CrossMark

# Mixed Poisson INAR(1) processes

**Wagner Barreto-Souza**[1] (ORCID)

**Abstract** Overdispersion is a phenomenon commonly observed in count time series. Since Poisson distribution is equidispersed, the INteger-valued AutoRegressive (INAR) process with Poisson marginals is not adequate for modelling overdispersed counts. To overcome this problem, in this paper we propose a general class of first-order INAR processes for modelling overdispersed count time series. The proposed INAR(1) processes have marginals belonging to a class of mixed Poisson distributions, which are overdispersed. With this, our class of overdispersed count models have the known negative binomial INAR(1) process as particular case and open the possibility of introducing new INAR(1) processes, such as the Poisson-inverse Gaussian INAR(1) model, which is discussed here with some details. We establish a condition to our class of overdispersed INAR processes is well-defined and study some statistical properties. We propose estimators for the parameters and establish their consistency and asymptotic normality. A small Monte Carlo simulation to evaluate the finite-sample performance of the proposed estimators is presented and one application to a real data set illustrates the usefulness of our proposed overdispersed count processes.

**Keywords** Autocorrelation · Count process · Markov chain · Poisson-inverse Gaussian distribution · Overdispersion

## 1 Introduction

Count time series arise naturally in many practical situations and due to this an increased interest in the modelling of such data have been observed. Some areas

✉ Wagner Barreto-Souza
   wagnerbs85@gmail.com; wagnerbs@est.ufmg.br

1  Departamento de Estatística, Universidade Federal de Minas Gerais, Pampulha, Belo Horizonte, MG 31270-901, Brazil

where we observe count time series data are economy, medicine, epidemiology, natural and social sciences, just to name a few. For example, in this paper we will deal with monthly counting of sex offences registered in the 21st police car beat in Pittsburgh from January 1990 and ending in December 2001.

The most common way for dealing with this type of data is to consider a ARMA-type model or an INteger-valued AutoRegressive (INAR) process. These models can be seen as parameter-driven and observation-driven models, respectively. We here focus on the observation-driven approach based on the INAR structure.

The Poisson INAR(1) (first-order INAR) model is the most popular INAR process for modelling count time series data. This process is based on the binomial thinning operator; this operator was introduced by Steutel and van Harn (1979). Pionering works on the Poisson INAR (in short PINAR) process are due to McKenzie (1985), Al-Osh and Alzaid (1987) and McKenzie (1988). After these works, some papers have dealt with inferential and forecasting aspects of the PINAR(1) process. For instance, see Freeland and McCabe (2004a), Freeland and McCabe (2004b) and Freeland and McCabe (2005).

An important variant of the PINAR(1) process was considered by Ristić et al. (2009). They introduced an INAR(1) process with geometric marginals by using a negative binomial thinning operator (Aly and Bouzar (1994)). The literature about INAR processes is very rich and the interest in this topic has increased in the last years. Some recent contributions on INAR processes are given by Nastić and Ristić (2012), Jazi et al. (2012), Ristić et al. (2013), Weiß (2013), Weiß and Kim (2013), Meintanis and Karlis (2014), Schweer and Weiß (2014), Andersson and Karlis (2014), Barreto-Souza and Bourguignon (2015), Barreto-Souza (2015), Weiß (2015), Bisaglia and Canale (2016), Nastić et al. (2016a), Nastić et al. (2016b), Nastić et al. (2016c), Yang et al. (2016) and Weiß et al. (2016). For a good review on INAR processes we recommend the papers by McKenzie (2003), Weiß (2008b) and Scotto et al. (2015).

Overdispersion is a phenomenon commonly observed in count time series. Since Poisson distribution is equidispersed, the PINAR process is not adequate for modelling overdispersed counts. Overdispersion can be caused by excess of zeros. In this direction, Jazi et al. (2012) and Barreto-Souza (2015) proposed INAR processes with zero-inflated Poisson innovations and zero-modified geometric marginals, respectively; both models are also able for dealing with underdispersion. Another important work on this issue is due to Schweer and Weiß (2014). They proposed a class of compound Poisson INAR(1) models and in particular obtained an overdispersion testing.

There are many different causes of overdispersion rather than excess of zeros. For example, the non-observation of significant covariates. A common way for treating overdispersion in count data is to use the mixed Poisson distributions, which is obtained by introducing a latent random effect on the mean of a Poisson distribution. For an account on mixed Poisson distributions, we recommend the paper by Karlis and Xekalaki (2005). With this in mind, in this paper we propose a class of overdispersed INAR(1) processes with marginals belonging to a general class of mixed Poisson distributions. With this approach, our class of INAR models includes the known negative binomial INAR(1) process as particular case and open the possibility of introducing new INAR(1) processes, such as the Poisson-inverse Gaussian INAR(1) model.

The paper is organized as follows. In Sect. 2 we define our class of mixed Poisson INAR(1) processes and derive some basic statistical properties. We also establish a condition to our class is well defined. In Sect. 3 we present the Poisson-inverse Gaussian INAR(1) model with some details. We propose estimators for the parameters and establish their consistency and asymptotic normality in Sect. 4. Further, we present a small Monte Carlo simulation to evaluate the finite-sample performance of the proposed estimators. An empirical illustration of our class of overdispersed INAR(1) processes is given in Sect. 5. Concluding remarks and future research are discussed in Sect. 6.

## 2 Definition and basic properties

In order to present our class of mixed Poisson INAR(1) processes and make this paper self-contained, we begin this section speaking about mixed Poisson distributions and self-decomposable random variables.

**Definition 1** A random variable $Y$ follows a mixed Poisson distribution if satisfies the stochastic representation $Y|Z = z \sim \text{Poisson}(\mu z)$, for $\mu > 0$, where $Z$ is some non-negative random variable.

The probability function of a mixed Poisson random variable $Y$ takes the form

$$P(Y = y) = \int_0^\infty \frac{e^{-\mu z}(\mu z)^y}{y!} dG_\phi(z),$$

for $y \in \mathbb{N}_0 \equiv \{0, 1, \ldots\}$, where $G_\phi(\cdot)$ is the distribution function of $Z$ and $\phi$ denotes the parameter vector associated to the distribution of $Z$; we denote $Y \sim \text{MP}(\mu, \phi)$. Let $\Psi_Z(t) = E(e^{tZ})$ be the moment generating function (mgf) of $Z$, for $t$ belonging to some interval containing the value zero.

We have that the mgf of $Y$, say $\Psi_Y(\cdot)$, can be expressed by

$$\Psi_Y(t) = \Psi_Z(\mu(e^t - 1)),$$

for $t$ belonging some interval containing the value zero. Assume that $Z$ has the finite second moment. Then, the mean and variance of $Y$ are given by $E(Y) = \mu E(Z)$ and $\text{Var}(Y) = \mu\{E(Z) + \mu \text{Var}(Z)\}$. As we can see, $Y$ is overdispersed for any choice of distribution for $Z$ (with finite second moment).

**Definition 2** A random variable $X$ is said to be self-decomposable if, for all $\alpha \in [0, 1)$, there is a random variable $X_\alpha$ independent of $X$ such that the following stochastic representation holds: $X \stackrel{d}{=} \alpha X + X_\alpha$, where "$\stackrel{d}{=}$" means equality in distribution.

*Remark 1* A non-degenerate random variable $X$ that is self-decomposable according Definition 2 is necessarily absolutely continuous.

**Definition 3** (Steutel and van Harn (1979)) A random variable $X$ assuming values on $\mathbb{N}_0$ is said to be discrete self-decomposable if, for all $\alpha \in [0, 1)$, there is a random

variable $X_\alpha$ independent of $X$ such that the following stochastic representation holds: $X \stackrel{d}{=} \alpha \circ X + X_\alpha$, where $\circ$ is the thinning operator defined by $\alpha \circ X = \sum_{i=1}^{X} B_i$, where $\{B_i\}_{i=1}^{\infty}$ is a sequence of iid Bernoulli random variables with $P(B_1 = 1) = 1 - P(B_1 = 0) = \alpha$ and $\alpha \circ 0 \equiv 0$.

In what follows, we present some results that will be important to define our class of overdispersed INAR processes.

