Filomat 31:13 (2017), 4009–4022
https://doi.org/10.2298/FIL1713009N

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

A Mixed Thinning Based Geometric INAR(1) Model

Aleksandar S. Nastić, Miroslav M. Ristić, Ana D. Janjić

University of Niš, Faculty of Sciences and Mathematics, Serbia

Abstract. In this article a geometrically distributed integer-valued autoregressive model of order one based on the mixed thinning operator is introduced. This new thinning operator is defined as a probability mixture of two well known thinning operators, binomial and negative binomial thinning. Some model properties are discussed. Method of moments and the conditional least squares are considered as possible approaches in model parameter estimation. Asymptotic characterization of the obtained parameter estimators is presented. The adequacy of the introduced model is verified by its application on a certain kind of real-life counting data, while its performance is evaluated by comparison with two other INAR(1) models that can be also used over the observed data.

1. Introduction

In the last few decades, there is a significant increase of interest in the integer-valued time series modeling. Some of the involved research was theoretically motivated, but most of it found its inspiration in real-life situations. After some results of using Markov chains by Cox and Miller [8] and the discrete ARMA models by Jacobs and Lewis [10–12], momentous results in the discrete valued time series analysis were achieved, independently of each other, by Al-Osh and Alzaid [2] and McKenzie [15]. Very soon, some of the first generalizations have been done in respect of marginal distribution and they can be found in [16], [1], [5] and [6]. However, they all used a binomial thinning operator based on the Bernoulli counting series in order to describe the autoregressive character of the dependance between the time series random variables. These models were highly adequate for modeling the periodical counting of random events or some population elements which could enter into or either survive or disappear from the observed system. Afterwards, responding to more demanding modeling requirements, there have been many modifications and generalizations of the thinning operator, such as in [3], [14] and [21, 22]. Also, Ristić et al. [17] introduced a new integer-valued autoregressive process based on the geometric counting sequence.

While, as mentioned above, binomial thinning based INAR models were ideal for counting the number of units which could enter or leave the observing area from one time interval to another, the negative binomial thinning was used for systems involving much more activity. Namely, it's defined using geometrically distributed random variables, so it might be applied for modeling elements or random events which could, by replicating themselves, affecting other elements, provoking other random events or in some other way,

---

2010 *Mathematics Subject Classification*. Primary 62M10
Keywords. Mixed thinning operator, Mixed thinning INAR(1) model, Binomial thinning, Negative binomial thinning, Geometric marginal distribution
Received: 23 February 2016; Accepted: 24 May 2016
Communicated by Dragan S. Djordjević
The first and the second author acknowledge the grant of MNTR 174013 for carrying out this research
Email addresses: anastic78@gmail.com (Aleksandar S. Nastić), miristic72@gmail.com (Miroslav M. Ristić)

contribute to the overall thinning sum by more than 1. In this way, we have obtained two thinning operators, completely different in terms of nature of the data which can be used on. But what happens if you deal with the data related to the counting of elements of variable character, ie. elements which during one period of time might be passive and in this way adding the total sum just 0 or 1, while already in the next counting interval, for some reasons, they might become active, contributing to the count-value in a much greater extent. Precisely for this reason, here we define a completely new thinning operator as a mixture of the binomial and the negative binomial operator. Moreover, this mixture is a probability pondered, thus is very capable of responding to requirements of a certain counting data. Likewise, there is a variety of real-life examples where this kind of the mixed thinning INAR models are more likely to be used than the INAR models based only on one of these two thinning operators.

Now, we provide the motivation for the introduction of a new INAR model with the mixed thinning operator through three very different examples from real life. First of all, the number of patients with certain diseases may be the subject of counting over time. Incubation period is the time elapsed between exposure to a pathogenic organism and a moment when symptoms become apparent. We focus our attention on the diseases of which the person is contagious during incubation period. Assume also that the counting time interval is grater or equal to incubation period. In this case, the probability ponder of negative binomial thinning directly corresponds to the ratio of the counting period and the disease incubation period. Namely, if the time of the incubation is significant or even almost equal to the observing interval, there are very high chances that an infected person with no symptoms will transmit the disease to someone else. However, as this ratio is smaller, the chances of transmitting the disease are less and the probability of using the binomial thinning is higher. On the other hand, if the patients are quickly quarantined, then the chances of transmission are very small. Therefore, the application of the Bernoulli distributed counting sequence becomes more frequent, affecting in the proper way the values of the mixture probability weights. Another application of the mixed thinning INAR models may be found in counting the number of bacteria. Bacteria are prokaryotic organisms that reproduce asexually. Bacterial reproduction most commonly occurs by a kind of cell division called binary fission. Under optimal conditions (temperature, humidity, sources of nutrients) bacteria can grow and divide extremely rapidly, so greater probability weight, corresponding to a geometric counting sequence, would be a proper choice. On the contrary, if the temperature is extremely law or the bacteria are exposed to certain antibiotics, then they will not reproduce, even more they will only be able to survive or die over time, in which case the higher binomial thinning probability will make the INAR model more appropriate. Finally, completely different from the previous applications, the mixed thinning model may be used in the social anomalies research, more precisely in the counting of crimes committed in a certain police district over time. For example, if you track the number of drug dealing in a specific urban area you can encounter with the real dealers who are able to commit a greater number of offenses, or only with "beginners" or addicts who commit during the counting period one or no crime. Here, as in the previous cases, it is very easy to see the opportunity for appropriate implementation of this new model.

The outline of the paper is as follows. In the next section a new mixed thinning operator is introduced and discussed. The third section is devoted to the introduction of the INAR model with geometric marginals generated by the new thinning operator. The model features are discussed including a k-step ahead conditional expectation and variance. Section 4 considers some procedures for estimating the unknown model parameters obtaining their asymptotic characterization. In the last section a possible application of the introduced model is discussed on some real-life data.

2. The Construction of the Thinning Operator

In this section we introduce a new thinning operator in the following way. Let $\{W_i\}_{i\in\mathbb{N}}$ be a sequence of the independent and identically distributed random variables given as

$$W_i = \begin{cases} B_i, & \text{with probability } p, \\ G_i, & \text{with probability } 1-p, \end{cases} \quad p \in [0,1], \tag{1}$$

where B_i and G_i represent random variables with Bernoulli(α) distribution and geometric$\left(\frac{\alpha}{1+\alpha}\right)$ distribution with probability mass function given by $P(G_i = x) = \frac{\alpha^x}{(1+\alpha)^{x+1}}$, $x = 0, 1, \ldots$, respectively, $\alpha \in [0, 1]$. We suppose that B_i and G_i are independent random variables. From (1) we can see that the distribution of the random variable W_i is a mixture of Bernoulli and geometric distribution. For $p = 0$ we obtain the geometric distributed random variable and for $p = 1$ we obtain the Bernoulli distributed random variable.

Before we are introducing a new thinning operator $\alpha \bullet_p$, we derive the probability generating function of the random variable W_i given by (1). Using the fact that the random variables B_i and G_i have Bernoulli and geometric distribution, respectively, we obtain that

$$\Phi_W(s) \equiv E\left(s^{W_i}\right) = p(1 - \alpha + \alpha s) + \frac{1-p}{1+\alpha-\alpha s} = \frac{1-\alpha^2 p(1-s)^2}{1+\alpha(1-s)}. \tag{2}$$

Now we are introducing the new thinning operator in the following way.

