |
|          | $\phi$    | 0.5885 (0.0255) |         |         |         |             |       |
| INARDP(1) | $\alpha$  | 0.1154 (0.0404) | 1565.50 | 1575.53 | 24.84   | 113.89      | 4.58  |
|          | $\mu$     | 21.976 (1.2204) |         |         |         |             |       |
|          | $\phi$    | 0.2001 (0.0195) |         |         |         |             |       |
| INARP(1)  | $\alpha$  | 0.1480 (0.0261) | 2016.54 | 2023.22 | 24.72   | 24.72       | 1     |
|          | $\mu$     | 21.063 (0.7087) |         |         |         |             |       |
| Empirical |           | 24.63           | 105.68  | 4.29    |         |             |       |

starting in January 1990 and ending in November 2001. The data set is obtained from the tsinteger package by data(violences).

The sample mean, variance and Fisher index of dispersion are 0.3846, 0.3369 and 0.8761, respectively, which indicates that the data are underdispersed. The first-order autocorrelation is 0.166. The series together with its sample autocorrelation and partial ACFs is displayed in Figure 2.
Figure 2. Plots of the time series, autocorrelation and partial ACFs for the family violence counts.

Table 6. Estimates of the parameters (MSE in parentheses), AIC, BIC and estimated quantities for the family violence counts.

| Model      | Parameter | CML estimate | AIC     | BIC     | $\mu_X$ | $\sigma^2_X$ | FL_X |
|------------|-----------|--------------|---------|---------|---------|--------------|------|
| INARGP(1)  | $\alpha$ | 0.1613 (0.0833) | 223.86  | 232.75  | 0.3887  | 0.3236       | 0.8325
|            | $\mu$    | 0.3632 (0.0627) |         |         |         |              |      |
|            | $\phi$   | -0.1142 (0.0527) |         |         |         |              |      |
| INARDP(1)  | $\alpha$ | 0.1924 (0.0893) | 223.64  | 232.53  | 0.3890  | 0.3204       | 0.8236
|            | $\mu$    | 0.3141 (0.0498) |         |         |         |              |      |
|            | $\phi$   | 1.2664 (0.1576)  |         |         |         |              |      |
| INARP(1)   | $\alpha$ | 0.1562 (0.0931)  | 224.98  | 230.91  | 0.3886  | 0.3886       | 1    |
|            | $\mu$    | 0.3279 (0.0566)  |         |         |         |              |      |
| Empirical  |           |              | 0.3846  | 0.3369  | 0.8761  |              |      |

Analyzing Figure 2, 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. So, we apply the INARGP(1), INARDP(1) and INARP(1) models to the data. Parameter estimates and their standard errors are summarized in Table 6. The AIC and BIC values are also provided.

Analyzing Table 6, note that $\hat{\phi}_{CML} < 0$ and $\hat{\phi}_{CML} > 1$ for the INARGP(1) and INARDP(1) models, respectively, which implies that the dispersion index of these models is less than 1. Based on AIC, we find that the INARGP(1) and INARDP(1) models are the best ones. Based
on BIC, we find that the INARP(1) is the best one. Within these three fitted models, the mean, variance and dispersion index are summarized in Table 6. All three models exhibit good fits of mean, but only the INARGP(1) and INARDP(1) models give a reasonable fit of variance and present underdispersed features. As a result, the INARP(1) model is not adequate for the data, for example, its variance 0.3886 is much larger than the empirical variance 0.3369, and also its dispersion index (1 vs. 0.8761). Based on this fact, we conclude that the INARGP(1) and INARDP(1) models capture more information of these data.

We test the null hypothesis $H_0 : \text{INARP}(1)$ against the alternative hypothesis $H_1 : \text{INARGP}(1)$, that is, $H_0 : \phi = 0$ against $H_1 : \phi \neq 0$ (with a significance level at 10%). The LR statistic to test the hypothesis is 3.123 ($p$-value is 0.0772). Furthermore, the LR statistic to test the hypothesis $H_0 : \text{INARP}(1)$ against the alternative hypothesis $H_1 : \text{INARDP}(1)$, that is, $H_0 : \phi = 1$ against $H_1 : \phi \neq 1$, is 3.342 ($p$-value is 0.0675). Thus, we reject the null hypothesis in favor of the INARGP(1) and INARDP(1) models. Therefore, the INARGP(1) and INARDP(1) models are significantly better than the INARP(1) model based on the LR statistic.

To conclude this section, let us briefly review the main characteristics of the proposed models. The question of which distribution to use for the innovation sequence may be somewhat subjective, and it may also depend on the specific situation we are dealing with. In the applications above, we note that the INARDP(1) model presented a good fitness by considering AIC and BIC. Moreover, for modeling count time series with underdispersion, we suggested to use this model because the GP distribution, when $\phi < 0$ (underdispersion), can be very problematic (see Section 2.2). Some discussions how to choose a specific model for the innovation sequence can be found in Bourguignon and Vasconcellos [7].