**Proposition 1** *Let $Y \sim MP(\mu, \phi)$. Then, $\alpha \circ Y \sim MP(\alpha\mu, \phi)$, for $\alpha \in [0, 1)$.*

*Proof* Using the fact that $\alpha \circ Y | Z = z$ follows a Poisson distribution with parameter $\alpha\mu z$, we obtain that

$$\Psi_{\alpha \circ Y}(t) \equiv E\left(\exp\{t(\alpha \circ Y)\}\right) = E[E(\exp\{t(\alpha \circ Y)\}|Z)] = E\left(\exp\left\{\mu\alpha Z\left(e^t - 1\right)\right\}\right)$$
$$= \Psi_Z\left(\mu\alpha\left(e^t - 1\right)\right),$$

that is the mgf of a MP($\alpha\mu, \phi$) random variable.                                                                              □

**Proposition 2** *Let $Z$ be a non-negative self-decomposable random variable according Def. 2 with mgf $\Psi_Z(\cdot)$ and $\Psi(\cdot)$ be a function defined as*

$$\Psi(t) = \frac{\Psi_Z\left(\mu\left(e^t - 1\right)\right)}{\Psi_Z\left(\alpha\mu\left(e^t - 1\right)\right)}, \tag{1}$$

*for t belonging some interval containing the value 0. Then $\Psi(t)$ is a proper mgf for all $\alpha \in [0, 1)$. Moreover, this is a mgf of a mixed Poisson distribution.*

*Proof* Since $Z$ is self-decomposable, there is a random variable $Z_\alpha$ independent of $Z$ such that $Z \stackrel{d}{=} \alpha Z + Z_\alpha$. Hence, we have that the function $\Psi_Z(t)/\Psi_Z(\alpha t)$ is a proper mgf for all $\alpha \in [0, 1)$; this function is the mgf of $Z_\alpha$. Let $W$ be a random variable with this mgf. We now show that the function (1) is a mgf of mixed Poisson random variable, denoted by $\epsilon$, which satisfies the following stochastic representation: $\epsilon | W = w \sim \text{Poisson}(\mu w)$, with $W$ as defined above. It follows that

$$E(\exp\{\epsilon t\}) = E\left[E\left(\exp\{\epsilon t\}|W\right)\right] = E\left[\exp\left\{\mu\left(e^t - 1\right)W\right\}\right]$$
$$= \frac{\Psi_Z\left(\mu\left(e^t - 1\right)\right)}{\Psi_Z\left(\alpha\mu\left(e^t - 1\right)\right)} = \Psi(t),$$

so proving the desired result.                                                                                                       □

With the results above, we now are ready to define our class of mixed Poisson INAR(1) processes as follows.

**Definition 4** We said that a sequence $\{X_n\}_{n=0}^{\infty}$ is a Mixed Poisson INAR(1) process, named in short by MPINAR(1) process, if $X_n \sim MP(\mu, \phi)$ for all $n$ and if it admits the following stochastic representation:

$$X_n = \alpha \circ X_{n-1} + \epsilon_n, \tag{2}$$

for $n \geq 1$, with $\{\epsilon_n\}_{n=1}^{\infty}$ being a sequence of iid random variables and $\epsilon_n$ independent of $X_j$, for $j < n$, for all $n$.

*Remark 2* Proposition 2 gives us that the self-decomposability of $Z$ is a sufficient condition to the MPINAR(1) process is well-defined. In fact, this condition also is necessary. Forst (1979) (see also Alamatsaz (1983)) showed that a mixed Poisson random variable is discrete self-decomposable if and only if its associated random effect (denoted in this paper by $Z$) is self-decomposable (in the sense of Def. 2). Therefore, in order to define an INAR process with mixed Poisson marginals, it is sufficient and necessary to restrict to the cases where $Z$ (the random effect) is self-decomposable.

An INAR(1) process $\{X_n\}_{n=0}^{\infty}$ satisfying (2) has 1-step transition probabilities given by

$$p_{l,k} \equiv P\left(X_n = k | X_{n-1} = l\right) = \sum_{j=0}^{\min(k,l)} \binom{l}{j} \alpha^j (1-\alpha)^{l-j} P(\epsilon_n = k - j), \quad (3)$$

for $k, l \in \mathbb{N}$. Further, the autocorrelation function is given by $\rho(k) = \alpha^k$. For more general properties on INAR processes, see Weiß (2008a).

The following result will be important to establish the asymptotic properties of the estimators which will be proposed in the Sect. 4.

**Proposition 3** *Let $\{X_n\}_{n=0}^{\infty}$ be a mixed Poisson INAR(1) process. If $E(X_n) < \infty$, then it is a geometrically ergodic, irreducible and aperiodic Markov chain.*

*Proof* The Markovian property of $\{X_n\}_{n=0}^{\infty}$ follows from (2). We now argue that $P(\epsilon_n = k) > 0$ for all $k \in \mathbb{N}$. This and (3) implies that the MPINAR(1) processes are irreducible Markov chains.

Since $\{X_n\}_{n=0}^{\infty}$ is a MPINAR(1) process, we have that the random latent effect $Z$ is an absolutely continuous random variable (see Remarks 1 and 2). With this, we have that $W$ defined in proof of Proposition 2 is an absolutely continuous random variable. Denote by $f(\cdot)$ its density function. Further, from proof of Proposition 2 we have that the innovations $\epsilon_n$ satisfies the stochastic representation $\epsilon_n | W = w \sim \text{Poisson}(\mu w)$. Therefore, we have that $P(\epsilon_n = k) = \int_0^\infty \frac{e^{-\mu w}(\mu w)^k}{k!} f(w) dw > 0$. The aperiodicity of $\{X_n\}_{n=0}^{\infty}$ follows from the fact (due to above arguments) $p_{k,k} > 0$ for all $k \in \mathbb{N}$.

Now, since $E(X_n) < \infty$, we obtain that $E(\epsilon_n) < \infty$. With this, by following the same steps of proof of Theorem 3.4.1 from Schweer and Weiß (2014), we conclude that $\{X_n\}_{n=0}^{\infty}$ is geometrically ergodic. $\square$

We now discuss about the joint and conditional mgfs and moments. Using standard manipulations, it can be checked that the conditional and joint mgfs are given by

$$\Psi_{X_n|X_{n-1}}(t) \equiv E(\exp\{t X_n\} | X_{n-1}) = \left(1 - \alpha + \alpha e^t\right)^{X_{n-1}} \frac{\Psi_Z(\mu(e^t - 1))}{\Psi_Z(\alpha \mu(e^t - 1))}$$

and

$$\Psi_{X_n, X_{n-1}}(t, s) \equiv E(\exp\{tX_n + sX_{n-1}\})$$
$$= \Psi_Z \left( \mu e^s \left( 1 - \alpha + \alpha e^t \right) - 1 \right) \frac{\Psi_Z \left( \mu \left( e^t - 1 \right) \right)}{\Psi_Z \left( \alpha \mu \left( e^t - 1 \right) \right)}, \qquad (4)$$

respectively, where $t$ belongs some interval containing the value 0 in both cases.

*Remark 3* From now on, we assume that $E(Z) = 1$ and $Var(Z) = \phi^{-1}$. The condition $E(Z) = 1$ is commonly adopted in mixed Poisson distributions in order to avoid non-identifiability problems.

The conditional mean and variance of $X_n$ given $X_{n-1}$ are respectively

$$E(X_n|X_{n-1}) = \alpha X_{n-1} + \mu(1 - \alpha)$$

and

$$\text{Var}(X_n|X_{n-1}) = \alpha(1 - \alpha)X_{n-1} + \mu(1 - \alpha)[1 + \mu\phi^{-1}(1 + \alpha)].$$

We now define the jump process by $J_n = X_n - X_{n-1}$, for $n \geq 2$. This process is generally used for diagnostic checking of fitted INAR models. For instance, see Weiß (2008b, 2009). The mgf of $J_n$ can be obtained from (4). More specifically, we have that $\Psi_{J_n}(t) \equiv E(\exp\{tJ_n\}) = \Psi_{X_n, X_{n-1}}(t, -t)$. From this, in particular, we obtain the two first moments of $J_n$, that are $E(J_n) = 0$ and $E(J_n^2) = 2\mu(1 - \alpha)(1 + \mu\phi^{-1})$.

In the subsequent section we introduce and study the Poisson-inverse Gaussian INAR(1) process with some details. Another process that belongs to our class of overdispersed INAR(1) models is the negative binomial INAR(1) process. This model is already known in the literature (under a different parametrization) and therefore we do not present it here; for instance, see McKenzie (1986) and Ristić et al. (2012). We call attention that the Poisson INAR(1) process is obtained as a limiting case of our proposed class by taking $\phi \to \infty$.

## 3 Poisson inverse-Gaussian INAR(1) process

We now present a particular model of our class that yet is unknown in the literature. Let $Z$ be a random variable with inverse Gaussian distribution with mean 1 and dispersion parameter $\phi > 0$. More specifically, assume that the moment generating function of $Z$ is given by

$$\Psi_Z(t) \equiv E(\exp\{tZ\}) = \exp\left\{ \phi \left[ 1 - \sqrt{1 - 2\phi^{-1}t} \right] \right\},$$

for $t < \phi/2$.

Let $X$ be a discrete random variable following a Poisson inverse-Gaussian (PIG) distribution with parameters $\mu > 0$ and $\phi > 0$. It is well known that $X$ satisfies

the following stochastic representation $X|Z = z \sim \text{Poisson}(\mu z)$, where $Z$ is defined above. We denote $X \sim \text{PIG}(\mu, \phi)$. The moment generating function of $X$ is given by

$$\Psi_X(t) \equiv E(\exp\{tX\}) = \exp\left\{\phi\left[1 - \sqrt{1 - 2\phi^{-1}\mu(e^t - 1)}\right]\right\},$$

for $t < \log(1 + \phi/(2\mu))$. The mean and variance of $X$ are $E(X) = \mu$ and $\text{Var}(X) = \mu(1 + \phi^{-1}\mu)$, respectively.