Definition 2.1. *An operator $\alpha \bullet_p$ defined as*

$$\alpha \bullet_p X = \sum_{i=0}^{X} W_i, \tag{3}$$

where $\alpha \in [0, 1]$ and X is a non-negative integer-valued random variable independent of the counting series $\{W_i\}_{i\in\mathbb{N}}$ given by (1) *is called a mixed thinning operator.*

The introduced thinning operator $\alpha \bullet_p$ is a mixture of the binomial and negative binomial thinning operators. This property is given in the following theorem.

Theorem 2.2. *Let X be a non-negative integer-valued random variable. Then for a given $X = x$, $x \in \mathbb{N}$, the distribution of the mixed thinning operator $\alpha \bullet_p$ given by* (3) *is a mixture of the binomial, negative binomial distribution and their convolution as*

$$\alpha \bullet_p X | \{X = x\} \stackrel{d}{=} \begin{cases} NB\left(x, \frac{\alpha}{1+\alpha}\right), & \text{with probability } (1-p)^x, \\ Bin(i,\alpha) + NB\left(x-i, \frac{\alpha}{1+\alpha}\right), & \text{with probability } \binom{x}{i} p^i (1-p)^{x-i},\ 1 \le i \le x-1, \\ Bin(x,\alpha), & \text{with probability } p^x, \end{cases}$$

where $NB\left(i, \frac{\alpha}{1+\alpha}\right)$, $i \ge 1$, represents a random variable with the negative binomial distribution with parameters i and $\alpha/(1+\alpha)$ and $Bin(i, \alpha)$, $i \ge 1$, represents a random variable with binomial distribution with parameters i and α.

Proof. To prove the theorem we need to derive the conditional probability generating function of the random variable $\alpha \bullet_p X$ for given $X = x$. Using the definition of the random variable W_i given by (1) and the independency of the random variables $\{W_i\}$, we obtain that

$$\begin{aligned} \Phi_{\alpha\bullet_p X|\{X=x\}}(s) &\equiv E\left(s^{\alpha\bullet_p X}\Big| X = x\right) = (\Phi_W(s))^x = \left(p(1-\alpha+\alpha s) + \frac{1-p}{1+\alpha-\alpha s}\right)^x \\ &= \sum_{i=0}^{x} \binom{x}{i} p^i (1-p)^{x-i} \frac{(1-\alpha+\alpha s)^i}{(1+\alpha-\alpha s)^{x-i}}. \end{aligned} \tag{4}$$

Finally, the proof follows from the fact that $(1-\alpha+\alpha s)^i$ represents the probability generating function of the random variable with Bin(i, α) distribution and $(1+\alpha-\alpha s)^{-(x-i)}$ represents the probability generating function of the random variable with NB$(x-i, \alpha/(1+\alpha))$ distribution. □

Corollary 2.3. *Let J be a random variable with the binomial distribution with parameters $x \in \mathbb{N}_0$ and $p \in [0, 1]$. Let $Bin(i, \alpha)$, $i \in \mathbb{N}_0$, $\alpha \in [0, 1]$, and $NB\left(k, \frac{\alpha}{1+\alpha}\right)$, $k \in \mathbb{N}_0$, $\alpha > 0$, be random variables with the binomial and negative* *binomial distribution, respectively, where* $Bin(0,\alpha) = 0$ *and* $NB\Big(0, \frac{\alpha}{1+\alpha}\Big) = 0$*. Let us suppose that* J, $Bin(i,\alpha)$ *and* $NB\Big(k, \frac{\alpha}{1+\alpha}\Big)$ *are independent random variables. Then*

$$\alpha \bullet_p X|\{X = x\} \stackrel{d}{=} Bin(J,\alpha) + NB\left(x - J, \frac{\alpha}{1+\alpha}\right) \stackrel{d}{=} \alpha \circ J + \alpha * (x - J).$$

Obviously, $\alpha \bullet_p X|\{X = 0\}$=0.

Some properties of the thinning operator $\alpha\bullet_p$ can be easily obtained from the conditional probability generating function given by (4). The first two derivatives of the probability generating function $\Phi_{\alpha\bullet_p X|\{X=x\}}(s)$ with respect to $s = 1$ are given respectively as $\Phi'_{\alpha\bullet_p X|\{X=x\}}(1) = \alpha x$ and $\Phi''_{\alpha\bullet_p X|\{X=x\}}(1) = \alpha^2 x(1 + x - 2p)$, which implies that the conditional expectation and variance of the random variable $\alpha \bullet_p X$ with respect to X are $E(\alpha\bullet_p X|X) = \alpha X$ and $Var(\alpha\bullet_p X|X) = \alpha(1-2\alpha p+\alpha)X$, respectively. These results yields that the unconditional expectation and variance of the random variable $\alpha \bullet_p X$ are

$$E(\alpha \bullet_p X) = \alpha E(X) \quad \text{and} \quad Var(\alpha \bullet_p X) = \alpha^2 Var(X) + \alpha(1 + \alpha - 2\alpha p)E(X),$$

respectively.

3. Stationary Time Series Model with the Geometric Marginals

In this section we construct a stationary time series model with the geometric marginals in the following way.

Definition 3.1. *The stationary sequence of random variables* $\{X_t\}_{t\in\mathbb{Z}}$ *is called a mixed thinning integer-valued autoregressive process of order 1 (*$MTINAR(1)$*) if it satisfies the equation*

$$X_t = \alpha \bullet_p X_{t-1} + \varepsilon_t, \tag{5}$$

where $\alpha\bullet_p$ *is a mixed thinning operator given by Definition* 2.1*,* $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ *is a sequence of i.i.d. random variables independent of the counting series* $\{W_i\}_{i\in\mathbb{N}}$ *and the random variables* X_{t-i} *and* ε_t *are independent for all* $i \geq 1$*.*

The time series model (5) contains two existing models as special cases: the time series model with geometric marginals generated by the binomial thinning operator for $p = 1$ and the NGINAR(1) time series model of Ristić et al. [17] with geometric marginals generated by the negative binomial thinning operator for $p = 0$.

First, we will consider the distribution of the random variable ε_t. It is given by the following theorem.

Theorem 3.2. *Let* $\{X_t\}_{t\in\mathbb{Z}}$ *be a stationary* $MTINAR(1)$ *time series model with geometric*$(\mu/(1+\mu))$ *marginals which satisfies the condition* $\mu \geq \alpha(1-\alpha p)/(1-\alpha)$*. Then the random variable* ε_t *has the distribution of the following form*

$$\varepsilon_t \stackrel{d}{=} \begin{cases} 0, & \text{with probability } \alpha p, \\ Geom\left(\frac{\alpha}{1+\alpha}\right), & \text{with probability } \frac{\alpha\mu(1-p)}{\mu-\alpha}, \\ Geom\left(\frac{\mu}{1+\mu}\right), & \text{with probability } \frac{\mu-\alpha(1+\mu-\alpha p)}{\mu-\alpha}. \end{cases}$$

Proof. Since the random variables $\{W_i\}$ are independent, we obtain from (5) and (2) that the probability generating function of the random variable ε_t is given by

$$\Phi_\varepsilon(s) = \frac{\Phi_X(s)}{\Phi_X(\Phi_W(s))} = \frac{1 + \alpha(1+\mu) + \alpha^2 p\mu - \alpha(1 + \mu + 2\alpha p\mu)s + \alpha^2 p\mu s^2}{(1 + \mu - \mu s)(1 + \alpha - \alpha s)}.$$