**Acknowledgments**

The authors are also grateful to two referees and an associate editor for helpful comments and suggestions.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

We gratefully acknowledge grants from CAPES and CNPq (Brazil). Research developed with the computational resources of the Center for Mathematical Sciences Applied to Industry (CeMEAI) financed by FAPESP.

**ORCID**

Manoel Santos-Neto http://orcid.org/0000-0002-9007-0680

**References**

[1] M.A. Al-Osh and A.A. Alzaid, *First-order integer-valued autoregressive (INAR(1)) process*, J. Time Ser. Anal. 8 (1987), pp. 261–275.

[2] M. Al-Osh and A.A. Alzaid, *Integer-valued moving average (INMA) process*, Statist. Papers 29 (1988), pp. 281–300.
[3] J. Andersson and D. Karlis, *A parametric time series model with covariates for integers in* $\mathbb{Z}$, Stat. Model. 14 (2014), pp. 135–156.

[4] W. Barreto-Souza and M. Bourguignon, *A skew INAR(1) process on* $\mathbb{Z}$, AStA Adv. Stat. Anal. 99 (2015), pp. 189–208.

[5] P. Borges, M. Bourguignon, and F.F. Molinares, *A generalised NGINAR(1) process with inflated-parameter geometric counting series*, Aust. N. Z. J. Stat. 59 (2017), pp. 137–150.

[6] M. Bourguignon, *Poisson-geometric INAR(1) process for modeling count time series with overdispersion*, Statist. Neerlandica 70 (2016), pp. 176–192.

[7] M. Bourguignon and K.L.P. Vasconcellos, *First order non-negative integer valued autoregressive processes with power series innovations*, Braz. J. Probab. Stat. 29 (2015), pp. 71–93.

[8] M. Bourguignon and K.L.P. Vasconcellos, *A new skew integer valued time series process*, Stat. Methodol. 31 (2016), pp. 8–19.

[9] M. Bourguignon and C.H. Weiβ, *An INAR(1) process for modeling count time series with equidispersion, underdispersion and overdispersion*, Test 26 (2017), pp. 847–868.

[10] K. Brännäs, J. Hellström, and J. Nordström, *A new approach to modelling and forecasting monthly guest nights in hotels*, Int. J. Forecast. 18 (2002), pp. 19–30.

[11] P.C. Consul, *Generalized Poisson Distribution: Properties and Applications*, Marcel Dekker, New York, 1989.

[12] P.C. Consul and E. Famoye, *Lagrangian Probability Distributions*, Birkhäuser, Basel, 2006.

[13] P.C. Consul and G.C. Jain, *A generalization of the Poisson distribution*, Technometrics 15 (1973), pp. 791–799.

[14] F.E.G. Cox, *Modern Parasitology: A Textbook of Parasitology*, Blackwell Science, Oxford, 1993.

[15] J. Del Castillo and M. Pérez-Casany, *Weighted Poisson distributions for overdispersion and underdispersion situations*, Ann. Inst. Stat. Math. 50 (1998), pp. 567–585.

[16] M. Eduarda Da Silva, and V.L. Oliveira, *Difference equations for the higher-order moments and cumulants of the INAR(1) model*, J. Time Ser. Anal. 25 (2004), pp. 317–333.

[17] B. Efron, *Double exponential families and their use in generalized linear regression*, J. Amer. Statist. Assoc. 81 (1986), pp. 709–721.

[18] A. Fernández-Fontelo, S. Fontdecaba, A. Alba, and P. Puig, *Integer-valued AR processes with hermite innovations and time-varying parameters: An application to bovine fallen stock surveillance at a local scale*, Stat. Model. 17 (2017), pp. 172–195.

[19] I.C.W. Hardy, *Sex Ratios. Concepts and Research Methods*, Cambridge University Press, Cambridge, 2002.

[20] J. Hellström, *Unit root testing in integer-valued AR(1) models*, Econom. Lett. 70 (2001), pp. 9–14.

[21] M.A. Jazi, G. Jones, and C.D. Lai, *First-order integer valued AR processes with zero inflated poisson innovations*, J. Time Ser. Anal. 33 (2012), pp. 954–963.

[22] M.A. Jazi, G. Jones, and C.D. Lai, *Integer valued AR(1) with geometric innovations*, J. Iranian Stat. Soc. 11 (2012), pp. 173–190.

[23] H. Joe and R. Zhu, *Generalized Poisson distribution: The property of mixture of Poisson and comparison with negative binomial distribution*, Biom. J. 47 (2005), pp. 219–229.

[24] N.L. Johnson, A.W. Kemp, and S. Kotz, *Univariate Discrete Distribution*, Wiley, New York, NY, 2005.