The probability function of $X$ can be expressed by

$$p(k) \equiv P(X = k) = \sqrt{\frac{2}{\pi}}[\phi(\phi + 2\mu)]^{-(k-1/2)/2}\frac{e^\phi(\mu\phi)^k}{k!}\mathcal{K}_{k-1/2}\left(\sqrt{\phi(\phi + 2\mu)}\right),$$

for $k = 0, 1, \ldots$, where $\mathcal{K}_\lambda(t) = \frac{1}{2}\int_0^\infty u^{\lambda-1}\exp\left\{-\frac{t}{2}\left(u + \frac{1}{u}\right)\right\}du$ is the modified Bessel function of third kind, which can be computed in softwares such as R program, Maple and Mathematica.

We have that the inverse-Gaussian distribution is self-decomposable; for instance, Pillai and Satheesh (1992) and Abraham and Balakrishna (2002). With this and using Proposition 2, we obtain that the Poisson inverse-Gaussian distribution is discrete self-decomposable. Hence, we can define our INAR process with PIG marginals as follows.

**Definition 5** We said that a sequence $\{X_n\}_{n=0}^\infty$ is a Poisson-inverse Gaussian INAR(1) process, named in short by PIGINAR(1) process, if $X_n \sim \text{PIG}(\mu, \phi)$ for all $n$ and if it admits the stochastic dynamics given in Definition 4.

The mgf of the innovations is given by

$$\begin{aligned}
\Psi(t) &= \frac{\Psi_Z(\mu(e^t - 1))}{\Psi_Z(\alpha\mu(e^t - 1))} \\
&= \exp\left\{\phi\left[\sqrt{1 - 2\phi^{-1}\mu\alpha(e^t - 1)} - \sqrt{1 - 2\phi^{-1}\mu(e^t - 1)}\right]\right\},
\end{aligned}$$

for $t < \log(1 + \phi/(2\mu))$. The first four cumulants of the innovations are

$$\begin{aligned}
E(\epsilon_n) &= \mu(1 - \alpha), \\
\text{Var}(\epsilon_n) &= \mu(1 - \alpha)[1 + \mu\phi^{-1}(1 + \alpha)], \\
E\left[(\epsilon_n - E(\epsilon_n))^3\right] &= 3\mu^3\phi^{-2}(1 - \alpha^3) + 3\mu^2\phi^{-1}(1 - \alpha^2) + \mu(1 - \alpha)
\end{aligned}$$

and

$$\begin{aligned}
E\left[(\epsilon_n - E(\epsilon_n))^4\right] = &\ 15\mu^4\phi^{-3}(1 - \alpha^4) + 18\mu^3\phi^{-2}(1 - \alpha^3) \\
&+ 7\mu^2\phi^{-1}(1 - \alpha^2) + \mu(1 - \alpha).
\end{aligned}$$

We are naturally interested in the distribution of the innovations since we need to simulate the $\epsilon_n$'s in order to construct simulated trajectories and to study the finite-sample behaviour of the estimators we will propose. On the other hand, it is very difficult to obtain an explicit form for the probability function of the innovations.

More specifically, our innovations satisfy the following stochastic representation: $\epsilon|W = w \sim \text{Poisson}(\mu w)$, where $W$ has the distribution of an innovation of an inverse-Gaussian AR(1) process [see Pillai and Satheesh (1992); Abraham and Balakrishna (2002)], which has not explicit density function. Therefore, we do not have an explicit form for the probability function of our innovations $\epsilon_n$.

To overcome this problem, we use a method for generating random variables based on the moment generating function given by Ridout (2009). With this, we generate observations of the distribution of $W$ and use the above stochastic representation for generating from the distribution of the innovations.

## 4 Estimation and asymptotic properties

We discuss here estimation methods of the parameter vector $\theta \equiv (\mu, \alpha, \phi)^\top$ and establish consistency and asymptotic normality. We begin with the conditional least squares (CLS) method. The function we need to minimize in order to find the conditional least squares estimator for $\theta$ is given by

$$Q(\theta) = \sum_{i=2}^{n}(X_i - E(X_i|X_{i-1}))^2 = \sum_{i=2}^{n}(X_i - \alpha X_{i-1} - \mu(1-\alpha))^2,$$

where $n$ is the sample size of the count time series. The above estimating equation does not depend on $\phi$ and therefore an alternative method to estimate it is necessary; this point will be discussed in the sequence. By solving the nonlinear system of equations $(\partial Q(\theta)/\partial\mu, \partial Q(\theta)/\partial\alpha)^\top = (0, 0)^\top$, we get

$$\widehat{\alpha}_{cls} = \frac{(n-1)\sum_{i=2}^{n} X_i X_{i-1} - \sum_{i=2}^{n} X_i \sum_{i=2}^{n} X_{i-1}}{(n-1)\sum_{i=2}^{n} X_{i-1}^2 - (\sum_{i=2}^{n} X_{i-1})^2}$$

and

$$\widehat{\mu}_{cls} = \frac{\sum_{i=2}^{n} X_i - \widehat{\alpha}_{cls} \sum_{i=2}^{n} X_{i-1}}{(n-1)(1-\widehat{\alpha}_{cls})}.$$

The above estimators based on the CLS method estimation assume the same form as the corresponding ones in the Poisson INAR(1) process; see Al-Osh and Alzaid (1987) and Freeland and McCabe (2005).

The additional equation for estimating $\phi$ is based on the $Q^*$ function given by

$$Q^*(\theta^*) = \sum_{i=2}^{n}(X_i^2 - E(X_i^2|X_{i-1}))^2$$

$$= \sum_{i=2}^{n}(X_i^2 - g(\widehat{\mu}_{cls}, \widehat{\alpha}_{cls}, X_{i-1}) - \phi^{-1}h(\widehat{\mu}_{cls}, \widehat{\alpha}_{cls}))^2,$$

where $\theta^* = (\widehat{\mu}_{cls}, \widehat{\alpha}_{cls}, \phi)^\top$, $g(\mu, \alpha, x) = (\alpha x + \mu(1 - \alpha))^2 + \alpha(1 - \alpha)x$ and $h(\mu, \alpha) = \mu^2(1 - \alpha^2)$.

The above method we using here, that is, by using two conditional least squares equations, it is known as two-step CLS estimation and was introduced by Karlsen and Tjostheim (1988). Solving the equation $\partial Q^*(\theta^*)/\partial\phi = 0$, we get

$$\widehat{\phi}_{cls} = \frac{(n-1)\widehat{\mu}_{cls}^2(1 - \widehat{\alpha}_{cls}^2)}{\sum_{i=2}^{n}\left\{X_i^2 - (1 - \widehat{\alpha}_{cls})(\widehat{\mu}_{cls} + \widehat{\alpha}_{cls}X_{i-1}) - (\widehat{\alpha}_{cls}X_{i-1} + \widehat{\mu}_{cls}(1 - \widehat{\alpha}_{cls}))^2\right\}}.$$

We obtain now the Yule-Walker estimators. Since $E(X_t) = \mu$, $\text{Var}(X_t) = \mu(1 + \mu/\phi)$ and $\text{corr}(X_{t-1}, X_t) = \alpha$, we get the following Yule-Walker estimators:

$$\widehat{\alpha}_{yw} = \frac{\sum_{i=2}^{n}(X_i - \bar{X}_n)(X_{i-1} - \bar{X}_n)}{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}, \tag{5}$$

$$\widehat{\mu}_{yw} = \bar{X}_n \tag{6}$$

and

$$\widehat{\phi}_{yw} = \frac{\bar{X}_n}{S_n^2 - \bar{X}_n}, \tag{7}$$

where $\bar{X}_n = \frac{1}{n}\sum_{t=1}^{n} X_t$ is the sample mean and $S_n^2 = \frac{1}{n}\sum_{t=1}^{n} X_t^2 - \bar{X}_n^2$ is the sample variance.

We perform now a small Monte Carlo simulation in order to study the finite-sample behaviour of the proposed estimators and to compare the CLS and Yule-Walker estimation methods. We generate the PIGINAR(1) process, which was introduced and studied with some details in the previous section, under four configurations. In the Configurations I, II, III and IV we set $(\mu, \alpha, \phi) = (0.5, 0.7, 0.2)$, $(\mu, \alpha, \phi) = (2, 0.5, 1)$, $(\mu, \alpha, \phi) = (1, 0.25, 0.5)$ and $(\mu, \alpha, \phi) = (0.5, 0.1, 1.5)$, respectively. We consider the sample sizes $n = 100, 200, 300, 400, 500, 1000$ and the number of loops of the Monte Carlo simulation equals to 10000.

The results based on the Configurations I, II, III and IV are presented in Tables 1, 2, 3 and 4, respectively. These tables give us the empirical mean of the estimates of the parameters and also the standard errors.