The last equation yields that the probability generating function of the random variable ε_t can be represented in the form

$$\Phi_\varepsilon(s) = \alpha p + \frac{\alpha\mu(1-p)}{\mu-\alpha} \cdot \frac{1}{1+\alpha-\alpha s} + \frac{\mu - \alpha(1+\mu-\alpha p)}{\mu-\alpha} \cdot \frac{1}{1+\mu-\mu s}.$$

Under the condition $\mu \geq \frac{\alpha(1-\alpha p)}{1-\alpha}$ we have that a distribution of the random variable ε_t is well-defined and that can be represented as a mixture of the following form

$$\varepsilon_t \stackrel{d}{=} \begin{cases} 0, & \text{with probability } \alpha p, \\ Geom\left(\frac{\alpha}{1+\alpha}\right), & \text{with probability } \frac{\alpha\mu(1-p)}{\mu-\alpha}, \\ Geom\left(\frac{\mu}{1+\mu}\right), & \text{with probability } \frac{\mu-\alpha(1+\mu-\alpha p)}{\mu-\alpha}. \end{cases}$$

□

Definition 3.3. *A time series $\{X_t\}_{t\in\mathbb{Z}}$ considered in Theorem* 3.2 *is called a mixed thinning geometrically distributed INAR*(1) *model* (*MTGINAR*(1)).

Corollary 3.4. *The mean and variance of the random variable ε_t are given respectively as $E(\varepsilon_t) = (1-\alpha)\mu$ and $Var(\varepsilon_t) = \mu(1-\alpha-2\alpha^2+\mu-\alpha^2\mu+2\alpha^2 p)$.*

Also we are in position to derive the transition probabilities of the process. Let j and x are some non-negative integer values, then

$$\begin{aligned} P(X_t = j|X_{t-1} = x) &= P(\alpha \bullet_p X_{t-1} + \varepsilon_t = j|X_{t-1} = x) = P\left(\sum_{i=0}^{x} W_i + \varepsilon_t = j\right) \\ &= \sum_{k=0}^{min\{j,x\}} P\left(\sum_{i=0}^{x} W_i = k\right) P\left(\varepsilon_t = j-k\right), \end{aligned} \tag{6}$$

where

$$\begin{aligned} P\left(\sum_{i=0}^{x} W_i = k\right) &= (1-p)^x P\left(NB\left(x, \frac{\alpha}{1+\alpha}\right) = k\right) + p^x P\left(Bin(x,\alpha) = k\right) \\ &+ \sum_{i=1}^{x-1} \binom{x}{i} p^i (1-p)^{x-i} P\left(Bin(i,\alpha) + NB\left(x-i, \frac{\alpha}{1+\alpha}\right) = k\right). \end{aligned} \tag{7}$$

We also have that

$$\begin{aligned} P\left(Bin(i,\alpha) + NB\left(x-i, \frac{\alpha}{1+\alpha}\right) = k\right) &= \sum_{l=0}^{min\{k,i\}} P\left(Bin(i,\alpha) = l\right) P\left(NB\left(x-i, \frac{\alpha}{1+\alpha}\right) = k-l\right) \\ &= \sum_{l=0}^{min\{k,i\}} \binom{i}{l} \binom{x-i+k-l-1}{x-i-1} \frac{\alpha^k(1-\alpha)^{i-l}}{(1+\alpha)^{k-l+x-i}}. \end{aligned}$$

Replacing this result in (7) we obtain the following

$$\begin{aligned} P\left(\sum_{i=0}^{x} W_i = k\right) &= (1-p)^x \binom{x+k-1}{x-1} \frac{\alpha^k}{(1+\alpha)^{k+x}} + p^x \binom{x}{k} \alpha^k (1-\alpha)^{x-k} \\ &+ \sum_{i=1}^{x-1} \binom{x}{i} \frac{p^i(1-p)^{x-i}\alpha^k(1-\alpha)^i}{(1+\alpha)^{k+x-i}} \sum_{l=0}^{\min\{k,i\}} \binom{i}{l} \binom{x-i+k-l-1}{x-i-1} \left(\frac{1+\alpha}{1-\alpha}\right)^l. \end{aligned} \tag{8}$$

Finally, replacing (8) in (6), we are able to calculate the process transition probability values.

Now, we derive and discuss some properties of the introduced model. From the definition of the model (5), the conditional mean of the random $\alpha \bullet_p X_t$ for given X_t and the result from Corollary 3.4, we obtain that the conditional mean of the random variable X_{t+1} for given X_t is given by $E(X_{t+1}|X_t) = \alpha X_t + (1-\alpha)\mu$, which means that our model has a first order conditional linear autoregressive structure, as it was discussed in [9]. This implies that the k-step ahead conditional mean is

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

If $k \to \infty$, then the k-step ahead conditional mean converges to the unconditional mean $E(X_t) = \mu$.

The k-step ahead conditional variance is given by

$$\begin{aligned} Var(X_{t+k}|X_t) &= \alpha^2 Var(X_{t+k-1}|X_t) + \alpha(1+\alpha-2\alpha p)(\alpha^{k-1}X_t + \mu(1-\alpha^{k-1})) + Var(\varepsilon_t) \\ &= \frac{\alpha^k(1-\alpha^k)(1+\alpha-2\alpha p)}{1-\alpha} X_t + \alpha(1+\alpha-2\alpha p)\mu\left(\frac{1-\alpha^{2k}}{1-\alpha^2} - \frac{\alpha^{k-1}(1-\alpha^k)}{1-\alpha}\right) + \frac{1-\alpha^{2k}}{1-\alpha^2} Var(\varepsilon_t). \end{aligned}$$

Letting $k \to \infty$, we obtain that the k-step ahead conditional variance converges to unconditional variance $Var(X_t) = \mu(1+\mu)$. Also, we can see that the k-step ahead conditional variance linearly depends on X_t.

Since the model $\{X_t\}_{t\in\mathbb{Z}}$ has the first order conditional linear autoregressive structure, it follows that the autocorrelation function of the process $\{X_t\}_{t\in\mathbb{Z}}$ is given by $Corr(X_{t+k}, X_t) = \alpha^{|k|}$, $k \in \mathbb{Z}$. Thus, the autocorrelation function is positive and decays exponentially, as k increases.

Finally, $\{X_t\}$ is a Markovian process of the first order with identical marginal distribution. So, following the proof of Theorem 2, in [19], its strict stationarity and ergodicity may be obtained.

4. Parameter Estimation

4.1. Conditional least squares

Let us consider the conditional least squares estimation of the unknown parameters α, μ and p. Since the conditional expectation $E(X_t|X_{t-1}) = \alpha X_{t-1} + (1-\alpha)\mu$ only depends on the first two parameters α and μ, we will use the two-step conditional least squares approach considered by Karlsen and Tjøstheim [13]. Thus, in the first step we will estimate the unknown parameters α and μ, and in the second step we will estimate the unknown parameter p by using the conditional least squares estimators of the parameters α and μ obtained in the first step.