[25] H. KarlSEN and D. Tjostheim, *Consistent estimates for the NEAR(2) and NLAR time series models*, J. R. Stat. Soc. B Methodol. 50 (1988), pp. 313–320.

[26] A.W. Kemp, *Polylogarithmic Distributions*, John Wiley & Sons, Ltd, New York, 2014.

[27] H. Kim and S. Lee, *On first-order integer-valued autoregressive process with Katz family innovations*, J. Stat. Comput. Simul. 87 (2017), pp. 546–562.

[28] G. King, *Variance specification in event count models: From restrictive assumptions to a generalized estimator*, Am. J. Pol. Sci. 33 (1989), pp. 762–784.

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

[30] E. McKenzie, *Some simple models for discrete variate time series*, J. Am. Water Resour. Assoc. 21 (1985), pp. 645–650.
Appendix

**Proof of Proposition 2.1:** The proof of Proposition 2.1 is omitted here since it is a straightforward consequence of an application obtained by Klimko and Nelson [29, p. 638].

**Proof of Proposition 2.2:** Let $X_1, \ldots, X_T$ be a sample of an INARGP(1) process. It can be verified that the regularity conditions given in Theorem 3.2 of Klimko and Nelson [29, p. 634], are satisfied by INARGP(1) process.

Consider the following quantities $E_{t | t-1} \equiv E(X_t | X_{t-1}) = \alpha X_{t-1} + \mu/(1 - \phi)$ and $d_{t | t-1} = \text{Var}(X_t | X_{t-1}) = \alpha(1 - \alpha)X_{t-1} + \mu/(1 - \phi)^2$, and calculate

$$
\frac{\partial E_{t | t-1}}{\partial \alpha} = X_{t-1}, \quad \frac{\partial E_{t | t-1}}{\partial \phi} = \frac{\mu}{(1 - \phi)^2}, \quad \frac{\partial^2 E_{t | t-1}}{\partial \alpha^2} = 0, \quad \frac{\partial^2 E_{t | t-1}}{\partial \phi^2} = \frac{2\mu}{(1 - \phi)^3}, \\
\frac{\partial^2 E_{t | t-1}}{\partial \phi \partial \alpha} = 0.
$$

Define the $2 \times 2$ matrix $V$ according to Equation (3.2) in Klimko and Nelson [29] as

$$
V = E \left( \begin{bmatrix} \frac{\partial E_{t | t-1}}{\partial \alpha} & \frac{\partial E_{t | t-1}}{\partial \phi} \\ \frac{\partial^2 E_{t | t-1}}{\partial \phi^2} & \frac{\partial^2 E_{t | t-1}}{\partial \phi \partial \alpha} \end{bmatrix} \frac{\partial E_{t | t-1}}{\partial \alpha} \frac{\partial E_{t | t-1}}{\partial \phi} \right) = \begin{bmatrix} E(X_{t-1}^2) & \frac{\mu}{(1 - \phi)^2}E(X_{t-1}) \\ \frac{\mu}{(1 - \phi)^2} \mu_X & \mu^2 \\ (1 - \alpha)\mu_x & \mu \sigma_x^2 \\ (1 - \alpha)\mu_x & \mu \sigma_x^2 \\ \mu \sigma_x^2 & \mu \sigma_x^2 \end{bmatrix}.
$$
and the $2 \times 2$ matrix $W$ according to Equation (3.5) in Klimko and Nelson [29] as

$$W = E \left( \begin{bmatrix} \frac{\partial E_t|_{t-1}}{\partial \alpha} \\ \frac{\partial E_t|_{t-1}}{\partial \phi} \end{bmatrix} d_{t|_{t-1}} \begin{bmatrix} \frac{\partial E_t|_{t-1}}{\partial \alpha} & \frac{\partial E_t|_{t-1}}{\partial \phi} \end{bmatrix} \right)$$

$$= \left( \begin{array}{cc} \alpha (1 - \alpha) \mu_{X, 3} + \sigma_e^2 \mu_{X, 2} & \mu_e \sigma_e^2 (\alpha \mu_e + \sigma_e^2) \\ \mu_e \sigma_e^2 (\alpha \mu_e + \sigma_e^2) & \lambda \left[ \alpha (1 - \alpha) \mu_{X, 2} + \mu_X \sigma_e^2 \right] / (1 - \theta)^2 \end{array} \right).$$

Hence, the estimators $\hat{\alpha}_{CLS}$ and $\hat{\phi}_{CLS}$ of $\alpha$ and $\phi$ have the following asymptotic distribution:

$$\sqrt{T} [ (\hat{\alpha}_{CLS}, \hat{\phi}_{CLS})^T - (\alpha, \phi)^T ] \overset{d}{\longrightarrow} N_2((0, 0)^T, V^{-1}WV^{-1}).$$