**Table 1** Empirical mean and standard errors (in parentheses) of the estimates of the parameters for $(\mu, \alpha, \phi) = (0.5, 0.7, 0.2)$ and some values of $n$ under the PIGINAR(1) process

| $n$ | $\widehat{\mu}_{cls}$ | $\widehat{\mu}_{yw}$ | $\widehat{\alpha}_{cls}$ | $\widehat{\alpha}_{yw}$ | $\widehat{\phi}_{cls}$ | $\widehat{\phi}_{yw}$ |
|---|---|---|---|---|---|---|
| 100 | 0.534 | 0.532 | 0.659 | 0.658 | 1.267 | 0.770 |
|     | (0.318) | (0.306) | (0.105) | (0.106) | (16.165) | (2.887) |
| 200 | 0.509 | 0.508 | 0.677 | 0.676 | 0.473 | 0.481 |
|     | (0.224) | (0.220) | (0.075) | (0.076) | (1.409) | (1.751) |
| 300 | 0.501 | 0.501 | 0.684 | 0.683 | 0.354 | 0.345 |
|     | (0.187) | (0.182) | (0.062) | (0.062) | (1.099) | (0.524) |
| 400 | 0.501 | 0.501 | 0.688 | 0.687 | 0.300 | 0.299 |
|     | (0.157) | (0.156) | (0.054) | (0.054) | (0.384) | (0.323) |
| 500 | 0.500 | 0.500 | 0.690 | 0.689 | 0.271 | 0.271 |
|     | (0.141) | (0.140) | (0.049) | (0.049) | (0.144) | (0.145) |
| 1000 | 0.501 | 0.501 | 0.694 | 0.694 | 0.235 | 0.235 |
|     | (0.099) | (0.099) | (0.036) | (0.036) | (0.078) | (0.078) |

**Table 2** Empirical mean and standard errors (in parentheses) of the estimates of the parameters for $(\mu, \alpha, \phi) = (2, 0.5, 1)$ and some values of $n$ under the PIGINAR(1) process

| $n$ | $\widehat{\mu}_{cls}$ | $\widehat{\mu}_{yw}$ | $\widehat{\alpha}_{cls}$ | $\widehat{\alpha}_{yw}$ | $\widehat{\phi}_{cls}$ | $\widehat{\phi}_{yw}$ |
|---|---|---|---|---|---|---|
| 100 | 2.005 | 2.005 | 0.474 | 0.474 | 1.391 | 1.354 |
|     | (0.430) | (0.425) | (0.093) | (0.093) | (1.000) | (0.857) |
| 200 | 2.001 | 2.001 | 0.486 | 0.486 | 1.170 | 1.159 |
|     | (0.302) | (0.300) | (0.067) | (0.068) | (0.407) | (0.400) |
| 300 | 1.999 | 1.999 | 0.491 | 0.491 | 1.117 | 1.110 |
|     | (0.246) | (0.245) | (0.056) | (0.056) | (0.317) | (0.313) |
| 400 | 1.998 | 1.998 | 0.493 | 0.493 | 1.085 | 1.081 |
|     | (0.213) | (0.213) | (0.049) | (0.049) | (0.262) | (0.260) |
| 500 | 1.999 | 2.000 | 0.495 | 0.495 | 1.063 | 1.060 |
|     | (0.191) | (0.190) | (0.044) | (0.044) | (0.228) | (0.227) |
| 1000 | 2.000 | 2.000 | 0.497 | 0.497 | 1.037 | 1.036 |
|     | (0.135) | (0.135) | (0.032) | (0.032) | (0.158) | (0.158) |

Looking at these tables, we see that the bias and standard error of the estimates of the parameters decreases as the sample size increases for all cases considered, as expected. We also observe that the parameters $\mu$ and $\alpha$ are in general well estimated and that the CLS and Yule-Walker estimators yield very similar results for these parameters. The explanation for this is that $(\widehat{\mu}_{cls}, \widehat{\alpha}_{cls})$ and $(\widehat{\mu}_{yw}, \widehat{\alpha}_{yw})$ are asymptotically equivalent. More specifically, we have that $(\widehat{\mu}_{cls} - \widehat{\mu}_{yw}, \widehat{\alpha}_{cls} - \widehat{\alpha}_{yw}) = o_p(n^{-1/2})$. This can be checked by following the same steps that in Proof of Theorem 3 by Freeland and McCabe (2005). We now concentrate our attention to the estimation of the parameter $\phi$. Under the Configurations II ($\phi = 1$) and III ($\phi = 0.5$), we observe that the

**Table 3** Empirical mean and standard errors (in parentheses) of the estimates of the parameters for $(\mu, \alpha, \phi) = (1, 0.25, 0.5)$ and some values of $n$ under the PIGINAR(1) process

| $n$ | $\widehat{\mu}_{cls}$ | $\widehat{\mu}_{yw}$ | $\widehat{\alpha}_{cls}$ | $\widehat{\alpha}_{yw}$ | $\widehat{\phi}_{cls}$ | $\widehat{\phi}_{yw}$ |
|---|---|---|---|---|---|---|
| 100 | 1.003 | 1.002 | 0.233 | 0.232 | 0.706 | 0.697 |
| | (0.228) | (0.226) | (0.101) | (0.100) | (0.546) | (0.544) |
| 200 | 1.001 | 1.001 | 0.240 | 0.239 | 0.592 | 0.589 |
| | (0.160) | (0.159) | (0.076) | (0.076) | (0.217) | (0.214) |
| 300 | 1.002 | 1.002 | 0.243 | 0.243 | 0.562 | 0.560 |
| | (0.130) | (0.130) | (0.063) | (0.063) | (0.163) | (0.163) |
| 400 | 0.999 | 0.999 | 0.244 | 0.244 | 0.548 | 0.548 |
| | (0.113) | (0.112) | (0.056) | (0.056) | (0.139) | (0.139) |
| 500 | 1.001 | 1.001 | 0.246 | 0.246 | 0.538 | 0.537 |
| | (0.101) | (0.101) | (0.050) | (0.050) | (0.122) | (0.121) |
| 1000 | 0.999 | 0.999 | 0.248 | 0.248 | 0.519 | 0.519 |
| | (0.072) | (0.072) | (0.036) | (0.036) | (0.085) | (0.085) |

**Table 4** Empirical mean and standard errors (in parentheses) of the estimates of the parameters for $(\mu, \alpha, \phi) = (0.5, 0.1, 1.5)$ and some values of $n$ under the PIGINAR(1) process

| $n$ | $\widehat{\mu}_{cls}$ | $\widehat{\mu}_{yw}$ | $\widehat{\alpha}_{cls}$ | $\widehat{\alpha}_{yw}$ | $\widehat{\phi}_{cls}$ | $\widehat{\phi}_{yw}$ |
|---|---|---|---|---|---|---|
| 100 | 0.509 | 0.509 | 0.124 | 0.124 | 4.423 | 2.897 |
| | (0.091) | (0.090) | (0.083) | (0.083) | (40.294) | (6.126) |
| 200 | 0.502 | 0.502 | 0.108 | 0.108 | 2.744 | 2.459 |
| | (0.064) | (0.064) | (0.066) | (0.066) | (7.079) | (5.276) |
| 300 | 0.501 | 0.501 | 0.103 | 0.103 | 2.291 | 2.112 |
| | (0.052) | (0.052) | (0.057) | (0.057) | (5.040) | (2.631) |
| 400 | 0.500 | 0.500 | 0.100 | 0.100 | 2.000 | 1.922 |
| | (0.045) | (0.045) | (0.050) | (0.050) | (3.224) | (2.041) |
| 500 | 0.500 | 0.500 | 0.098 | 0.099 | 1.832 | 1.801 |
| | (0.041) | (0.041) | (0.047) | (0.047) | (1.711) | (1.522) |
| 1000 | 0.500 | 0.500 | 0.098 | 0.098 | 1.623 | 1.614 |
| | (0.029) | (0.029) | (0.035) | (0.035) | (0.492) | (0.486) |

estimators $\widehat{\phi}_{cls}$ and $\widehat{\phi}_{yw}$ yield similar results with a slightly advantage for the Yule-Walker estimator that presents smaller bias and standard error with respect to the CLS estimator. Now, with respect to the Configurations I ($\phi = 0.2$) and IV ($\phi = 1.5$), we see that there is a considerable bias under both CLS and Yule-Walker estimation methods for estimating $\phi$ with $n = 100, 200, 300$. Anyway, this bias is smaller under the Yule-Walker approach. We also see that the standard error of the Yule-Walker estimator is smaller than that one based on the CLS estimation. For $n \geq 400$, both methods work similarly. Overall, we see that the Yule-Walker estimation method works better that the CLS one for our mixed Poisson INAR(1) process considered. With this,

we suggest the use of the Yule-Walker approach. This method will be considered in our application to a real dataset.

We establish now the consistency and asymptotic normality of the estimators (5), (6) and (7).