Let us provide more details for each step of the estimation. In the first step we minimize the function $S_1(\alpha,\mu)$ given by

$$S_1(\alpha,\mu) = \sum_{t=2}^{n}[X_t - E(X_t|X_{t-1})]^2 = \sum_{t=2}^{n}[X_t - \alpha X_{t-1} - (1-\alpha)\mu]^2.$$

Then, minimizing the function $S_1(\alpha,\mu)$ with respect to the unknown parameters α and μ, we obtain the conditional least squares estimators of these parameters, given by

$$\begin{aligned} \hat{\alpha}_{cls} &= \frac{(n-1)^{-1}\sum_{t=2}^{n} X_t X_{t-1} - (n-1)^{-2}\sum_{t=2}^{n} X_t \sum_{t=2}^{n} X_{t-1}}{(n-1)^{-1}\sum_{t=2}^{n} X_{t-1}^2 - (n-1)^{-2}\left(\sum_{t=2}^{n} X_{t-1}\right)^2} \\ \hat{\mu}_{cls} &= \frac{\sum_{t=2}^{n} X_t - \hat{\alpha}_{cls}\sum_{t=2}^{n} X_{t-1}}{(n-1)(1-\hat{\alpha}_{cls})}. \end{aligned}$$

The derived conditional least squares estimators of the unknown parameters α and μ are strongly consistent as an immediate consequence of the ergodic theorem, since our novel time series model is strictly stationary and ergodic. The asymptotic distribution of the estimators is established in the following theorem.

Theorem 4.1. *The estimator $(n-1)^{\frac{1}{2}}(\hat{\alpha}_{cls}-\alpha, \hat{\mu}_{cls}-\mu)^T$ has asymptotic bivariate normal distribution with mean $(0,0)^T$ and covariance matrix*

$$D = \begin{bmatrix} \frac{\alpha(1+\alpha-2\alpha p)(1+4\mu+3\mu^2)}{\mu(1+\mu)^2} - \frac{z}{1+\mu} & \frac{\alpha(1+\alpha-2\alpha p)}{(1-\alpha)} \\ \frac{\alpha(1+\alpha-2\alpha p)}{(1-\alpha)^2} & \frac{\mu\left(\alpha(1-\alpha-2\alpha p)-z\right)}{(1-\alpha)^2} \end{bmatrix}$$

where $z = 1-\alpha+\mu+\alpha^2(2(p-1)-\mu)$.

Proof. First, note that we have introduced z as above for the sake of simplicity. In order to prove the stated theorem, we will check whether the conditions of Theorem 3.1. and Theorem 3.2. of [20] are satisfied. Denote $\mathbf{a} \equiv (a_1, a_2) = (\alpha, \mu)$ and $g(\mathbf{a}, X_{t-1}) \equiv E(X_t|X_{t-1}) = \alpha X_{t-1} + (1-\alpha)\mu$. It immediately follows that

$$\begin{aligned} E\left|\frac{\partial g}{\partial \alpha}\right|^2 &= E(X_{t-1}-\mu)^2 = Var(X_{t-1}) = \mu(1+\mu) < \infty, \\ E\left|\frac{\partial g}{\partial \mu}\right|^2 &= E(1-\alpha)^2 = (1-\alpha)^2 < \infty \end{aligned}$$

and obviously $E\left|\frac{\partial^2 g}{\partial\alpha\partial\mu}\right|^2 < \infty$, so the condition $C1$ of Theorem 3.1. is confirmed. Similarly the condition $C3$ of the same theorem holds. Moreover, let c_1 and c_2 be the arbitrary real numbers, such that the following equation holds

$$E\left|c_1\frac{\partial g}{\partial \alpha} + c_2\frac{\partial g}{\partial \mu}\right|^2 = 0.$$

Now, $E\left(c_1(X_{t-1}-\mu) + c_2(1-\alpha)\right)^2 = c_1^2 Var(X_{t-1}) + c_2^2(1-\alpha)^2 = 0$ and it is obvious that the last equation implies $c_1 = c_2 = 0$. Thus, the condition $C2$ is fulfilled and now all the conditions of Theorem 3.1. are checked. Note that we have in this way once more established strict stationarity of the estimators. Next, we will check whether the condition $D1$ of Theorem 3.2. holds and for this purpose let's introduce

$$f_{t|t-1} \equiv Var(X_t|X_{t-1}) = \alpha(1+\alpha-2\alpha^2 p)X_{t-1} - \mu z,$$

where z is defined above in the statement of the theorem. Using simple calculations, we get

$$R \equiv E\left(\frac{\partial g}{\partial \mathbf{a}} f_{t|t-1} \frac{\partial g^T}{\partial \mathbf{a}}\right) = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix},$$

where

$$\begin{aligned} r_{11} &= \alpha(1+\alpha-2\alpha p)\mu(1+4\mu+3\mu^2) - \mu^2(1+\mu)z \\ r_{12} &= r_{21} = \alpha(1-\alpha)\mu(1+\mu)(1+\alpha-2\alpha p) \\ r_{22} &= \mu(1-\alpha)^2\left(\alpha(1+\alpha-2\alpha p) - z\right). \end{aligned}$$

It is easy to conclude that $R < \infty$, which implies the condition $D1$. Furthermore, it can be easily derived that

$$U = \left[E\left(\frac{\partial g}{\partial a_i}\frac{\partial g}{\partial a_j}\right)\right]_{2\times 2} = \begin{bmatrix} \mu(1+\mu) & 0 \\ 0 & (1-\alpha)^2 \end{bmatrix}.$$

Finally, from Theorem 3.2. immediately follows that the estimator $(n-1)^{\frac{1}{2}}\left(\hat{\alpha}_{cls}-\alpha, \hat{\mu}_{cls}-\mu\right)^T$ has asymptotic bivariate normal distribution with mean $(0,0)^T$ and covariance matrix $U^{-1}RU^{-1}$. After some calculations we obtain $D \equiv U^{-1}RU^{-1}$ and the proof is now complete. □

In the second step we estimate the unknown parameter p in the following way. Let us introduce a random variable $Y_t = [X_t - E(X_t|X_{t-1})]^2 = [X_t - \alpha X_{t-1} - (1-\alpha)\mu]^2$. Now we are minimizing the function $S_2(p)$ defined as

$$\begin{aligned} S_2(p) &= \sum_{t=2}^{n}[Y_t - E(Y_t|X_{t-1})]^2 = \sum_{t=2}^{n}[Y_t - Var(X_t|X_{t-1})]^2 \\ &= \sum_{t=2}^{n}[Y_t - \alpha(1+\alpha-2\alpha p)X_{t-1} - \mu(1-\alpha-2\alpha^2+\mu-\alpha^2\mu+2\alpha^2 p)]^2 \\ &= \sum_{t=2}^{n}[2\alpha^2 p(X_{t-1}-\mu) - Z_t]^2, \end{aligned}$$

where $Z_t = -Y_t + \alpha(1+\alpha)X_{t-1} + \mu(1-\alpha-2\alpha^2+\mu-\alpha^2\mu)$. Now, minimizing the above function $S_2(p)$ with respect to the unknown parameter p we obtain that the conditional least squares estimator of the parameter p is given by

$$\hat{p}_{cls} = \frac{\sum_{t=2}^{n} \hat{Z}_t(X_{t-1} - \hat{\mu}_{cls})}{2\hat{\alpha}_{cls}^2 \sum_{t=2}^{n}(X_{t-1} - \hat{\mu}_{cls})^2},$$

where $\hat{Z}_t = -\hat{Y}_t + \hat{\alpha}_{cls}(1+\hat{\alpha}_{cls})X_{t-1} + \hat{\mu}_{cls}(1-\hat{\alpha}_{cls}-2\hat{\alpha}_{cls}^2+\hat{\mu}_{cls}-\hat{\alpha}_{cls}^2\hat{\mu}_{cls})$ and $\hat{Y}_t = [X_t - \hat{\alpha}_{cls}X_{t-1} - (1-\hat{\alpha}_{cls})\hat{\mu}_{cls}]^2$.

Next, we will discuss the asymptotic normality of the previously obtained conditional least squares estimator of the unknown parameter p. In particular, we will consider two cases, the case when the parameters α and μ are known and the case when they are not. The results are stated within the following theorems.

Theorem 4.2. *Suppose that parameters α and μ are known. The conditional least squares estimator of the unknown parameter p has asymptotic normal distribution.*

Proof. In order to prove the statement of the theorem, based on the derived form of conditional least squares estimate of parameter p, it follows that

$$(n-1)^{\frac{1}{2}}\left(\hat{p}_{cls}(\alpha,\mu) - p\right) = \frac{(n-1)^{-\frac{1}{2}} \sum_{t=2}^{n} \left(Z_t - 2\alpha^2 p(X_{t-1}-\mu)\right)(X_{t-1}-\mu)}{(n-1)^{-1} 2\alpha^2 \sum_{t=2}^{n}(X_{t-1}-\mu)^2}.$$

Now, as stated earlier, $\{X_t, t\in\mathbb{Z}\}$ is a strictly stationary and ergodic sequence, so it follows that $\{(X_{t-1}-\mu)^2, t\in\mathbb{Z}\}$ and $\{\left(Z_t - 2\alpha^2 p(X_{t-1}-\mu)\right)(X_{t-1}-\mu), t\in\mathbb{Z}\}$ are also strictly stationary and ergodic sequences as continuous transformations of $\{X_t, t\in\mathbb{Z}\}$. Next, involving ergodic theorem, we conclude that $(n-1)^{-1}2\alpha^2\sum_{t=2}^{n}(X_{t-1}-\mu)^2$ is convergent with probability 1, which implies the convergence in probability, too. Further, denote as F_{t-1} a σ-field generated with random variables $X_{t-i}, i \geq 1$. After some calculations it can be easily shown that $E(Z_t|F_{t-1}) = 2\alpha^2 p(X_{t-1}-\mu)$, which implies that $E\left((Z_t - 2\alpha^2 p(X_{t-1}-\mu))(X_{t-1}-\mu)|F_{t-1}\right) = 0$. Thus, we conclude that $\{\left(Z_t - 2\alpha^2 p(X_{t-1}-\mu)\right)(X_{t-1}-\mu), t\in\mathbb{Z}\}$ is an ergodic, strictly stationary martingale difference sequence. An application of central limit theorem for ergodic, strictly stationary martingale difference sequences provides the convergence in distribution of $(n-1)^{-\frac{1}{2}}\sum_{t=2}^{n}\left(Z_t - 2\alpha^2 p(X_{t-1}-\mu)\right)(X_{t-1}-\mu)$ to a normal distribution with mean 0 and variance $E\left(\left(Z_t - 2\alpha^2 p(X_{t-1}-\mu)\right)(X_{t-1}-\mu)\right)^2$. Finally, using Slutsky's theorem, we are able to conclude the asymptotic normality of $\hat{p}_{cls}(\alpha,\mu)$. $\square$

Theorem 4.3. *Suppose that parameters α and μ are unknown. The conditional least squares estimator of the unknown parameter p has asymptotic normal distribution.*

Proof. We will consider the following decomposition

$$\begin{aligned}(n-1)^{\frac{1}{2}}\left(\hat{p}_{cls}(\hat{\alpha}_{cls},\hat{\mu}_{cls}) - p\right) &= (n-1)^{\frac{1}{2}}\left(\hat{p}_{cls}(\hat{\alpha}_{cls},\hat{\mu}_{cls}) - \hat{p}_{cls}(\alpha,\hat{\mu}_{cls})\right) + (n-1)^{\frac{1}{2}}\left(\hat{p}_{cls}(\alpha,\hat{\mu}_{cls}) - \hat{p}_{cls}(\alpha,\mu)\right) \\ &\quad + (n-1)^{\frac{1}{2}}\left(\hat{p}_{cls}(\alpha,\mu) - p\right). \end{aligned} \tag{9}$$

Notice that within the previous theorem we have already proved the asymptotic normality of the third summand of (9), i.e. $(n-1)^{\frac{1}{2}}\left(\hat{p}_{cls}(\alpha,\mu) - p\right)$. In the following we shall prove the same result for the first and the second summand.

First, we consider $(n-1)^{\frac{1}{2}}\left(\hat{p}_{cls}(\hat{\alpha}_{cls},\hat{\mu}_{cls}) - \hat{p}_{cls}(\alpha,\hat{\mu}_{cls})\right)$. In this case, we have

$$\hat{p}_{cls}(\hat{\alpha}_{cls},\hat{\mu}_{cls}) - \hat{p}_{cls}(\alpha,\hat{\mu}_{cls}) = \frac{\sum_{t=2}^{n}(X_{t-1}-\hat{\mu}_{cls})\left(\alpha^2\hat{Z}_t(\hat{\alpha}_{cls},\hat{\mu}_{cls}) - \hat{\alpha}_{cls}^2\hat{Z}_t(\alpha,\hat{\mu}_{cls})\right)}{2\alpha^2\hat{\alpha}_{cls}^2\sum_{t=2}^{n}(X_{t-1}-\hat{\mu}_{cls})^2}.$$

Having in mind an earlier introduced form of Z_t, after some calculations we obtain

$$\begin{aligned}\alpha^2\hat{Z}_t(\hat{\alpha}_{cls},\hat{\mu}_{cls})-\hat{\alpha}_{cls}^2\hat{Z}_t(\alpha,\hat{\mu}_{cls}) &= (\hat{\alpha}_{cls}-\alpha)\Big((\hat{\alpha}_{cls}+\alpha)X_t^2-2\alpha\hat{\alpha}_{cls}X_tX_{t-1}-2\hat{\mu}_{cls}X_t(\hat{\alpha}_{cls}+\alpha-\alpha\hat{\alpha}_{cls})\\ &\quad +\alpha\hat{\alpha}_{cls}X_{t-1}(2\mu-1)-\hat{\mu}_{cls}(\hat{\alpha}_{cls}+\alpha-\alpha\hat{\alpha}_{cls}+2\alpha\hat{\alpha}_{cls}\hat{\mu}_{cls})\Big)\\ &= (\hat{\alpha}_{cls}-\alpha)V_t,\end{aligned}$$

where V_t is just a shortened notation for a corresponding expression. So, we have the following result

$$(n-1)^{\frac{1}{2}}\left(\hat{p}_{cls}(\hat{\alpha}_{cls},\hat{\mu}_{cls})-\hat{p}_{cls}(\alpha,\hat{\mu}_{cls})\right)=\frac{(n-1)^{\frac{1}{2}}(\hat{\alpha}_{cls}-\alpha)\sum_{t=2}^{n}(X_{t-1}-\hat{\mu}_{cls})V_t}{2\alpha^2\hat{\alpha}_{cls}^2\sum_{t=2}^{n}\left(X_{t-1}-\hat{\mu}_{cls}\right)^2}.$$