**Proposition 4** *Assume that* $E(X_n^4) < \infty$. *Then, the estimators* $\widehat{\mu}_{yw}$, $\widehat{\phi}_{yw}$ *and* $\widehat{\alpha}_{yw}$ *are strongly consistent for* $\mu$, $\phi$ *and* $\alpha$ *respectively and satisfy the asymptotic normality*

$$\sqrt{n}\{(\widehat{\mu}_{yw}, \widehat{\phi}_{yw}, \widehat{\alpha}_{yw}) - (\mu, \phi, \alpha)\} \xrightarrow{d} N(0, B\Sigma B^\top),$$

*as* $n \to \infty$, *where*

$$B = \begin{pmatrix} 1 & 0 & 0 \\ \mu\dfrac{2(\kappa_2 + \mu^2) - \mu}{(\kappa_2 - \mu)^2} & -\dfrac{\mu^2}{(\kappa_2 - \mu)^2} & 0 \\ -2\mu\dfrac{1 - \alpha}{\kappa_2} & -\dfrac{\alpha}{\kappa_2} & \dfrac{1}{\kappa_2} \end{pmatrix}$$

*and* $\Sigma$ *being a* $3 \times 3$ *symmetric matrix with entries given by*

$$\Sigma_{11} = \kappa_2 \frac{1 + \alpha}{1 - \alpha},$$

$$\Sigma_{12} = \kappa_3 + 2\mu\kappa_2 + (\kappa_3 - \kappa_2)\frac{\alpha^2}{1 - \alpha^2} + \frac{\alpha}{1 - \alpha}(\kappa_2 + 4\mu\kappa_2 + \kappa_3),$$

$$\Sigma_{13} = \alpha(1 + \alpha)(\kappa_3 - \kappa_2) + 2\kappa_2(\alpha + \mu(1 + \alpha)) + \frac{\alpha^3}{1 - \alpha^2}(\kappa_3 - \kappa_2)$$
$$+ \frac{\alpha}{1 - \alpha}\{\alpha^2(\kappa_3 - \kappa_2) + 2\kappa_2[\mu + \alpha(1 + \mu)]\},$$

$$\Sigma_{22} = \mu_4 - \mu_2^2 + \frac{2\alpha^2}{1 - \alpha^2}\left\{\kappa_4 + \kappa_3(2\mu - 1) - \kappa_2^2 - 2\mu\kappa_2\right\}$$
$$+ \frac{2\alpha}{1 - \alpha}(1 + 2\mu)(\kappa_3 + 2\mu\kappa_2),$$

$$\Sigma_{23} = \alpha(1 + \alpha^2)(\kappa_4 - 3\kappa_3 + \kappa_2(2 - 3\kappa_2)) + (\kappa_3 - \kappa_2)\{3\alpha(1 + \alpha)(1 + \mu) + 2\mu\}$$
$$+ 2\mu^2\kappa_2 + 4\kappa_2^2\alpha + 2\kappa_2(1 + \mu)(\alpha + \mu(1 + 2\alpha))$$
$$+ \frac{\alpha^3}{1 - \alpha^2}\left\{\kappa_4 - 3\kappa_3 + \kappa_2(2 - 3\kappa_2) + 2(\kappa_3 - \kappa_2)(1 + \mu) + 2\kappa_2^2\right\}$$
$$+ \frac{\alpha^2}{1 - \alpha^2}\left\{\alpha^3(\kappa_4 - 3\kappa_3 + \kappa_2(2 - 3\kappa_2)) + \alpha^2(\kappa_3 - \kappa_2)(2 + \mu)\right.$$
$$+ 2\alpha\kappa_2^2 + \mu(\kappa_3 - \kappa_2)\Big\}$$
$$+ \frac{\alpha}{1 - \alpha}\left\{(\kappa_3 - \kappa_2)\left(\alpha^2(1 + 2\mu) + \alpha(1 + \mu) + \mu\right)\right.$$
$$+ 2\alpha\kappa_2(1 + \mu)^2 + 2\mu\kappa_2(\mu + (1 + \mu)(1 + \alpha))\Big\}$$

*and*

$$
\Sigma_{33} = \frac{2\alpha^3}{1-\alpha^2} \left\{ \alpha^3(\kappa_4 - 3\kappa_3 + \kappa_2(2 - 3\kappa_2)) + \alpha^2(\kappa_3 - \kappa_2)(2 + \mu) \right.
$$

$$
\left. + 2\alpha\kappa_2^2 + \mu(\kappa_3 - \kappa_2) \right\}
$$

$$
+ \frac{2\alpha}{1-\alpha} \left\{ \alpha^3(\kappa_3 - \kappa_2)(1 + \mu) + \alpha^2\mu(\kappa_3 - \kappa_2) + \alpha^2\kappa_2(1 + \mu)^2 \right.
$$

$$
\left. + 2\alpha\mu\kappa_2(1 + \mu) + \mu^2\kappa_2 \right\}
$$

$$
+ \alpha^2(1 + 2\alpha^2) \{\kappa_4 - 3\kappa_3 + \kappa_2(2 - 3\kappa_2)\}
$$

$$
+ \alpha(\kappa_3 - \kappa_2) \left\{ \alpha + (1 + \alpha)(1 + 4\mu) + 2\alpha^2(3 + 2\mu) \right\} + \alpha\kappa_2(1 + \mu)(5\mu
$$

$$
+ 2\alpha(1 + \mu) + 1) + (4 + \alpha)\mu^2\kappa_2 + 5\alpha^2\kappa_2^2 + \kappa_2(\kappa_2 - 2\alpha\mu^2).
$$

*Proof* Since $E(X_n) < \infty$ by assumption, it follows from Proposition 3 that the stationary process $\{X_n\}_{n=0}^{\infty}$ is ergodic. With this, we can apply the Law of Large Numbers for stationary and ergodic processes to obtain the strong consistency of the estimators; for instance, see Proposition 7.1 by Hamilton (1994).

In order to show the asymptotic normality of the estimators, we define $Y_n = \sum_{t=1}^{n} X_t$, $Z_n = \sum_{t=1}^{n} X_t^2$ and $W_n = \sum_{t=1}^{n} X_{t-1} X_t$, for $n \geq 1$. Note that the estimators $(\widehat{\mu}_{yw}, \widehat{\phi}_{yw}, \widehat{\alpha}_{yw})$ are asymptotically equivalent to $g(Y_n/n, Z_n/n, W_n/n)$, where $g(y, z, w) = \left( y, \frac{y}{z - y^2 - y}, \frac{w - y^2}{z - y^2} \right)$, for $(y, z, w) \in \mathbb{R}^+$, $z \neq y(y + 1)$ and $z \neq y^2$.

By using the Central Limit Theorem for stationary and ergodic processes (see Chapter 7 from Anderson (1971)), we obtain that $\sqrt{n}\{(Y_n/n, Z_n/n, W_n/n) - (\mu, \mu^2 + \kappa_2, \mu^2 + \alpha\kappa_2)\} \xrightarrow{d} N(0, \Sigma)$ as $n \to \infty$, where $\mu = E(X_n)$, $\kappa_2 = \mathrm{Var}(X_n)$ and $\Sigma$ is some asymptotic covariance matrix. With the results above and by applying the Delta Method with the function $g(\cdot, \cdot, \cdot)$ as defined above, we obtain that $\sqrt{n}\{(\widehat{\mu}_{yw}, \widehat{\phi}_{yw}, \widehat{\alpha}_{yw}) - (\mu, \phi, \alpha)\} \xrightarrow{d} N(0, B\Sigma B^\top)$, where $B = \partial g(\theta)/\partial \theta$. The point that remains to be proven is the form of the covariance matrix $\Sigma$, which is done in the Appendix. $\qquad\square$