Further reasoning is very similar to the one used in the previous theorem. Thus, as $\{X_t, t\in\mathbb{Z}\}$ is a strictly stationary and ergodic sequence, it follows that $\{(X_{t-1}-\hat{\mu}_{cls})V_t, t\in\mathbb{Z}\}$ and $\{(X_{t-1}-\hat{\mu}_{cls})^2, t\in\mathbb{Z}\}$ are also strictly stationary and ergodic sequences as continuous transformations of $\{X_t, t\in\mathbb{Z}\}$. Moreover, based on ergodic theorem, we may conclude that the sums $(n-1)^{-1}\sum_{t=2}^{n}(X_{t-1}-\hat{\mu}_{cls})V_t$ and $(n-1)^{-1}\sum_{t=2}^{n}(X_{t-1}-\hat{\mu}_{cls})^2$ are convergent with probability 1. At the end, having in mind that $(n-1)^{\frac{1}{2}}(\hat{\alpha}_{cls}-\alpha)$ has asymptotic normal distribution with mean 0 as a result of Theorem 4.1 and making use of Slutsky's theorem again, we conclude that $(n-1)^{\frac{1}{2}}\left(\hat{p}_{cls}(\hat{\alpha}_{cls},\hat{\mu}_{cls})-\hat{p}_{cls}(\alpha,\hat{\mu}_{cls})\right)$ has also asymptotic normal distribution with mean 0.

Now, we take in consideration $(n-1)^{\frac{1}{2}}\left(\hat{p}_{cls}(\alpha_{cls},\hat{\mu}_{cls})-\hat{p}_{cls}(\alpha,\mu)\right)$ and establish its asymptotic normality. Based on the derived form of $\hat{p}_{cls}$, we obtain

$$\begin{aligned}\hat{p}_{cls}(\alpha_{cls},\hat{\mu}_{cls})-\hat{p}_{cls}(\alpha,\mu) &= \frac{\sum_{t=2}^{n}(X_{t-1}-\mu)^2\sum_{t=2}^{n}\hat{Z}_t(\alpha,\hat{\mu}_{cls})(X_{t-1}-\hat{\mu}_{cls})}{2\alpha^2\sum_{t=2}^{n}(X_{t-1}-\mu)^2\sum_{t=2}^{n}(X_{t-1}-\hat{\mu}_{cls})^2}\\ &\quad -\frac{\sum_{t=2}^{n}(X_{t-1}-\hat{\mu}_{cls})^2\sum_{t=2}^{n}Z_t(\alpha,\mu)(X_{t-1}-\mu)}{2\alpha^2\sum_{t=2}^{n}(X_{t-1}-\mu)^2\sum_{t=2}^{n}(X_{t-1}-\hat{\mu}_{cls})^2}\\ &= \frac{A}{2\alpha^2\sum_{t=2}^{n}(X_{t-1}-\mu)^2\sum_{t=2}^{n}(X_{t-1}-\hat{\mu}_{cls})^2}.\end{aligned}$$

After some calculations we obtain the following

$$\begin{aligned}A &= (\hat{\mu}_{cls}-\mu)\Bigg[\sum_{t=2}^{n}X_{t-1}^2\left(\sum_{t=2}^{n}X_{t-1}U_1(t)-\sum_{t=2}^{n}U_2(t)\right)-2\sum_{t=2}^{n}X_{t-1}\left(\sum_{t=2}^{n}X_{t-1}U_3(t)-\mu\hat{\mu}_{cls}\sum_{t=2}^{n}U_1(t)\right)\\ &\quad +(n-1)\left(\sum_{t=2}^{n}X_{t-1}U_4(t)-\sum_{t=2}^{n}U_2(t)\right)\Bigg],\end{aligned}$$

where

$$\begin{aligned}U_1(t) &= (1-\alpha)\left(2(X_t-\alpha X_{t-1}+\alpha(\hat{\mu}_{cls}+\mu))+1\right)-2\alpha^2,\\ U_2(t) &= -X_t^2+2\alpha X_tX_{t-1}-\alpha^2X_{t-1}+\alpha(1+\alpha)X_t+1-\alpha-2\alpha^2+2(\hat{\mu}_{cls}+\mu)(1-\alpha)(X_t+\alpha X_{t-1})\\ &\quad +2\alpha(1-\alpha)(\hat{\mu}_{cls}^2+\mu\hat{\mu}_{cls}+\mu^2)\\ U_3(t) &= X_t^2-2\alpha X_tX_{t-1}+\alpha^2X_{t-1}-\alpha(1+\alpha)X_t-2\alpha(1-\alpha)\mu\hat{\mu}_{cls}\\ U_4(t) &= (\hat{\mu}_{cls}-\mu)(X_t^2-2\alpha X_tX_{t-1}+\alpha^2X_{t-1}-\alpha(1+\alpha)X_t)-\mu\hat{\mu}_{cls}(2(1-\alpha)(X_t+X_{t-1})+1-\alpha-2\alpha^2).\end{aligned}$$

Using the introduced notation, we get the following equation

$$
\begin{aligned}
(n-1)^{\frac{1}{2}}\hat{p}_{cls}(\alpha_{cls},\hat{\mu}_{cls})-\hat{p}_{cls}(\alpha,\mu) \quad = \quad & \frac{(n-1)^{\frac{1}{2}}(\hat{\mu}_{cls}-\mu)}{2\alpha^2\left(\frac{1}{n-1}\sum_{t=2}^{n}(X_{t-1}-\hat{\mu}_{cls})^2\right)\left(\frac{1}{n-1}\sum_{t=2}^{n}(X_{t-1}-\mu)^2\right)}\\
& \times\Bigg[\frac{1}{n-1}\sum_{t=1}^{n}X_{t-1}^2\left(\frac{1}{n-1}\sum_{t=2}X_{t-1}U_1(t)-\frac{1}{n-1}\sum_{t=2}^{n}U_2(t)\right)\\
& -\frac{2}{n-1}\sum_{t=2}^{n}X_{t-1}\left(\frac{1}{n-1}\sum_{t=2}^{n}X_{t-1}U_3(t)-\frac{\mu\hat{\mu}_{cls}}{n-1}\sum_{t=2}^{n}U_1(t)\right)\\
& +\frac{1}{n-1}\sum_{t=2}^{n}X_{t-1}U_4(t)-\frac{1}{n-1}\sum_{t=2}^{n}U_2(t)\Bigg]
\end{aligned}
$$

As in the previously considered case, using the fact that $\{X_t, t\in\mathbb{Z}\}$ is a strictly stationary and ergodic sequence, all sequences obtained as its continuous transformations are also strictly stationary and ergodic sequences. Notice that all the relevant sequences used in the preceding relation are in fact continuous transformations of $\{X_t, t\in\mathbb{Z}\}$. So, the corresponding sums are convergent with probability 1, based on the ergodic theorem. Since $(n-1)^{\frac{1}{2}}(\hat{\mu}_{cls}-\mu)$ has asymptotic normal distribution with mean 0, as derived in Theorem 4.1, using Slutsky's theorem once more, we establish that $(n-1)^{\frac{1}{2}}\,(\hat{p}_{cls}(\alpha_{cls},\hat{\mu}_{cls})-\hat{p}_{cls}(\alpha,\mu))$ has also asymptotic normal distribution with mean 0.

Finally, having in mind the decomposition (9) we have begun with and the fact that we have established asymptotic normality with 0 mean for all the summands of (9), we may conclude that the conditional least square estimator of the unknown parameter p, when the parameters α and μ are also unknown, has asymptotic normal distribution. □

4.2. Yule-Walker estimation

In this section we provide Yule-Walker estimators and derive some of their features. As mentioned earlier, $E(X_t)=\mu$ and $Corr(X_t,X_{t-1})=\alpha$. It directly follows that parameters μ and α can be estimated using sample mean and sample autocorrelation, respectively. Thus,

$$
\begin{aligned}
\hat{\mu}_{yw} &= \frac{1}{n}\sum_{t=1}^{n}X_t\equiv\bar{X}_n,\\
\hat{\alpha}_{yw} &= \frac{\sum_{t=2}^{n}(X_t-\bar{X}_n)(X_{t-1}-\bar{X}_n)}{\sum_{t=1}^{n}(X_t-\bar{X}_n)^2}.
\end{aligned}
$$

In order to derive Yule-Walker estimator of the unknown parameter p, we need the following auxiliary result that can be obtained after some simple calculations. It follows that

$$
\begin{aligned}
E\left(\alpha\bullet_p X_{t-1}\right)^2 &= E(X_{t-1})Var(W)+E(X_{t-1}^2)(E(W))^2,\\
E\left(\left(\alpha\bullet_p X_{t-1}\right)^2X_{t-1}\right) &= E(X_{t-1}^2)Var(W)+E(X_{t-1}^3)(E(W))^2.
\end{aligned}
$$

Using the above results we obtain that

$$
\begin{aligned}
Cov(X_t^2,X_{t-1}) &= Cov\left(\left(\alpha\bullet_p X_{t-1}\right)^2,X_{t-1}\right)+2Cov\Big(\left(\alpha\bullet_p X_{t-1}\right)\varepsilon_t,X_{t-1}\Big)\\
&= Var(X_{t-1})Var(W)+(E(W))^2(E(X_{t-1}^3)-E(X_{t-1}^2)E(X_{t-1}))+2E(\varepsilon_t)E(W)Var(X_{t-1})\\
&= -2p\alpha^2\mu(1+\mu)+y,
\end{aligned}
$$

where $y = \alpha\mu\left((1+\mu)(1+\alpha+2\mu(1-\alpha)) + \alpha(1+5\mu+4\mu^2)\right)$. Finally, we have the required estimator as

$$\hat{p}_{yw} = \frac{\hat{y}_{yw} - \widehat{Cov}_{yw}(X_t^2, X_{t-1})}{2\hat{\alpha}_{yw}^2\hat{\mu}_{yw}(1+\hat{\mu}_{yw})},$$

where

$$\begin{aligned}\widehat{Cov}_{yw}(X_t^2, X_{t-1}) &= \frac{1}{n-1}\sum_{t=2}^{n} X_t^2 X_{t-1} - \left(\frac{1}{n-1}\sum_{t=2}^{n} X_t^2\right)\left(\frac{1}{n-1}\sum_{t=2}^{n} X_{t-1}\right),\\ \hat{y}_{yw} &= \hat{\alpha}_{yw}\hat{\mu}_{yw}\left((1+\hat{\mu}_{yw})(1+\hat{\alpha}_{yw}+2\hat{\mu}_{yw}(1-\hat{\alpha}_{yw})) + \hat{\alpha}_{yw}(1+5\hat{\mu}_{yw}+4\hat{\mu}_{yw}^2)\right).\end{aligned}$$

Again, as in the previous section, strong consistency of the obtained estimates follows directly from the ergodic theorem. In addition we give useful result. The proof is straightforward and is omitted here but can be found in [18], since the forms of the relevant estimators are the same.

Theorem 4.4. *Yule-Walker and the conditional least squares estimators of the unknown parameters α and μ satisfy the following equations:*

(i) $\hat{\alpha}_{cls} - \hat{\alpha}_{cls} = o(n^{-\frac{1}{2}})$ *with probability 1*

(ii) $\hat{\mu}_{cls} - \hat{\mu}_{cls} = o(n^{-\frac{1}{2}})$ *with probability 1*

Using the result of this theorem together with the Proposition 6.3.3. of [7], we conclude that the asymptotic distribution of Yule-Walker estimators of the unknown parameters α and μ is the same as the one obtained in Theorem 4.1 for the conditional least squares estimators.

4.3. Maximum likelihood estimation

Using the additional information on the process marginal distribution, i.e. starting from the sample of the observed process introduced by Definition 3.3, we are able to derive the maximum likelihood estimators of the unknown model parameters μ, α and p. As usually, these estimators are obtained by maximizing the log-likelihood function

$$\log L(x_1,\ldots,x_N;\mu,\alpha,p) = x_1\log\mu - (x_1+1)\log(1+\mu) + \sum_{t=2}^{N}\log P_t(x_t,x_{t-1};\mu,\alpha,p),$$

where $P_t(x_t,x_{t-1};\mu,\alpha,p) = P(X_t = x_t|X_{t-1} = x_{t-1})$, which is given by (6) and (8). Since the maximum likelihood estimators can not be easily obtained as analytical solution of the the system of equations $\frac{\partial \log L}{\partial\mu} = 0$, $\frac{\partial \log L}{\partial\alpha} = 0$ and $\frac{\partial \log L}{\partial p} = 0$, these estimators are rather obtained by using some of the numerical minimizing functions implemented in most of the statistical software packages.

5. Application

In this section we apply a geometrically distributed integer-valued autoregressive process based on a newly introduced mixed thinning operator on the counting data series from real life. In order to demonstrate an appealing quality or even an advantage of usage of this new thinning operator over the exclusive application of only one of its building operators, i.e. the binomial and the negative binomial thinning operator, we compare here the introduced Mixed Thinning Geometrically distributed INAR(1) model (MTGINAR(1)) with Geometric INAR(1) model (GINAR(1)) defined in [4] and New Geometric INAR(1) model (NGINAR(1)) of [17]. For this purpose, we considered a data series representing a monthly counting of committing a light criminal activity, public drunkenness, in a period from January 1990 to December 2001, constituting a sequence of 144 observations. This data set, (PubDrunk-22) is created by the 22*nd* police car beat of Pittsburgh and can be downloaded from the website Forecasting Principles (http://www.forecastingprinciples.com). The sample mean, variance and autocorrelation of the PubDrunk-22 are respectively, 1.34, 10.1 and 0.761. The plots of the time series, their autocorrelations and partial autocorrelation functions are given in Figure 1.