## 5 Empirical illustration

In this section we illustrate the usefulness of the proposed class of mixed Poisson INAR(1) processes. We consider a sex offence count time series data, which can be found at site www.forecastingprinciples.com. For this count time series, an observation corresponds to a monthly count of sex offences reported in the 21st police car beat in Pittsburgh. These data contain 144 observations starting in January 1990 and ending in December 2001 and was firstly used by Ristić et al. (2009). The plots of the data and the associated ACF and PACF are presented in Fig. 1. From these plots, we see that the first sample autocorrelation is more significant than the other and there exists
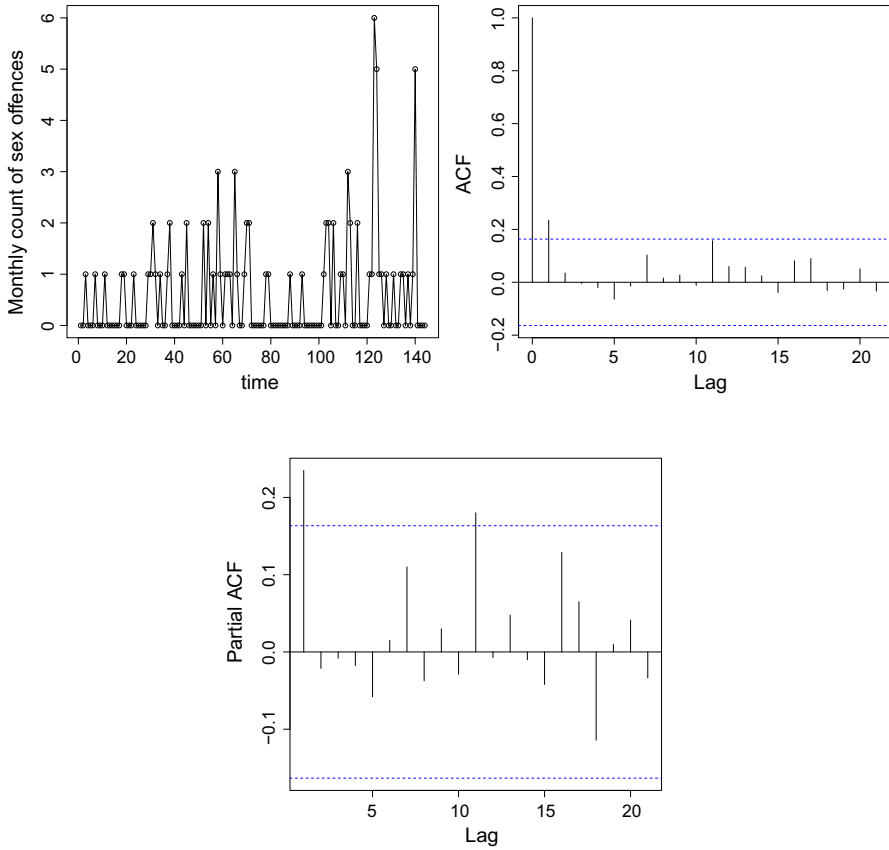
**Fig. 1** Plots of the monthly count of sex offences reported in the 21st police car beat in Pittsburgh from January 1990 up to December 2001 and the associated ACF and PACF

a cutoff after lag 1 in the partial autocorrelations. These facts indicate that an INAR process of order 1 can be suitable for modelling the monthly count of sex offences.

We begin this application performing an overdispersion test proposed by Schweer and Weiß (2014), which the test statistic is based on the empirical index of dispersion $\widehat{I}_d \equiv S_n^2/\bar{X}_n$. The null hypothesis is $H_0: X_1, \ldots, X_n$ *come from a Poisson INAR(1) process* against the alternative hypothesis $H_1: X_1, \ldots, X_n$ *come from an overdispersed INAR(1) process*. For the count time series considered here, we have that $\widehat{I}_d = 1.7394$. The associated p-value is $2.55 \times 10^{-12}$, that is, by using any usual significance level (for instance at 5%) we reject the null hypothesis in favor of the alternative hypothesis that states that an overdispersed INAR(1) process is more adequate for modelling this count data.

In Table 5 we present the estimates of the parameters based on the Yule-Walker estimation method given in Sect. 4. We also present in this table the associated standard errors and confidence intervals for the parameters with significance level at 5%. We consider the PIGINAR(1) and NBINAR(1) (negative binomial) processes and call

**Table 5** Estimates of the parameters and their associated standard errors under the MPINAR(1) processes based on the Yule-Waker estimation method for the monthly count of sex offences time series data

| Model | Parameter | Estimate | Stand. error | Inf. bound | Sup. bound |
|---|---|---|---|---|---|
| PIGINAR(1) | $\mu$ | 0.5903 | 0.1073 | 0.3800 | 0.8006 |
| | $\alpha$ | 0.2354 | 0.1062 | 0.0272 | 0.4435 |
| | $\phi$ | 0.7983 | 0.4183 | 0.0000 | 1.6180 |
| NBINAR(1) | $\mu$ | 0.5903 | 0.1073 | 0.3800 | 0.8006 |
| | $\alpha$ | 0.2354 | 0.1042 | 0.0311 | 0.4397 |
| | $\phi$ | 0.7983 | 0.3406 | 0.1308 | 1.4658 |

**Table 6** Comparison among the models based on empirical and estimated quantities under the PIG, NB and Poisson INAR(1) processes

| Quantity | Empirical | PIG | NB | Poisson |
|---|---|---|---|---|
| $\kappa_1$ | 0.5903 | 0.5944 | 0.5944 | 0.5944 |
| $\kappa_2$ | 1.0268 | 1.0370 | 1.0370 | 0.5944 |
| Skew. | 2.5990 | 2.7565 | 2.4444 | 1.2971 |
| Kurt. | 11.6804 | 12.3732 | 8.4806 | 1.6824 |
| $I_d$ | 1.7394 | 1.7446 | 1.7446 | 1.0000 |
| $p_0$ | 0.6250 | 0.6305 | 0.6413 | 0.5519 |
| $\mu(1)$ | 0.5903 | 0.5974 | 0.5974 | 0.4932 |

attention that the estimates of the parameters are the same in both case since the estimation depends only on the first moments that are equal for these models. We see that the standard error for the estimates of $\mu$ are equal for both models (as expected) and that the standard errors of the estimates of $\alpha$ and $\phi$ are smaller under the negative binomial assumption with respect to the PIG-based model.

We also fit a Poisson INAR(1) process and compare with our fitted mixed Poisson INAR(1) processes. In Table 6 we provide some empirical and estimated quantities based on the Poisson, negative binomial and Poisson-inverse Gaussian INAR(1) models. We consider the following quantities: mean, variance, skewness, kurtosis, index of dispersion (ratio of the variance to the mean), probability of zero and the first higher-order moment, denoted by $\kappa_1$, $\kappa_2$, skew., kurt., $I_d$, $p_0$ and $\mu(1)$, respectively.

From the results presented in Table 6, it is very clear that the Poisson INAR(1) process is not adequate for modelling the monthly count of sex offences considered here. On the other hand, we observe a good agreement among the empirical quantities and the estimated quantities based on our mixed Poisson INAR(1) processes, with a slight advantage of the PIGINAR(1) process over the NBINAR(1) process.

We now discuss the goodness-of-fit of the models based on the residuals $R_t \equiv X_t - \widehat{E}(X_t|X_{t-1}) = X_t - \widehat{\alpha}_{yw} X_{t-1} - \widehat{\mu}_{yw}(1 - \widehat{\alpha}_{yw})$ and on the jumps $J_t$ defined in the final of Sect. 2, for $t = 2, \ldots, n$. In the Fig. 2, we present the ACF of the residuals based on the MPINAR(1) process. This figure indicates that the residuals $R_2, \ldots, R_n$ are not correlated, so our fitted models seem to have captured well the dependence
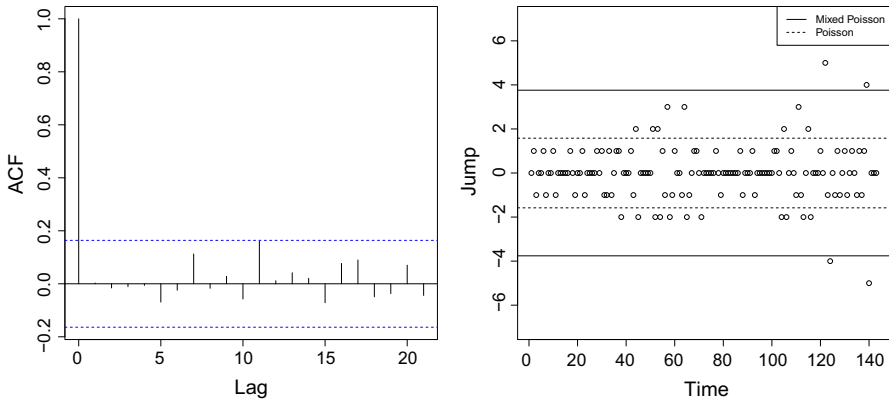
**Fig. 2** Sample autocorrelation function (ACF) of the residuals and jumps against time with bounds based on the mixed Poisson INAR(1) and Poisson INAR(1) processes

of the count time series. In this figure we also plot the jumps against the time with $\pm 3\sigma_J$ bounds chosen as benchmark chart, where $\sigma_J = \sqrt{\text{Var}(J_t)}$; this graphic is also known in the literature as Shewhart control chart and was proposed by Weiß (2009) as a diagnostic checking of fitted INAR(1) models. We plot the bounds based on our mixed Poisson INAR processes and also the corresponding to the Poisson INAR process. We see that the bounds based on the mixed INAR process accommodate adequately the points (around 97% of them) in contrast with the benchmark chart related to the Poisson INAR process, where many points are out of the bounds (around 16% of them).