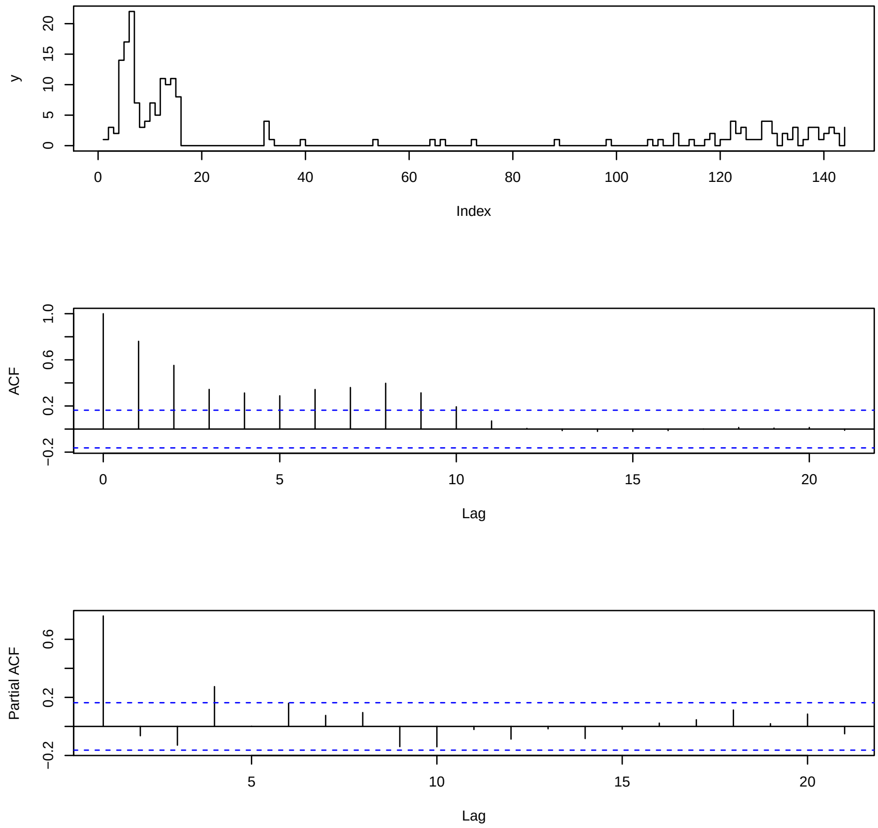

Figure 1: PubDrunk-22 series, autocorrelations and partial autocorrelations

Based on this figure, we conclude that it is possible to use the corresponding INAR(1) modeling. Although, some of the higher order INAR models might have been a preferable choice, we have not considered their usage, since in this article we are only dealing with the introduction of the INAR model of the first order, based on a newly defined mixed thinning operator. To carry out the quality comparison of our MTGINAR(1) model with GINAR(1) and NGINAR(1), first we have calculated the maximum likelihood estimations of the parameters of the considered models and then using these estimations we have obtained as a comparison criteria the Root Mean Squares of differences between the observations and predicted values (RMS). The RMS is used as a standard quality measure of the model fitting performance, expecting smaller values for more appropriate models. All the results are presented in Table 1.

Table 1: ML parameter estimates and RMS values for INAR(1) modeling of the PubDrunk-22 counts data.

Model	MLE	RMS
GINAR(1)	$\hat{q} = 0.6030(0.0434)$	
	$\hat{\alpha} = 0.4800(0.0466)$	2.248
NGINAR(1)	$\hat{\mu} = 1.3729(0.2958)$	
	$\hat{\alpha} = 0.5720(0.0525)$	2.146
MTGINAR(1)	$\hat{\mu} = 1.4544(0.3425)$	
	$\hat{p} = 0.2234(0.1644)$	
	$\hat{\alpha} = 0.6167(0.0638)$	2.111

Obvious conclusion, based on the RMS values, is that the MTGINAR(1) model shows better performance than the other two models in modeling the observed counting. Moreover, in case of the mixed thinning model, the obtained probability estimate $\hat{p}$ is much more closer to its limit value 0, than to the 1, meaning that the negative binomial thinning operator is much more often realized than the binomial one. In this regard, for the observed series, NGINAR(1) model, based on the negative binomial thinning, is much more appropriate than the GINAR(1), while our MTGINAR(1) is notably the most persuasive among all three observed models. The best results of newly introduced mixed thinning model can be explained by the nature of the observed events. Namely, public drunkenness is an offence which may happen occasionally by one or two persons with no special reason, but when there are some social events, like street festivals, sport matches, etc., it would be expected to see more drunk persons in the streets than usual. Therefore, if these events occur frequently, then the value of parameter p should be smaller, since it is a binomial thinning weight which is more preferable to use during calmer period of time. So, it can be concluded that the mixed thinning operator is designed for exactly this kind of counting processes, i.e. processes which from time to time, depending on social circumstances, move from passive to active state and back. This might be noticed in time series plot in Figure 1, where during approximately the first and the last 40 months we have a more frequent occurrence of crime, than in between, giving sufficient reasons for the successful implementation of MTGINAR(1) model to this type of data.

References

[1] Al-Osh, M.A. and Aly, E.E.A.A. (1992) First order autoregressive time series with negative binomial and geometric marginals, Commun. Statist. Theory Meth. 21, 2483–2492.

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

[3] Aly, E.E.A.A., Bouzar, N. (1994) On Some Integer-Valued Autoregressive Moving Average Models, Journal of Multivariate Analysis 50, 132–151.

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

[5] Alzaid, A. A., Al-Osh, M. A. (1993). Some autoregressive moving average processes with generalized Poisson marginal distributions. Ann. Inst. Statist. Math. 45, 223–232.

[6] Bakouch, H.S., Ristić, M.M. (2010) Zero Truncated Poisson Integer Valued AR(1) Model, Metrika 72(2), 265-280.

[7] Brockwell, P. J., Davis, R. A. (1987). Time series: Theory and Methods. New York: Springer-Verlag

[8] Cox, D.R. Miller, H.D. (1965) The Theory of Stochastic Processes, London: Methuen.

[9] Grunwald, G. K., Hyndman, R. J., Tedesco, L., Tweedie, R. L. (2000) Theory and Methods: Non-Gaussian Conditional Linear AR(1) Models, Australian and New Zealand Journal of Statistics 42(4), 479–495.

[10] Jacobs, P.A., Lewis, P.A.W. (1978a) Discrete Time Series generated by Mixtures I: Correlation and Runs Properties. J.R. Statist. Soc., (B) 40, 94–105.

[11] Jacobs, P.A., Lewis, P.A.W. (1978b) Discrete Time Series generated by Mixtures II: Asymptotic Properties. J.R. Statist. Soc., (B) 40, 222–228.

[12] Jacobs, P.A., Lewis, P.A.W. (1978c) Discrete Time Series generated by Mixtures III: Autoregressive Processes (DAR(p)). Naval Postgraduate School Technical Report, NPS55Lw 73061A.

[13] Karlsen, H., Tjøstheim, D. (1988) Consistent Estimates for the NEAR(2) and NLAR(2) Time Series Models, Journal of the Royal Statistical Society B, 313–320.

[14] Latour, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive process. J. Time Ser. Anal. 19, 439–455.

[15] McKenzie, E. (1985) Some simple models for discrete variate time series, Water Resour. Bull. 21, 645–650.
[16] McKenzie, E. (1986) Autoregressive moving-average processes with negative binomial and geometric distributions, Adv. Appl. Prob. 18, 679–705.
[17] Ristić, M. M., Bakouch, H. S., Nastić, A. S. (2009) A new geometric first-order integer-valued autoregressive (NGINAR(1)) process, J. Statist Plann. Infer. 139, 2218–2226.
[18] Ristić, M. M., Nastić, A. S., and Bakouch, H. S. (2012). Estimation in an integer-valued autoregressive process with negative binomial marginals (NBINAR (1)). Communications in Statistics-Theory and Methods, 41(4), 606-618.
[19] Ristić, M. M., Nastić, A. S. (2012) A mixed INAR(p) model, Journal of Time Series Analysis, 33, 903–915.
[20] Tjøstheim, D. (1986) Estimation in nonlinear time series models, Stochastic Processes and their Applications 21(2), 251–273.
[21] Zheng, H., Basawa, I.V., Datta, S. (2006) Inference for pth-order random coefficient integer-valued autoregressive processes, J. Time Ser. Anal. 27, 411–440.
[22] Zheng, H., Basawa, I.V.,Datta, S. (2007) First-order random coefficient integer-valued autoregressive processes, J. Stat. Plann. Inf. 137, 212–229.