As suggested by one referee, we also use the standardized Pearson residual (for instance, see Harvey and Fernandes (1989) and Jung and Tremayne (2011))

$$
\begin{aligned}
r_t &\equiv \frac{X_t - \widehat{E}(X_t|X_{t-1})}{\sqrt{\widehat{\text{Var}}(X_t|X_{t-1})}} \\
&= \frac{X_t - \widehat{\alpha}_{yw} X_{t-1} - \widehat{\mu}_{yw}(1 - \widehat{\alpha}_{yw})}{\sqrt{\widehat{\alpha}_{yw}(1 - \widehat{\alpha}_{yw})X_{t-1} + \widehat{\mu}_{yw}(1 - \widehat{\alpha}_{yw})[1 + \widehat{\mu}_{yw}\widehat{\phi}_{yw}^{-1}(1 + \widehat{\alpha}_{yw})]}},
\end{aligned}
$$

for $t = 2, \ldots, n$, in order to check the variance, which should be close to 1. According to Harvey and Fernandes (1989), a sample variance greater than 1 indicates overdispersion with respect to the model that is being considered. We compute the sample variance of the standardized Pearson Residual and obtain that it is equal to 0.9539, so giving evidence that our mixed Poisson INAR(1) processes have captured with success the overdispersion of the count time series considered.

With the results and discussion presented above, we conclude that the INAR process based on the Poisson assumption is not adequate for modelling the present count time series data and that the mixed Poisson INAR processes provide an adequate approach.

# 6 Conclusions and future research

A common way for treating overdispersion in count data is to use the mixed Poisson distributions, which is obtained by introducing a latent random effect on the mean of a Poisson distribution. With this motivation, we introduced a class of INAR(1) processes with mixed Poisson marginals. We established a condition to our class is well-defined and discussed statistical properties. The INAR(1) process with Poisson-inverse Gaussian marginals was introduced and presented with some details. We discussed estimation of the parameters and gave conditions to have consistency and asymptotic normality of the estimators. Simulated results showed that the estimators work well for the scenarios considered. An empirical example illustrated the importance in considering our overdispersed INAR processes for analyzing count time series data.

Possible points of future research are: (a) multivariate extension of our mixed Poisson INAR(1) process; (b) to define and study a mixed Poisson INAR process of $p$-order; (c) to explore more particular cases of our class of overdispersed INAR processes such as an INAR model with Poisson-log normal marginals.

## Appendix

We here derive the asymptotic covariance matrix of the weak convergence given in Proposition 4. For this, we will use joint moments, which are defined by

$$\mu(s_1, \ldots, s_r) \equiv E(X_t X_{t+s_1} \ldots X_{t+s_r}),$$

where $0 \leq s_1 \leq \ldots \leq s_r$ and $r$ belongs to $\mathbb{N}$. Expressions of joint moments up to fourth order for INAR(1) processes were obtained by Schweer and Weiß (2014) (Theorem 3.3.1). We will use these expressions for computing the desired asymptotic covariance matrix.

Let $Y_n$, $Z_n$ and $W_n$ be as defined in Sect. 4 and define $g_n(i, j) = n - i - \frac{\alpha^j}{1-\alpha^j}\big(1 - \alpha^{j(n-i)}\big)$, for $i, j \in \mathbb{N}$. Using joint moments and their expressions given in Theorem 3.3.1 from Schweer and Weiß (2014), after some manipulations we obtain that

$$\text{Var}(Y_n) = n\text{Var}(X_t) + 2\sum_{t=1}^{n}\sum_{k=t+1}^{n}\text{cov}(X_k, X_t) = n\kappa_2 + 2\sum_{t=1}^{n}\sum_{k=t+1}^{n}\left\{\mu(k-t) - \mu^2\right\}$$

$$= n\kappa_2\left\{1 + 2\frac{\alpha}{1-\alpha} - 2\alpha^2\frac{1-\alpha^n}{n(1-\alpha)^2}\right\},$$

$$\text{Var}(Z_n) = n\text{Var}(X_t^2) + 2\sum_{t=1}^{n}\sum_{k=t+1}^{n}\text{cov}\left(X_k^2, X_t^2\right)$$

$$= n(\mu_4 - \mu_2^2) + 2\sum_{t=1}^{n}\sum_{k=t+1}^{n}\left\{\mu(0, k-t, k-t) - \left(\kappa_2 + \mu^2\right)^2\right\}$$

$$= n\left(\mu_4 - \mu_2^2\right) + 2\frac{\alpha^2}{1-\alpha^2}g_n(1,2)\left\{\kappa_4 + \kappa_3(2\mu-1) - \kappa_2^2 - 2\mu\kappa_2\right\}$$

$$+ 2\frac{\alpha}{1-\alpha}g_n(1,1)(1+2\mu)(\kappa_3 + 2\mu\kappa_2),$$

$$\text{Var}(W_n) = n\text{Var}(X_{t-1}X_t) + 2\sum_{t=1}^{n-2}\sum_{k=t+2}^{n}\text{cov}(X_{k-1}X_k, X_{t-1}X_t)$$

$$+ 2\sum_{t=1}^{n-1}\text{cov}(X_{t-1}X_t, X_tX_{t+1})$$

$$= n\left\{\mu(0,1,1) - \mu(1)^2\right\} + 2\sum_{t=1}^{n-2}\sum_{k=t+2}^{n}\left\{\mu(1, k-t, k-t+1) - \mu(1)^2\right\}$$

$$+ 2\sum_{t=1}^{n-1}\left\{\mu(1,1,2) - \mu(1)^2\right\}$$

$$= 2g_n(2,2)\frac{\alpha^3}{1-\alpha^2}\left\{\alpha^3(\kappa_4 - 3\kappa_3 + \kappa_2(2 - 3\kappa_2))\right.$$

$$+ \alpha^2(\kappa_3 - \kappa_2)(2+\mu) + 2\alpha\kappa_2^2 + \mu(\kappa_3 - \kappa_2)\Big\}$$

$$+ 2g_n(2,1)\frac{\alpha}{1-\alpha}\left\{\alpha^3(\kappa_3 - \kappa_2)(1+\mu) + \alpha^2\mu(\kappa_3 - \kappa_2)\right.$$

$$+ \alpha^2\kappa_2(1+\mu)^2 + 2\alpha\mu\kappa_2(1+\mu) + \mu^2\kappa_2\Big\}$$

$$+ \alpha^2\{\kappa_4 - 3\kappa_3 + \kappa_2(2 - 3\kappa_2)\}(n + 2\alpha^2(n-1))$$

$$+ \alpha(\kappa_3 - \kappa_2)\{n(\alpha + (1+\alpha)(1+2\mu))$$

$$+ 2(n-1)(\mu(1+\alpha) + \alpha^2(3+2\mu))\} + \alpha\kappa_2\{n(1+\mu)(1+3\mu)$$

$$+ 2(n-1)(1+\mu)(2\mu + \alpha(1+\mu))\}$$

$$+ \mu^2\kappa_2(n(4+\alpha) - 2) + \alpha^2\kappa_2^2(5n-2) + n\kappa_2(\kappa_2 - 2\alpha\mu^2),$$

$$\text{cov}(Y_n, Z_n) = nE(X_t^3) + \sum_{t=1}^{n-1}\sum_{k=t+1}^{n}E(X_tX_k^2)$$

$$+ \sum_{k=1}^{n-1}\sum_{t=k+1}^{n}E(X_tX_k^2) - n^2E(X_t)E(X_t^2)$$

$$= n\mu_3 + \sum_{t=1}^{n-1}\sum_{k=t+1}^{n}\mu(k-t, k-t) + \sum_{k=1}^{n-1}\sum_{t=k+1}^{n}\mu(0, t-k) - n^2\mu\mu_2$$

$$= n(\kappa_3 + 2\kappa_2\mu) + (\kappa_3 - \kappa_2)\frac{\alpha^2}{1-\alpha^2}g_n(1,2)$$

$$+ \frac{\alpha}{1-\alpha}g_n(1,1)(\kappa_2 + 4\mu\kappa_2 + \kappa_3),$$

$$\mathrm{cov}(Y_n, W_n) = nE(X_{t-1}X_t^2) + (n-1)E(X_{t-1}^2 X_t) + \sum_{t=1}^{n-2}\sum_{k=t+2}^{n} E(X_t X_{k-1} X_k)$$

$$+ \sum_{k=1}^{n-1}\sum_{t=k+1}^{n} E(X_t X_{k-1} X_k) - n^2 E(X_t)E(X_{t-1}X_t)$$

$$= n\mu(1,1) + (n-1)\mu(0,1) + \sum_{t=1}^{n-2}\sum_{k=t+2}^{n}\mu(k-t-1, k-t)$$

$$+ \sum_{k=1}^{n-1}\sum_{t=k+1}^{n}\mu(1, t-k+1) - n^2\mu\mu(1)$$

$$= \alpha(n\alpha + n - 1)(\kappa_3 - \kappa_2) + (2n-1)\kappa_2(\alpha + \mu(1+\alpha))$$

$$+ (\kappa_3 - \kappa_2)\frac{\alpha^3}{1-\alpha^2}g_n(2,2)$$

$$+ \kappa_2\frac{\alpha}{1-\alpha}g_n(2,1)(\mu + \alpha(1+\mu))$$

$$+ \frac{\alpha}{1-\alpha}g_n(1,1)\{\alpha^2(\kappa_3 - \kappa_2) + \alpha\kappa_2(1+\mu) + \mu\kappa_2\},$$

and

$$\mathrm{cov}(Z_n, W_n) = nE(X_{t-1}X_t^3) + (n-1)E(X_{t-1}^3 X_t) + \sum_{t=1}^{n-2}\sum_{k=t+2}^{n} E(X_t^2 X_{k-1} X_k)$$

$$+ \sum_{k=1}^{n-1}\sum_{t=k+1}^{n} E(X_{k-1}X_k X_t^2) - n^2 E(X_t^2)E(X_{t-1}X_t)$$

$$= n\mu(1,1,1) + (n-1)\mu(0,0,1) + \sum_{t=1}^{n-2}\sum_{k=t+2}^{n}\mu(0, k-t-1, k-t)$$

$$+ \sum_{k=1}^{n-1}\sum_{t=k+1}^{n}\mu(1, t-k+1, t-k+1) - n^2\mu_2\mu(1)$$

$$= \alpha(n\alpha^2 + n - 1)\{\kappa_4 - 3\kappa_3 + \kappa_2(2 - 3\kappa_2)\}$$

$$+ \alpha(n\alpha + n - 1)(\kappa_3 - \kappa_2)(3 + 2\mu)$$

$$+ \mu\{n(\alpha(1+\alpha) + 2) - (1+\alpha)\}(\kappa_3 - \kappa_2)$$

$$+ (2n-1)\kappa_2\{\mu^2 + 2\alpha\kappa_2 + (1+\mu)(\alpha + \mu(1+2\alpha))\}$$

$$+ \frac{\alpha^3}{1-\alpha^2}g_n(2,2)\{\kappa_4 - 3\kappa_3 + \kappa_2(2 - 3\kappa_2)$$

$$+ 2(\kappa_3 - \kappa_2)(1+\mu) + 2\kappa_2^2\}$$

$$+ \frac{\alpha}{1-\alpha}g_n(2,1)\{(\kappa_3 - \kappa_2)(\mu + \alpha(1+\mu))$$

$$+ \alpha\kappa_2(1+\mu)^2 + \kappa_2\mu(1+\mu)(1+\alpha) + \mu^2\kappa_2\}$$

$$+ \frac{\alpha^2}{1 - \alpha^2} g_n(1, 2)\{\alpha^3(\kappa_4 - 3\kappa_3 + \kappa_2(2 - 3\kappa_2))$$
$$+ \alpha^2(\kappa_3 - \kappa_2)(2 + \mu) + 2\alpha\kappa_2^2 + \mu(\kappa_3 - \kappa_2)\}$$
$$+ \frac{\alpha}{1 - \alpha} g_n(1, 1)\{\alpha^2(\kappa_3 - \kappa_2)(1 + 2\mu) + \alpha\kappa_2(1 + \mu)^2$$
$$+ \mu\kappa_2(\mu + (1 + \mu)(1 + \alpha))\},$$

where $\mu = E(X_n)$, $\kappa_j = E((X_n - \mu)^j)$ and $\mu_j = E(X_n^j)$, for $j = 2, 3, 4$.

Let $\Sigma_n$ be the covariance matrix with the above terms. Hence, it follows that the covariance matrix $\Sigma$ of the Proposition 4 can be obtained by $\Sigma = \lim_{n \to \infty} \Sigma_n / n$.

# References

Abraham B, Balakrishna N (2002) Inverse gaussian autoregressive models. J Time Ser Anal 20:605–618

Alamatsaz MH (1983) Completeness and self-decomposability of mixtures. Ann Inst Stat Math 35:355–363

Aly EEAA, Bouzar N (1994) Explicit stationary distributions for some galton-watson processes with immigration. Stoch Models 10:499–517

Al-Osh MA, Alzaid AA (1987) First-order integer valued autoregressive (INAR(1)) process. J Time Ser Anal 8:261–275

Anderson TW (1971) The statistical analysis of time series. Wiley, New York

Andersson J, Karlis D (2014) A parametric time series model with covariates for integers in $\mathbb{Z}$. Stat Model 14:135–156

Barreto-Souza W (2015) Zero-modified geometric INAR(1) process for modelling count time series with deflation or inflation of zeros. J Time Ser Anal 36:839–852

Barreto-Souza W, Bourguignon M (2015) A skew INAR(1) process on $\mathbb{Z}$. Adv Stat Anal 99:189–208

Bisaglia L, Canale A (2016) Bayesian nonparametric forecasting for INAR models. Comput Stat Data Anal 100:70–78

Forst G (1979) A characterization of self-decomposable probabilities in the half-line. Zeit Wahrscheinlichkeitsth 49:349–352

Freeland RK, McCabe BPM (2004a) Analysis of low count time series data by Poisson autoregression. J Time Ser Anal 25:701–722

Freeland RK, McCabe BPM (2004b) Forecasting discrete valued low count time series. Int J Forecast 20:427–434

Freeland RK, McCabe BPM (2005) Asymptotic properties of CLS estimators in the Poisson AR(1) model. Stat Prob Lett 73:147–153

Hamilton JD (1994) Time series analysis. Princeton University Press, Princeton

Harvey AC, Fernandes C (1989) Time series models for count or qualitative observations. J Bus Econ Stat 7:407–417

Jazi MA, Jones G, Lai CD (2012) First-order integer valued processes with zero inflated poisson innovations. J Time Ser Anal 33:954–963

Jung RC, Tremayne AR (2011) Useful models for time series of counts or simply wrong ones? Adv Stat Anal 95:59–91

Karlis D, Xekalaki E (2005) Mixed Poisson distributions. Int Stat Rev 73:35–58

Karlsen H, Tjostheim D (1988) Consistent estimates for the NEAR(2) and NLAR(2) time series models. J R Stat Soc Ser B 50:313–320

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

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

McKenzie E (1988) Some ARMA models for dependent sequences of Poisson counts. Adv Appl Probab 20:822–835

McKenzie E (2003) Discrete variate time series. In: Rao CR, Shanbhag DN (eds) Handbook of statistics. Elsevier, Amsterdam, pp 573–606

Meintanis SG, Karlis D (2014) Validation tests for the innovation distribution in INAR time series models. Comput Stat 29:1221–1241

Nastić AS, Ristić MM (2012) Some geometric mixed integer-valued autoregressive (INAR) models. Stat Probab Lett 82:805–811

Nastić AS, Ristić MM, Djordjević MS (2016a) An INAR model with discrete Laplace marginal distributions. Braz J Probab Stat 30:107–126

Nastić AS, Laketa PN, Ristić MM (2016b) Random environment integer-valued autoregressive process. J Time Ser Anal 37:267–287

Nastić AS, Ristić MM, Janjić AD (2016c) A mixed thinning based geometric INAR(1) model. Filomat

Pillai RN, Satheesh S (1992) $\alpha$-inverse Gaussian distributions. Sankhya A 54:288–290

Ridout MS (2009) Generating random numbers from a distribution specified by its Laplace transform. Stat Comput 19:439–450

Ristić MM, Nastić AS, Ilić AVM (2013) A geometric time series model with dependent Bernoulli counting series. J Time Ser Anal 34:466–476

Ristić MM, Bakouch HS, Nastić AS (2009) A new geometric first-order integer-valued autoregressive (NGINAR(1)) process. J Stat Plan Inference 139:2218–2226

Ristić MM, Nastić AS, Bakouch HS (2012) Estimation in an integer-valued autoregressive process with negative binomial marginals (NBINAR(1)). Commun Stat 41:606–618

Schweer S, Weiß CH (2014) Compound Poisson INAR(1) processes: stochastic properties and testing for overdispersion. Comput Stat Data Anal 77:267–284

Scotto MG, Weiß CH, Gouveia S (2015) Thinning-based models in the analysis of integer-valued time series: a review. Stat Model 15:590–618

Steutel FW, van Harn K (1979) Discrete analogues of self-decomposability and stability. Ann Probab 7:893–899

Weiß CH (2008a) Thinning operations for modeling time series of counts-a survey. Adv Stat Anal 92:319–341

Weiß CH (2008b) Serial dependence and regression of Poisson INARMA models. J Stat Plan Inference 138:2975–2990

Weiß CH (2009) Controlling jumps in correlated processes of Poisson counts. Appl Stoch Models Bus Ind 25:551–564

Weiß CH (2013) Integer-valued autoregressive models for counts showing underdispersion. J Appl Stat 40:1931–1948

Weiß CH (2015) A Poisson INAR(1) model with serially dependent innovations. Metrika 78:829–851

Weiß CH, Homburg A, Puig P (2016) Testing for zero inflation and overdispersion in INAR(1) models. Stat Pap

Weiß CH, Kim HY (2013) Binomial AR(1) processes: moments, cumulants, and estimation. Statistics 47:494–510

Yang K, Wang D, Jia B, Li H (2016) An integer-valued threshold autoregressive process based on negative binomial thinning. Stat Pap